Deliverable D1.3

Design for systems and interfaces for slice operation v1

Editor:	Jorge Carapinha, Altice Labs
Deliverable nature:	Report (R)
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	30/06/2019
Actual delivery date:	06 July 2019
Suggested readers:	ICT-19 projects, industry verticals seeking to conduct Testing on 5G systems
Version:	1.0
Total number of pages:	90
Keywords:	5G, Network Slicing, Management, Orchestration

Abstract

Network slicing is one of the main differentiating and innovative features of 5G by enabling the partitioning of the network infrastructure to support specific use cases, services, individual customers or vertical industries. This feature will be implemented, evaluated and validated by the 5G-VINNI end-to-end Facility though the deployment of a wide range of 5G trials, according to the requirements defined by end users. This document is focused on operation, management and orchestration of network slicing and draws upon the current state of the art in standardization and R&D projects, as well as two previous 5G-VINNI deliverables, D1.1 and D1.2.

[End of abstract]

Disclaimer

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 815279.

Impressum

Full project title	5G Verticals Innovation Infrastructure
Project acronym	5G-VINNI
Number and title of work-package	WP1 Architecture and Design of 5G-VINNI End-to-End Platform
Number and title of task(s)	T1.3 Design of a system to support management and orchestration of slices, and slice operations (by a slice instance owner)
Document title	Design for systems and interfaces for slice operation v1
Editor: Name, company	Jorge Carapinha, Altice Labs
Work-package leader: Name, company	Dan Warren, Samsung

Copyright notice

Executive summary

The 5G-VINNI project is aimed at developing an advanced 5G E2E facility that is able to validate 5G KPIs, supporting the execution of vertical use case trials, demonstrating the value of 5G solutions and ultimately fostering the widespread adoption of 5G technologies. This document is the third in a series of documents that are expected to lay the foundations of the 5G-VINNI E2E facility design, based on the latest available 5G technologies including network slicing, core network, radio access, backhaul/fronthaul, virtualization infrastructure (NFV, SDN), service orchestration and other relevant 5G components. It draws upon two previous 5G-VINNI deliverables – D1.1 “Design of infrastructure architecture and subsystems v1” and D1.2 “Design of network slicing and supporting systems v1”.

This document addresses the operation, management and orchestration of network slicing, which constitutes one of main differentiating and innovative features of 5G by enabling the partitioning of the network infrastructure to support specific use cases, services, individual customers or vertical industries. In addition, the network slicing impacts and requirements are analysed in relation to the several 5G infrastructure components. Apart from the 5G stringent requirements to accomplish 5G target KPIs, the deployment of network slicing, on its own, raises a number of new challenges to operators and service providers (e.g. isolation) which are relevant to all network infrastructure segments (e.g. radio, transport, core) and in some cases call for new solutions.

A key target of this document are the internal 5G-VINNI activities to be conducted with a view to materializing and operationalizing 5G network slicing at the facility sites. Nonetheless, the information contained in this document should be also of interest to network architects and engineers aiming at deploying a 5G experimentation facility for the purposes of technological evaluation and validation.

To a large extent, 5G-VINNI will build on existing solutions, components and architectural frameworks produced by open source communities, standardization bodies and industry initiatives. An important part of this document is devoted to the description of the state of the art and the identification of relevant results related to 5G network slicing, which should be taken into account in the context of 5G-VINNI. This includes aspects such as management, orchestration, interfaces, integration of multiple domains (both technological and administrative).

An evolution roadmap is required to enable to gradual implementation of 5G innovative features, including network slicing. In this respect, the 3-phase evolution plan outlined by 5G-VINNI (i.e. 4G slicing, 5G slicing, Network slice as a Service) is generally in line with the industry and standardization roadmap.

The lessons learned from the practical implementation, deployment and operation in 5G-VINNI facility sites will be reflected in the second iteration of this document, to be produced in month 30.

List of authors

Author	Company / Affiliation
Adrián Gallego Sánchez	Universidad Carlos III de Madrid
Ahmed Elmokashfi	Simula
Andres J. Gonzalez	Telenor
Carlos Parada	Altice Labs
Carmen Guerrero	Universidad Carlos III de Madrid
Christos Politis	SES
Christos Tranoris	University of Patras
Dimitrios Kritharidis	ICOM
Foivos Michelinakis	Simula
Ishan Vaishnavi	Huawei
João Rodrigues	Nokia
Jorge Carapinha	Altice Labs
Jose Ordonez-Lucena	Telefonica
Konstantinos Liolis	SES
Konstantinos Stamatis	Intracom Telecom
Konstantinos V. Katsaros	Intracom Telecom
Pål Grønsund	Telenor
Robert Lagerholm	Ericsson
Vasileios Theodorou	Intracom Telecom
Wint Yi Poe	Huawei
Yue Wang	Samsung

Table of Contents

List of Figures

List of Tables

Abbreviations

3GPP	3rd Generation Partnership Project
5G	Fifth Generation (mobile/cellular networks)
5G PPP	5G Infrastructure Public Private Partnership
5QI	5G QoS Indicator
AAI	Active and Available Inventory
AF	Application Function
AGF	Access Gateway Function
AMF	Access and Mobility Management Function
AN	Access Network
API	Application Programming Interface
APN	Access Point Name
BBF	Broadband Forum
BSS	Business Support System
CAPEX	Capital Expenditure
CFS	Customer Facing Service
CN	Core Network
CP	Control Plane
CPRI	Common Public Radio Interface
C-RAN	Cloud-RAN (also referred to as Centralized-RAN)
CRUD	Create, Read, Update, Delete
CSC	Communication Service Customer
CSMF	Communication Service Management Function
CSP	Communication Service Provider
DECOR	Dedicated Core
DetNet	Deterministic Networking
DRB	Data Radio Bearer
DSCP	Differentiated Services Code Point
E2E	End-to-End
eDECOR	Enhanced Dedicated Core
eMBB	enhanced Mobile Broadband
EMS	Element Management System
EN-DC	E-UTRA-NR Dual Connectivity
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
FCAPS	Fault, Configuration, Accounting, Performance and Security
FEC	Forwarding Equivalence Class
FW	Firewall
GEO	Geostationary Orbit
gNB	next generation NodeB
GST	General Service Template
GTP	GPRS Tunnelling Protocol
GW	Gateway
HAPS	High Altitude Platform Systems
HLR	Home Location Register
HSS	Home Subscriber Service
ICM	Infrastructure Control and Management
IETF	The Internet Engineering Task Force

IM	Information Model
IoT	Internet of Things
IP	Internet Protocol
IPS	Interface Profile Specification
KPI	Key Performance Indicator
LAN	Local Area Network
LCM	Lifecycle Management
LSO	Lifecycle Services Orchestration
LTE	Long Term Evolution
MANO	Management and Orchestration
MEC	Multi-access Edge Computing
MEF	Metro Ethernet Forum
mIoT	massive Internet of Things
MME	Mobility Management Entity
mMTC	massive Machine Type Communications
MNO	Mobile Network Operator
MPLS	Multiprotocol Label Switching
MPLS-TE	Multiprotocol Label Switching - Traffic Engineering
MQTT	Message Queuing Telemetry Transport
MTNSI	Mobile-Transport Network Slice Interface
NBI	Northbound Interface
NEF	Network Exposure Function
NEST	Network Slicing Task Force
NF	Network Function
NFV	Network Function Virtualization
NFVI	Network Functions Virtualization Infrastructure
NFVO	Network Function Virtualization Orchestrator
NGMN	Next Generation Mobile Networks
NMS	Network Management System
NOP	Network Operator
NR	New Radio
NRF	NF Repository Function
NSA	Non-Standalone
NSaaS	Network Slicing as a Service
NSD	Network Service Descriptor
NSE	Network Slicing Engine
NSI	Network Slice Instance
NSI-ID	NSI Identifier
NSMF	Network Slice Management Function
NSSAI	Network Slice Selection Assistance Information
NSSF	Network Slice Selection Function
NSSI	Network Slice Subnet Instance
NSSMF	Network Slice Subnet Management Function
NSST	Network Slice Subnet Template
NST	Network Slice Template
NTN	Non-Terrestrial Network
OAM	Operations, Administration and Management
ONAP	Open Network Automation Platform
OPEX	Operational Expenditure
OSM	Open Source MANO

OSS/BSS	Operations Support System /Business Support System
PCEP	Path Computation Element Communication Protocol
PCF	Policy Control Function
PCRF	Policy and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PHB	Per-Hop Behaviour
PLMN	Public Land Mobile Network
PNF	Physical Network Function
PoC	Proof of Concept
PoP	Point of Presence
QCI	QoS Class Identifier
QoE	Quality of Experience
QoS	Quality of Service
R&D	Research and Development
RAN	Radio Access Network
RAT	Radio Access Technology
REST	Representational State Transfer
RFS	Resource Facing Service
RG	Residential Gateway
RO	Resource Orchestration
SatCom	Satellite Communication
SD	Slice Differentiator
SDC	Service Design and Creation
SDK	Software Development Kit
SDM-C	Software-Defined Mobile Network Controller
SDM-O	Software-Defined Mobile Network Orchestrator
SDM-X	Software Defined Mobile Network Coordinator
SDN	Software Defined Network
SDO	Standards Developing Organization
SD-WAN	Software Defined Wide Area Network
SGSN	Serving GPRS Support Node
SLA	Service Level Agreement
SMF	Session Management Function
SNMP	Simple Network Management Protocol
SNO	Satellite Network Operator
S-NSSAI	Single-NSSAI
SO	Service Orchestrator
SO&M	Service Operations and Management
SP	Service Platform
SPID	Service Profile Identifier
SST	Slice/Service Type
TALENT	Terrestrial Satellite Resource Coordinator
TDM	Time Division Multiplexing
TMF	TeleManagement Forum
TN	Transport Network
TOSCA	Topology and Orchestration Specification for Cloud Applications
T-SDN	Transport SDN
UDM	Unified Data Management
UE	User Equipment
UE-AMBR	UE Aggregated Maximum Bit Rate(for non-GBR traffic)

UML	Unified Modeling Language
uRLLC	ultra Reliable Low Latency Communications
V&V	Verification & Validation
VIM	Virtualized Infrastructure Manager
VIoTIM	Virtualized IoT Infrastructure Management
VLAN	Virtual LAN
VNF	Virtual Network Function
VNFD	VNF Descriptor
VPN	Virtual Private Network
VPN+	Enhanced Virtual Private Network
vPP	virtual Packet Processor
VS	Vertical Slicer
VSB	Vertical Service Blueprint
VSD	Vertical Service Descriptor
VTN	Virtual Tenant Network
WID	Work Item Description
WIM	WAN Infrastructure Manager
ZSM	Zero-touch network and Service Management

1 Introduction

1.1 Objective of this document

This document constitutes the first deliverable of 5G-VINNI Task 1.3 “Design of a system to support management and orchestration of slices, and slice operations (by a slice instance owner)”. According to the project description [1], the objective of T1.3 is to design a supporting system for slice operations and service orchestration, including aspects such as intra/inter-slice management, multi-domain, and relevant APIs/protocols.

It should be noted that this deliverable reflects the vision of 5G-VINNI at the end of project month 12. Lessons learned from the implementation of network slicing-based 5G services, as well as the evolution of relevant standards and ongoing activities in this area may lead to updates or changes, which will be reflected in the second release of this document, due in project month 30.

1.2 Scope and relationship with other deliverables

The present document is closely related with two previous 5G-VINNI deliverables:

- D1.1 “Design of infrastructure architecture and subsystems v1” [2], which defines the architecture for the overall 5G-VINNI facility, as well as for the generalised 5G-VINNI site, based on requirements for support of the 5G-VINNI Release 0, 1, 2 and 3 implementations. D1.3 builds on D1.1 as starting point, but is mainly focused on network slicing, especially from a management and orchestration perspective.
- D1.2 “Design of network slicing and supporting systems v1” [3], which describes the network slicing design and supporting systems for 5G-VINNI Facility Sites, based on 3GPP specifications and other standardization bodies. D1.2 is primarily concerned about architectural issues, whereas D1.3 is focused on management and orchestration. Furthermore, in relation to some topics, there is a relationship of complementarity between the two documents. Whenever applicable, the relevant D1.2 reference is indicated, in order to minimize overlaps and misalignments between the two documents. In addition, D1.3 addresses slicing-related topics that were not covered in D1.2 – section 6, focused on mMTC slicing, is an example.

In addition, the present document bears a very close relationship with 5G-VINNI deliverable D3.1 [4]. Management and orchestration of network slice-based services is a key topic in both documents, but D1.3 is focused on the resource-facing issues, whereas D3.1 is mainly related to the customer-facing perspective.

1.3 Document structure

The rest of the document is structured as follows:

- Section 2 provides an overview of the state of the art and pre-existing work in areas related to the 5G network slicing, including standardization efforts, industry initiatives and open source platforms.
- Section 3 is focused on network slice lifecycle management and orchestration and provides the basic 5G-VINNI slice design principles.
- Section 4 analyses the requirements to build network slices, which are applicable to the different network segments and technologies.
- Section 5 addresses the management interfaces and APIs, including interfaces for slice lifecycle management, interfaces for intra-slice management and inter-domain management. The network slicing information model is also covered.

- Section 6 is focused on mMTC slicing and describes different slicing models derived by varying degrees of integration of the actors of the ecosystem.
- Finally, section 7 provides an evolution roadmap for network slicing in the context of the deployment of 5G-VINNI infrastructure.
- Additionally, Annex A provides an overview of three 5G-PPP Phase 2 projects, considered relevant to 5G-VINNI, especially in relation to the topics of network slicing, management and orchestration and 5G/satellite integration.

2 State of the Art

2.1 Introduction

In section 2, a general overview of potentially relevant activities and results from standardization, industry alliances, open source communities and R&D projects is provided.

In terms of 5G standardization, and specifically 5G network slicing, 3GPP should be seen as the main reference, having produced a general framework for management and orchestration and defined the main management functions. Other Standards Developing Organizations (SDOs) play a relevant role, including ETSI (anything related to virtualization of network functions), BBF (for transport related matters), TM Forum and MEF (in relation to inter-domain interfaces).

A selection of open source platforms are analysed in this section, some of which are already planned to be incorporated in 5G-VINNI facility sites (e.g. OSM, OpenBaton, SONATA), or under evaluation (e.g. ONAP). Details on the deployment of these tools in 5G-VINNI facility sites can be found in other 5G-VINNI deliverables, including D2.1 [5].

Likewise, relevant 5G-PPP phase 2 projects have been monitored, with a view to evaluating the potential applicability of components or architectural models. For reasons of economy of space, only three projects, considered especially relevant, are described in this section – 5G-TRANSFORMER, 5GTANGO, SaT5G.

Industry initiatives such as NGMN and 5G-PPP are also analysed, especially the respective architectural frameworks, which should provide a solid foundation to 5G-VINNI. 5G slicing should serve a variety of different use cases, with very heterogeneous requirements.

It should be noted that this section is not intended to be exhaustive and only the most relevant (from a 5G-VINNI perspective) SDOs, industry initiatives, open source platforms and projects have been included in this analysis.

2.2 Standardization

2.2.1 3GPP SA5

3GPP SA5 is the key SDO that will drive the work in Task 1.3. It released R15 specifications towards the end of 2018. The primary set of specifications, TS 28.530 [6], TS 28.531 [7] and TS 28.532 [8], look into the stage 1 and stage 2 of the Network slicing specification. TS 28.530 presents the business use cases and requirements while TS 28.531 further specifies technical use cases, requirements and initial set of solutions. TS 28.532 proposes solution in terms of the generic services required for provisioning, performance and fault management of the management architecture. TS 28.533 explains the concepts and theory behind the construction management services detailing the possible architectural frameworks in which the other management services can be used.

The key new concept relating to slicing in R15 is the ability to *provision slices* (TS 28.531) as opposed to configuration management of hardware prior to R14. The provisioning (as well as all other services) exist in three main hierarchies: the Network Function (NF) level, either Physical or Virtual (i.e. VNF/PNF), the Network Slice Subnet Instance (NSSI) level and the Network Slice Instance (NSI) level. The NSSI level is recursive in nature. An NSSI could be composed of multiple other NSSI. This is represented in the information models in TS 28.540 [9] and TS 28.541 [10]required to support network slicing and other management services. The key information model to understand slicing is shown in Unified Modeling Language (UML) diagram in Figure 2-1 [11].

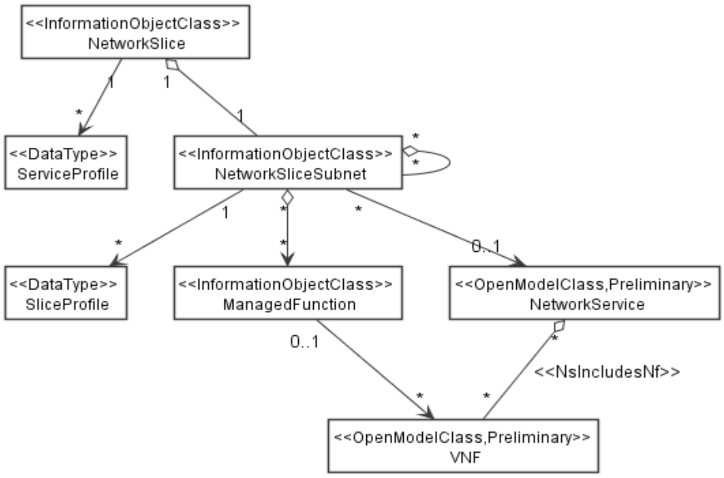

Figure 2-1: UML diagram showing the network slice, network slice subnet and NF hierarchies

In addition, the UML diagram also shows the relationship to the network functions (via ManagedFunction class) defined in the 3GPP SA2 specification (Control plane architecture) [12], as well as the ETSI NFV (via the NetworkService) class.

Other relevant services are specified in performance assurance (TS 28.550 [13]) and Fault supervision (TS 28.545 [14]). These services are based on the measurements and Key Performance Indicators (KPIs) specified in TS 28.552 [15] and TS 28.554 [16], respectively.

In order to manage Network Slicing entities, 3GPP defines three basic management functions in TR 28.801 [17]. Figure 2-2 illustrates the relationship of management functions and managed entities defined by 3GPP.

- Communication Service Management Function (**CSMF**): This function manages Communication Services, translates the communication service related requirements to network slice related requirements and communicates with the NSMF.
- Network Slice Management Function (**NSMF**): This function manages and orchestrates Network Slice Instances (NSI), derives network slice subnet related requirements from network slice related requirements and communicates with the NSSMF and CSMF.
- Network Slice Subnet Management Function (**NSSMF**): This function manages and orchestrates the NSSI and communicate with the NSMF.

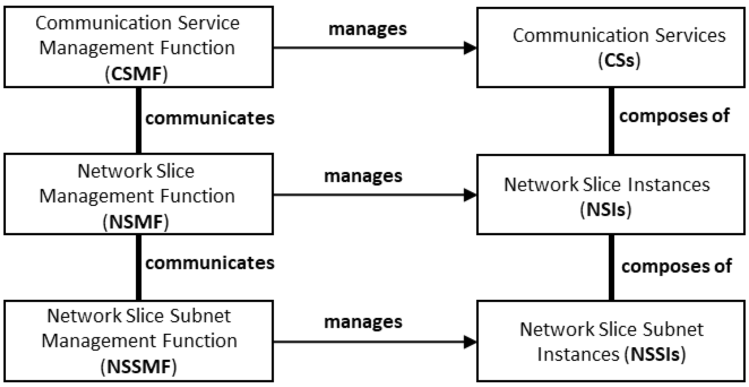

Figure 2-2: 3GPP defined Network Slicing management functions and managed entities

A new 3GPP Release 16 SA5 Study Item "Study on Management and Orchestration aspects with integrated satellite components in a 5G network" (Reference: "FS_5G_SAT_MO") was approved in 3GPP TSG-SA5 #123 in January 2019 (SP 190138). The results of this ongoing study item are reflected in 3GPP TR 28.808 [18].

Its objective is to identify the main key issues associated with business roles, service and network management and orchestration of a 5G network with integrated satellite element(s) (whether as NG-RAN or non-3GPP access, or for transport) and to study the associated solutions. This study aims at minimising the impacts and the complexity of the satellite integration in the existing business model, management and orchestration of the current 5G network.

At the time of writing, the following three use cases have been defined in the associated 3GPP TR 28.808, whereas the next steps are to define the key issues associated to these use cases.

- Network slicing with Satellite elements ,
- Network slicing with Satellite RAN sharing
- Network Slicing and handling of latencies of satellite latencies.

2.2.2 ETSI NFV

The reference architectural framework originally specified by the ETSI ISG NFV in ETSI GS NFV-MAN 001 [19], and more specifically the NFV Management and Orchestration (MANO) stack, provides architectural foundations that allow consistency and uniformity for deployment and operation of network services in any virtualization environment. As a great part of the components of a network slice can be virtualized, it is widely accepted in the research community that both information model and functional blocks taking part in NFV MANO can be effectively reused and (when needed) extended for the management and orchestration of those components, thus enabling network slicing support. To study and assess how network slicing can benefit from the capabilities provided by NFV MANO, the Evolution and Ecosystem (EVE) working group initiated a work item with a twofold purpose: the analysis of use cases related to network slicing as defined in SDOs and industry fora, and the description of how these use cases can be mapped by the NFV MANO. The ideas resulting from this work item are summarized in ETSI NFV-EVE 012 [20]. Although visions on network slicing from NGMN and ONF are discussed in this technical report, their relevance is significantly lower than 3GPP, which is taken as the reference body for the definition and realization of the network slicing concept.

The main scope of ETSI NFV EVE 012 is to identify how NFV could assist the 3GPP in the management and orchestration of network slices, specifically when they are deployed over virtualized network infrastructures. For this end, two main issues arise. First, the conciliation of ETSI and 3GPP terminology in network slicing context, finding a match between a NFV network service and a 3GPP network slice. This will allow understanding the role that NFV plays in the deployment and operation activities that need to be done over 3GPP network slices. Secondly, the definition of a unified management framework where 3GPP management system (see Section 2.2) and NFV MANO can coexist, identifying potential reference points for their interplay. These two issues are addressed in the technical report.

ETSI states that 3GPP network slices can leverage the capabilities offered by network services to satisfy the network requirements of the communication services they may accommodate. Based upon this premise, ETSI NFV-EVE 012 establishes the correspondence between 3GPP conceptual outline on network slicing (i.e. network slice, network slice subnet, and network function) and the ETSI concepts (i.e. network service and V/PNFs), thus aligning terminology from both SDOs. As seen from Figure 2-3, ETSI claims that a network service can be regarded as the resource-facing view of a network slice subnet, for the cases where at least one of the network slice subnet's NFs is deployed as a VNF.

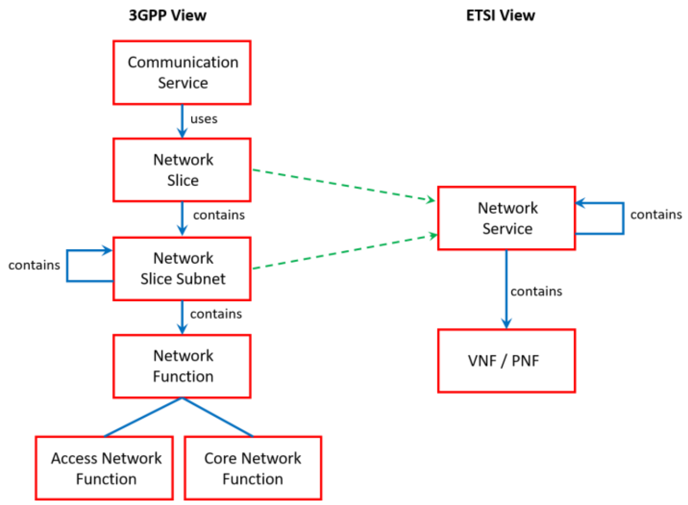

Figure 2-3: Relating 3GPP and ETSI models in the context of network slicing [20]

This means that a network slice may consist of one or more network services. The mapping between these two concepts depend on three main factors:

1. *The internal composition of the network slice*: the number and type of the network slice subnets that take part in the network slice.
2. *The internal composition of the network service*: a NFV network service could be simple or composite. A simple network service includes one or more VNFs/PNFs and virtual links providing connectivity across them. However, in search of modularity and recursion, the design of a network service can also include one or more nested network services, as defined in ETSI NFV 003 [21], resulting in a composite network service. In the example depicted in Figure 2-4, the composite network service consists of two VNFs and one simple network service.
3. *The scope of network service functionality*: the functionality of a network service could focus on a single network domain (i.e. service providing CN functionality), or span across various domains (i.e. service providing the entire End-to-End (E2E) functionality, covering Access Network (AN), Transport Network (TN), Core Network (CN).

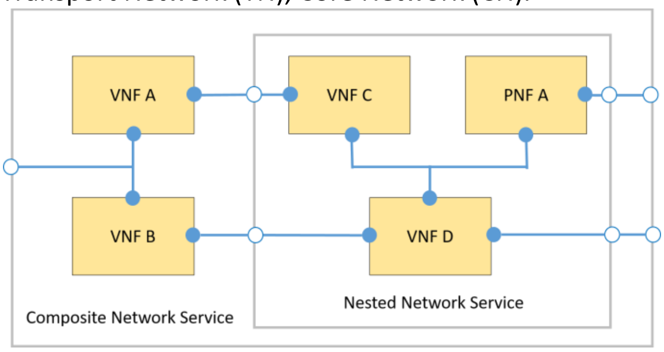

Figure 2-4: An example of a composite NFV network service (adapted from [22])

These three factors can be flexibly combined, resulting in multiple scenarios when instantiating a 3GPP network slice in NFV-ready environments. Indeed, an NSI can be deployed as an instance of a single or composite network services, or as a concatenation of any number of such instances.

The next step is to define a unified framework for network slicing management and orchestration, where both NFV MANO and 3GPP network slicing management system (see Section 2.2.1) can coexist. Their coexistence is necessary, due to their complementary scopes. On one side, NFV MANO deals with virtualization-specific management and orchestration issues at VNF and network service, including management activities over their descriptors, i.e. VNF Descriptors (VNFDs) and Network

Service Descriptors (NSDs), and performance/fault/lifecycle management operations over instances resulting from those descriptors. On the other hand, 3GPP management functions deal with application-aware management over entities at multiple levels (i.e. NFs, network slice subnet, network slices, communication services), without notions on how these entities are deployed in the virtualized infrastructure.

As seen from the above-referred description, 3GPP can handle all the activities that go beyond virtualization, and hence are out of scope of NFV MANO. These activities include (i) the deployment and operation of PNFs, (ii) VNF application layer configuration and management, and (iii) application-aware network service configuration and management. To carry out these activities, 3GPP network slicing management functions need to consume data and operations from the reference points exposed by NFV MANO stack, so an interaction between two frameworks are required. Taking into account the relationship between a 3GPP network slice and a network service as discussed earlier, ETSI NFV-EVE 012 [20] proposes in Figure 2-5 a first approach on how the intended interaction could take place, showing the reference points that can be potentially (re)used for this end: Ve-vnfm-em [23] and Os-Ma-Nfvo [24].

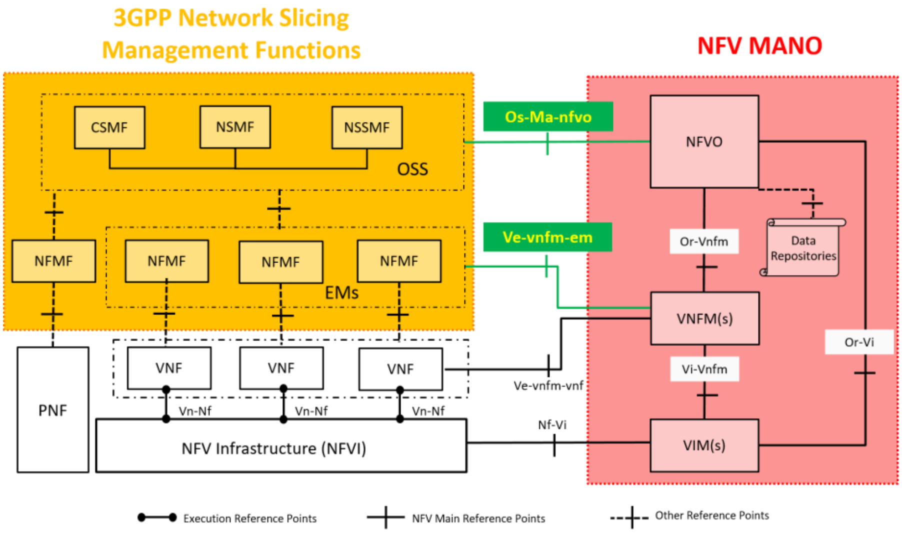

Figure 2-5: 3GPP network slicing management in NFV framework (adapted from [20])

On one hand, the NF Management Function (NFMF), taking the role of EM, makes use of the interfaces exposed in the Ve-Vnfm-em reference point to exchange information with the VNFM, including application-aware VNF performance and fault data. These data could also combined with VNF information at virtualized resource level (e.g. CPU usage) to trigger appropriate scaling operations or reconfiguration over one or more VNF instances of the NSSI/NSI.

On the other hand, NSSMF/NSMF (taking the role of OSS) makes use of the interfaces exposed in the Os-Ma-nfvo reference point to exchange information with the NFVO. The specific 3GPP network slicing management function that directly interfaces with the NFVO depends on the internal composition of the network slice. Considering the case where a network slice consists of more than one network slice subnet, it is the NSSMF the one that interfaces with the NFVO. Based on the NSSI resource requirements sent from the NSMF, the NSSMF may need to determine the type of network service that satisfy those requirements, maintaining an association between the Network Slice Subnet Template (NSST) and the NSD with applicable deployment flavour identifiers, as well as an association between NSSI and network service instance identifiers. Different NSSIs can use instances of the same type of network service (i.e. they are instantiated from the same NSD) with the same or different deployment flavours. Alternatively, different NSSIs can use instances of different types of network services. The first approach can be used if the NSSIs share the same types of NFs (or a large common subset) but differ in terms of the performance expected from these NFs and/or the number of instances to be deployed for each of them. If NSSIs differ more significantly, mapping to different network service instances, each with its own NSD, can be considered. Note that the same mapping principles might apply for the NSI.

2.2.3 ETSI ZSM

Zero touch network and Service Management (ZSM) is a next-generation management system, aimed at addressing the challenges introduced by the deployment of new network architectures and foundations such as NFV and 5G. ZSM leverages the NFV and SDN principles and is designed for the new cloud-based network infrastructures and functions. ZSM is based on cloud-native principles to accomplish zero-touch (fully automated) management and operation [25].

NV and 5G have triggered the need to accelerate network transformation and radically change the way networks and services are managed and orchestrated. Networks are being transformed into programmable, software-driven, service-based and holistically-managed infrastructures.

The ZSM architecture is shown in Figure 2-6 [25].

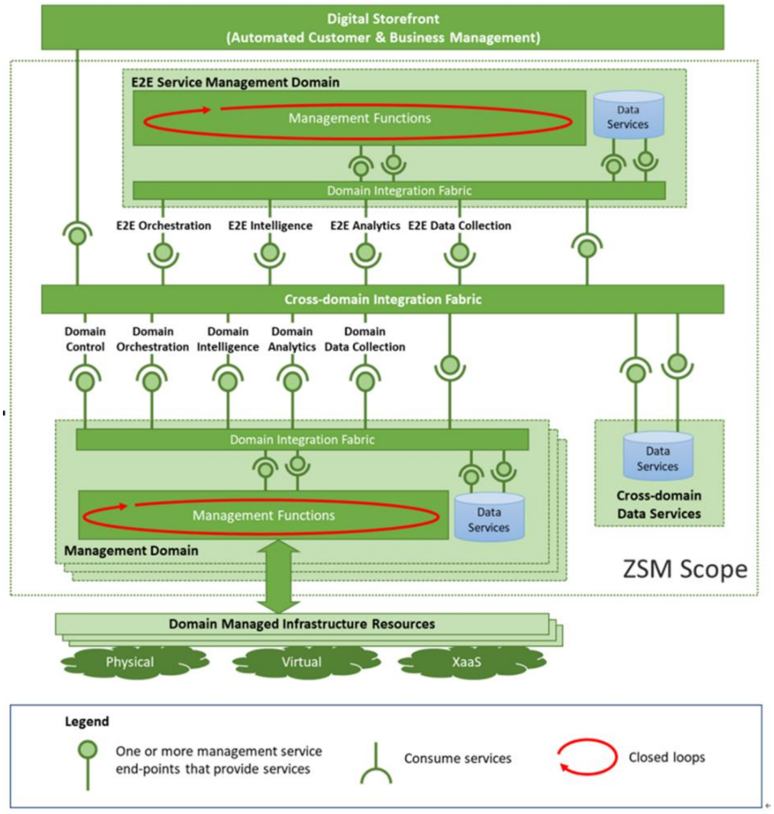

Figure 2-6: ZSM architecture proposes coordination across multiple domains [25]

ZSM proposes a framework of standard set of interfaces (management service) that are exposed externally by every management domain to enable inter-domain interaction and automation. A ZSM compliant domain then would simply need to translate those management services to the SDO specific management plane instructions within the domain. The E2E coordination of the slice instance is done by the E2E management domain.

2.2.4 ETSI SmartM2M

In order to realize mMTC slicing, the appropriate virtualization techniques and mechanisms need to employ, abstracting IoT infrastructural resources. In this respect, this subsection provides a brief description of the IoT Virtualization architecture defined in TR 103 527, “One High-Level Architecture for IoT Virtualization” [26].

Figure 2-7 introduces an example of a functional architecture structured into layers (and sublayers) with an indication of the main functions that are expected to be provided in each of the layers and sublayers. In addition, two vertical functions are added related to cross-layer functionality: security and management.

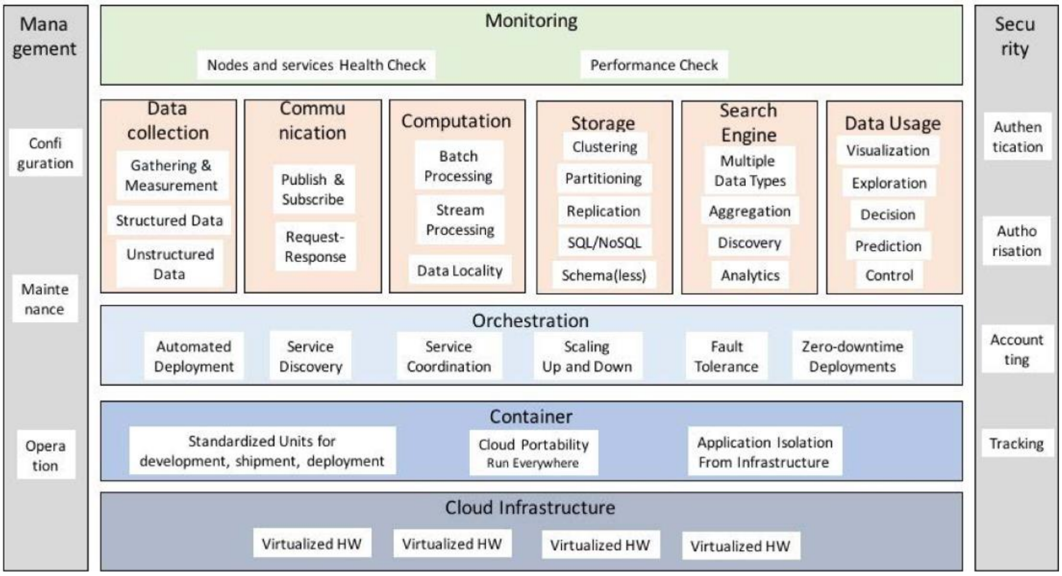

Figure 2-7: A High-Level Architecture for IoT Virtualization [26]

The functions described in the layers and sublayers are used for the identification of potential Open Source Components that can support the implementation of the IoT microservices. It should be noted that the realization of functions within an IoT resources entity is optional and thus, software units (e.g., VNFs) that represent such entities can be arbitrarily enriched with (platform) functionalities.

This architecture clearly highlights the isolation between applications and infrastructure and indicates the orchestration functionalities, which can be used as reference for the definition of mMTC slicing orchestration mechanisms. These include automated deployment, (auto-)scaling, resilience, as well as monitoring of health and performance of different services. Performance of virtualized resources can relate among others, to Edge requirements (e.g., low latency).

The Orchestration and Monitoring layers can include components (i.e., agents) to act as the interfacing between the virtualized IoT resources and an Orchestration Entity (e.g., the NFVO in the NFV MANO case).

2.2.5 Other SDOs – Broadband Forum, TM Forum, MEF

In addition to 3GPP and ETSI, the work of other SDOs should be also taken into account, namely BBF, TMF, MEF. The outcomes of these three SDOs are not covered in this section in detail, but will be handled later in relation to the transport network (BBF) and inter-domain interfaces (TMF, MEF).

Broadband Forum: The work done in BBF related to 5G has been mainly focused on two different, although inter-related, aspects: 5G fixed-mobile convergence and 5G fixed access and transport. The BBF has established the 5G project stream [27], focused on the architectural and functional impact on the fixed broadband system of wireline access integration, including N1, N2, N3 5G reference points, as well as aspects such as slicing, multi-tenancy, CPE management/configuration and QoS. This project stream is also intended to be used as a framework for the cooperation with 3GPP to steer 5G fixed-mobile convergence. The scope of the project stream includes providing recommendations about 5G system architectural and functional integration, as well as devising strategies and developing specifications for interworking of existing fixed access subscribers and deployed equipment with a 5G core. Two potentially relevant projects for 5G evolution are currently active – SD-406 “End-to-End Network Slicing Study” and SD-407/SD-420 “5G Fixed Mobile Convergence Study”, both of which are analysed in Section 4.5.

TM Forum: TM Forum [28] is the global industry association that drives collaboration and collective problem-solving to maximize the business success of communication and digital service providers and their ecosystem of suppliers. TMF provides an open, collaborative environment and practical support to enable CSPs and suppliers to rapidly transform their business operations, IT systems and ecosystems to capitalize on the opportunities presented in a rapidly evolving digital world.

MEF: Originally dedicated to Carrier Ethernet networks and services, MEF has redefined its focus to handle assured services orchestrated across a global ecosystem of automated networks [29]. MEF scope includes Optical, Carrier Ethernet, IP, SD-WAN Services and Cloud Services, as well as orchestration of the service lifecycle. In particular, LSO (Lifecycle Service Orchestration) is the set of MEF-defined specifications enabling standardized service orchestration based on standardized lifecycles of end-to-end connectivity services across one or more network service domains. LSO, in combination with SDN and NFV, is designed to enable MEF 3.0. LSO is an agile approach to streamlining and automating the service lifecycle in a sustainable fashion for coordinated management and control across all network domains responsible for delivering an end-to-end Connectivity Service (e.g., Carrier Ethernet, IP VPN, MPLS, etc.).

2.3 5G Industry Initiatives

The Next Generation Mobile Network Alliance’s (NGMN) architectural vision [30] advocates a flexible softwarized network approach. This views network slicing as a necessary means for allowing the coexistence of different verticals over the same physical infrastructure. Initial proposals were limited to slicing the CN, but NGMN has argued for an end-to-end (E2E) scope that encompasses both the RAN and CN. To realize this and provide the needed context awareness, both parts need to be flexibly sliced into several overlaid instances serving different types of users, devices, and use cases. This whole process needs to be orchestrated by an end-to-end Slice Orchestration and Management (SO&M) entity that has a central role in the architecture.

The overall NGMN architecture is split into three layers: infrastructure resource, business enablement, and business application, as depicted in Figure 2-8. Realizing a service in this proposal follows a top-down approach via a network slice blueprint that describes the structure, configuration, and work- flows for instantiating and controlling the network slice instance for the service during its life cycle. The service/slice instance created based on the blueprint may be composed of several subnet- work instances, each in turn comprising a set of network functions and resources to meet the requirements stipulated by the service in question.

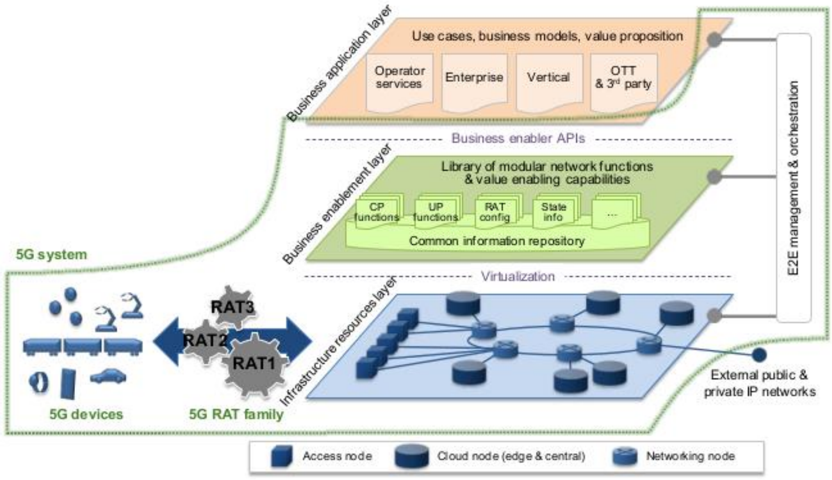

Figure 2-8: 5G network slices as seen by NGMN [30]

5G-PPP defines network slice as "a composition of adequately configured network functions, network applications, and the underlying cloud infrastructure (physical, virtual or even emulated resources, RAN resources etc.), that are bundled together to meet the requirements of a specific use case, e.g., bandwidth, latency, processing, and resiliency, coupled with a business purpose". Network Slicing is seen as an end-to-end concept covering all network segments including radio networks, wired access, core, transport and edge networks.

The overall 5G-PPP architecture, based on four layers, is shown in Figure 2-9 [31]. The Management and Orchestration layer incorporates the ETSI NFV MANO, including the VIM, the VNF Manager and the NFVO functions, as well as 3GPP network management. The Software-Defined Mobile Network Orchestrator (SDM-O) is able to set up slices and merge slices following the network slice templates.

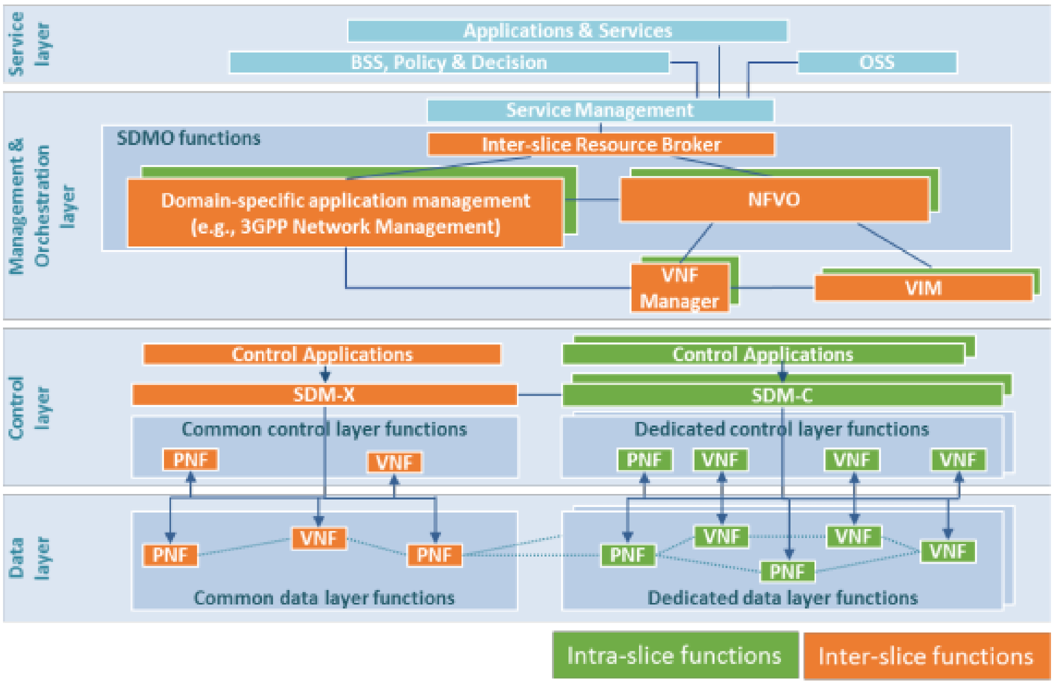

Figure 2-9: 5G-PPP 5G architecture [31]

At the control layer, dedicated and shared resources are handled by the Software-Defined Mobile Network Controller (SDM-C) (one instance per slice) and the Software Defined Mobile Network Coordinator (SDM-X), respectively. For each network slice, the SDM-C is responsible for managing the network slice resources and building the paths to join the network functions taking into account the received requirements and constraints gathered by the QoS/QoE Monitoring and Mapping module. The SDM-C, based on received QoS/QoE information, may adjust the network slice configuration by reconfiguring some of the VNFs in a network slice or the data paths. If the requirements cannot be met by reconfiguring VNFs or data paths, the SDM-O can perform a slice reshaping e.g., by adding more resources to the network slice.

5G-PPP's architectural vision offers a more elaborate examination of the roles and relationships between different parts of the 5G network. Overall, 5G-PPP shares the NGMN view that a potential 5G architecture must support softwarization natively and leverage slicing for supporting diverse use cases. 5G-PPP's architectural proposal is divided into five layers: infrastructure, network function, orchestration, business function, and service layers. Relating this to the NGMN proposal, while both are built on infrastructure and network function (business enablement) layers, there are a couple of differences: the orchestration/MANO is viewed as a separate layer in the 5G-PPP proposal; and the business application layer in the NGMN proposal is divided into two layers (business function and service) in the 5G-PPP case.

2.4 Open Source platforms

Open source technologies are expected to play an important role in multiple aspects of 5G services. Several projects have recently emerged with M&O tools, such as OSM, ONAP, OpenBaton and SONATA, which are briefly described below.

2.4.1 OSM

The Open Source MANO (OSM) project [32] is an E2E Network Service Orchestrator aligned with ETSI NFV specification. It is an ETSI-hosted initiative, which is currently formed by more than 110 contributors [33], that aims to develop a solution that facilitates the use and maturation of NFV technologies, gives access to a huge ecosystem of VNF vendors and allows testing and monitoring between the orchestrator and the rest of the elements (NFVI, VIM, VNFs, PNFs...). To achieve this goal, OSM focuses on minimizing the integration effort through three key points (see Figure 2-10):

- **Unified Northbound Interface (NBI):** Based on NFV SOL005 [34], allows the management of the lifecycle of NS and NSI.
- **Defined Information Model (IM):** Models and automates the full lifecycle of NF, NS and NSI and maintains an aligned view with ETSI NFV.
- **Extended concept of Network Service:** Allows that a NS spans across the different domains (virtual, physical and transport).

Moreover, OSM's latest release (release FIVE), makes use of a micro-service architecture, extends its orchestration beyond virtual domains (includes physical and hybrid elements) and provides full support of 5G Network Slices. On the other hand, at a user experience perspective, new levels of responsiveness and improved metrics, logs and alarms have been achieved.

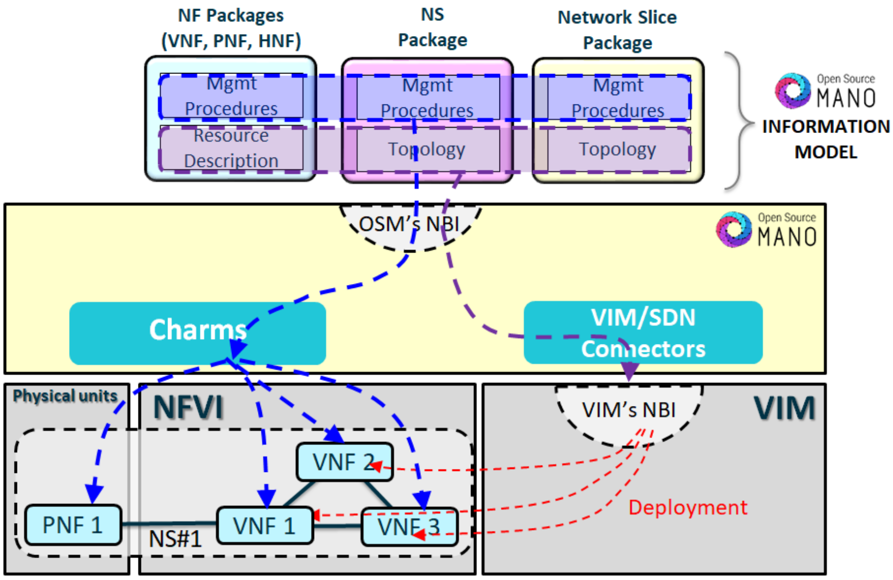

Figure 2-10: IM and NS operation via NBI [35]

2.4.2 SONATA

SONATA [36] is an open source framework, initially developed in the scope of the SONATA 5G-PPP phase-1 project and extended in the 5GTANGO 5G-PPP phase-2 project (see Annex A). The framework has three main subsystems, described in Annex A in relation to the 5G-TANGO project: (1) Software Development Kit (SDK); (2) Service Platform (SP); (3) Validation and Verification (V&V). In particular, the SP, a powerful MANO tool that enables operators, developers and customers to flexibly manage and orchestrate Network Services, is the focus of this section. The following is a list of the most relevant SONATA Release 4.0 features.

- *MANO basic functions*, enabling the on-boarding and lifecycle management of VNFs and NSs (VNFM/NFVO functions) by different roles (operators, developers, customers);
- *Catalogue-driven*, enabling the storage of VNFs and NSs, in a centralized Catalogue database (using package files, where software images, descriptors, etc. are included), to be later used for lifecycle management operations (e.g. instantiation, termination, etc.);
- *Records Repository*, stores runtime data in a Repository databased, where information about instances currently running as well as historic data can be obtained;
- *Multi-VIM/WIM*, enabling the support multiple VIM and WIM technologies, by using a plug-in approach that abstracts infrastructure technologies, facilitating the introduction of new ones (Openstack, VTN are supported in this version);
- *Function Specific Manager* for VNFs and *Service Specific Managers* for NSs, to create tailored behaviour for lifecycle operations, according to VNF/NS specificities;
- *Policy-driven*, enabling the autonomous management (zero-touch service management), through policies that take lifecycle management decision (e.g. deploying, scaling, healing) based on predefined rules, triggered by monitoring events (thresholds);
- *SLAs Assurance*, enabling the commitment of SLAs to NSs; checking violations;
- *Monitoring*, enabling the collection of monitoring data from VNFs and NS, feeding Policy, SLA and other components;
- *RESTful APIs*, exposing ETSI NFV aligned APIs for an external access to the SP capabilities, namely to on-board VNFs/NSs, instantiate/terminate NSs, get Catalogue data, get Repository records, manage Policies and SLAs, among others;
- *Management Portal*, enabling the management by multiple roles (operators, developers, customers), on-boarding and managing NS and VNF lifecycle, managing Policies and SLAs.

SONATA Release 5.0, bringing significant improvements, is expected in July 2019. The final architecture of the SP subsystem is depicted in the Figure 2-11, where the main internal components (developed as microservices) and the interactions are shown. This release will support features such as *Network Slicing,* enabling the deployment of NSIs, created from Network Slice Templates (NST), which use a combination of NSs, either deployed in a single PoP or multi-PoP, to build the Network Slice; *Deployment Flavours*, enabling the design of different VNF and NS flavours (e.g. Gold, Silver, Bronze); *Kubernetes as VIM*, enabling the deployment of VNFs/NSs on top of Kubernetes-based infrastructures (support also hybrid scenarios, e.g. mixed Openstack/Kubernetes VNFs); *Quality of Service*, enabling the specification and enforcement of different levels of QoS, both within PoP environments and within WAN segments; *Ingress and Egress endpoint*, enabling NSs to have ingress and egress WAN segments before the first/last VNF, allowing an NS to expand towards the customer premises, ensuring end-to-end QoS; *Licensing*, enabling the control of VNFs/NSs usage (support public, trial and private licenses); *Portal Enhancements*, extending the number of operation supported via Portal.

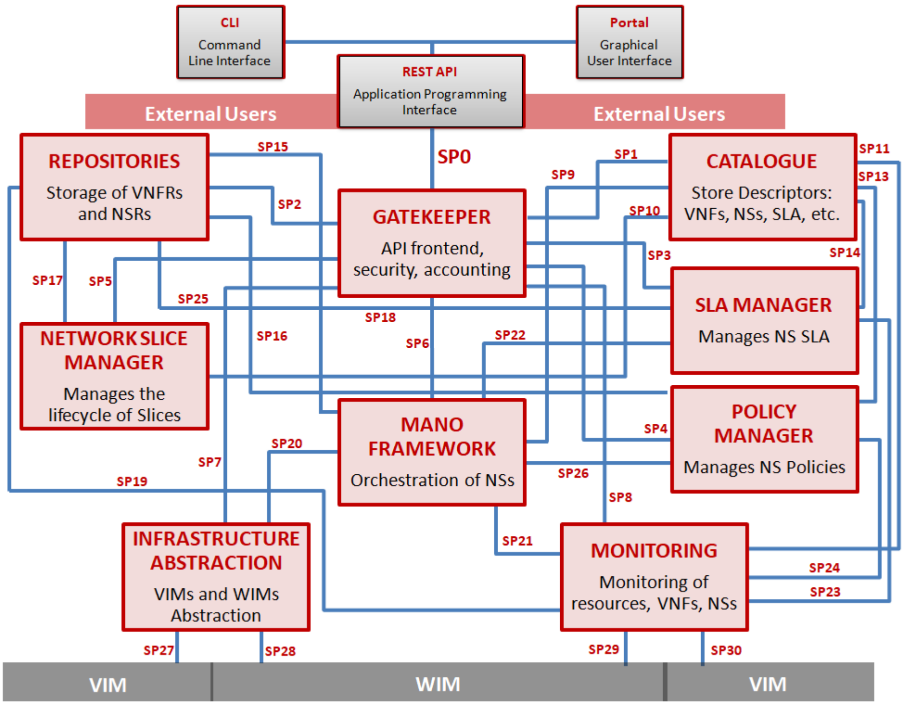

Figure 2-11: SONATA Service Platform Architecture [36]

2.4.3 ONAP

The Open Network Automation Platform (ONAP) [37] is based on a unified architecture and implementation to deliver an open platform enabling end users to create their own VNFs. The platform aims at automating, orchestrating and managing VNFs and network services. The vision of ONAP in terms of isolation and multi-tenancy was described in D2.1 [5], while in the current deliverable the vision of ONAP in the network slicing is briefly presented hereinafter.

Network slicing management in ONAP is done by extending the cloud sharing network/compute/storage to share network functions and services implemented across PNFs & VNFs. Figure 2-12 illustrates the 5G network slice architecture based on ONAP management. ONAP enhances Service Design and Creation/Active and Available Inventory (SDC/AAI) modeling and service definition, lifecycle management of slices. Slices will have state, metrics and scaling procedures different from the network functions (e.g. service/slice across multiple NFs from different providers):

- Enhance SDC to model & define slice segments (RAN, Transport, core), E2E slice, and Mobility services
- Enhance Service Orchestrator (SO) Instantiate and lifecycle manage nested slice segments, E2E slicing and slicing services
- Enhance AAI to inventory and store state of slice segments, slice instances, and 5G slicing services
- Integrated design (e.g. complex service composition)
- Support Service Aggregation (e.g. complex service hierarchy)
- Support Service Chaining (e.g. across multiple clouds, service path)
- Support Service Modification Capability (e.g. modifying without downtime)
- Monitoring and management of complex nested slice segments, slice instances and slice services.

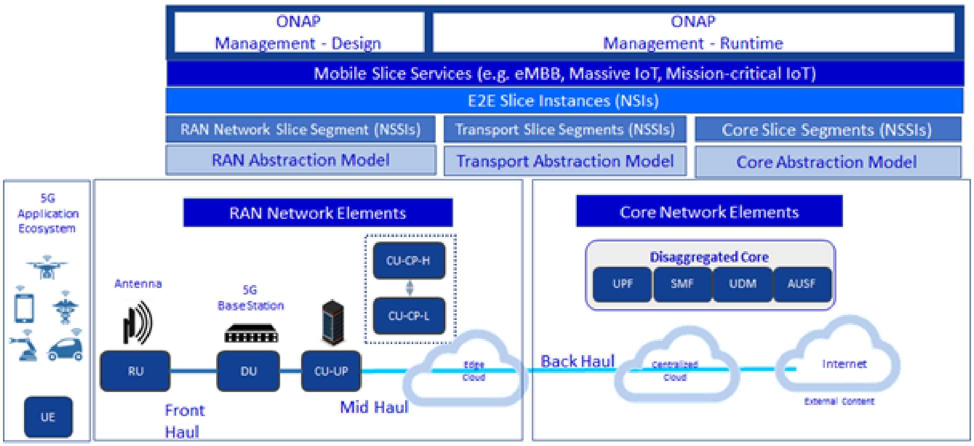

Figure 2-12: 5G network slice - network architecture with ONAP management [38]

2.4.4 OpenBaton

Open Baton [39] is an implementation of the ETSI NFV MANO specification. It is an open source project developed by Fraunhofer FOKUS and TU Berlin. It follows a modular approach with the main components being an NFV orchestrator, a generic VNF manager and VIM drivers to support different types of VIMs (e.g. OpenStack and Docker). Open Baton allows network service description in both the ETSI NFV model and TOSCA.

Beyond the basic MANO functionality, OpenBaton was extended with capabilities to enable the support of additional management functionality addressing the needs of software networks. One of its components is a Network Slicing Engine (NSE) [40], depicted in Figure 2-13, which instantiates rules on physical networks for allocated bandwidth as per Network service specific requirements. In a nutshell this component ensures QoS configuration defined in descriptors provided by the NFVO. It runs as an external component and communicates with NFVO via OpenBaton's SDK.

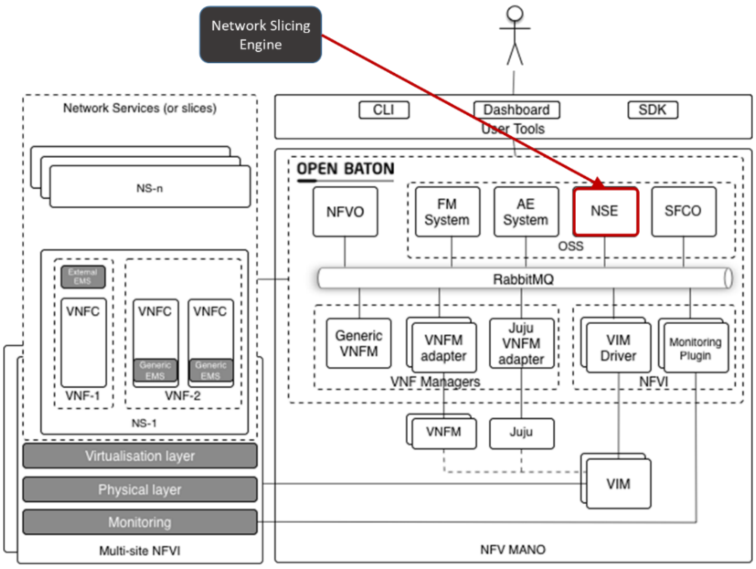

Figure 2-13: OpenBaton Architecture [39]

3 Network slice management and orchestration

3.1 Introduction

Network slicing in the context of 5G is distinguished from other forms of slicing that have been considered in the past by its end-to-end nature, the requirement to express a service through a high-level description and to flexibly map it to the appropriate infrastructural elements and network functions. This observation regarding the operation of slicing in the context of 5G naturally leads to two new high-level concepts:

- A service layer that is directly linked to the business model behind the creation of a network slice
- Network slice orchestration to handle a slice lifecycle.

Due to the novelty that this layer introduces in terms of concepts and ideas, the related research in this domain naturally focuses on answering fundamental questions regarding network slicing architectures. More specifically, the topics considered are related to the way services should be described and how they should be mapped to the underlying network components, and the architecture of network slicing managers and orchestrators.

Regarding the service layer and the way that the business model of a service should be described in high-level terms, different proposals have been put forward by bodies such as the NGMN, ITU-R and 5G PPP. The service level description (manifest) may be simply a set of traffic characteristics, SLA requirements (e.g., for performance related aspects like throughput and latency), and additional services (e.g., localization service) [41]. Alternatively, the service description may be more detailed in the sense that it can identify specific functions or RATs that are bundled together and should be used for the creation of the slice [42] that provides a specific service (slice as an application). The main difference lies in the way that the network slice is generated. In the first case, the NFV orchestrator is assigned the more complex task of identifying the appropriate functions and technologies that will guarantee the fulfilment of the requirements described in the slice manifest, while in the second case things are simplified since the required slice building blocks are already identified in its description. However, the second approach can be less efficient as it leaves less flexibility to the slice orchestrator to tune the slice components.

The exact form that the network slicing MANO entity should have is still unclear with different works presenting different ideas. Some proposals envision that network slicing will come through an evolution of the current 3GPP standards and therefore propose enhancements in terms of interfaces and functionalities for the existing mobile architecture [43]. Other proposals envision a more radical clean-slate approach where the slice management and orchestration will be implemented as an application over an SDN controller, which will oversee both the wired and wireless domains [44] [45] [46]. Concrete implementations of MANO reference frameworks such as OSM are already appearing, which enable experimental studies on end-to-end 5G network slicing. Another very important issue is how to map and stitch together the components that are available to the various layers of the architecture in order to compose an end-to-end slice. Two types of mapping have been considered, each at a different abstraction level:

- The functional/SLA mapping of the service requirements to network functions and infrastructure types;
- The mapping of selected network functions and infrastructure types to vendor implementations [41] [42].

The first type of mapping refers to the way that MANO chooses appropriate high-level network elements which are required to create a slice for a given service in order to meet its functional requirements and SLA. For example, if a slice has a need to cover devices over a wide area without any capacity concerns, choosing a deployment with macro-cells might be a good option. For this mapping, it has been proposed that the available infrastructural elements and network functions should reveal their capabilities to the MANO in a form of metadata, describing the types of services that they can support [41].

Once the type of functions and infrastructural elements required for the slice have been identified, there is a need for a further mapping of these elements to concrete implementations. Depending on the implementation of a function by a vendor, different levels of services can be offered. For example, alternative software implementations of a gNB could provide support for a different number of users, with different performance guarantees or even different capabilities (e.g., flexible modification of the MAC scheduler). Here too, the high-level solution to this problem seems to be the use of metadata in the elements provided by the vendors of the functions and the infrastructure. Such metadata could describe both the capabilities of the vendor-specific functions and hardware [41] as well as their deployment and operational requirements (connectivity, supported interfaces, and infrastructural KPI requirements) [41] [42], providing the MANO with sufficient information to perform the best possible configuration for the slice.

A good approach to handle the mapping of high-level service description to a concrete slice in terms of infrastructure and network functions is to develop domain-specific description languages that allow the expression of service characteristics, KPIs, and network element capabilities and requirements in a comprehensive manner while retaining a simple and intuitive syntax. Nevertheless, one of the most critical factors on creating E2E Slice Orchestration is to be compliant with the Information Model defined at this level, so all other data models can link with it.

These languages should inherently provide features such as flexibility/extensibility to accommodate new network elements that may appear in the future (e.g., new network functions, new RATs) and the applicability to be used in multi-vendor environments. A desirable feature would also be the capability to compose complex rules and expressions out of simpler ones, introducing abstraction layers in the expression of service requirements.

As noted before, concrete MANO frameworks like OSM have emerged in recent years. While such platforms are essential to flexibly realize network slices end-to-end and as needed, there is a more significant challenge that is only starting to be addressed. This concerns holistic orchestration of different slices so that each meets its service/SLA requirements while at the same time efficiently utilizing underlying resources. This calls for a sophisticated end-to-end orchestration and management plane. Such a plane should not be limited to trivial slice generation that does mapping of slices to network components and statically allocates them resources. Instead, it should be adaptive, ensuring that the performance and resiliency requirements of the deployed services are met. To achieve this, it should efficiently and holistically manage resources by making decisions based on the current state of slices as well as their predicted state/demands in the near future.

Such issues have been thoroughly investigated in the context of cloud computing and data centres, where many concrete solutions have already been proposed (e.g., [47]). While underlying principles from these other contexts can be leveraged, mechanisms targeting 5G network slicing should be suitably adapted and extended considering additional types of resources. Specifically, not just the resources found in cloud environments (processing, storage, network), but also radio resources need to be included, considering their correlation and how adjusting one resource type could have a direct effect on the efficiency of another, and therefore on the overall service quality. The problem of meeting requirements of different services while efficiently managing underlying network resources in the 5G network slicing context is also somewhat analogous to quality of service (QoS) provisioning in the Internet.

Finally, in addition of the mapping of the high-level service description with concrete infrastructure and network functions, one of the main requirements in the E2E slice orchestration and management is the existence of the right interfaces in order of being able to understand and interact with the different network and administrative domains, needed to deliver the service, as it will be described in Section 5.

3.2 Overall 5G-VINNI architecture and components

5G-VINNI deliverable D1.1 [2] defined the 5G-VINNI reference architecture, in which three different levels were identified. The network domain level at the bottom, where the RAN, CORE and Transport infrastructure and PNFs/VNFs are allocated. Then, the second level includes the NFV MANO and Network Domain Controllers. Finally, at the top is the E2E Service Operations and Management (E2E SO&M) level, which has a holistic perspective of the slices and service offered.

This section is focused on the network slice lifecycle management and orchestration, and hence, the two levels above are fundamental and will be described, as illustrated in Figure 3-1.

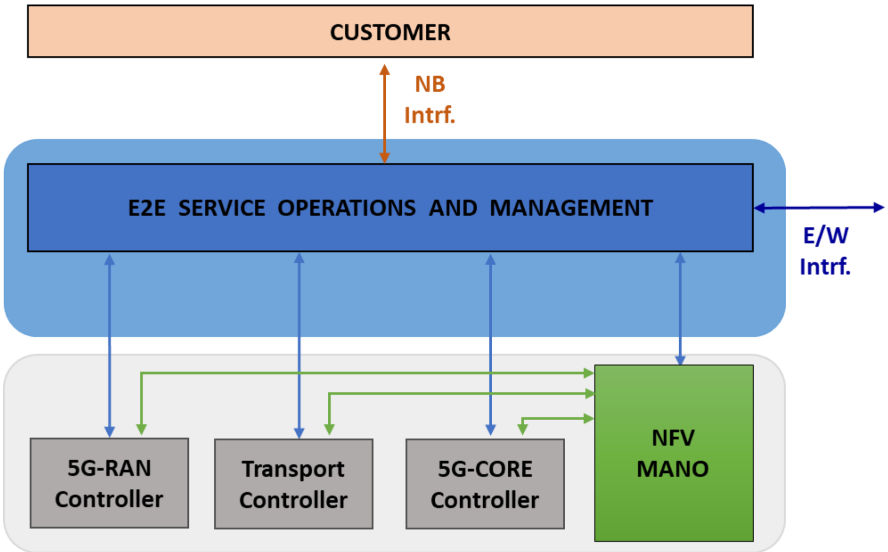

Figure 3-1: Architectural Components for the slice life cycle management and orchestration

Management and Orchestration is a key network slice capability. Management and orchestration in the context of network slicing should include: first, the duties needed in order to provide and administrate the respective isolation of the shared resources, which will be denominated as Network Slice Provider management and orchestration duties; second, the management and orchestration within the isolated slice resources, which will be referred to as Network Slice Tenant management and orchestration duties. More extensive details on shared and isolated resources in network slicing can be found in the 5G-VINNI deliverable D1.2 [3].

The remaining of this section will explain first the E2E Service Operations and Management top layer, followed by the Network Domain Controllers and NFV MANO layer.

3.2.1 E2E Service Operations and Management (SO&M)

To understand this key 5G-VINNI architectural component, three important concepts should be considered: 1) A customer-close component able to see a slice and a service as holistic elements; 2) Multi network domain environments that goes beyond the orchestration of VNFs; 3) Multi administrative domains that enable the establishment of slices across different 5G-VINNI facility-sites. E2E Service Operation and Management, within the 5G-VINNI architecture, encompasses the management and operations of E2E services that span multiple domains. It translates service definition / Service Design into configuration of resources (physical and virtualized) needed for service establishment, using orchestration to coordinate between domain controllers, the NFV-MANO, and other administrative domains.

The E2E SO&M can be associated with the 3GPP Network Slice Management Function (NSMF), since it is also in charge of the lifecycle of a network slice, it uses different network domains to create the end-to-end slice, and it has full visibility and control to the end-to-end slice and its performance.

One of the goals of the E2E SO&M is to offer business, or customer facing, services northwards via standardised APIs to fulfil the service requirements of customers and verticals [4]. The E2E SO&M translates these business services into their component network services and requests the services from the appropriate domain. The E2E SO&M interfaces to the network controllers and NFV MANO via southbound interfaces to request the network service. The E2E SO&M interfaces with other facility-sites via east/west to request the service from other service provider. In this way the E2E SO&M hides how the business services are realised at the resource level, thereby allowing verticals to focus on their service requirements, Service Level Agreements (SLAs) etc, rather than upon service implementation.

There is no single standard definition of E2E SO&M, although its functions are spread in multiple standards including 3GPP, ETSI and industry associations (e.g. TMF, GSMA). The E2E SO&M within the 5G-VINNI reference architecture builds upon these references, in special the MEF LSO Reference Architecture, TMF Open API and the OSM release 5 Architecture, in which the E2E SO&M can be mapped into the upper level functionality of E2E Network Service Orchestrator, where the information elements of network slice and network slice subnets are defined on top of network services. Figure 3-2 presents the most relevant architectural components of those important references. A deeper level of detail on those components is presented in 5G-VINNI deliverable D.1.1 [2]; some additional details regarding the interfaces needed will be provided in section 5.

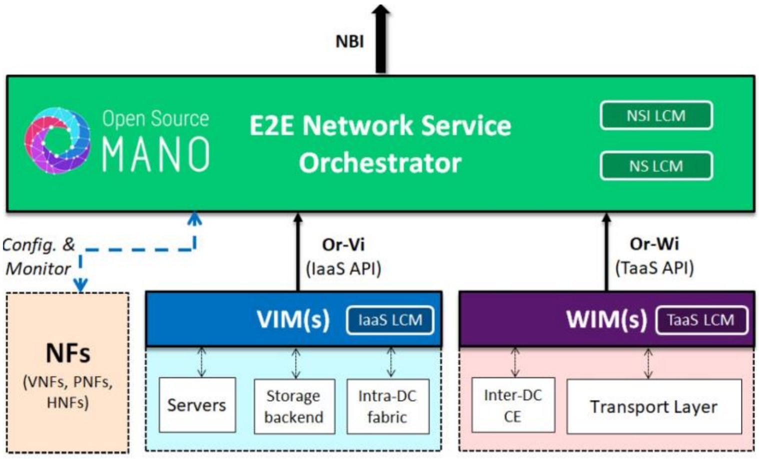

Figure 3-2: OSM Architectural Components

3.2.2 NFV MANO

From a resource management point of view, 5G-VINNI services are deployed and operated as NFV Network Services NFV-NSs). This makes NFV MANO a key component of the 5G-VINNI facility. With NFV MANO, facility operators are able to deploy and execute 5G-VINNI services in the NFVI with great agility and flexibility. The NFVI in 5G-VINNI facility span across seven facility sites (four main sites and three experimental sites), and consists of different PoPs for VNF deployment and execution, ranging from cell sites to edge/regional data centres.

Each facility site will have its own NFV MANO stack to manage and orchestrate NFV-NSs within its administrative domain. NFV MANO exposes a set of SBIs and NBIs to facilitate the interaction with the rest of facility site components. On one hand, NFV MANO makes use of the SBI to manage (and collect information from) the site's NFVI resources on top of which 5G-RAN and 5G-Core VNFs run. On the other hand, NFV MANO uses the exposed NBIs to provide/receive information and operations that have an impact on the NSs under its management. NBIs facilitate the interplay between the

NFVO and the rest of management blocks, including 5G-RAN and 5G-Core controllers and E2E Service Operations and Management. From the perspective of NFV MANO, these management blocks take the role of Network Management System (NMS) in the NFV reference framework, so NBIs can take the role of Os-Ma-nfvo and Ve-Vnfm-em reference points (see Figure 2-5 in Section 2.2.2).

3.2.3 Network-Domain Controllers

3.2.3.1 RAN and Core Domain Controllers

The Domain Controllers at the RAN and Core are in charge of managing the different NFs at the application level (independently of their deployment), and in general to provide control on all the non-virtualization-related operations. Among others, they are also aware of the alarms on the system with the respective reaction, in a manual or automatic way, and they can get access to dashboard reports and real time and historical statistics from the respective network domain (RAN or Core).

Core & RAN domain controllers can be associated with the 3GPP defined, Network Slice Subnet Management Function (NSSMF), and during the first 5G-VINNI phases they can be mapped to the EMS (Element Management System) function group of FCAPS (Fault, Configuration, Accounting, Performance and Security management) for the functional/application part of the underneath deployed PNFs and VNFs. In this way, we can say that these domain controllers perform changes at the application level in a “vertical” way on functions “horizontally” deployed by the NFVO. For instance, signalling issue at the mobile Core, abnormal releases at the RAN, or in general any potentially anomaly detected in terms of accessibility, retainability, integrity, availability, or mobility as defined in ETSI TS 132 450 [48], will be addressed by the RAN and Core Domain Controllers. Finally, the RAN domain controllers must have the capacity to translate RAN slice requirements into radio resource allocations and RAN resource management, as well as the ability to adjust dynamically, including changes in the radio conditions or the deployment of new slices.

3.2.3.2 Transport Domain Controllers

The Domain Controllers at the transport include components such as SDN controllers or MPLS management and control components. In 5G-VINNI the Transport network will provide advanced backhaul functionality, to interconnect for instance different locations from a common site.

MPLS networks for instance are well-known solutions used for several years in the transport network. The process of sending a packet through the MPLS network is the following: First, the Forwarding Equivalence Class (FEC) is defined. The FEC is a group of packets with similar characteristics (IP address source, IP address destination, Port, etc). Second, a path between the source and the destination is found using routing algorithms. This path is distinguished on each router through the use of one specific label. Finally, the first router on the MPLS domain labels each of the incoming packets in order that they can be clearly identified and sent through the predefined path. MPLS allows the visualization of the E2E path instead of a hop by hop vision.

In SDN, the CP and the data plane are separated from each other. In addition, the CP is logically centralized in a software-based controller defined as the “network brain”, while the data plane is composed of network devices (“network arms”) that forward packets. The CP includes both northbound and southbound interfaces. The northbound interface provides a network abstraction to network applications, and the southbound interface standardizes the information exchange between the control and data planes.

3.3 Network slice lifecycle management

5G-VINNI plans to offer Network Slicing as a Service (NSaaS) as a means to support a wide range of use cases to vertical industries and their applications [3]. A 5G-VINNI service (instance) is the set of capabilities of a network slice that are exposed to a customer via the external-facing interfaces provided by the 5G-VINNI facility [4]. This section describes network slice lifecycle from the perspective of the layers below the E2E service operation and management layer. Therefore, as opposed to D3.1 [4], where the lifecycle provided there is customer-focused, this section describes the lifecycle management of a slice, from the network management perspective.

According to 3GPP [6], the lifecycle of a NSI is split into four phases (see Figure 3-3), including the Preparation phase, Commissioning phase, Operation phase, and Decommissioning phase. In the following, the operations of slice lifecycle management for each phase are briefly described.

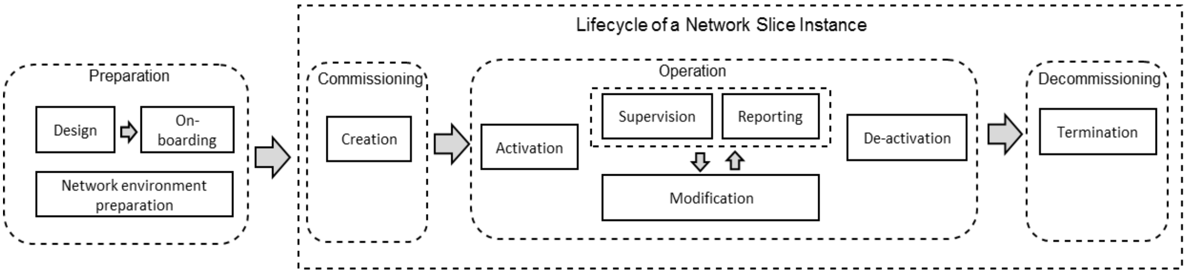

Figure 3-3: 3GPP vision on NSI lifecycle [6]

1. **Preparation**

In the preparation phase the NSI does not exist. According to [6], the preparation phase includes network slice template design, network slice capacity planning, on-boarding and, evaluation of the network slice requirements, preparing the network environment and other necessary preparations required to be done before the creation of an NSI.

2. **Commissioning**

According to [6], NSI provisioning in the commissioning phase includes creation of the NSI. During NSI creation all needed resources are allocated and configured to satisfy the network slice requirements. The creation of an NSI can include creation and/or modification of the NSI constituents. The operations specified in [17] for the creation of an NSI include: a) create a NSI; b) create a NSI with shared and non-shared CN network functions; and c) create Network Slice Instance with shared NSSIs. Details of the operations listed above can be found in [17].

3. **Operation**

The Operation phase includes the activation, supervision, performance reporting (e.g. for KPI monitoring), resource capacity planning (resource calculation based on an NSI provisioning, performance monitoring, and generation of modification policies) , modification, and deactivation of an NSI. In particular [6]:

- Activation makes the NSI ready to support communication services.
- Resource capacity planning includes any actions that calculates resource usage based on an NSI provisioning, and performance monitoring and generates modification polices as a result of the calculation.
- NSI modification could be including e.g. capacity or topology changes. The modification can include creation or modification of NSI constituents. NSI modification can be triggered by receiving new network slice requirements or as the result of supervision/reporting
- The deactivation includes actions that make the NSI inactive and stops the communication services.
- Network slice provisioning actions in the operation phase involves activation, modification and de-activation of an NSI.

4. **Decommissioning**

The decommissioning phase includes decommissioning of non-shared constituents if required and removing the NSI specific configuration from the shared constituents. After the decommissioning phase, the NSI is terminated and does not exist anymore. The operations of the decommissioning phase of an NSI [17] include, for example, terminating a Network Slice Instance with shared and non-shared CN network functions and shared AN.

3.4 5G-VINNI network slice composition

As described in Section 3.3, the preparation phase includes the design and all procedures required to create a Network Slice, containing all the functionality needed to deliver a specialized service to the customer (e.g. radio, mobile core, IMS, transport VPN, application functions). A Network Slice is composed of Network Subnetwork Slices, each of which is in turn mapped to a NFV-NS.

In general, there are common network functions (e.g. UDM, NSSF, NRF, and PCF). Other network functions can be flexibly deployed at different parts of the network, for example, at the edge, or the core, to meet the different requirements for each type of slice.

In 5G-VINNI, the design of network slices is anchored on three basic principles, as follows (following NSA-based example, illustrated in Figure 3-4):

- Reduce complexity by reducing the number of NS to be managed by E2E SO/OSS: this means that not every VNF should have its own dedicated NS, although this might be the case for some VNFs. The main goal is to simplify modelling of E2E services
  - NFV-NS to be managed as a single entity, e.g.
    - Lifecycle Management (e.g. instantiate, scale, heal)
    - Day-2 operations (e.g. add customer, add APN, change QoS)
- Make NFV-NS reusable across Network Slices, as illustrated in Figure 3-4. Examples are:
  - EPC (mobile data) and IMS (mobile multimedia) are reusable NFV-NSs across multiple slices
  - The Policy and Charging Rules Function (PCRF) might be used by one slice and not another, hence should be a separate NFV-NS
  - VPN and FW might and might not be used by a Network Slice, hence should be a separate NS
  - All other customer specific apps or services not necessarily reusable should be a separate Network Service.

    Note: it is possible to create NS based on a combination of such VNFs (e.g. NS consisting of EPC, VPN, FW) if this is a NS that is instantiated many times and simplifies operation.

- Enable sharing of NSs across Network Slices / services

In cases where these design criteria might trade-off (e.g. I and III), the above principles must be balanced.

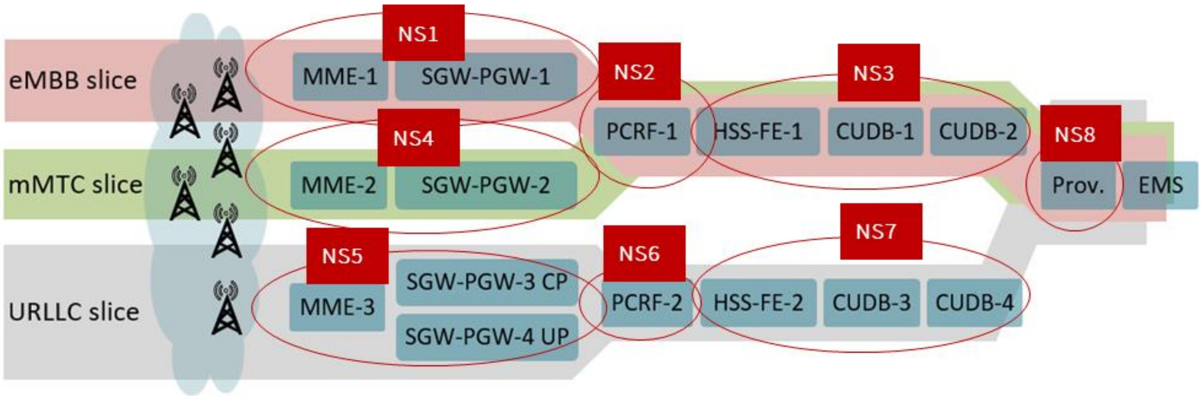

Figure 3-4: Composition of Network Slices based on Network Services

3.5 Multi-domain network slicing

Multi-domain network slice deployment refers to the ability of a slice to exist and operate across multiple-domains where domains refer to any delineation of resources such as due to technology, administration, legal, business or geographic. Each delineation presents a different set of challenges as mentioned in the sub-sections that follow. The focus is on the key separations due to technology and due to administration.

3.5.1 Domains based on technology

A network slice instance may exist across various technology domains each with a respective technology management system, proprietary or standardized. In the control plane the 3GPP slice architecture shows a signalling interface to the access and the data network domains. However, this is the *logical* control plane interface across domains. In practice, each of the logical interface could be configured over much longer physical links crossing many types of technology domains. The NFs defined in 3GPP SA2 are often hosted in VMs that are managed according to ETSI NFV. The NFs themselves (i.e. NFs software applications) are managed by 3GPP, which requires a coordination across the 3GPP management and the ETSI NFV management domains. In 3GPP SA5 the interaction of the 3GPP management system with an ETSI NFV domain (since R14) and with transport domains (since R15) has been studied. Figure 3-5 shows the interaction between 3GPP management and ETSI NFV reflected in 3GPP.

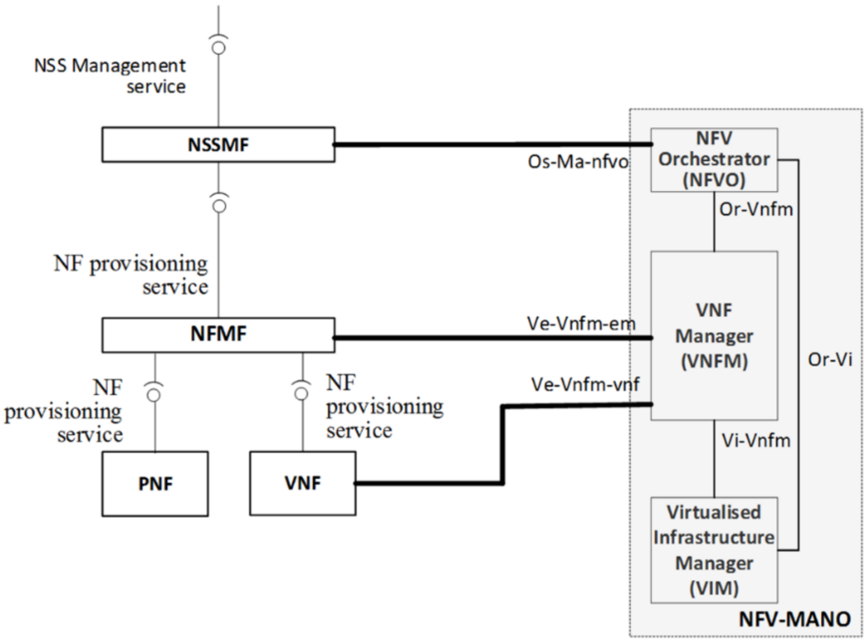

Figure 3-5: Example of 3GPP ETSI NFV interaction from TS 28.533 [49]

5G-VINNI proposed implementation architecture, as described in Section 3.2, is aligned with this view.

A generalization of the inter-domain integration is currently being studied in ETSI ZSM, as reported in Section 2.2.3.

3.5.2 Domains based on business and administration

A domain separation based on business and administration issues involves adds further complexities to the multi-technology domain issue. Primarily, it introduces issues relating to charging across the business domains and the creation of trust. As an example: one operator may be a wholly owned subsidiary of another operator. If there is a trust relationship between operators, without legal issues with regard to data sharing, the inter-domain scenario is simplified. However, in case of trust and or legal issues the level of management interfaces and the information exposed across operators must also be managed. Static implementations of this exposure from management domain are not scalable in terms of supporting slicing where different other operators with different trust relationships could use the slicing paradigm. Instead the exposure must re-expose internal management services and data at different (abstraction) levels. One solution to do this is using a new service that is able to configure this exposure on per customer per slice base, let's call it, exposure configuration service. The service exists as a management service exposed from a management domain and can be used by the operator or the 3rd party to configure and request, respectively a change in the management services and data that is exposed to that specific user. Figure 3-6 shows an example sequence diagram where a 3rd party (another operator) requests changes in the exposed management services from the operator.

1. The operator using the exposure configuration service is enabled to configure an exposure for a given 3rd party
2. (+ 3.) The 3rd party is able to obtain a list of management service from the exposure service or the operator domain. Under normal use the 3rd party could then use those services, till it has a change in the 3rd party service it uses from the domain or requires new management entities to be exposed
4. The 3rd party may then request this change to the exposure service (or the domain).
5. On reception of a request or notification the operator may authorize a new set of exposure for the 3rd party. In case of a notification of change in the 3rd part service the change in exposure is activated by the operator.

Conditional: If authorization succeeds (at least partially)

6. A new set of exposed management entities is provided to the 3rd party
7. The changes are reported internally to the operator's other management systems, such as charging and billing

Conditional: if the authorization fails completely

8. The change request is reported to the operator and/or other management service in the operator's domain
9. The operator and/or the other management service may configure a new exposure of management entities
10. If so, the new exposure of management entities is reported to the 3rd party

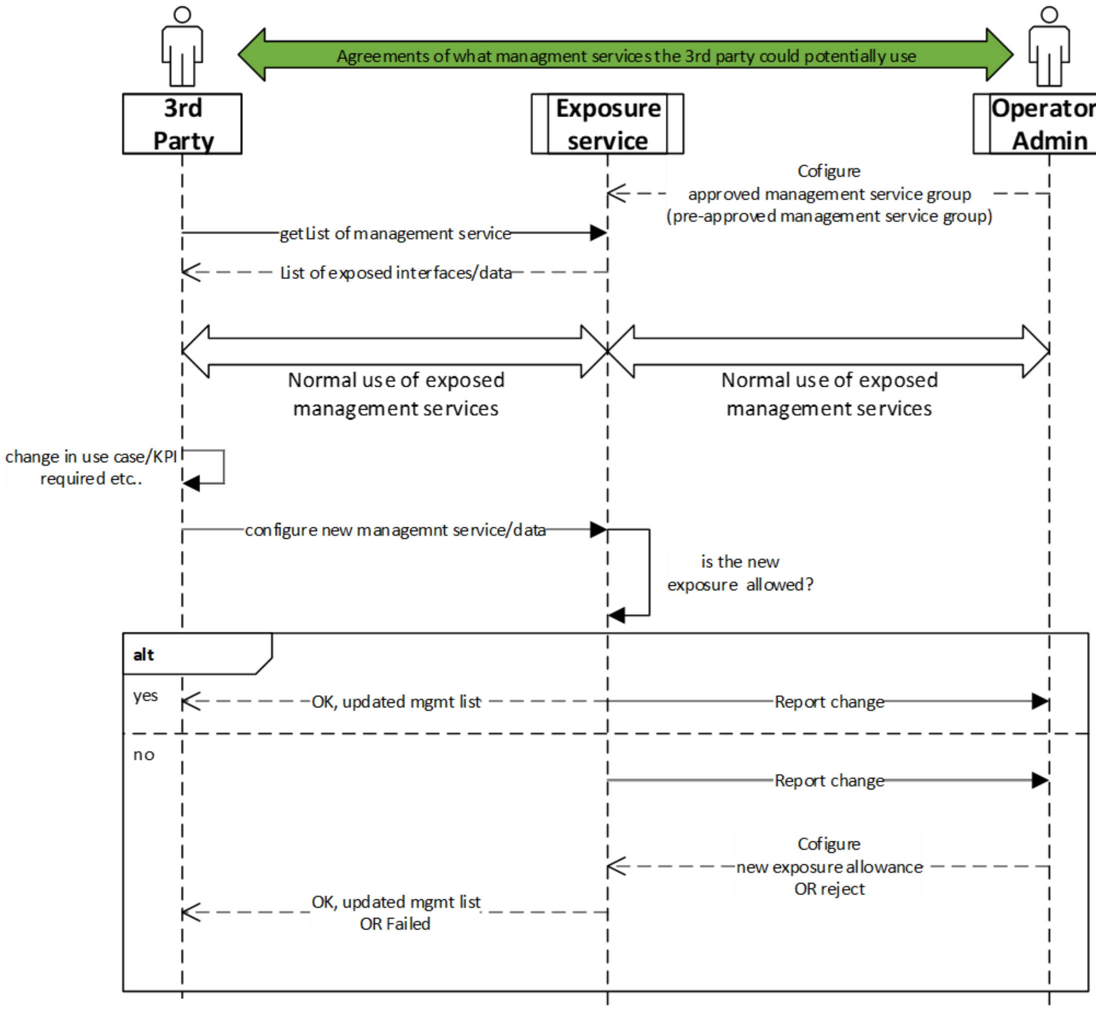

Figure 3-6: Exposure to 3rd parties for multi-operator interaction

With such an exposure configuration service the right level of management services and data could be exposed to the 3rd party (another operator) to enable multi-operator-domain interaction.

4 Network slicing requirements on infrastructure components

4.1 Introduction

5G network slicing should provide logically isolated end-to-end networks/network slice instances with a set of customized network capabilities, including access network, transmission network and core network. There are general requirements and characteristics for 5G network slices, which calls for customisation, according to different network capabilities, performance, and service requirements such as SLAs. In addition, for security and service assurance purposes, slices dedicated to different service scenarios need to be isolated from each other.

This section provides an overview of requirements related to different infrastructure components (radio access network, wireline access network, 5G core, transport network, data network), with a direct impact on the fulfilment of 5G KPIs, such as latency, throughput, reliability, or slicing characteristics such as isolation.

In terms of content, this section should be seen as complementary to 5G-VINNI Deliverable D1.2 [3], which already identifies the requirements of different network slice types, as specified by 3GPP. D1.2 also defines a list of high-level network slice requirements to be considered in 5G-VINNI. This section is intended to look at the several 5G components and see how the corresponding candidate technologies are likely to be impacted by network slicing.

This section is also complementary to the information provided in 5G-VINNI Deliverable D3.1 [4] about network slicing requirements from a customer perspective and end to end KPIs that 5GVINNI can offer to its customers. This section is intended to provide a resource-facing counterpart, by focusing on network slice requirements from an infrastructure perspective. D1.3 identifies four basic service types exposed to customers (eMBB, uRLLC, mIoT, Customised) and defines a set of service parameters, organized in three main groups (performance, functionality, network optimization). To build each of these service types, specific network slices instances should be used, following the 3GPP vision, as described in TS 28.530 [6]. Therefore, service requirements will be mapped into network slice requirements, which ultimately will be translated into requirements on the underlying technological components, including access, transport and core networks. This section is intended to provide an overview of how each of the network segment is impacted by network slicing requirements.

From the point of view of 5G-VINNI facility sites, as stated in D1.1 [2], all 5G-VINNI facility sites must support a core set of functional components and capabilities, however there is also a wide range of optional elements to allow each site to be unique, support specific use cases and KPIs, and allow specific tailoring and customisation to meet specific requirements. As shown in this section, for every major infrastructure component, there is a wide range of options that will ultimately affect the way targeted requirements and KPIs can be fulfilled by individual 5G-VINNI facility sites.

4.2 Radio Access Network

Support for slicing for LTE and NR NSA is based on RAN specific slice identifiers i.e. Public Land Mobile Network (PLMN) or Service Profile Identifier (SPID). The Radio Resource functionality enables the configuration of predefined shares of radio resources. The partitions are based on PLMN for operators sharing the radio network. For specific groups of UEs, the partitions can be based on SPID values.

The impact of network slicing to the 5G radio access network is further discussed in [50].

QoS handling within E2E network slicing is enforced in two ways:

- With bearer QoS on an E2E perspective

- With the QoS transport traffic management concept on a single enforcement point, such as a transport node.

Bearer QoS enables QoS handling in LTE and NR Non-Standalone (NSA) networks. It is the result of QoS negotiation between the User Equipment (UE), the RAN, and the packet core network. It defines the end user experience of the different services in both non-congested and congested networks.

From an E2E perspective, bearer QoS can be enforced in a static or dynamic manner. Static QoS is defined according to the static subscriber QoS. The static subscriber QoS is configured in the Home Subscriber Server (HSS) or Home Location Register (HLR). It can also be configured through a predefined QoS profile in the Serving GPRS Support Node (SGSN) Mobility Management Entity (MME). Dynamic QoS means that the Policy and Charging Rules Function (PCRF) evaluates the QoS. Its evaluation depends on the service, session, location, and subscriber parameters, as well as different trigger conditions. A high-level overview of the role each node has in E2E QoS is illustrated in Figure 4-1.

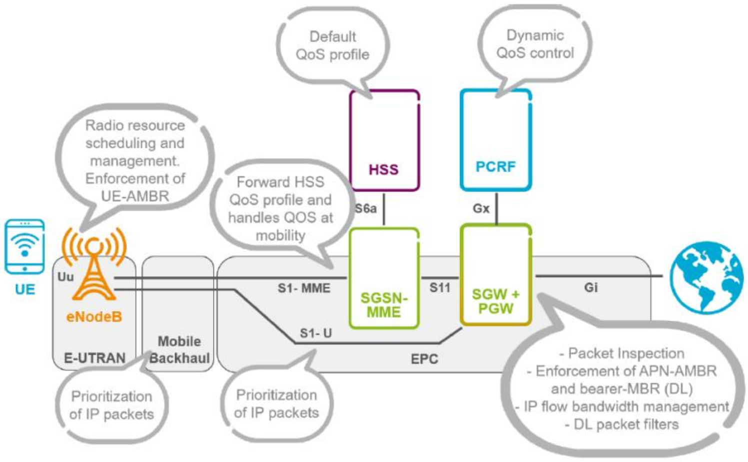

Figure 4-1: E2E EPS Nodes QoS Roles (Overview)

Radio bearer traffic is mapped to the Differentiated Services Code Point (DSCP) which enables the intermediate transport network to prioritize traffic. In transport networks with limited capacity it is especially important that the mapping is done at the egress ports of the related radio and core nodes. A DSCP value defines a Per-Hop Behavior (PHB), that is, a set of packet forwarding properties, given to an IP packet transported through the network.

E2E QoS design uses a blend of the 3GPP and IETF QoS architectures, refer to Figure 4-2.

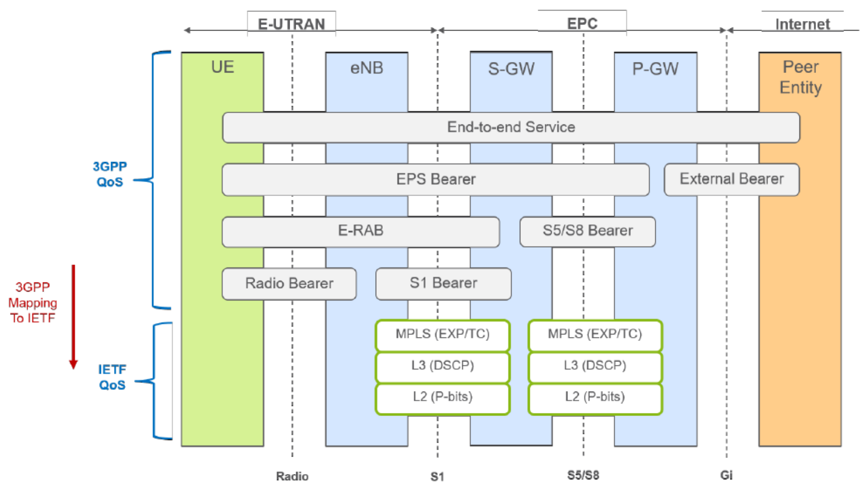

Figure 4-2: E2E QoS design uses a blend of the 3GPP and IETF QoS architectures

Mapping from DiffServ to MPLS or p-bits, or the other way around, is a function of the IETF QoS principles and not 3GPP. 3GPP only defines mapping bearer-level QoS to IETF DSCP transport QoS.

Dual Connectivity with LTE and NR is introduced with NSA option 3x. The supporting UE is simultaneously associated with LTE (eNodeB) and NR (gNodeB). This is E-UTRA-NR Dual Connectivity (EN-DC) RAN, see Figure 4-3.

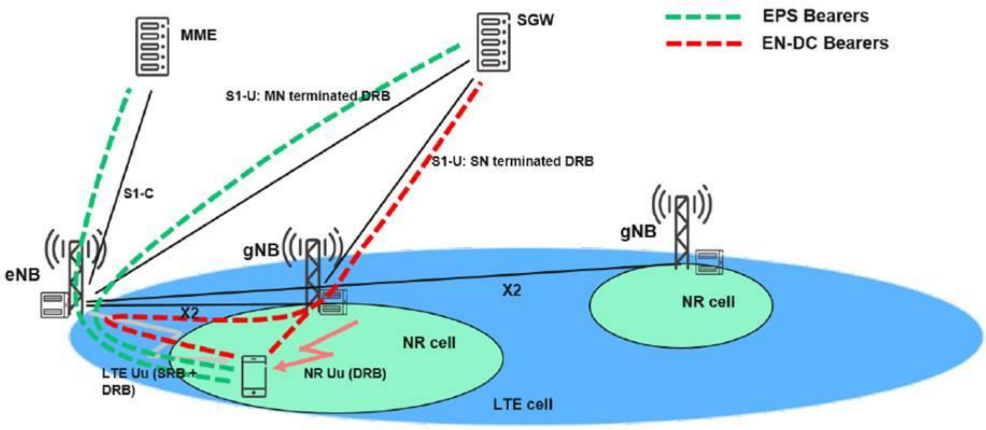

Figure 4-3: Conceptual Architecture for EPS with EN-DC (Option 3x)

The LTE EPS Bearer QoS framework defined in 3GPP TS 36.300 [51] also applies for NR NSA option 3x.

In the EN-DC architecture, the control plane and user plane traffic are separated into the following categories:

Control Plane Traffic: The eNodeB serves as the UE anchor point and handles related signaling traffic for both LTE and NR services. The eNodeB acts as the RAN master node towards the gNodeB over an X2-C interface. Communication with MME is over S1-C interface.

User Plane Traffic: The user Data Radio Bearer (DRB) is set up either as split-bearer (using both LTE and NR radio resources) or LTE-only bearer (using only LTE radio resources).

Split-bearer is a 3GPP function related to Dual Connectivity to enable division of a bearer between LTE and NR for optimized session performance. With the Split DRB the traffic is as follows:

Uplink User Plane Traffic – Split DRB: Uplink user plane traffic is configurable and controlled by the operator. It is done by the LTE leg only (default) or by the NR leg only.

Downlink User Plane Traffic – Split DRB: Downlink user plane traffic can be sent over one of NR or LTE RAT types or sent simultaneously over both RAT types. From QoS perspective the QCI (QoS Class Identifier) and ARP (Allocation and Retention Priority) values define if a bearer can be moved in split-bearer mode by the gNodeB.

The X2 link in the new 5G EN-DC architecture can under certain conditions carry the peak LTE RAT capacity of the connected master eNodeB for data traffic only, not voice (voice traffic is still carried over LTE DRBs only). This is the case when the UE is still in split-bearer configuration but the gNodeB decides to use only LTE resources in downlink. So, QoS implementation on X2 links becomes vital with respect to a pure LTE network. Both S1-AP/X2-AP is mapped to DSCP (Signaling), while S1-UP/X2-UP follows QCI mappings.

E2E network slicing allows differentiation of groups of users and will be further developed for RAN architectures based on 5G Core such as NR standalone. It will allow the radio access network to further facilitate differentiated handling of users. For RAN, a network slice provides additional information on the policy for handling traffic in addition to the regular QoS for radio bearers. A gNB may serve several network slices and each network slice may (or may not) contain users with different QoS requirements. RAN resources may be separated per network slice or shared between the network slices (e.g. spectrum). The decision on which resources to isolate and which to share between the slices impacts how the RAN requirements for the network slice can be fulfilled.

As further outlined below E2E network slicing will be introduced step-wise in NR-SA.

Requirements on RAN support for E2E network slicing are still being explored, functionality will evolve over time dependent on use case scenarios for network slicing.

The following key components provide support for E2E network slicing in RAN for NR-SA architecture options.

- Slice-aware CN instance selection
- Slice-aware performance monitoring
- Slice-aware resource management
- Slice Life Cycle Management

Slice-aware CN instance selection is about selecting the right core network instance for a particular UE, this can be done by using different mechanisms as shown in Figure 4-4:

- Based on PLMN ID like in legacy LTE for traditional network sharing cases.
- DECOR is the solution for Dedicated Core Network (DCN) that can be used for legacy terminals
- eDECOR requires new Rel-14 UEs that will convey the DCN-ID to RAN for core network instance selection

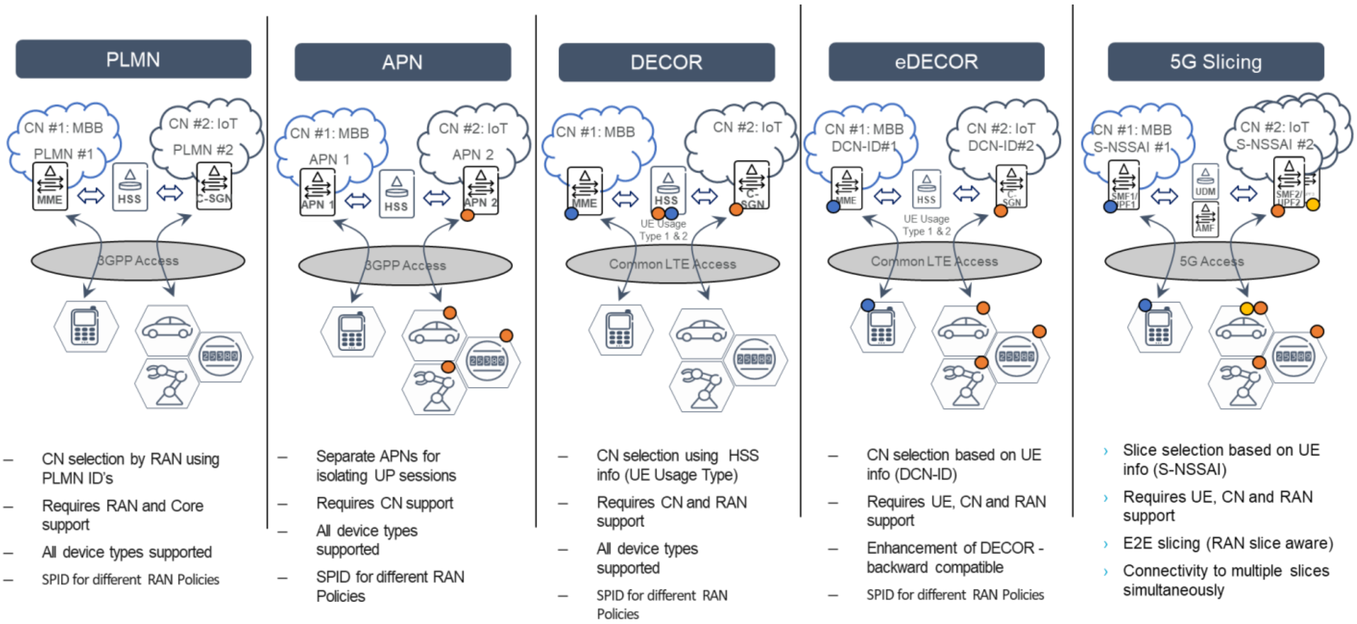

Figure 4-4: Slice-aware CN instance selection

For 5G slicing in Figure 4-4**,** S-NSSAI is the abbreviation for Single Network Slice Selection Assistance Information and identifies a Network Slice. The S-NSSAI consist of:

- SST: Slice/Service type, which refers to the expected Network Slice behaviour in terms of features and services (8bits); three standardized SST exist:
    - SST 1 = eMBB
    - SST 2 = URLLC (Ultra-Reliable Low Latency Communications)
    - SST 3 = Massive IoT

- SD: Slice Differentiator, which is optional information that complements the SST(s) to differentiate amongst multiple Network Slices of the same Slice/Service Type (24 bits).

Network Slice Selection Assistance Information, NSSAI denotes a collection of maximum 8 S-NSSAIs. Further details about S-NSSAI are available in D1.2 [3].

Slice aware performance monitoring is mainly about observability to support business management (SLAs) and resource management of a particular slice. In legacy LTE RAN there is the Flexible counter framework that can be used to instantiate KPIs and counters based on for example PLMN ID.

In order to be able to do adjustments in RAN for E2E SLAs, SLA observability/reporting is required on:

- PLMN level
- S-NSSAI level

When there is a need to also slice/partition the RAN network, observability on "RAN slice" level is also needed. The main idea is that similar framework for both observability and resource management can be used independent of the "slice identifier" being used. Slice aware performance monitoring is illustrated in Figure 4-5.

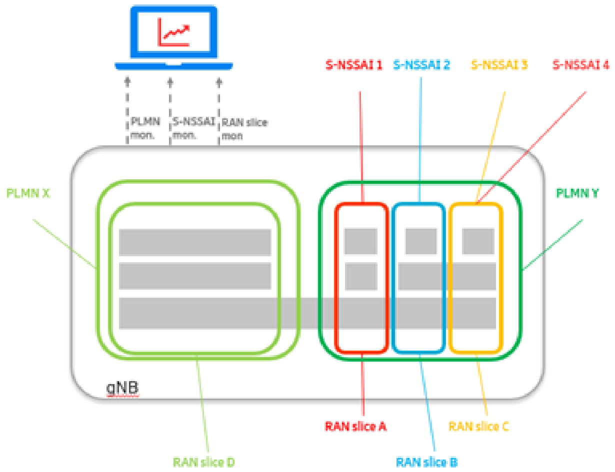

Figure 4-5: Illustration of slice aware performance monitoring

To fulfil an SLA, SLA reporting on a network slice is needed. The SLA reporting is done by measuring the slicing KPIs between the defined reference points PDCP to PDCP (Packet Data Convergence Protocol). Based on the envisioned use cases for slicing it is believed that the already existing KPIs on Integrity, Availability, Accessibility and Retainability can be used as baseline for slice-aware performance reporting. However, in order for RAN to support slice-aware performance monitoring, a framework for observability and slicing needs to be defined and should lead to a set of required RAN functionality. This is ongoing standardization work in 3GPP.

Slice aware resource management is about handling the trade-off between isolation and pooling of resources to provide different possibilities to give a group of users "special treatment". Isolation is needed to make sure that different slices do not impact each other, for the radio access it is particularly important in congestion situations caused, and resource pooling is most efficient due to multiplexing gains. Dynamic resource partitioning should be considered where resources are shared dynamically between all slices and yet certain resource levels can be reserved for each slice in case of contention.

A differentiation of user handling per slice should be made by mapping 5G QoS Indicator (5QI) to parameters.

Separation can be made by Instantiation/selection of virtual Packet Processor (vPP) resources per slice and/or instantiation/selection of UE-handler resources.

4.3 Wireline Access Network

There are multiple motivations for 5G wireline access network and Fixed-Mobile convergence (FMC), including the need for a seamless service experience independent from the type of access (i.e. fixed, wireless and cellular customer access) and the improvement of infrastructure efficiency through a common core network architecture and a common resource management. This topic has been addressed by 3GPP and BBF, from different perspectives, with the ultimate goal of

3GPP TR 23.716 [52], a Release 16 study on the Wireless and Wireline Convergence for the 5G system architecture, is aimed at enhancing the common 5G core network defined in TS 23.501 [53] and TS 23.502 [54] in order to support wireline access networks and trusted non-3GPP access networks.

The focus areas include the support of wireline access networks, the definition of architecture for the trusted non-3GPP scenario, as well as an investigation on whether enhancements are needed to interfaces (e.g. N1, N2 and N3) used to connect the trusted non-3GPP access network to converged 5G core; study the impact of the common framework for authentication and security, policy and QoS, Network Slicing and investigation on whether enhancements are needed. The support of CPE/RG capable of connecting simultaneously via both NG RAN and wireline access to 5GC is also under investigation.

Several scenarios are studied, taking into account different capabilities of the CPE/Residential Gateway, including the support of 5G interfaces.

This 3GPP study takes into account information provided by the Broadband Forum on wireless and wireline convergence. BBF has studied 5G FMC scenarios and provides recommendations on using a converged network for wireline subscribers and services [27]. The converged network includes a 5G core network, enhanced to support wireline access. Based on the review of 3GPP specifications TS23.501 [53]and TS23.502 [54], a number of issues have been identified in relation to the support of fixed access networks by the 5G core that potentially require additional specification work. Two (not mutually exclusive) converged scenarios are envisaged, as represented in Figure 4-6:

- Integration scenario – the converged 5G core network is used to deliver functions traditionally offered by the wireline core network. This scenario assumes some modifications to the Residential Gateway (5G RG) and a new function to mediate between the wireline access network and the converged core network, named 5G Access Gateway Function (5G AGF).
- Interworking scenario – the wireline core network continues to provide subscriber management and IP functions, but an interworking function (5G Fixed Mobile Interworking Function – 5G FMIF) enables some form of service convergence by linking the wireline core network to the converged core network.

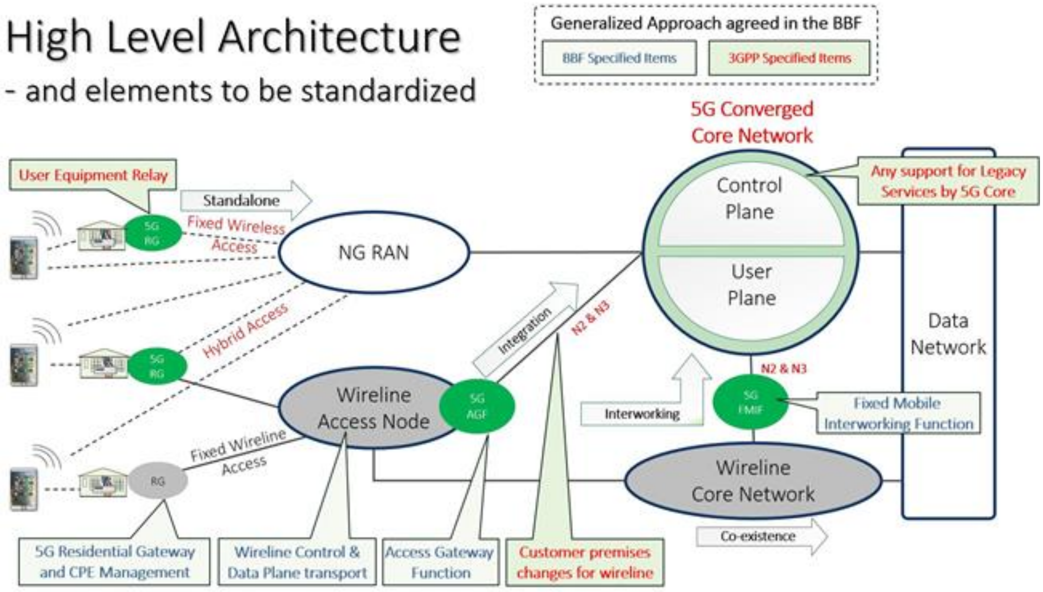

Figure 4-6: BBF 5G FMC architecture and elements to be standardized [55]

BBF aims at investigating common interfaces for the Access Network and Core Network, to support converged wireline-wireless networks that use the 5G core network. It studies the N1, N2 and N3 interfaces in order to provide detailed feedback to 3GPP about protocols, information models or procedures that need to be supported.

To adapt fixed access onto the 5G core, the project specifies a 5G AGF, and several architectural deployment options, including the underlying infrastructure sharing aspects. It also devises strategies and develops specifications to address operator requirements for interworking of existing fixed access subscribers and deployed equipment into a 5G core.

SD-420 includes the examination of a number of issues and aims at identifying respective solutions, as follows: Registration and Connection Management Procedures, Transport and Encapsulation in the Wireline Access of Control Plane (e.g. NAS) and User Plane Traffic Exchanged with 5G Core, Regulatory Requirements, Operational Requirements, Resource Management in the Access, Session Management, Addressing for IPv4 and IPv6, Home LAN Support, IPTV and Multicast, Network Slice Selection, QoS and Policy Management. These topics are currently under study.

4.4 5G Core

This section provides the network slice requirements towards NFs and the overall 5G Core Network. It should be noted that at this stage these requirements should be seen as guidelines mainly, not a firm commitment that they will be supported in a specific 5G-VINNI release. A network slice, depending on multi-domain and multi-tenant requirements, could include an appropriately split instance of the 5G Core Network, tailored for the specific use case or designed to accommodate multiple use cases. To achieve such a solution, the requirements to be considered are the following:

Virtualization: The 5G Core Network components shall support virtualization in order to be deployed in network slices in a dynamic and flexible manner. Albeit possible without virtualization, most of the advantages of slicing come from the possibility to run in parallel multiple customized network addressing the needs of the use cases. Depending on 5G use case and service requirements as well as component specifics the trade-off between deployment efficiency and performance has to be considered when choosing between VNFs and PNFs. (E.g. control plane components in environments with low expected load do not need the performance of specialized hardware, while UPFs processing high throughput user traffic do.)

Configurability: The 5G Core Network components shall provide high degree of configurability for addressing different use cases. By configurability it is understood not only having different values for the same parameters, but also being able to support subscribers with different classes of services e.g. QoS classes). It also presumes functional configurability, in the sense that multiple slices will have different functionality which is run, through this practically having different complexity in the network functions themselves across the different slices. With this, the deployed slices are different enough to motivate that two networks have to run in parallel.

On the 5G Core level configurability also includes support for configuring and enforcing slice and service level policies by translating them to network level choices and component level configurations. This type of functionality for instance could be provided through the Network Slice Selection Function and the Policy Control Function.

Programmability: To be able to quickly adapt to changing conditions and allow fast reaction to network events as well as to new requirement, the 5G Core system components have to be dynamically programmable. By this, it is understood that the components and component behaviours within a slice may change during the lifetime cycle of the service. This programmability is best achieved by the replacement of old components with newer versions responding better to the service requirements. During the upgrade or replacement procedure care has to be taken to ensure sufficient service continuity depending on the availability and reliability requirements of the hosting slice and served services. This includes the implementation of component-specific load and state transfer procedures.

Parallelization and scalability: A slice and the encompassed components should be able to scale depending on the momentary requirements of the subscribers as well as based on the predicted usage of the network. (e.g. in case the management of the network functions is becoming aware that at a specific time a higher level of resources is needed this should trigger scaling operations proactively instead of waiting reactively for the load to increase. This is highly important in IoT use cases where a very high number of devices may communicate a large amount of data at fixed moments in time. The 5G Core has to implement the load balancing methods on component level, depending on the functionalities and supported procedures of specific components. Based on the isolation capabilities supported by the infrastructure, the 5G Core system shall support parallel deployments on multiple slices. By this, it is understood that multiple core networks running in different slices are able to use in parallel the same infrastructure resources as well as to integrate transparently with the applications and radio components.

Infrastructure agnostic: To enable deployment in every possible scenario regardless of the underlying hardware resources, 5G Core VNFs are expected to be implemented in an infrastructure agnostic manner. This presumes that they are capable of using the available virtual resources, which are offered by the local system without major performance penalties (e.g. components should be compiled for the specific virtual CPUs within the offered operating system for a good performance. Alternatively, the VNFs are deployed with their preferred virtual CPUs and Operating System on top of generic resources).

Isolation: NFs are required to implement the necessary capabilities to allow isolation of NF instances belonging to different slices as well as the isolation of information belonging to different slices in cases when the same NF is used by multiple slices. Isolation is related to the privacy of the communication within a specific slice such as the private information in the slice is not visible to other slices or to the infrastructure providers. Similarly, by isolation it is understood that the owner of a slice will perceive its own slice as a complete and independent network abstracting the underling resources. Furthermore, the isolation presumes that the resources allocated to a slice are carefully protected by the infrastructure provider as to avoid side effects such as noisy neighbour where parallel slices are racing for infrastructure resources, resulting into unwanted performance related side effects.

Slice stitching: For the cases when the merging of slices or inter-slice communication is to be enabled, 5G NFs are required to possess the ability to open communication channels to respective components of other slices as well as to support synchronisation of relevant information. The slice stitching presumes also the maintenance of the privacy of the internal information within a slice i.e. protecting the privacy of the deployment of a slice from the other slices to which it is stitched to. In certain cases, e.g for an SMF-SMF interface, inter-slice communication could be considered a roaming-like behaviour and handled accordingly.

Management and orchestration: Connected to configurability and programmability, 5G Core components are required to expose the appropriate northbound interfaces to allow the selected MANO solutions to access or influence slice deployment and runtime configurations. Furthermore, the components within the slice should have their own management system executing the FCAPS operations as well as any OAM operations. The slice management components should be able to interact with the selected MANO of the domain where the slice is deployed in order to transmit indications on scaling and network function placement as to be able to serve the service within the SLA requirements. Furthermore, the MANO components have to possess sufficient authority and view of the network to be able to make slice and policy selection decisions and enforce them through the corresponding lower level management components (e.g NSSF or PCF).

Monitoring and discovery: To enable more informed automated or human decision making, the 5G Core components are required to advertise certain state information and expose certain runtime metrics towards the management layer. The monitoring should include not only the 5G core specific metrics but also at least the access to the infrastructure metrics in order to be able to correlate the service SLAs with the reserved and the used resources. This enables the management of the system to make appropriate decisions concerning resources consumed and to request scaling operations when needed.

Multi-tenancy: For the cases when the same network slice is deployed as a multi-tenant environment, the 5G Core components have to be able to isolate tenant-specific information and processing either as a capability of a single NF instance or implemented using multiple NF instances.

Network function exposure: Certain 5G Core NF functionalities have to be available for external applications, taking into account the network slice the application is attempting to use. Exposure is possible through the network exposure function, and the Network Exposure Function (NEF) in the appropriate slice is to be selected for the specific application scenario. In dynamic deployed slices, the NEF functionality should be also dynamically exposed to the interested third parties in order to be able to use the specific communication.

Authentication and Authorization: Connected to function exposure, the 5G Core NFs or the slice management entities have to implement the necessary procedures to authenticate users and applications and authorize only limited access to resources based on their ownership and access rights in case of external or internal access attempts. At the same time the 5G Core has to provide its functionalities in a transparent manner towards devices. For instance, a device does not need to know whether the AMF (Access and Mobility Management Function) is in a common slice or not.

Reliability: The 5G core network components should reach the reliability expected from the slice owner. This presumes that the service will remain reliable towards the subscribers no matter on the existing or lacking reliability at the infrastructure level. A slice can benefit from infrastructure level reliability, however it cannot assume that it will be available in all systems where it will be deployed. Reliability in addition, includes the requirement towards each 5G core component to be able to accurately interpret, represent and act on the trust, security, isolation, etc… requirements requested by the MANO layer based on the slice and service specifications.

Automation: Connected to monitoring and management, the 5G Core components may implement interfaces to support automation, by exposing real-time management functionalities, which makes it possible to execute specific actions on the component automatically (e.g. life-cycle management or low level policy changes). This includes specifically operations related to fault, performance and security. Entities driving automation could reside on different levels of the network and execute actions of different complexity ranging from slice-level actions initiated by slice management functions to NF-level actions initiated by element managers.

- in case of a failure, the management of the network functions should be able to transmit indications to network functions as well as to MANO components on how to adapt the system to maintain the SLA levels towards the subscribers;
- in case of a performance decrease, the management of the network functions should be able to determine from the monitoring the performance decrease and to trigger actions within the network functions as well as to the MANO in order to adapt the system to the expected performance level
- in case of a security attack, the management of the network should be able to isolate the attack or the compromised network functions and to maintain the function of the system towards non-malicious users.

4.5 Transport Network

The transport network has a critical role to play in 5G by providing the backhaul and new fronthaul /midhaul infrastructure required to enable the stringent requirements, such as high bandwidth and ultra-low latency. 5G, and particularly network slicing, raise new requirements to the transport network: isolation, automated lifecycle management, customization, adaptability, programmability.

4.5.1 Transport Network Management

Although the Transport Network (TN) is out of 3GPP scope, it represents a crucial component of a network slice. In order to guarantee end to end SLAs and KPIs, the transport network plays an important role and needs to be sliced as part of services bound to the different slices. Figure 4-7 illustrates the potential roles of TN in the context of an NSI.

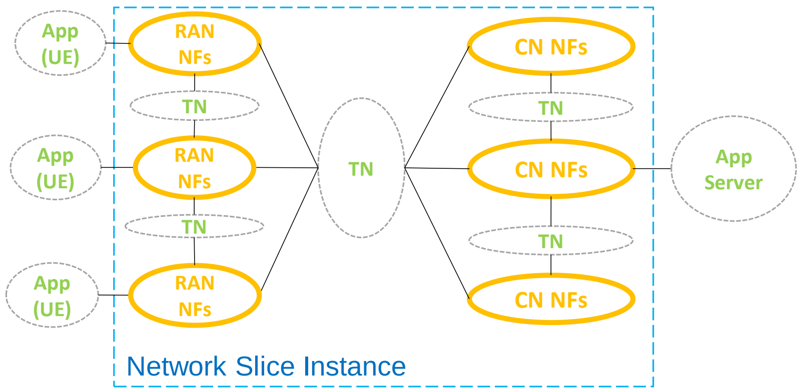

Figure 4-7: Example of an NSI [6]

Figure 4-8 illustrates the relationship between the 3GPP management system and the transport network segment, as defined by 3GPP in [6]. As reported in section 2.2.5, the BBF has been particularly focused on the interface between the 3GPP Network Slice Management Function and the Transport Network Slice Management. The management aspects of the network slice instance are represented by the management of the CN and AN parts (directly managed by the 3GPP management system) and the management of non-3GPP part (not directly managed by the 3GPP management system), which includes the TN, providing connectivity within and between CN and AN segments.

Also, the ITU-T Technical Report GSTR-TN5G “Transport network support of IMT-2020/5G” [56] defines requirements on 5G transport networks and specifically on the relevant interfaces (e.g. F1, Fx, Xn, NG). The support of IMT 2020/5G network slicing (data plane and control plane) and Control/Management interfaces are also addressed.

The 3GPP management system maintains the network topology and related QoS requirements and is supposed to request the configuration of Managed Network Slice Subnet Instances (MNSSIs) to support several 5G services (e.g. uRLLC, eMBB, etc.) and define the characteristics (e.g. bandwidth, type of isolation, latency) of each MNSSI required to address the needs of each 5G service. The 3GPP management system identifies the requirements on involved network domains, including RAN, CN and non-3GPP parts of a slice by deriving them from the customer requirements to the services supported by the network slice. The derived requirements are sent to the corresponding management systems. The coordination may also include related management data exchange between those management systems and the 3GPP management system.

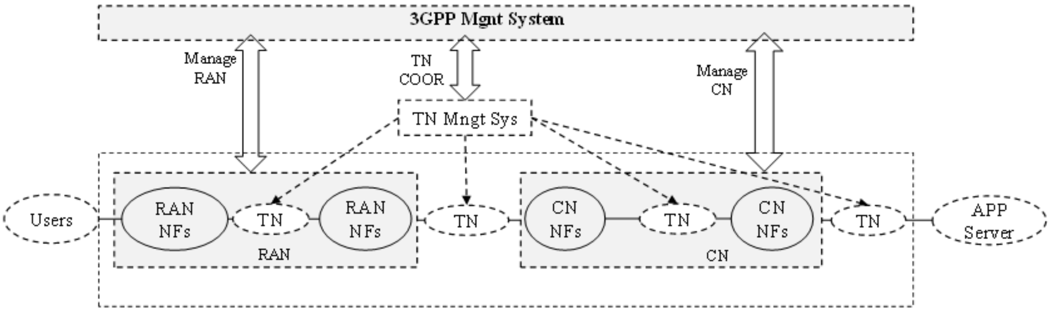

Figure 4-8: Coordination between 3GPP and TN management systems [56, 6]

Therefore, in order to ensure the performance of a communication service according to the business requirements, the 3GPP management system has to coordinate with the management systems of the non-3GPP parts (e.g., TN, MANO system) when preparing a network slice instance for this service. This coordination may include obtaining capabilities of the non-3GPP parts and providing the slice specific requirements and other requirements on the non-3GPP parts.

The BBF, focused on requirements for end-to-end network slicing, including the identification of umbrella use cases and relevant business entities, uses the model described in 3GPP TR28.801 [17] (as described in Section 2.2) as starting point and further develops the bottom layer, as illustrated in Figure 4-9. The CSMF and the NSMF follow the TR 28.801 definitions, whereas the NSSMF, as seen by 3GPP, is split in 3 separate entities, namely the Access Network Slice Management (ANSM), which takes care of the slice life-cycle management of the access NSSI, the Core Network Slice Management (CNSM), which takes care of the slice lifecycle management of the core NSSI and the Transport Network Slice Management (TNSM), which takes care of the slice lifecycle management of the transport network NSSI and provides the capability exposure of the transport network towards the network slice management function, while it also provides the mapping of the 3GPP mobile network requirements to the corresponding transport network.

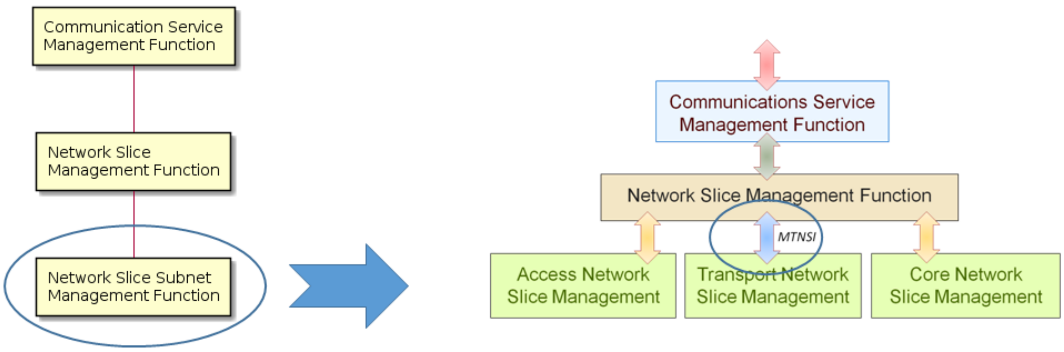

Figure 4-9: 3GPP network slice management functions and BBF focus

The Mobile-Transport Network Slice Interface (MTNSI) handles the transport network capabilities exposure towards the NSMF, as well as slice requests received from the NSMF, indicating the required parameters such as latency, delay variation, loss ratio, maximum and minimum bit rate, etc. The MTNSI interface is also responsible for performing the network slice mapping procedures from the 3GPP mobile network towards the transport network, in a sequence of basic steps, as illustrated in Figure 4-10.

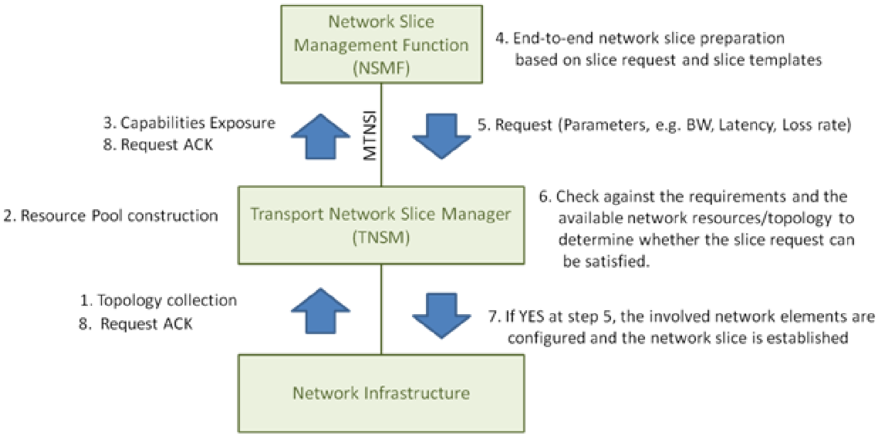

Figure 4-10: NSMF - TNSM mapping procedures

The BBF describes the network slice management operations over the MTNSI interface identified in Figure 4-10 and the parameters exchanged in each phase of the life-cycle management process, namely the Instantiate phase, the run-time phase and the terminate phase.

4.5.2 gNB centralized deployment

In 5G, the transport network has gained an increasing prominence as it provides the connectivity between the components of the disaggregated RAN, with a direct impact on performance and cost. Several factors have contributed to intensify this trend, including the expected increase of traffic load in 5G, driven by increase of capacity of the radio access network and the growth of density of connected devices, as well as the virtualization and centralization of gNB functional components as illustrated in the Figure 4-11 [57].

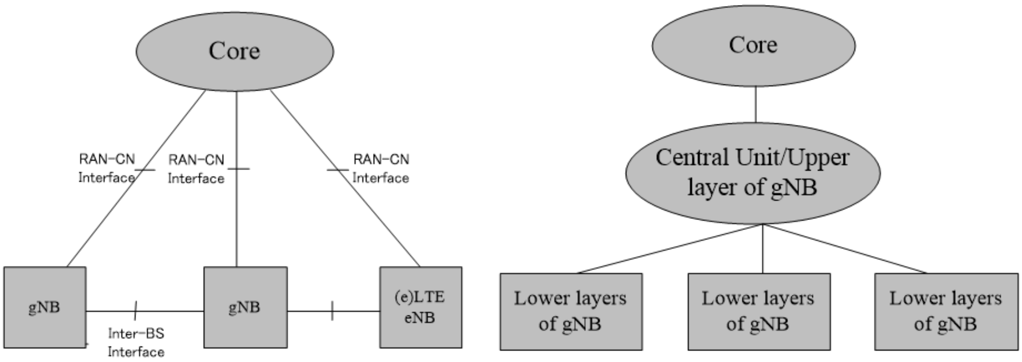

Figure 4-11: Non-centralized vs. centralized RAN deployment [57]

In 5G, the gNB functionality is split into three functional modules hosting signal processing functions, which can be deployed in multiple combinations: the centralized unit (CU), the distributed unit (DU) and the radio unit (RU), the combination of which correspond to what was previously known as BBU and RRH. As a result of this functionality split, two new RAN transport segments should be considered: the "fronthaul" and the "midhaul" networks. As shown in Figure 4-12 [56], the fronthaul network connects the RU to the DU (if located at different sites, i.e., when the two are not co-located at the cell site), the midhaul connects the DU to the CU and the backhaul connects the CU (usually co-located with edge computing components) to the core network.

Unlike 4G, where the backhaul represents the bulk of the access and aggregation networks, in the 5G disaggregated architecture, the transport role is played by the fronthaul and midhaul segments to a large extent. Consequently, in 5G, the transport network may contain fronthaul, midhaul and backhaul networks. However, different deployment scenarios can be used, as follows:

1. Independent RRU, CU and DU locations. In this scenario, there are fronthaul, midhaul and backhaul networks. The distance between an RRU and DU is in the range of 0-20 km while the distance between the DU and CU is up to tens of km.
2. Co-located CU and DU. In this scenario, the CU and DU are located together, consequently there is no midhaul.
3. RRU and DU integration. In this scenario, the RRU and DU are deployed close to each other, typically hundreds of meters, for example in the same building. In order to reduce cost, an RRU is connected to a DU just through straight fibre and no transport equipment is needed. In this case, there are midhaul and backhaul networks.
4. RRU, DU and CU integration. There is only backhaul in this case. This network structure may be used for small cell and hotspot scenarios.

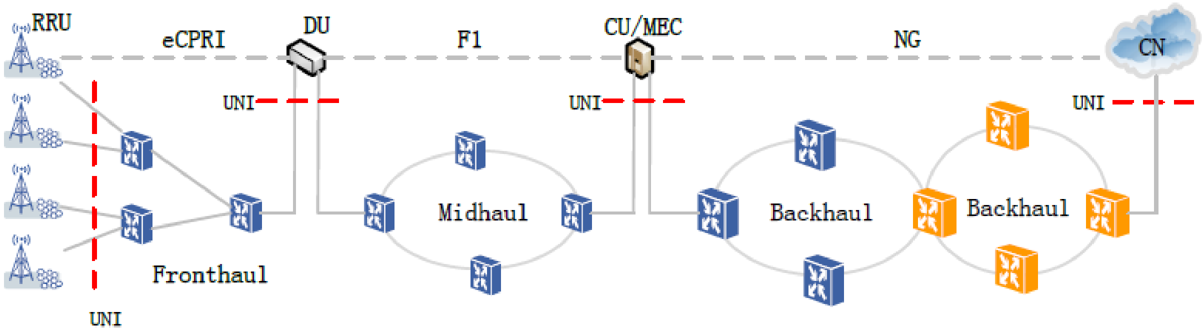

Figure 4-12: Transport network architecture for independent CU and DU deployment [56]

For each of these network segments, multiple options are available for the physical layer, e.g. dark fiber, active wavelength division multiplexing (WDM), next-generation passive optical network (NG-PON), or wireless.

Network slicing brings an additional layer of complexity to this equation, by requiring one physical network to support several logical networks with potentially very heterogeneous requirements on top. For example, eMBB requires high capacity, i.e. downlinks rates up to 1 Gbit/s. In mMTC applications data rates to individual sensors are usually very low (in the order of kbit/s), but the number of connected devices may scale into extremely large numbers. Finally, uRLLC mission-critical and extreme precision applications may require end-to-end latency around 1 ms, jitter below 1 µs, and reliability in the order of six nines.

In terms of RAN disaggregation, a key issue to be taken into account is the functional splitting of the functions of the protocol stack, as illustrated in Figure 4-13 [56]. The bottom part of the figure represents the possible options for the location of split points between functions in 5G and how they compare to the approach followed in 4G. Essentially, three options are available for 5G – a) high layer split (CU/DU F1 interface); b): low layer split (DU/RU FX interface); c) cascaded split.

The optimal choice for functional split should take into account the trade-offs between throughput, latency, and functional centralization. Different functional splits have different pros and cons.

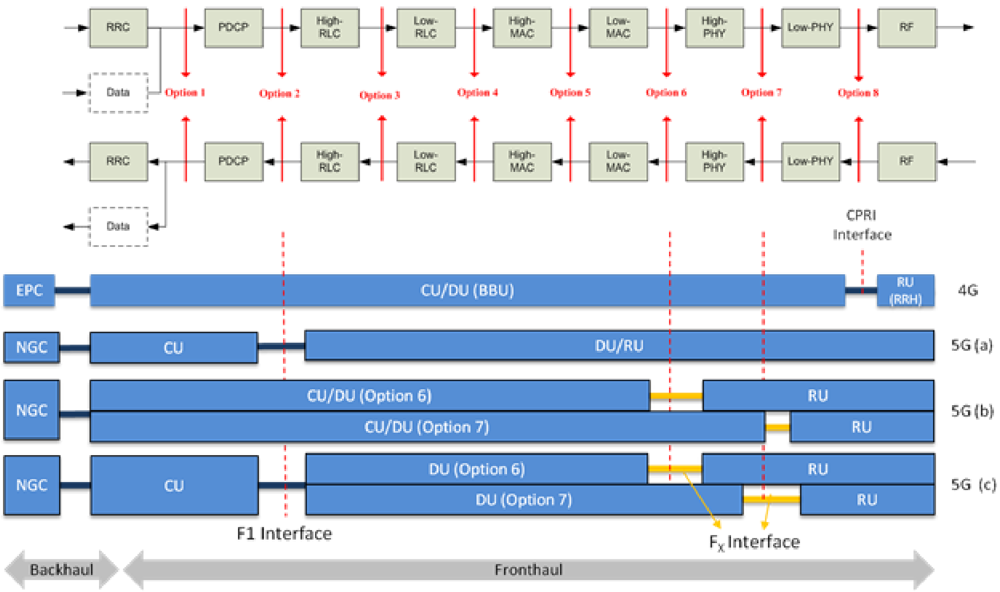

Figure 4-13: C-RAN functional split [56]

In the following, a brief overview of the challenges related to each of the three transport network segments is provided, including those related to network slicing.

4.5.3 Fronthaul, Midhaul and Backhaul

The performance of a fronthaul network depends on whether a circuit-switched or packet-switched solution is used. The fronthaul is traditionally based on circuit-switched Common Public Radio Interface (CPRI), but the use of packet switching (e.g. Ethernet) has gained increasing acceptance by the industry, particularly in scenarios that require network virtualization and efficient use of resources. On the other hand, a major constraint to be taken into account in the fronthaul is the high sensitiveness to latency. The fulfilment of strict timing and synchronization requirements is especially challenging for packet-based fronthaul implementations. In [58], the challenges of Cloud-RAN (C-RAN) in relation to network slicing are analysed, with a particular focus on the potential advantages of packet-switched fronthaul.

The functional split choice directly affects the bit rate in the fronthaul segment (the more to the right on Figure 4-13 the higher the bit rate) and whether this bit rate is constant or variable with the traffic load, depending on where in the physical layer the PHY split is located. Traditional fronthaul using Option 8 (CPRI) requires continuous bitrate transport regardless of the volume of user traffic. With a higher layer split the latency and bandwidth requirements can be relaxed, but the price to pay is the lower number of processing functions that can be centralized, which means reduced economies of scale.

Due to cellular coordination requirements, the physical separation of RU and DU requires a low-layer functional split, with low-layer PHY processing at the RU and high PHY and all higher-layer processing at the DU. Standards are being developed to make packet networks suitable for fronthaul latency and synchronization restrictions, as reported in Section 4.5.4.

From a network slicing perspective, functional splits that enable a variable fronthaul bitrate and packet-switched solutions are advantageous. In terms of the slice types, eMBB benefits from a variable bitrate on the fronthaul link. As huge amounts of data are expected, centralized processing is beneficial to enable a more efficient allocation of resources. In [59], three C-RAN functionality split options (PDCP-RLC, MAC-PHY, and intra-PHY) are analysed in terms of their impact on delivering uRLLC traffic, particularly in relation to latency and jitter. The results show that the MAC-PHY split is the most suitable split option for uRLLC as being able to guarantee lowest delay and jitter due to the synchronization needed between MAC and PHY layers.

Compared to fronthaul, the higher-layer functional split used by midhaul and backhaul allows a higher tolerance to delay. In addition, network slicing requires a single physical network to support several logical networks with different requirements, which means that a single physical network must be prepared to handle extreme scenarios.

4.5.4 Data Plane isolation

As mentioned before, the ITU-T Technical Report GSTR-TN5G [56] defines requirements on 5G transport networks, including the support of IMT 2020/5G network slicing. The main requirements are related to the need to provide isolation between different 5G transport network slice instances, in such a way that the requirements of multiple verticals or business customers can be satisfied.

The forwarding plane must ensure isolation between virtual networks, of which two different types are defined, hard isolation (the traffic loading one VN has no impact on the traffic in any other VN, including QoS effects) and soft isolation (The traffic loading of one VN may have an impact on the QoS provided to the traffic in other VNs), as described in [56] and also reported in [3]. For either hard or soft isolation, a number of solutions are available for the midhaul and backhaul segments. Hard isolation can be guaranteed through independent circuit switched connections (e.g. dedicated wavelength, dedicated TDM time slot) for the exclusive use of a single virtual network. Resources including routers, control planes, and links may be physically partitioned, if needed. For soft isolation, resources are shared. Using traffic statistical multiplexing of two or more virtual networks using packet-based technologies (e.g., Ethernet VLAN, MPLS tunnel) is usually the solution. The mutual impact on QoS of virtual networks sharing the same infrastructure may be mitigated by traffic engineering including, for example, limiting the statistical multiplexing ratio, traffic policing on each virtual network.

The use of network slicing implies that the data plane would conform to the desired SLA requested by a service provider, vertical or 3rd party. As it has already been stated in Section 4.2, QoS transport traffic management in combination with bearer QoS is essential in order to create E2E network slices. To achieve this, UPF and 5G RAN perform transport level packet marking in DL and UL respectively, e.g. setting the DiffServ Code point in outer IP header [53]. Other identifiers such as P-bits, VLANs, EXP field etc, can also be used in order to differentiate traffic.

In the context of transport networks, network slicing can be supported with technologies which provide network virtualization. Network virtualization is really referring to the creation of virtual topologies and sometimes these virtual topologies are overlays, sometimes they are forms of multiplexing, and sometimes they are a combination of the two. Tunnel overlays over public Internet are, by definition, a best effort service, as the data network has no visibility over the end-to-end connection requirements. For this reason, this solution is not amenable to provide any type of performance or reliability guarantees, let alone isolation, to traffic belonging to different slices. However, in practice, this kind of solution is often considered good enough, especially in the cases where very strict performance/reliability is not required. Slicing can also be provided with VPNs, including BGP/MPLS VPN (e.g. RFC 4364 [60]) or MEF Ethernet-based service (e.g. MEF E-Line, E-LAN [61]). Performance and reliability levels can be expressed by specific SLA service attributes. The automated management and control of the service lifecycle are usually beyond the scope.

According to ACG Research who conducted an independent primary research to identify the trends [62], opportunities and challenges in 5G mobile transport, the main technologies to, are layer 2 and layer 3 VPNs, Segment Routing, EVPNs, OTN slicing, FlexE and wavelength slicing and are shown in Figure 4-14.

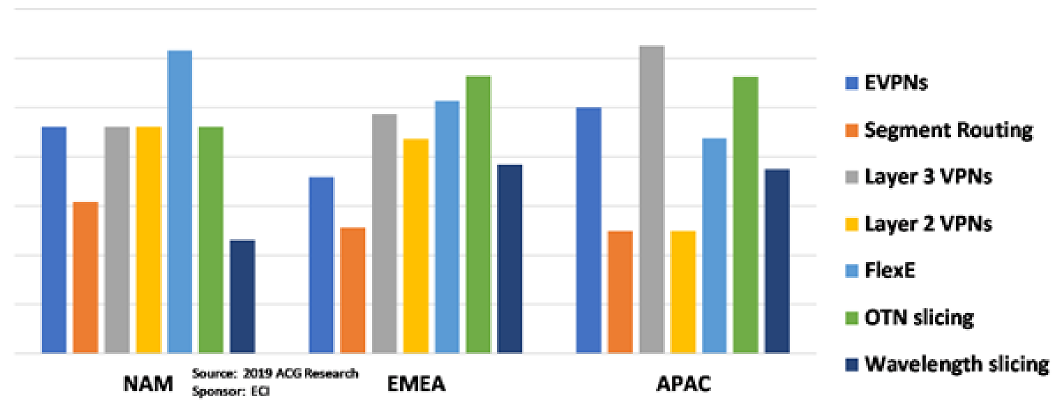

Figure 4-14: Top Transport Technologies to Implement Network Slicing [62]

Some parts of emerging 5G services, e.g. uRLLC with stringent latency requirements, may not though be supported in an efficient way by the current VPN/VLAN technologies. New emerging solutions can be used to optimize the performance and isolation solution in the transport network.

On one hand, link layer encompasses L1.5/L2 technologies that provides isolation in the data plane. Examples of link layer solutions that can be used in search of the desired pragmatic isolation in 5G-VINNI facility TN include:

- *Flex Ethernet (FlexE)*. FlexE [63] emerges as a new transport technology that creates a PtP Ethernet with a specific fixed bandwidth. Basically, FlexE decouples the MAC binary rates from the underlying PHY rate thanks to the insertion of a new, thin layer called FlexE shim between traditional MAC and PCS layers of Ethernet, as originally defined in IEEE 802.3. Standardized by the Optical Internetworking Forum (OIF), the FlexE shim provides the ability to multiplex multiple asynchronous channels (i.e. FlexE channels, assigned to different virtual links) in the MAC layer over one or more synchronous PHY layer in a way that provides high degree of isolation on interface level, allowing separation between different channels on a shared interface. Depending on how MAC layer speed is related to PHY layer speed, different schemes can be used for the time-based scheduling, including (i) bonding, (ii) sub-rating, and (iii) channelization.
- *Dedicated Queuing.* The use of dedicated queues seeks for steering traffic to dedicated input and output queues to avoid negative performance influence between competing virtual links. With sophisticated queuing systems (e.g., split of one physical interfaces into multiple virtual sub-interfaces, each with dedicated queueing and buffer resources), this approach can solve the scalability issues presented by conventional DiffServ based queuing system, as defined in IETF RFC 2475 [64] and IETF RFC 4594 [65].
- *Time Sensitive Networking (TSN)*. TSN [66] is an IEEE project within the IEEE 802.1 aiming at designing a method to carry time sensitive information over Ethernet. TSN allows differentiation between high priority packet streams (i.e. packets from a delay sensitive application) and low priority packet streams (the rest of packets), applying packet scheduling over the former, giving them a scheduled time slot, thereby assuring they experience no queuing delay, and hence a reduced latency. However, when no scheduled packet arrives, its reserved time slot is handed over to low priority packet streams, thereby improving the network economics [67]. The use of TSN for slicing in the TN domain helps in the task of scheduling resources across virtual links from different NSIs when a uRLLC NSI comes into play, allowing the former to achieve its strict latency demands.

On the other hand, network layer deals with the problems of differentiation and resource representation of virtual networks in the overlay. Examples of network layer solutions that could be used in 5G-VINNI facility TN include:

- *Multiprotocol Label Switching - Traffic Engineering (MPLS-TE).* MPLS-TE allows the reservation of E2E bandwidth for a Traffic Engineering – Label Switched Paths (TE-LSPs), thereby ensuring throughput guarantee for the different virtual links, each running over a dedicated TE-LSP.
- *Deterministic Networking (DetNet).* Discussed in [68], DetNet technology provides deterministic communication on selected flows, allowing them to be delivered with extremely low data loss and bounded end-to-end latency. The QoS in DetNet can be expressed in terms of (i) upper bounds on latency and jitter, (ii) packet loss ratio, and (iii) tolerance for out-of-ordered packet delivery. To satisfy these QoS dimensions, different mechanisms can be applied, including congestion protection, service protection, and explicit routes. The use of DetNet technology for slicing in TN domain allows configuring non-overlapping resource between virtual networks from different NSIs, representing a virtual network providing connectivity to a NSI with deterministic behavior (e.g. uRLLC NSI) as a set of DetNet flows.
- *Segment Routing (SR).* SR [69] is a technology that provides packet instructions at the entry node and optionally at some intermediate nodes, allowing the packets to be routed on alternative paths (strict or loose paths, different from shortest paths) for various TE goals, while removing per-path state. In [70] a number of potential advantages are identified, including the possibility to leverage network slicing, improved scalability characteristics, simplified automation and easy SDN integration.
- *Enhanced Virtual Private Networks (VPN+)* [67] is aimed at optimizing packet-based transport networks to handle network slicing requirements, particularly a degree of isolation and performance that traditionally can only be satisfied by dedicated networks. Unlike traditional VPNs, an enhanced VPN is supposed to achieve greater isolation with strict guaranteed performance. The solutions proposed in [67] can be applied to a wide range of L2/L3 VPN services and are supported by mechanisms in the data plane (e.g. FlexE, TSN, Dedicated Queues, DetNet, Segment Routing), control plane and management plane.

Other solutions can be considered, such as the adoption of Hierarchical QoS (HQoS) [71]. HQoS uses multiple levels of queues to further differentiate service traffic, and provides uniform management and hierarchical scheduling for transmission objects such as users and services.

4.5.5 Control/Management Plane

SDN enables the smooth transforming of several transport networks into a single end-to-end fully automated network. Transport networks have features usually not present in computer networks where the SDN paradigm arose, like resilience (protection or restoration mechanisms to assure Service Level Agreements, sometimes implemented in coordination between data and control plane), more sophisticated architectures and heterogeneous technologies that need to be taken into account when applying the concepts introduced by SDN. In this sense, transport SDN (T-SDN) is an extension of SDN introducing different abstractions, interfaces, protocols, and control plane elements to cope with transport networks peculiarities [72]. The cross-layer interaction of the SDN controller is enabled by the north-bound interfaces to the orchestration and management plane and by the south-bound interfaces to the network devices.

As it has already been stated in D1.1 [2], a single SDN controller, comprising multiple network nodes featuring diverse technologies, provided by different vendors, is not realistic. The SDN controller supports the communication with network devices per transport domain, with different protocols such as NETCONF, PCEP, P4Runtime, OpenFlow, CLI and SNMP. These protocols provide a set of basic functions such as collection of network attributes, collection of events such as link failures or status changes and the configuration of network devices.

Furthermore, the WIM, as introduced in D1.1, allows service orchestration by integrating the wireless, optical and satellite backhaul domains with the orchestration layer, thus allowing efficient cross-domain network slicing and service automation. The interface of the SDN controller with the WIM and the overall orchestration system consequently, is based on REST APIs that support CRUD (Create, Read, Update, Delete) operations on network-related resources (Figure 4-15).

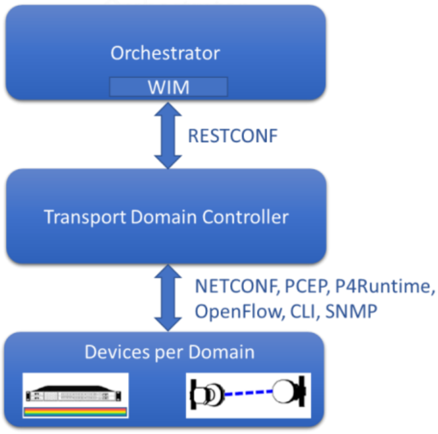

Figure 4-15: Transport SDN controller NBIs and SBIs

4.5.6 Satellite Transport Network Slicing

Terrestrial and satellite networks today are managed by independent systems, i.e. each segment holds its own OSS/BSS and NMS systems, although several attempts for system integration have been made in the community for different aspects. When using SDN technologies, resource abstractions have enabled initial wired / wireless full convergence both in operation and management components in access network. In parallel, the virtualization of specific network functions provides a cloud-like view for the management system, adding both the traditional NMS and the lifecycle service and orchestration functionalities associated to the original cloud-based system. Meanwhile, orchestrators like OSM or ONAP provide a solution for management and orchestration of the network resources and services.

Network Slicing is a key enabler for the future 5G networks, which involves several domains (5G core network (CN), Satellite network, RAN and UE) and implies a cross-domain resource slicing handling. Figure 4-16 depicts these domains, which have to support network slicing from the satellite perspective.

This section is focused on network slicing in the case where satellite is used as transport network.

Figure 4-16: Cross-domain slice

4.5.6.1 Satellite as Transport Network

A satellite system can be used as a transport network (TN) within the 5G network in order to provide connectivity between areas. The satellite system does not necessarily need to be managed by the same entity, which manages the 5G network and is responsible for the 3GPP management system. The transport network can have its own dedicated management system.

In that sense, two TN-based implementation options have been specified regarding the provided flexibility and the level of exposure of interfaces at different planes between the 5G network and the satellite as transport network, in particular the management plane and the control plane [73], [74]. These two implementation options, which are both considered in 5G-VINNI, are:

- Satellite Transport network based on 3GPP system specifications as shown in Figure 4-17;

- Satellite Transport network not based on 3GPP system specifications as shown in Figure 4-18.

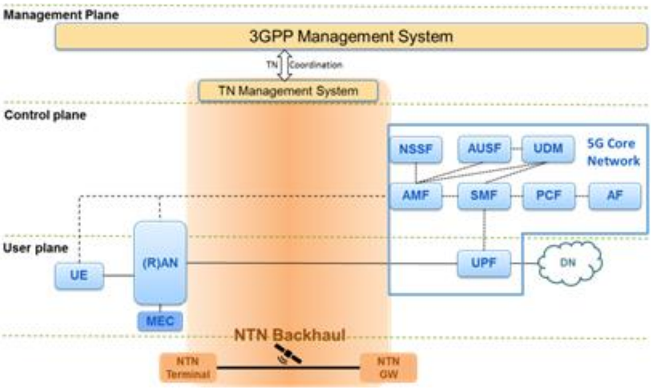

Figure 4-17: Satellite transport network based on 3GPP system specifications [75]

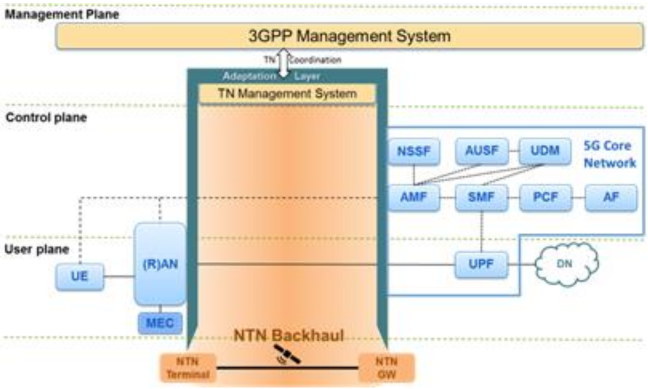

Figure 4-18: Satellite transport network not based on 3GPP system specifications [75]

In the above figures, NTN stands for Non-Terrestrial Network, which includes Satellite Networks and/or High Altitude Platform Systems (HAPS). In this document, the focus is only on Satellite Networks and so, NTN and Satellite Network terms can be used interchangeably.

4.5.6.2 Management and orchestration in satellite transport networks for Network Slicing

The management of network slicing is key issue for the satellite integration in 5G systems, to be supported by NMS. In the frame of SaT5G, two NMSs are envisaged: 3GPP NMS and SatCom NMS. 3GPP NMS is foreseen to manage UE, (R)AN and 5GC; whereas SatCom NMS is foreseen to manage the satellite system. However, depending on the implementation option performed, the management architecture can be provided through two distinct ways [76]:

- 3GPP NMS only: In this case, the 3GPP management (e.g. the MNO) also manages the SatCom system. This case is illustrated in Figure 4-17 above.
- 3GPP NMS + SatCom NMS: this case is essentially the transport network case where a Satellite Network Operator (SNO) manages the SatCom system accordingly to the requirements (i.e. slice requirements) provided by 3GPP Management (e.g., the MNO). This case is illustrated in Figure 4-18 and is further elaborated in Figure 4-19 below.

These ways of network management lead in any case to the definition of slice expressed by means of the slice requirements. Some slice definition criteria for integrated satellite terrestrial 5G networks are provided in [76].

Figure 4-19: Slicing management in satellite transport network [76]

If satellite is used as transport network supporting the mobile network, the link limitations could have a big impact on the performance of the mobile network. 3GPP TS 28.533 [49] has defined an example in the informative annex, which shows that there could be a TN Domain manager that communicates directly to the 3GPP management system. However the TN Domain manager itself is out of scope of 3GPP – this will be handled in more detail in Section 4.5. The interface between 3GPP management systems and the TN Domain manager (see Figure 4-20) should be in scope of 3GPP, but currently this has not been addressed in 3GPP. This interface is particularly relevant for the scenarios in which the satellite network is used as the transport network for backhaul (as described above) where a close coordination between the satellite management system and the 3GPP management system is required to provide functionalities.

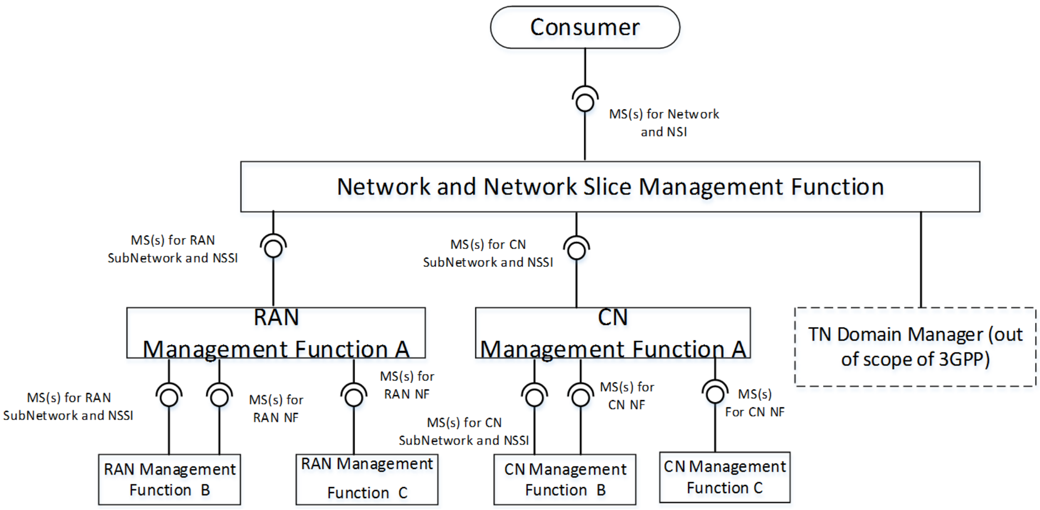

Figure 4-20: A possible interaction with a TN manager [49]

Reaching across a network to provision an end-to-end service is certainly challenging in any network with diverse control across a multi-layer and heterogeneous service network. However, in satellite networks there is the constant challenge of variable performance and reliability combined with greater latencies of GEO and Non-GEO satellite links that can add hundreds of milliseconds of latency. This creates exceptional challenges for the management and control planes because critical control operations can be interrupted by intermittent link availability. Centralized Orchestration or SDN Controller architectures are vulnerable to a high frequency of command exchanges across the space segment. Furthermore, synchronized message exchanges suffer on high latency and variable availability links. These demanding operating conditions need a resilient and recoverable control architecture that provides durability in the face of variable performance that a physically centralized architecture does not sustain. A distributed and possibly federated SDN Controller architecture puts greater intelligence and autonomy at the edge beyond the space segment, minimizing the dependency and use of space segment resources using distributed intelligence to maintain a rich control interface with local devices where terrestrial networks are typically available [77]. Figure 4-21 illustrates specialized controllers for various layers and segments of the network where each controller can support either the centralization or distribution of network control functions. These functions can typically be implemented as virtualized network functions that are placed at the network edge in devices that support a virtualization infrastructure.

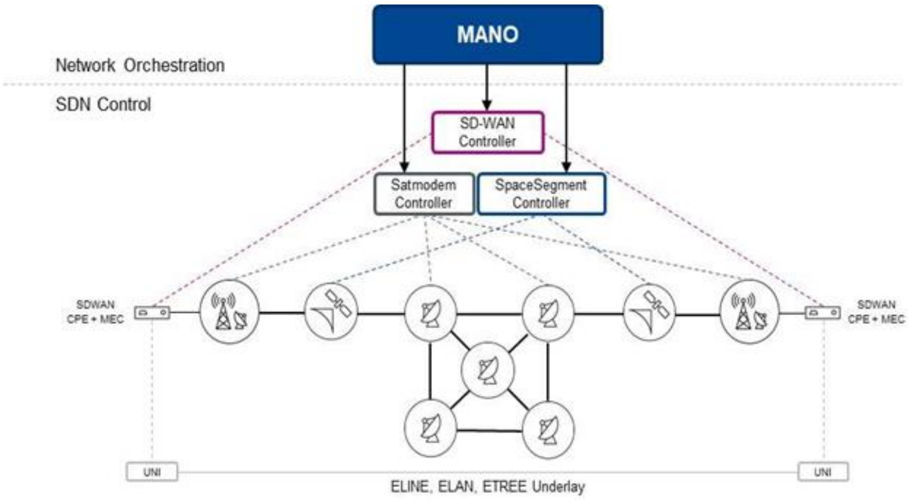

Figure 4-21: Distributed SDN Controller Architecture [77]

Finally, a reference VNF data model presenting the satellite communication systems' capabilities based on the orchestration framework defined in ETSI and adapted to 5G by 3GPP is currently being investigated in the ETSI Work Item [78]. As shown in Figure 4-22, the objective of this reference VNF data model is to allow 3rd party (e.g. MNO) to orchestrate any satellite communication system along with any other access network. This will enable the network slicing concept over satellite transport networks.

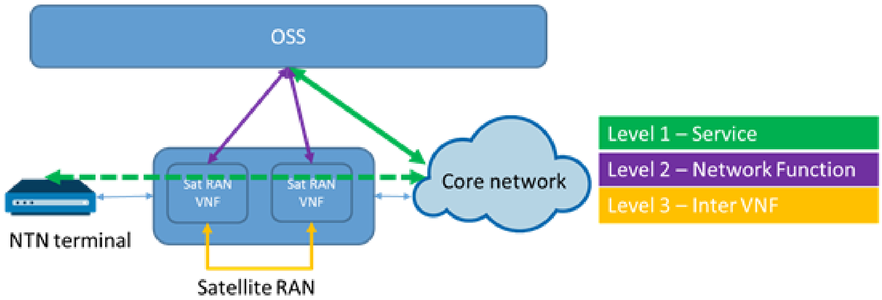

Figure 4-22: Enabler for network slicing in a satellite communication system [78]

4.6 Data Network

Although 3GPP usually views the data network as an external domain, thus outside the scope of specifications, it is clear that an end-to-end view of the network infrastructure requires the data network segment to be taken into consideration whenever the communication path extends beyond the 5G core domain (i.e. beyond the N6 interface). The same applies to network slices that contain components located beyond the data network segment. From the user equipment to the server, an end-to-end communication path supported by a network slice will involve multiple infrastructure domains, including the Data Network. This is the case whenever a network slice is composed of resources (e.g. applications, network functions) that are reachable through one or multiple data network domains. A typical example is a network slice composed of resources that belong to multiple mobile network domains (e.g. a slice owned by a vertical spanning multiple operator domains).

Figure 4-23 illustrates the protocol stack for the user plane transport. Basically, it is an extension of the model described in [53], by adding the data network segment on the right hand side.

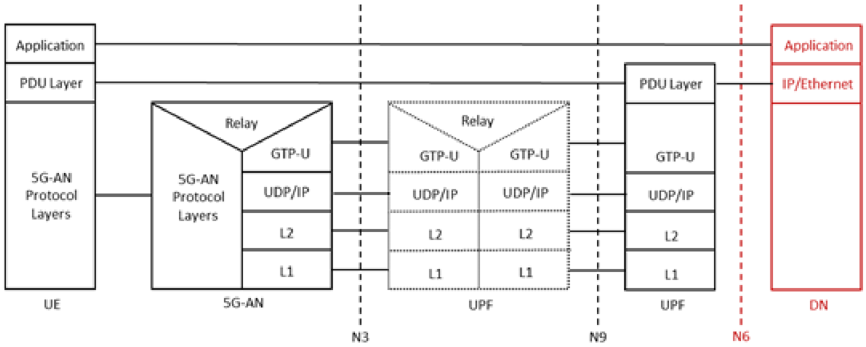

Figure 4-23: User plane protocol stack

As integral part of a network slice, the data network component must be able to fulfil any applicable requirement. The network slice data plane should comply with the applicable SLA.

The concept of building multiple virtual networks over a common shared infrastructure is not new and has been part of different service models based on technologies such as MPLS or Carrier Ethernet. However, the strict isolation requirements required by network slicing raise new challenges. The solutions available for slicing the data network should be essentially the same as those already presented in relation to the backhaul segment, as discussed in section 4.5. When multiple administrative domains are involved, the scenario complexity may grow, if the solutions proposed to tackle network slicing requirements (e.g. FlexE, TSN, DetNet, Segment Routing) are not easily deployable in a multi-domain environment.

4.7 User Equipment

In general, the topic of UE is outside of the scope of 5G-VINNI. However, as this will be an indispensable component to connect end users in a practical scenario, it will be briefly analysed in this section. The device can be authenticated and attached to a diverse set of network slices for specific applications. A given device can access different slices at the same time, for different application scenarios for the user equipment. 3GPP defined that a given device can support up to eight different slices with a common AMF for all the slices and a SMF per slice [79].

In the SBA, the UE is connected to AMF through the N1 interface. The UE provides in a Registration Request the requested NSSAI. The network verifies the Requested NSSAI against the Subscription Information [53]. When the UDM updates the Subscribed S-NSSAI(s) to the serving AMF, based on configuration in this AMF, the AMF itself or the NSSF determines the mapping of the Configured NSSAI for the Serving PLMN and/or Allowed NSSAI to the Subscribed S-NSSAI(s). The serving AMF then updates the UE with the above information, through UE NSSAI configuration and NSSAI storage.

UE Network Slice configuration [6]

The Network Slice configuration information contains one or more Configured NSSAI(s). The UE may be pre-configured with the Default Configured NSSAI. The UE may be provisioned/updated with the Default Configured NSSAI. When providing a Requested NSSAI to the network upon registration, the UE in a given PLMN only includes and uses S-NSSAIs applying to this PLMN. Upon successful completion of a UE's Registration procedure over an Access Type, the UE obtains from the AMF an Allowed NSSAI for this Access Type, which includes one or more S-NSSAIs and, if needed, their mapping to the HPLMN S-NSSAIs. These S-NSSAIs are valid for the current Registration Area and Access Type provided by the AMF the UE has registered with and can be used simultaneously by the UE (up to the maximum number of simultaneous Network Slices or PDU Sessions).

Update of UE Network Slice configuration [6]

At any time, the AMF may provide the UE with a new Configured NSSAI for the Serving PLMN. The AMF can provide an updated Configured NSSAI. A UE for which the Configured NSSAI for the Serving PLMN has been updated and has been requested to perform a Registration procedure, shall initiate a Registration procedure to receive a new valid Allowed NSSAI.

If the UE receives indication from the AMF that Network Slicing subscription has changed, the UE locally deletes the network slicing information it has for all PLMNs, except the Default Configured NSSAI (if present). It also updates the current PLMN network slicing configuration information with any received values from the AMF.

5 Network slicing management interfaces

5.1 Interfaces for slice lifecycle management

The interfaces exposed externally for life cycle management are presented in multiple SDO venues: TMF, GSMA, ETSI ZSM as well as 3GPP SA5.

GSMA NEST (Network Slicing Task Force) focuses on the design of a model-based template network slices called General Service Template (GST). GST aims at bringing automation in the lifecycle management of end-to-end NSIs scoping eMBB, mMTC and uRLLC service types, from their deployment to their decommissioning, so underlying multi-SDO management systems can operate NSIs as required in an agile and flexible manner, with zero-touch fulfilment. Currently, NEST has proposed a (non-public yet) information model for GST, and send it to 3GPP to get some feedback via a liaison statement.

Both ETSI ZSM and 3GPP SA5 architectures are based on the concept of management services. Thus, lifecycle management is done by existing management services within the management domain (core, RAN, TN or E2E) as presented in Section 3. The idea is that instead of re-creating and re-implementing interfaces exposed externally for lifecycle management, there should be a way to manage the management services and the respective management data that can be accessed externally via an API. In TS 28.533 [49], 3GPP proposes a management function that can perform exposure consuming internal management services and re-exposing them to external consumers, as shown in Figure 5-1.

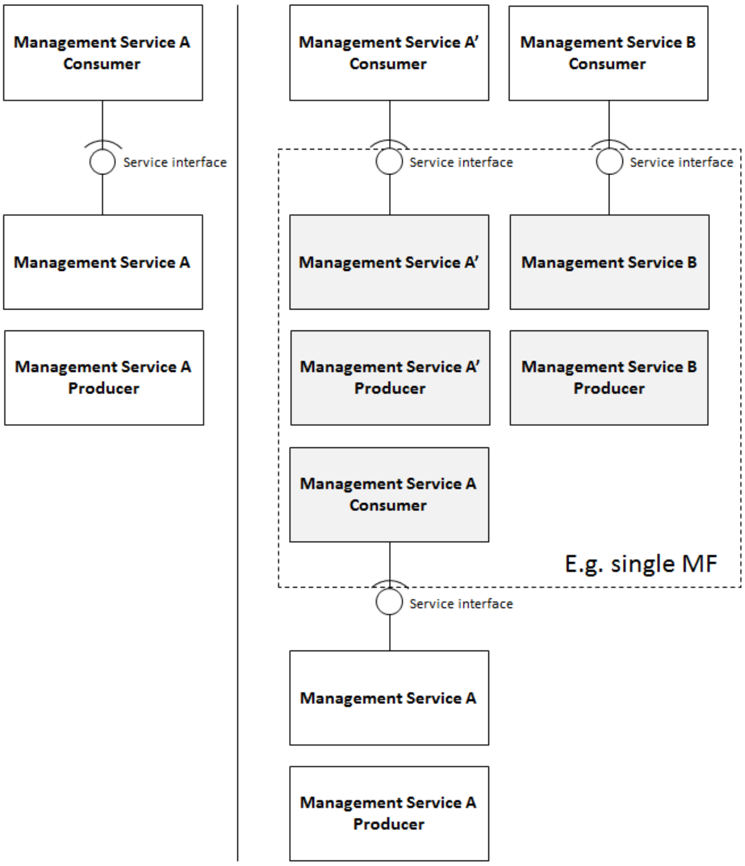

Figure 5-1: Exposure as conceptualized in item 4.4 of TS28.533 in 3GPP SA5 [49]

Similarly, ETSI ZSM is in the process of discussing (proposed by 5G-VINNI partners) an Exposure management service as shown in Figure 5-2. The steps in Figure 5-2 can be explained as:

Precondition: the operator using the exposure service configures which authorization is able to access which service from the management domain.

1. In step 1 the exposure service can register services/capabilities/end points/management data (via the registration service) including those for lifecycle management with the integration fabric;
2. The service consumer discovers management services offered by the MD (via the Discovery service);
3. The service consumer may then request for service/capability use arrives, example a provisioning service (via the Broker service);
4. The exposure service verified if show a request is accessible to the said consumer, if yes, management service access is granted.

In conclusion, the API exposed from an operator for lifecycle management are a subset of the APIs the operator has internally for lifecycle management. The most important of those are specified by 3GPP SA5 in the provisioning Work Item Description (WID) in TS 28.531 [7].

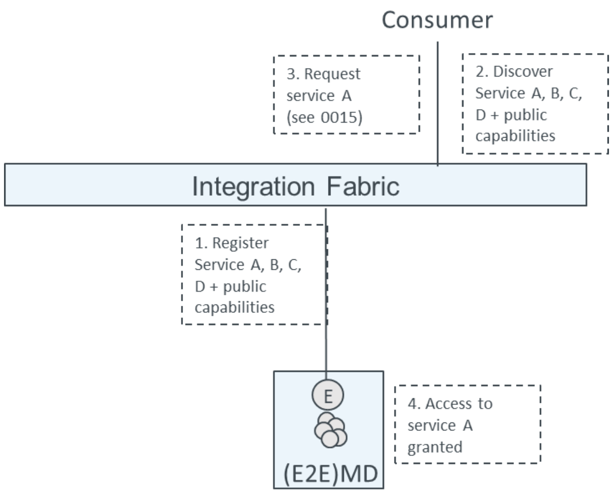

Figure 5-2: Concept under discussion in ETSI ZSM on exposure of management services

5.2 Interfaces for intra-slice management

The performance and functionality of an NSI depends on the configuration and resources assigned to the CP and UP NFs building up the NSI. From the perspective of NSI application (related to NSI functionality), these NFs are 3GPP RAN NFs and 3GPP CN NFs, arranged into one RAN-NSSI and one CN-NSSI, respectively. In some cases, NSI could include non-3GPP NFs providing value-added functionality (e.g. AF, MEC platform). From an NSI deployment viewpoint (related to NSI performance), these NFs can be deployed as PNFs and/or VNFs, arranged into one or more NFV-NSs. Figure 5-3 provides an example of the internal composition of an NSI.

The expected behavior of an NSI is defined by the NSMF, with the specification of the desired NSI state. This state can be defined by the combination of the states and attribute values of the NSI's constituent NFs (and their possible relationships with other NSIs, e.g. in NF sharing scenarios) that make NSI to behave as expected in terms of functionality and performance.

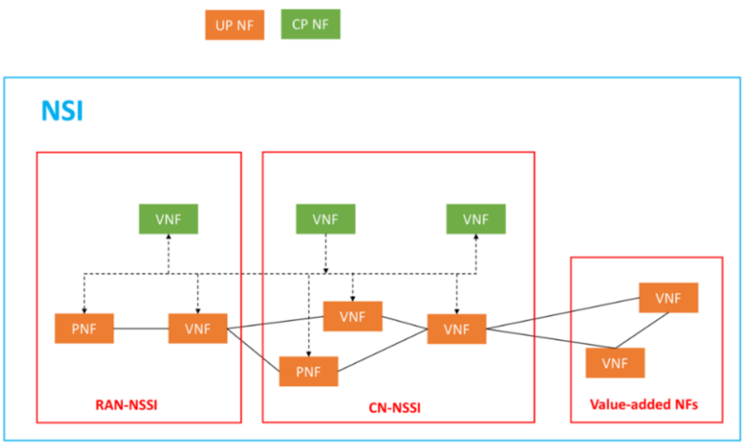

Figure 5-3: Internal view on a NSI

Intra-slice management defines an environment where management functions take control of the in-NSI CP and UP NFs, configuring and dynamically operating them in an optimal manner, leading the NSI to the desired state in force at any time. For this end, feedback control is key, in which three main steps are involved:

- Calculate the actual NSI state. For this end, performance and fault management data from NSI's NFs (and their underlying resources) shall be first collected and then analyzed.
- Calculate delta between actual and desired states for the NSI.
- Determine which modifications are required in the NSI to make the delta as close as possible to zero. For this end, this delta shall be compared against an optimization policy.

The feedback control as the core functionality of intra-slice management is aligned with ONF TR-521 (SDN Architecture Issue 1.1) [80] (see Figure 5-4), where it is claimed that the essence of an SDN controller is its participation as the active entity in a feedback loop.

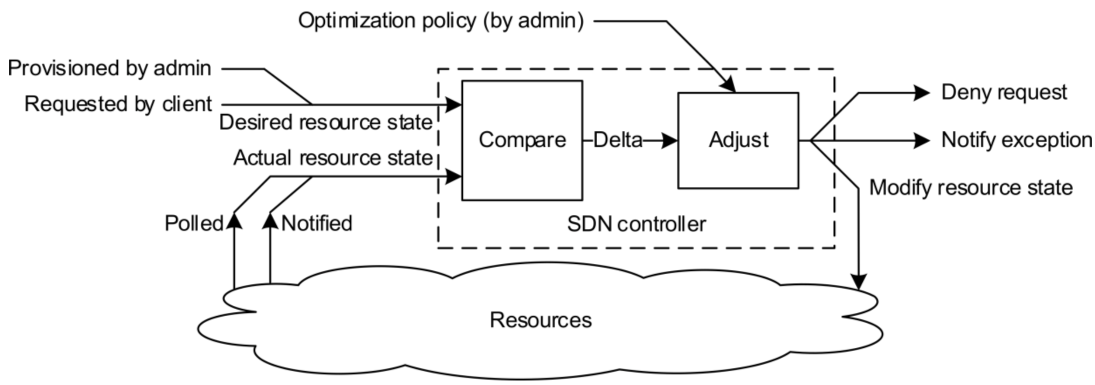

Figure 5-4: Control as feedback [80]

To keep the NSI in the desired state throughout its lifecycle, modifications within NSI could be required at multiple layers. In 5G-VINNI system, intra-NSI management can be executed by:

- NSSMF from the 5G-RAN controller, if the required in-NSI modification means modifying one or more UP/CP NFs from the RAN NSSI at the application level. In such a case, NSSMF exposed interfaces [7] can be used for intra-NSI management.
- NSSMF from the CORE controller, if the required in-NSI modification means modifying one or more UP/CP NFs from the CN NSSI at the application level. The interfaces are the same as described above but applied to the CN domain.

- Transport controller, if the required in-NSI modification means modifying connectivity across the two NSSIs. In such a case, Transport controller exposed interfaces (e.g. BBF's MTNSI [27] – see Section 4.5]) can be used for intra-NSI management.
- NFVO, if the required in-NSI modification means modifying one or more UP/CP NFs from any NSSI at NFV level (e.g. scale NFV-NS, scale VNF). In such a case, NFVO exposed interfaces [81] can be used for intra-NSI management.
- VIM, if the required in-NSI modification means modifying the infrastructure providing NSI virtualized resources. In such a case, VIM exposed interfaces [82] can be used for intra-NSI management.

In 5G-VINNI, intra-NSI management is a task usually shared between the roles of NOP/CSP (i.e. 5G-VINNI facility) and CSC (i.e. vertical customer). The CSC's degree of implication in intra-NSI management activities depends on the set of exposed interfaces that the CSC can consume, being this set determined by the selected exposure level. The relationships between exposure levels defined in D3.1 and intra-NSI management capabilities granted to CSC are summarized in Table 5-1.

Table 5-1: CSC-driven intra-NSI management and exposure levels

Exposure Level	Exposed interfaces made available to the CSC	Description
Level 1	NSMF exposed interfaces	The CSC is unable to carry out intra-NSI management
Level 2	NSSMF and Transport controller exposed interfaces	The CSC is able to only carry out intra-NSI management at the application level, but not at the NFV level.
Level 3	NSSMF, Transport controller and NFVO exposed interfaces	The CSC is able to carry out intra-NSI management at both application and NFV levels.
Level 4	NSSMF, Transport controller, NFVO and VIM exposed interfaces	The CSC takes full control of intra-NSI management activities

In 5G-VINNI facility, intra-NSI management could be effectively handled in most cases by applying the approach described so far. However, there might exist some scenarios requiring NSIs with very particular requirements in their internal management. For those cases, an NSI could be optionally provided with a dedicated, customized management stack. Examples of management functions that could take part in this stack are NFMFs (3GPP), VNFMs (NFV) and a tenant controller (SDN). Described in [83], the tenant SDN controller allows on-demand programmability in NF composition. Figure 5-5 illustrates an example of in-NSI management stack involving the three types of management functions.

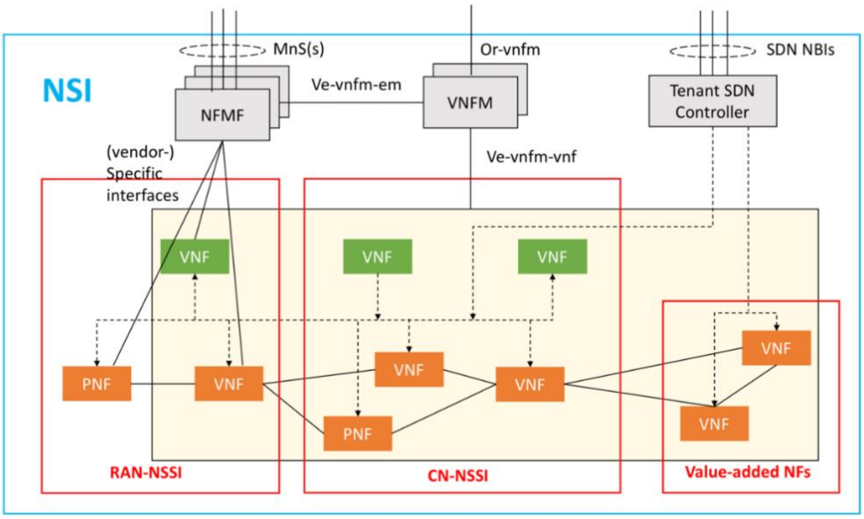

Figure 5-5: In-NSI management stack

5.3 Inter-domain interfaces

In 5G-VINNI it will exist two different types of inter domain interfaces, based on the two different domain concepts: First, Network Domains that make reference to RAN, Core, and transport domain. Second, Administrative domains that refer to each of the 5G-VINNI facility sites. 5G-VINNI inter-domain interfaces must be implemented in such a way that achieve interoperability especially since the facility sites adopt solutions by multiple vendors under different licensing schemes (e.g. Open Source). According to the 5G-VINNI architecture, these interfaces are those defined between the E2E SO&M and the NFV MANO, RAN, Transport and core controllers. Once an NSI is created, there will be a set of NSSIs linked (which can be limited per administrative domain), Radio, Core or Transport. The Service Provider can optionally choose different Vendors to cooperate, in addition slices may involve more than one administrative domain, which can be accessed by a multi-domain orchestrator for CFS (Customer Facing Services) and RFS (Resource Facing Services). Based on the classification previously presented, how those interoperable interfaces will look like according to the structure provided in Figure 5-6 will be described here.

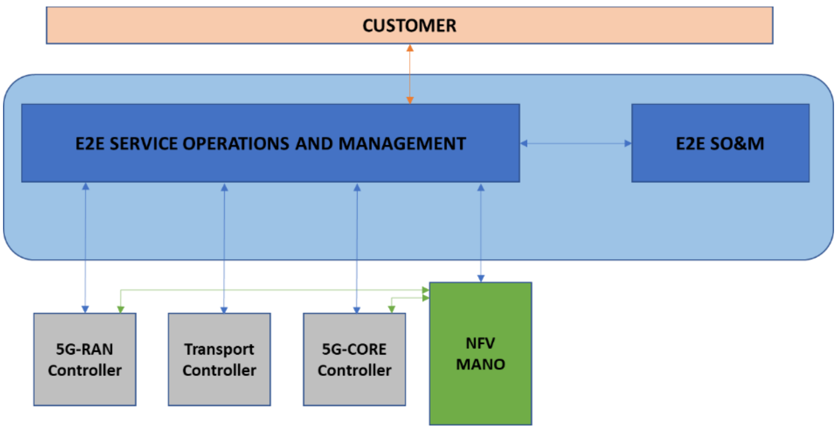

Figure 5-6: 5G-VINNI Inter Domain Interfaces

5G-VINNI needs to adopt and implement inter-domain interfaces that achieve interoperability, since multiple solutions will be deployed on each facility site. Thus, the adopted technologies must be open and based on well-known standards. Most noted efforts are the following:

ETSI and the SOL specifications having OSM implementing most of these specifications.

The interaction with the Transport and virtual domain is achieved via the following interfaces.

- OSM Or-Wi Interfaces: This interface is used in the context of Software-Defined Network Platforms, each managed by a WAN Infrastructure Manager (WIM) (often a SDN Controller). The WIMs would expose the reference point Or-Wi northbound to provide connectivity as a service among the VIMs (even if they required inter-datacentre connections). This reference point would be accessible for OSM to serve to the NS/NSI creation and lifecycle.
- SOL005 and NetSlice NBI: OSM's NBI provides a superset of ETSI NFV SOL005 API calls with the addition of E2E NS operation capabilities and the ability to handle Network Slices. Alike the IM, the latest official version of OSM's NBI is openly available in OpenAPI format, and can be used as the authoritative reference for interoperability northbound.

TMF and its defined Open APIs having ONAP and various vendors implementing them

The TM Forum is working actively on building a view of the current OSS/BSS solutions evolution, and making sure that it integrates and supports existing standard and supporting released frameworks (e.g. ONAP). TMF states and emphasizes that the way to create an ecosystem which can be used by Partners and Service Provider and allow different Vendors to cooperate is to develop such an ecosystem based upon an ODA (Open Digital Architecture). Therefore, their main investment in terms of industry is to provide and push an alignment among industry standard through the Open API Layer, where specification for each API Rest and respective information framework (SID Service Model) can be found. There are components from the TM Forum blueprint which are relevant for the inter-domain interfaces:

- TMF ecosystem supports the creation of network slices, and subsequently the new use cases brought by 5G technology (eMBB, mMTC, uRLLC)
- TMF also invests on several catalysts like ZOOM (Zero-touch Orchestration, Operations and Management)
- MEF and TM Forum are teaming up with multiple service providers working to standardize LSO application programming interfaces (APIs) to orchestrate services across multiple networks. TMF Open APIs may be used to deliver MEF orders to Service Providers

TMF defined two type of services which can be accessible through the APIs in OpenAPI layer and can be used on a network slicing offering and for inter-operability use cases:

- CFS (Customer Facing Service): Services which can be obtained by an end-user/Customer (e.g. Software, Hardware, Service). It can be used to operate the network slice lifecycle
- RFS (Resource Facing Service): Services which access resources (e.g. network elements, IT Systems, transport/core/radio domains) , required by a CFS.

The following CFSs are available, for inter administrative domain interfaces, and can be considered as the most promising specifications, because those are being used across several frameworks (e.g. ONAP):

- TMF633 Service Catalogue Management API [84]
- TMF641 Service Ordering API [85].

MEF Lifecycle Services Orchestration Architecture

From the MEF side, the interaction with the network domains is specified in a more generic way, assuming general domains and exposure capabilities and interaction between the Orchestration functionality and the Infrastructure and Element Control and Management as follows:

- Presto (SOF:ICM) Service Orchestration interface to Infrastructure Control and Management (ICM) of each domain needed to deliver a Connectivity service for instance. This is another interface which has a long list of Interface Profile Specification (IPS). ICM provides the information related to resources and capabilities that can be used in a service. The IPS are under discussion, among MEF members, which is the reason why the current APIs available may change until those are "officially" released by MEF

- Adagio (ICM:ECM) This interface manages directly the resources, and the directly related management functions, which are available from the resource layer on a network slicing. There is active contribution from MEF member, reason it's expected to see different APIs until the IPS are officially released.

At the moment, there are not widely used standards in the direction of inter administrative domain interfaces. However, the INTERLUDE interfaces provided by MEF are one of the most solid approaches at the moment. It is important to highlight that this kind of interfaces still remains as one open challenges, and hence the importance of beginning some test in this regard.

- Interlude interface (SOF:SOF): inter domain management reference integration point of LSO between a partner domain and the service provider orchestration functionality. It covers the operations that are defined in the policies and linked to Services (it can be seen as the RFS). This a very complex integration point, because it includes a long list of interface profile specifications which provides an outline of the use cases, and leads to API development

By using a model-driven approach along with abstractions representing products, services, and resources, LSO ensures an agile approach to streamlining and automating the entire service lifecycle in a sustainable fashion. LSO orchestrates connectivity services across all internal and external network domains from one or more network operators, including all communications service providers, data centre operators, enterprises, wireless network operators, virtual network operators, and administrative domains supplying or consuming components of the service. LSO encompasses all network domains to provide coordinated end-to-end management and control of connectivity services.

5.4 Network slicing information model

As mentioned by Y. Tina Lee in [86], an information model can be defined as follows: "A representation of concepts, relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse". Currently, there is a huge amount of emerging technologies, global competition and product diversification, even in just one single project; this raises a potential and clear need of new ways for information transfer in an efficient, reliable and time-adjusted way.

It is important to remark that there is the need to use a protocol-neutral language such as *YAML or UML* (some examples will be described below) in order to develop an accurate IM. Furthermore, apart from considering what an Information Model (IM) is and its need of usage, it is vital to comprehend and link together this model with the current needs of the telecom industry, operators and SDOs, which, in summary, are to converge into a single and standardized IM, at least, on service orchestration and NFV/SDN management. As shown in Figure 5-7, currently this convergence does not exist:

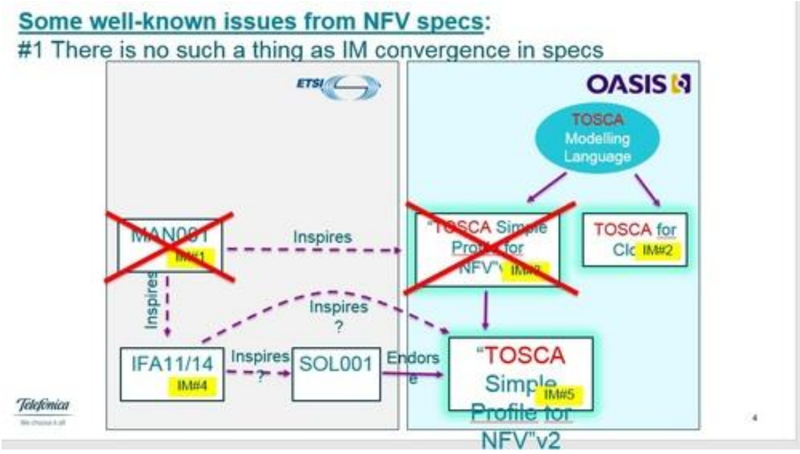

Figure 5-7: Multiple IM un-convergence [87]

This situation results in a scenario where different vendors produce new IM or new IM versions, which tend to be without any interoperability or compliance between them, lacking a clear path to progress the IM. There are a couple of SDOs working and developing a common IM for different areas, the best and most advanced examples are mentioned below:

- **3GPP:** TS 28.541 [10] specifies the IM and Solution Set for the Network Resource Model (NRM) definitions of NR, NG-RAN, 5G Core Network (5GC) and network slice with the objective of fulfilling the requirements identified in TS 28.540 [9]. In this IM model the protocol-neutral language selected has been UML; for instance, in 3GPP TS 38.401 [88], NR node (gNB) is defined to support three functional split options, so that in the appropriate IM (NR NRM IM) various UML diagrams show the relationship between each gNB split option and the corresponding Information Object Class (IOC). Although it has a limited content, mainly about 3GPP standards such as TS 38.401 [88], TS 28.622 [89] or TS 28.708 [90] related concepts, it could be a good starting point to define the bases of the 5G-VINNI IM.
- **ETSI-NFV:** ETSI NFV ISG has built an information model, introduced in NFV IFA 015 [91], based on information elements developed in other specifications such as NFV IFA 011 [92] (VNFD Information Elements) and NFV IFA014 [93] (NSD IE), among others, translating them into a full-UML NFV IM (for more information about UML Modeling Guidelines see NFV IFA 017 [94]). Unlike 3GPP's IM, NFV's IM aims at providing a complete and consolidated view of the IE that are part of the interface specifications, mainly from a deployment point of view. This IM follows a clear model Structure organized in an NFV Core Model (main functional and basic components) and extensions modules which broaden the NFV Core Model for concrete requirements. In Figure 5-8, the extension module NFV Interface IM expanding the NFV Core Model is shown:

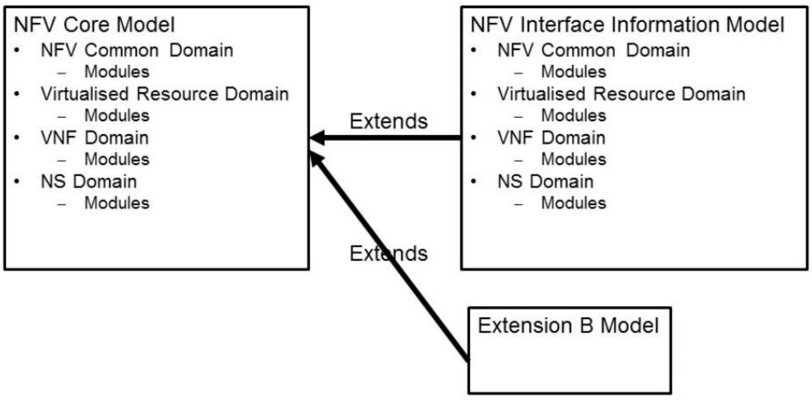

Figure 5-8: NFV Core Model extensions

Currently, each model is structured in four domains and each domain is composed of the following modules: NFV Common Domain, Virtualised Resource Domain, VNF Domain, NS Domain.

Figure 5-9 and Figure 5-10 show a logical and deployment overview of the VNFD and NSD structure IE (Information Elements) respectively:

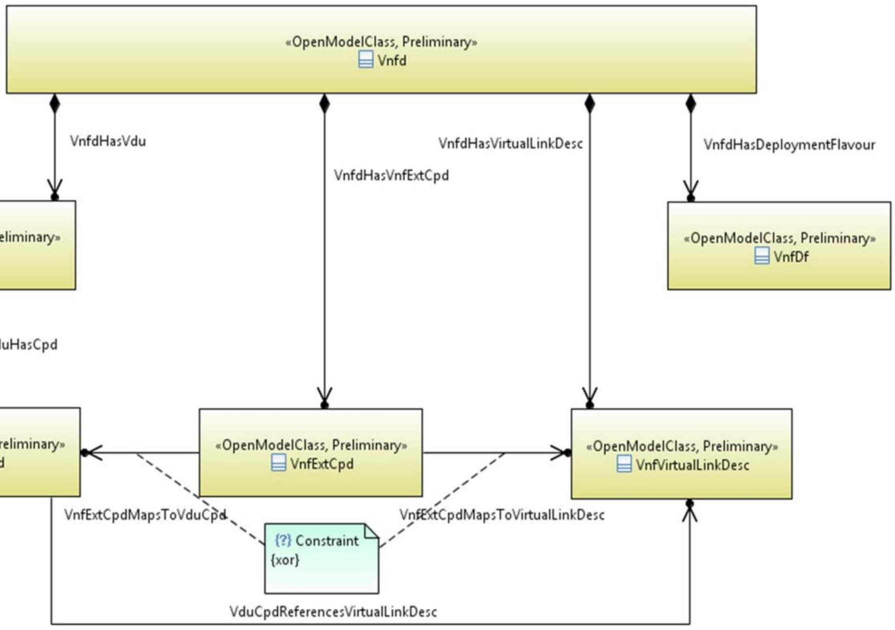

Figure 5-9: VNFD High-level structure

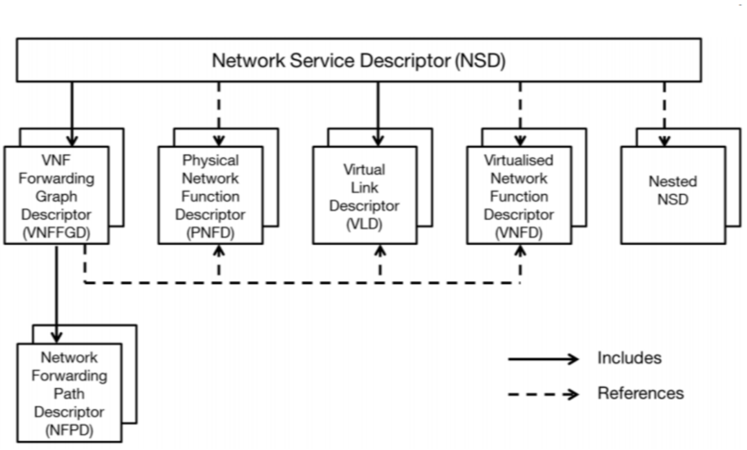

Figure 5-10: NSD IE overview

Finally, it is important to mention that, although this IM is highly focused on the realization at infrastructure level, it includes two types of views:

- **Logical View:** The majority of the classes that form the IM belong to this view, which is concerned about the functionality given by the system to end-users.
- **Deployment View:** All the descriptor classes (VNFD, NSD...) are part of this view, which is primarily concerned about requirements to deploy the system needed by end users.

As mentioned before, although ETSI NFV IM contemplates both, logical and deployment viewpoints, it is clearly focused on the realization at infrastructure level and 3GPP IM, on the counterpart, aims at a more logical/application point of view. Both, logical and deployment points of view are vital and required in any 5G Network Slicing IM, so in order to generate a new IM at 5g-VINNI, IE from 3GPP and ETSI-NFV should be retrieved.

Moreover, there are non-SDO IM defined from projects such as ETSI-OSM. ETSI-OSM IM is able to provide a single entry point to drive full automation of NS, retrieving IE and content from 3GPP (S-NSSAI, 5QI parameters and network slice subnet information element) and ETSI-NFV (NSDs mapped into network slice subnets and Virtual Link Descriptors allowing inter-subnet connectivity). This IM can be accessed in different formats (YANG Model, Navigable Version and Plain Test version) in [95].

Despite all these characteristics, ETSI-OSM IM still lacks a lot of capabilities and is in need of further improvements, particularly in the stack domain-specific configuration. Nonetheless, ETSI-OSM IM could be used as a first input/baseline for 5G-VINNI. An example of this IM can be seen in Figure 5-11.

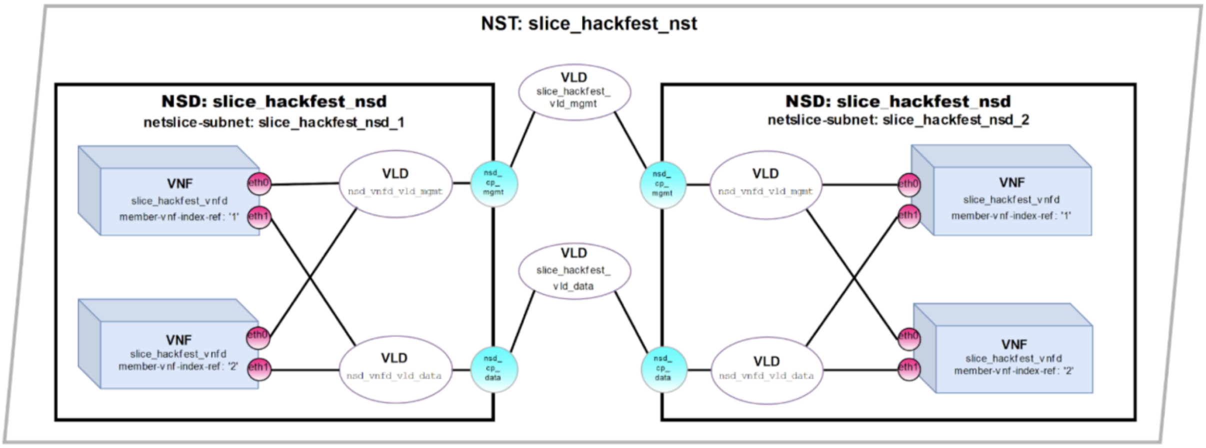

Figure 5-11: NST information model for OSM release 5, from OSM 5th Hackfest

On the other hand, 5G-VINNI IM may take inputs from other 5G-PPP Phase 2 projects in order to complement and expand it. It is strongly recommended to use 5G-TRANSFORMER IM for Network Slice description and IM for Vertical Service Description, SLICENET IM and 5G-TANGO inputs to ETSI-OSM model; while keeping track on SDO updates and upgrades on their own IMs, especially:

- **BBF:** Transport network issues, including intra/inter-subnet connectivity.
- **3GPP:** RAN and CN slice subnets (application level).
- **ETSI-NFV:** Virtualised infrastructure and components of a slice.

At some point it will be useful also to be able to federate 5G-VINNI IM with other external models, gr_NFV-IFA024 could be of great help as they stablish touchpoints with external ETSI-NFV models such as: ONF Core Model, TM Forum Service Model, TM Forum Resource Model, 3GPP Generic Network Resource Model and 3GPP Network Slicing Model.

Finally, to direct the further design of the 5G-VINNI IM, gr_NFV-IFA017 and gr_NFV-IFA016 are good guides/references if UML is used as modeling language and Papirus Open Source Tool is used as IM editing tool.

5.5 Testing interfaces

A high-level overview of the testing architecture of 5G-VINNI can be found in Section 4.9 of D1.1 [2]. The specific details of testing interfaces are part of 5GVINNI's future work and have not been designed yet. They will be described in detail in a future 5G-VINNI deliverable.

6 mMTC Slicing

D1.2 [3] provides a general description of the three broad 5G use case families envisaged by ITU – Enhanced Mobile Broadband (eMBB), Ultra-reliable and low latency communications (uRLLC) and Massive Machine-type communications (mMTC). This section goes a bit deeper on mMTC slicing as a special case of 5G slicing. Other types will be handled in future 5G-VINNI deliverables.

The distinct responsibilities that can be undertaken by different or by the same actors, depending on the level of exposure of infrastructure and IoT platform services, are identified below. Actors in this ecosystem can include Telecom Operators, IoT Platform operators, civilians and potentially any entity able to register/connect their infrastructural/management resources via open, standardized APIs.

- **IoT device provisioning**: offering & registration of physical infrastructure, i.e., smart devices and edge equipment to appropriate slice(s).
- **Infrastructure Operation & Management**: management of the physical & virtual IoT infrastructure, including operation and maintenance of physical/virtual resources.
- **Virtualization & Slicing**: interaction with IoT resources, virtualization and mediated provisioning, definition and enforcement/control of the interaction boundaries and QoS.
- **mMTC Slice ownership**: mediated ownership of virtualized underlying devices and platforms for the development & deployment of IoT services, including configuration and management primitives within the slice limits.

In what follows, notable slicing models are described, derived by varying degrees of integration of the above-mentioned responsibilities in the formed ecosystem.

6.1 Baseline mMTC Slicing Model

As a baseline mMTC network slicing approach, this section considers the ability of the 5G network to dedicate wireless and wired network, as well as compute and storage resources, to the communication needs of individual sets of IoT devices (UEs) owned, deployed and operated by prospective IoT platform operators. Prospective IoT platform operators deploy their IoT devices relying on the establishment of mMTC network slices for the support of the corresponding communication needs. As a result, IoT devices are presented with network slice instances to select from, so as to be integrated in the 5G facility control and data plane e.g., device authentication, data forwarding.

Network resources are allocated by each slice for the transmission and delivery of IoT data, either from or to the IoT devices (e.g., in case of actuators). Compute and storage resources are also allocated for the support of 5G Core functions required for this integration, covering IoT Gateway (GW) functionality (if not realized as a dedicated PNF). The described model is depicted in Figure 6-1, where a single Slice Network Service (NS) instance maps to each slice. The exclusivity of resources per slice simplifies management, as well as the enactment of actuation logic and allows tenants to seamlessly onboard, manage and configure their own IoT devices.

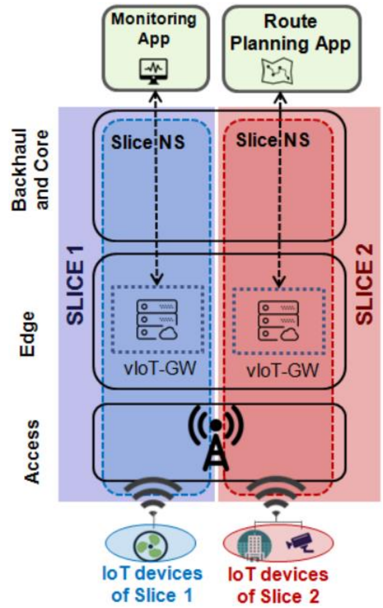

Figure 6-1: Baseline mMTC Slicing Model

Each tenant is provided with a slice, which has its own dedicated IoT Infrastructural resources. The transmission and delivery of data from the smart devices within a slice are end-to-end isolated and independent from the existence of other slices and the control/management functionality extends uninterrupted from the tenants all the way down to the IoT devices.

6.2 Extended mMTC Slicing Model

Building on the baseline mMTC Slicing approach, an extended mMTC Slicing model is further targeted, where emphasis is put on the sharing of existing IoT resources. Namely, the objective is to allow the sharing of the same IoT devices by different Slice owners, the former being owned and deployed by IoT infrastructure stakeholders, as illustrated in Figure 6-2. Example stakeholders include network infrastructure operators that extend their network infrastructure with IoT devices (sensors, actuators) or even IoT Platform operators operating mMTC network slices.

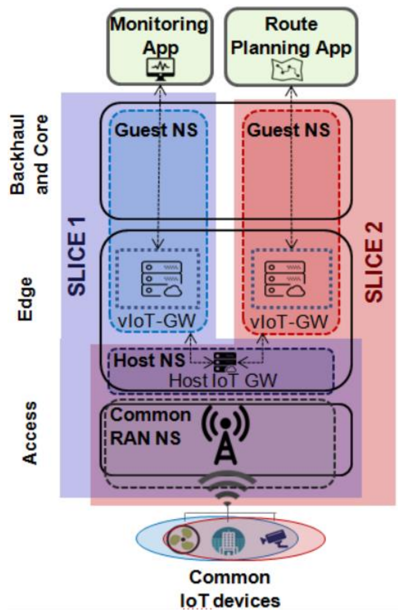

Figure 6-2: Extended mMTC Slicing Model

While the support of IoT device virtualization appears as the straightforward approach, the intended scale of mMTC slices (i.e., 1 M devices/km^2) presents significant scalability and complexity concerns, when considering the creation of one virtual IoT device instance per instantiated mMTC slice. Bypassing these concerns, in the extended IoT sharing model, IoT devices become a shared commodity on the IoT GW level, where mediation is provided for the coordinated communication with devices. This virtualization at the GW level allows for the separation of concerns between different stakeholders and thus, the streamlining of IoT services on-boarding and deployment.

To this end, this Slicing model is compatible with a PaaS sharing approach, with virtualized IoT GW instances (vIoT-GW) dedicated to the support of the corresponding IoT slices. vIoT-GWs can be realized as VNFs, each delivered to a single slice of the extended mMTC slice type.

A north-bound interface (NBI) of such virtualized instance can provide application support, in an IoT Slicing-agnostic manner i.e., facilitating data delivery (e.g., MQTT [Message Queuing Telemetry Transport] based); exposure of IoT devices, their capabilities and configuration interfaces; Rule Management; User Management, as well as traditional FCAPS tasks. Notice that in this model, RAN and Host NSs are shared by different slices, however with the capability of customized management for each slice. A south-bound interface (SBI) handles the interconnection with an underlying IoT GW (Host IoT-GW), common for different slices and operated by the Telecom Operator. The Host IoT-GW is responsible for the management of the devices in the field, supporting the physical interface to the IoT infrastructure in an IoT Slicing-orthogonal manner. At the same time it realizes the IoT GW virtualization primitives by mediating access of IoT GW VNFs to the underlying infrastructure e.g., delivering sub-streams of data and realizing conflict resolution for actuators. This corresponds to the realization of a Virtualized IoT Infrastructure Management (VIoTIM) system, in an analogy to VIM for compute and storage resources, however bearing a PaaS flavour.

Hence, vIoT-GWs are responsible for exposing virtual views of the IoT device data, meta-data and control command interfaces within the corresponding slices, without incorporating any IoT device (UE) control plane functionality. Essentially, this corresponds to the virtualization of the Host IoT-G, operated by the infrastructure provider. As shown in Figure 6-2, the extended mMTC slice boundaries are located at the Edge of the network, where the virtualization of the Host IoT-GW takes place.

In spite of management overhead, the main benefit of unified mediation is the efficient reuse of existing IoT infrastructure (IoT devices, gateways, southbound mechanisms/configurations, management), lowering CAPEX/OPEX costs and removing entry barriers for IoT application owners. In addition, it allows the ``Plug & Play'' deployment of applications built for legacy IoT systems, since slice owners are presented with a virtualization-agnostic view of the underlying hardware, obsolescing the use of wrapper layers and glue code for interoperability.

6.3 Hybrid models

Considering variations of the above-described models, it is possible to identify combinatory models constructed by allowing different degrees of unified mediation and management over underlying IoT infrastructure – introduced in the extended model – and resources exclusivity per slice---described in the baseline model. As an example, the extended model is compatible with the provisioning of dedicated IoT resources per slice, however with the management, control and data planes actually passing through the shared Host IoT-GW---which in this case should be responsible for guaranteeing the (virtual) isolation.

Besides, the two extreme models, as well as their variations can physically coexist, with the infrastructure operator adopting versatile composition mechanisms to support the symbiosis of different slicing schemes.

6.4 Extended mMTC slicing management interfaces

The lifecycle management (LCM) of mMTC slices is a special case of network slice LCM. Figure 6-3 illustrates the management interfaces for the case of extended mMTC slicing, in which there are two different types of network services: the Host IoT GW network service, which is instantiated once for a given pool of IoT devices shared among slices, and the Guest GW network service, which is instantiated once for every different mMTC slice.

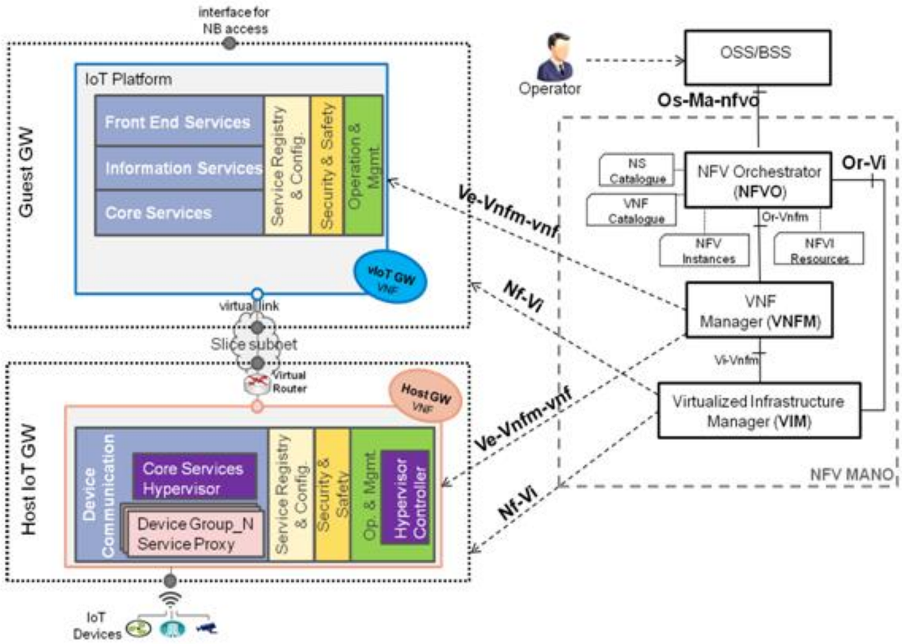

Figure 6-3: Interfaces for mMTC slice lifecycle management

The management interfaces that are involved with the LCM of the mMTC slices are realized through the NFV-MANO reference points as follows:

- Os-Ma-nfvo: In the case of mMTC slicing, this reference point supports the propagation of all relevant mMTC slice LCM requests (slice instantiation, update, slice instance querying, scaling and termination) from the OSS/BSS—which acts as the Operator's proxy—to the NFV Orchestrator. This LCM can take place at the coarse granularity of mMTC slices, as well as at the level of specific VNF components and their relations to NFVI resources. Moreover, it includes the propagation of relevant policy management and enforcement on one direction, and forwarding of events and accounting information on the other.
- Or-Vi: This reference point supports NFVI resource provisioning and configurations for different mMTC slices. In this respect, it involves the onboarding/removal of software images for the Guest GW VNF, as well as for the Host IoT GW VNF in the case it is not available as a PNF. In addition, it supports the management for reservation of resources (e.g., CPU, memory, storage, network bandwidth) for each Guest GW VNF and the VM in which the software image is instantiated. On the opposite direction, it facilitates the reporting to the NFV Orchestrator about the usage of NFVI resources by different mMTC slice instances, as well as about the instances' health, i.e., whether they are up and running, utilizing the prescribed resources.
- Nf-Vi: This reference point is used for exchanges between Virtualisation Infrastructure Manager and NFV Infrastructure, essentially acting as the control interface for ensuring that the management interactions through the Or-Vi, as described above are realized on the actual NFVI resources, at the (virtualized) hardware level. On the opposite direction towards the VIM, it supports the reporting of failure events, measurement results and usage data

about the resources that are allocated to mMTC slices, both for Host GW and for Guest GW VNF instances.

- Ve-Vnfm-vnf: This reference point is used for exchanges between both the Host and Guest GW VNFs and the VNF Manager, mainly supporting the initialization of mMTC slices, given that the Host GW is at an operational state and the Guest GW VNF for a specific mMTC slice has been instantiated. The Ve-Vnfm-vnf supports the activation of day-2 configuration primitives for preparing on one hand the Host GW VNF about the new slice and the IoT information that should be delivered to it and on the other hand, for configuring the Guest GW VNF to properly link to the Host GW with the appropriate information and credentials. Furthermore, it supports day-2 configuration primitives related to configuration updates on the running VNFs, as well as the management of monitored information (which could be arbitrarily aggregated and customized) and their delivery from the VNFs to the VNFM, also supporting the triggering of relevant functions on the Host and Guest GW VNFs based on observations.

7 Evolution roadmap

In the 5G-VINNI deliverable D1.2 [3] the roadmap of network slicing was defined, based on the evolution approach followed by the 3GPP 5G architecture in Release 15 and Release 16. In addition to the roadmap, all relevant technical details concerning the non-standalone and standalone architectural options were presented, providing a clear map of the evolution roadmap in terms of 3GPP. Therefore, this section will be focused on the complementary/different aspects to those presented in D1.2.

The mentioned complement/differences in the context of 5G-VINNI can be summarized in: 1) In terms of management capabilities and its gradual implementation. 2) In some challenges that remain in the radio part and that some of them will be gradually addressed in 5G-VINNI. Figure 7-1 illustrates the concepts to be described in this section.

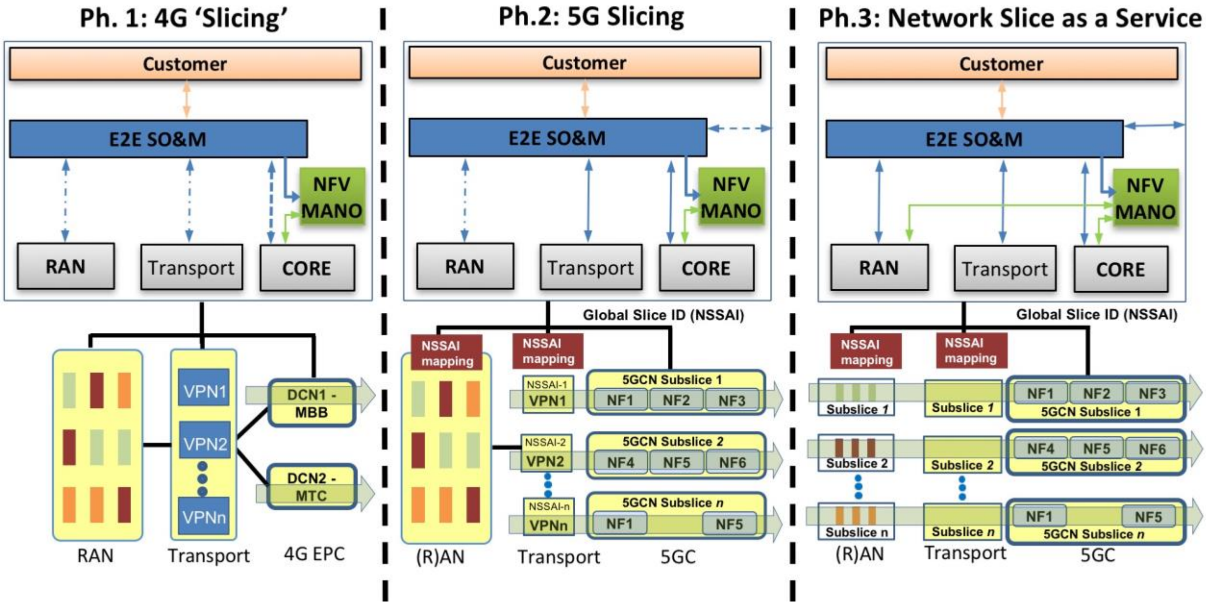

Figure 7-1: 5G-VINNI Evolution Roadmap (LCM and Orchestration emphasis)

As explained in deliverable D1.2 [3], at the initial phase denominated as *"4G Slicing"*, it is expected to have 5G NR to 4G based dedicated core networks (DECOR) to leverage new spectrum and 5G radio capabilities following a NSA implementation. The description provided here applies for the 5G-VINNI release 0. From the management point of view there are some details that need to be mentioned.

First, regarding the NFV environment, although the use of Cloud-RAN is possible in some facility sites, it is expected that most of the facilities will only use VNFs (and the respective virtualized environment relevant) at the Core Network Domain, and hence the role of the NFV-MANO is key in this network domain from the very beginning. This is indicated in Figure 7-1 with the green arrow that connects the NFV MANO and the Core.

The Core and RAN Domain Controllers at phase 0 are expected to be implemented as the conventional EMS, not very different from those used for the current 4G networks. For the Core, slicing will be introduced by using the 3GPP defined, dedicated core networks (DECOR) concept. In the transport network, it is expected to use conventional VPNs manually mapped to the different cores obtained from DCORE. The RAN will be configured manually and at this initial stage is not expected an NSSAI mapping. In conclusion, for the domain controllers a system with several manual procedures is expected, especially regarding to the integration between network domains.

From the E2E SO&M point of view, the functions concerning mapping of customers request will be implemented in an initial step, constrained by the technical limitations of the NSA architecture. On the other hand, regarding the interaction with the network domains controllers goes as follows. The dashed line between the E2E SO&M and the RAN Controllers means that there is not automatic execution of functions between these two components. However, the E2E SO&M needs to have knowledge on the RAN capabilities in order to provide network slice services to customers. Therefore, such RAN capabilities have to be transmitted/configured using manual procedures at the E2E SO&M. A similar approach is followed by the dashed line showed between the E2E SO&M and the Transport Controllers, where manual VPN channels will be configured, but the E2E SO&M needs to be aware of the capabilities that are configured there. Regarding the Core controllers, some initial level of interaction with the E2E is expected, using for instance the day 2 capabilities provided by OSM. However, this interaction features may have some limitations, and therefore it is presented in Figure 7-1 still as a dashed line, just with some enhancements. Finally at this stage it is not expected that there will be any kind of interaction across facility sites (different administrative domains).

The next phase denominated as *"5G-Slicing"* is expected to follow the 3GPP Release 15 and 16 Standalone (SA). The description provided here applies for a gradual evolution during the 5G-VINNI releases 1 and 2. One of the main features here is that it includes E2E slice selection based on global slice ID which is S-NSSAI or NSI-ID, and Core network is 5GC. RAN supports slicing with S-NSSAI and Transport supports slicing via VPNs mapped according to S-NSSAI. The role of the NFV-MANO remains the same as in the previous phase, although some facility sites may have a more dynamic deployment of VNFs in edge nodes. Regarding the E2E SO&M, the interaction with the customer keep the same properties but it includes now the possibilities offered by using the standalone architecture. Concerning the RAN controllers, the setting remains very similar, where several manual procedures are needed. Since slicing will be supported via VPNs mapped according to NSSAI, a more automated interaction is expected between the E2E SO&M and the Transport Network controllers. Regarding the Core controller, a better and enhanced level of interaction is expected with the E2E SO&M. Finally, at this phase the first implementation for the interaction between facility sites is expected by using the interfaces mentioned in Section 5.3, which may be implemented using an Open API approach.

Finally, Phase 3 will be called *"Network Slice as a Service"*. Here, fully automated, adaptive and self-organizing slices and sub-slices are expected. RAN, transport, and core are all sliced, and slices capable of hosting 3rd party NFs (non-3GPP) in addition to 3GPP NFs. As explained in the 5GVINNI deliverable 1.2 [3], this phase includes some aspects that are more experimental, and hence not all the facility sites will need to deliver all the mentioned features. However, the implementation is expected to be completed during the 5G-VINNI releases 3 and Final. For the NFV-MANO point of view, with Phase 3 a bigger number of VNFs is expected beyond the Core network, which includes larger virtualized deployments at the edge and RAN. Therefore, the functionalities of the NFV MANO are expected to be more advanced in order to satisfy a higher level of demands. In this phase, it is also expected that all the domain controllers have a full interaction with the E2E SO&M, by exposing all the capabilities needed in order to automate all the end-to-end life cycle management actions needed by the slices to be provided. Finally, fully operational east/west interfaces are expected at this stage, in order to enable the instantiation of federated slices across different facility sites.

References

[1] 5G-VINNI, “Grant Agreement 815279 - 5G Verticals INNovation Infrastructure,” 2018.

[2] 5G-VINNI, “Deliverable D1.1 - Design of infrastructure architecture and subsystems v1,” 2018.

[3] 5G-VINNI, “Deliverable D1.2 - Design of network slicing and supporting system v1,” 2019.

[4] 5G-VINNI, “Deliverable D3.1 - Specification of services delivered by each of the 5G-VINNI facilities,” 2019.

[5] 5G-VINNI, “Deliverable D2.1 - 5G-VINNI Solution Facility-sites High Level Design (HLD),” 2019.

[6] 3GPP, “3GPP TS 28.530 V15.1.0; Technical Specification Group Services and System Aspects; Management and orchestration; Concepts, use cases and requirements (Release 15),” 2018.

[7] 3GPP, “TS 28.531 V16.1.0, Technical Specification Group Services and System Aspects; Management and orchestration; Provisioning;,” 2019.

[8] 3GPP, “TS 28.532 V15.2.0, Technical Specification Group Services and System Aspects; Management and orchestration; Generic management services(Release 15),” 2019.

[9] 3GPP, “TS 28.540 V15.1.0 Technical Specification Group Services and System Aspects; Management and orchestration; 5G Network Resource Model (NRM); Stage 1 (Release 15),” 2018.

[10] 3GPP, “TS 28.541 V16.0.0; Technical Specification Group Services and System Aspects; Management and orchestration; 5G Network Resource Model (NRM); Stage 2 and stage 3 (Release 16),” 2019.

[11] 3GPP, “TS 28.542 V1.0.0; Technical Specification; Management and orchestration of networks and network slicing; 5G Core Network (5GC) Network Resource Model (NRM); Stage 1 (Release 15),” 2018.

[12] 3GPP, “SA2 - Architecture,” 2019. [Online]. Available: https://www.3gpp.org/specifications-groups/sa-plenary/sa2-architecture. [Accessed May 2019].

[13] 3GPP, “TS 28.550 V16.1.0 Management and orchestration; Performance assurance (Release 16),” 2019.

[14] 3GPP, “TS 28.545 V15.1.0 Management and orchestration; Fault Supervision (FS) (Release 15),” 2018.

[15] 3GPP, “TS 28.552 V16.2.0 Management and orchestration; 5G performance measurements (Release 16),” 2019.

[16] 3GPP, “TS 28.554 V16.1.0 Management and orchestration; 5G end to end Key Performance Indicators (KPI) (Release 16),” 2019.

[17] 3GPP, “TR 28.801 Technical Report Technical Specification Group Services and System Aspects; Telecommunication management; Study on management and orchestration of network slicing for next generation network (Release 15),” 2018.

[18] 3GPP, “TR 28.808 - Study on management and orchestration aspects with integrated satellite components in a 5G network, V0.1.0,” 2019.

[19] ETSI NFV ISG, "GS NFV-MAN 001, "Network Functions Virtualisation (NFV); Management and Orchestration", v1.1.1," 2014.

[20] ETSI NFV ISG, "GR NFV-EVE 012, "Network Functions Virtualisation (NFV) Release 3; Evolution and Ecosystem; Report on Network Slicing Support with ETSI NFV Architectural Framework", v3.1.1," 2017.

[21] ETSI NFV ISG, "ETSI GS NFV 003, "Network Functions Virtualisation (NFV); Terminology for Main Concepts in NFV", v3.1.1," 2018.

[22] J. Ordones-Lucena, O. Adamuz-Hinojosa, P. Ameigeiras, P. Muñoz, J. Ramos-Muñoz, J. Chavarria and D. Lopez, "The Creation Phase in Network Slicing: From a Service Order to an Operative Network Slice," in *European Conference on Networks and Communications (EuCNC)*, Ljubljana, Slovenia, 2018.

[23] ETSI NFV ISG, "GS NFV-IFA 008, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Ve-Vnfm reference point – Interface and Information Model Specification", v3.1.1," 2018.

[24] ETSI NFV ISG, "ETSI GS NFV-IFA 012, "Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; GS IFA012 Os-Ma-nfvo reference point – Interface and Information Model Specification", v3.1.1," 2018.

[25] ETSI ZSM ISG, "Zero touch network & Service Management (ZSM)," 2018. [Online]. Available: https://www.etsi.org/technologies/zero-touch-network-service-management. [Accessed May 2019].

[26] ETSI SmartM2M, "ETSI TR 103 527, One High-Level Architecture for IoT Virtualization, v1.1.1," 2018.

[27] Broadband Forum, "BBF 5G Project Stream," [Online]. Available: https://wiki.broadband-forum.org/display/BBF/5G+Project+Stream. [Accessed April 2019].

[28] TM Forum, "TM Forum," 2019. [Online]. Available: https://www.tmforum.org/. [Accessed May 2019].

[29] MEF, "MEF," 2019. [Online]. Available: http://www.mef.net/. [Accessed May 2019].

[30] NGMN Alliance, "5G White Paper," 2015. [Online]. Available: https://www.ngmn.org/fileadmin/ngmn/content/images/news/ngmn_news/NGMN_5G_White_Paper_V1_0.pdf. [Accessed April 2019].

[31] 5GPPP, "5GPPP Architecture Working Group, "View on 5G Architecture", Version 2.0," 2017.

[32] OSM, "Open Source MANO," 2019. [Online]. Available: https://osm.etsi.org/ . [Accessed April 2019].

[33] OSM, "List of OSM Members," 2019. [Online]. Available: https://portal.etsi.org/TBSiteMap/OSM/ListofOSMMembers.aspx. [Accessed April 2019].

[34] ETSI NFV, "ETSI GS NFV-SOL 005 V2.4.1, Protocols and Data Models; RESTful protocols specification for the Os-Ma-nfvo Reference Point," 2018.

[35] OSM End User Advisory Group, "OSM Scope, Functionality, Operation and Integration Guidelines, Issue 1," 2019.

[36] SONATA, "SONATA Open Source Github repositories," 2019. [Online]. Available:

https://github.com/sonata-nfv. [Accessed May 2019].

[37] ONAP, "ONAP - Open Network Automation Platform," 2019. [Online]. Available: https://www.onap.org/. [Accessed April 2019].

[38] ONAP 5G Use Case Team, "ONAP 5G RAN Network Slicing," 2018. [Online]. Available: https://wiki.onap.org/download/attachments/10784151/ONAPNetworkSlicingv3_25Jan2018.pptx. [Accessed April 2019].

[39] Open Baton, "An ETSI NFV compliant Management and Orchestration (MANO) Framework," [Online]. Available: https://github.com/openbaton. [Accessed April 2109].

[40] Open Baton, "Network Slicing Engine," 2017. [Online]. Available: https://openbaton.github.io/documentation/network-slicing/. [Accessed April 2019].

[41] X. Zhou, "Network Slicing As A Service: Enabling Enterprises' Own Software-Defined Cellular Networks," *IEEE Communications Magazine,* 2016.

[42] N. Nikaein, "Network Store: Exploring Slicing in Future 5G Networks," in *10th ACM Int'l. Wksp. Mobility in the Evolving Internet Architecture*, 2015.

[43] K. Samdanis, X. Costa-Perez and V. Sciancalepore, "From Network Sharing to Multi-Tenancy: The 5G Net- work Slice Broker," *IEEE Communications Magazine,* vol. 54, no. 7, 2016.

[44] F. Akyildiz, P. Wang and S. Lin, "SoftAir: A Software Defined Networking Architecture for 5G Wireless Systems," *Computer Networks,* vol. 85, 2015.

[45] P. Rost, "Mobile Network Architecture Evolution Toward 5G," *IEEE Communications Magazine,* vol. 54, no. 5, 2016.

[46] A. Banchs, "A Novel Radio Multiservice Adaptive Network Architecture for 5G Networks," in *81st IEEE VTC-Spring'15*, 2015.

[47] D. Shue, M. J. Freedman and A. Shaikh, "Performance Isolation and Fairness for Multi-Tenant Cloud Storage," in *Proc. 10th USENIX Symp. Operating Systems Design and Implementation*, 2012.

[48] 3GPP, "TS 132 450 V14.0.0 Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Key Performance Indicators (KPI) for Evolved Universal; Terrestrial Radio Access Network (E-UTRAN): Definitions," 2017.

[49] 3GPP, "TS 28.533 V15.1.0, Technical Specification Group Services and System Aspects; Management and orchestration; Architecture framework (Release 15)," 2018.

[50] L. Da Silva, A. Kaloxylos and G. Zimmermann, "On the impact of network slicing on 5G radio access networks," in *European Conference on Networks and Communications (EuCNC)*, Athens, Greece, 2016.

[51] 3GPP, "TS 36.300 V15.5.0 - Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 15)," 2019.

[52] 3GPP, "TR 23.716 V16.0.0; Technical Specification Group Services and System Aspects; Study on the Wireless and Wireline Convergence for the 5G system architecture (Release 16)," 2018.

[53] 3GPP, "TS 23.501, Technical Specification Group TS 23.501, Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2, v15.4.0," 2018.

[54] 3GPP, “TS 23.502 V16.0.2 Technical Specification Group Services and System Aspects; Procedures for the 5G System; Stage 2 (Release 16),” 2019.

[55] Broadband Forum, “MR-427 “Fixed-Mobile Convergence Marketing Report”,” 2018.

[56] ITU-T, “Technical Report GSTR-TN5G Transport network support of IMT-2020/5G,” 2018.

[57] 3GPP, “TR 38.801 V14.0.0, Study on new radio access technology: Radio access architecture and interfaces,” 2017.

[58] L. Larsen, M. Berger and H. Christiansen, “Fronthaul for Cloud-RAN Enabling Network Slicing in 5G Mobile Networks”,” *Wireless Communications and Mobile Computing,* vol. 2018, 2018.

[59] G. Mountaser, M. Condoluci , T. Mahmoodi, M. Dohler and I. Mings, “Cloud-RAN in Support of URLLC,” in *2017 IEEE Globecom Workshops (GC Wkshps)*, Singapore, 2017.

[60] E. Rosen and Y. Rekhter, “IETF RFC 4364 BGP/MPLS IP Virtual Private Networks (VPNs),” 2006. [Online]. Available: https://tools.ietf.org/html/rfc4364.

[61] Metro Ethernet Forum, “Carrier Ethernet Services,” [Online]. Available: https://www.mef.net/carrier-ethernet/carrier-ethernet-services. [Accessed April 2019].

[62] ACG Research, “Challenges and Opportunities in 5G Transport,” 2016. [Online]. Available: https://www.acgcc.com/challenges-and-opportunities-in-5g-transport. [Accessed April 2019].

[63] Optical Internetworking Forum, “IA # OIF-FLEXE-02.0 - Flex Ethernet 2.0 Implementation Agreement,” 2018. [Online]. Available: https://www.oiforum.com/wp-content/uploads/2019/01/OIF-FLEXE-02.0-1.pdf.

[64] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang and W. Weiss, “IETF RFC 2475 An Architecture for Differentiated Services,” 1998. [Online]. Available: https://www.rfc-editor.org/info/rfc2475.

[65] J. Babiarz, F. Baker and K. Chan, “IETF 4594 Configuration Guidelines for DiffServ Service Classes,” 2006. [Online]. Available: https://www.rfc-editor.org/info/rfc4594.

[66] IEEE 802.1, “Time-Sensitive Networking (TSN) Task Group,” 2019. [Online]. Available: https://1.ieee802.org/tsn/. [Accessed April 2019].

[67] J. Dong, S. Bryant, Z. Li, T. Miyasaka and Y. Lee, “draft-ietf-teas-enhanced-vpn-01, A Framework for Enhanced Virtual Private Networks (VPN+) Service (work in progress),” 2019. [Online]. Available: https://tools.ietf.org/html/draft-ietf-teas-enhanced-vpn-01.

[68] N. Finn, P. Thubert, B. Varga and J. Farkas, “IETF Draft - DetNet, Deterministic Networking Architecture (work in progress),” 2019. [Online]. Available: https://www.ietf.org/id/draft-ietf-detnet-architecture-12.txt.

[69] C. Filsfils, S. Previdi, L. Ginsberg, B. Decraene, S. Litowski and R. Shakir, “IETF RFC 8402 Segment Routing Architecture,” 2018. [Online]. Available: https://www.rfc-editor.org/info/rfc8402.

[70] S. Perrin, “Converged Packet Transport for 5G”, Heavy Reading white paper,” 2018. [Online]. Available: https://img.lightreading.com/downloads/Cisco-Converged-Packet-Transport-for-5G.pdf. [Accessed April 2019].

[71] Huawei, “Huawei HQoS Technical White Paper,” 2010.

[72] R. Alvizu, G. Maier, N. Kukreja, A. Pattavina, R. Morro and A. Capello, “Comprehensive Survey on T-SDN: Software-Defined Networking for Transport Networks,” *IEEE Communications Surveys & Tutorials,* vol. 19, no. 4, 2017.

[73] SaT5G Consortium, “Deliverable D3.1 - Integrated SaT5G General Network Architecture,” 2018. [Online]. Available: http://sat5g-project.eu/.

[74] ETSI, “ETSI TR 103 611, Seamless integration of satellite and/or HAPS (High Altitude Platform Station) systems into 5G system and related architecture options,” 2018.

[75] SaT5G, “H2020 5G PPP project “SaT5G” (Satellite and Terrestrial Network for 5G),” 2017. [Online]. Available: http://sat5g-project.eu/.

[76] SaT5G Consortium, “Deliverable D3.2 - Integrated SaT5G Detailed Network Architecture,” 2018. [Online]. Available: http://sat5g-project.eu/.

[77] P. Sayyad, P. Khodashenas, C. Fernandez, D. Guija, K. Liolis, C. Politis, G. Atkinson, J. Cahill, R. King, M. Kavanagh, B. Jou and O. Vidal, “Benefits and Challenges of Software Defined Satellite-5G Communication,” in *IEEE/IFIP WONS 2019*, Wengen, Switzerland, 2019.

[78] ETSI Satellite Earth Stations and Systems (SES), “DTR/SES-00446 Work Item; Reference Virtualised Network Functions data model for satellite communication systems,” 2018. [Online]. Available: https://portal.etsi.org/webapp/WorkProgram/Report_WorkItem.asp?wki_id=54042.

[79] IEEE Softwarization, “Network Slicing and 3GPP Service and Systems Aspects (SA) Standard,” 2017. [Online]. Available: https://sdn.ieee.org/newsletter/december-2017/network-slicing-and-3gpp-service-and-systems-aspects-sa-standard. [Accessed April 2019].

[80] Open Networking Foundation, “TR-521 - SDN Architecture Issue 1.1,” 2016. [Online]. Available: https://www.opennetworking.org/wp-content/uploads/2014/10/TR-521_SDN_Architecture_issue_1.1.pdf. [Accessed April 2019].

[81] ETSI NFV ISG, “GS NFV-IFA 013, Management and Orchestration; Os-Ma-Nfvo reference point – Interface and Information Model Specification, v3.2.1,” 2016.

[82] ETSI NFV ISG, “GS NFV-IFA 005, “Management and Orchestration; Or-Vi reference point – Interface and Information Model Specification”, v3.2.1,” 2016.

[83] ETSI NFV ISG, “GS NFV-EVE 005 V1.1.1; Report on SDN Usage in NFV Architectural Framework,” 2015.

[84] TM Forum, “TMF633 Service Catalog Management API REST Specification R16.5.1,” 2019. [Online]. Available: https://www.tmforum.org/resources/specification/tmf633-service-catalog-management-api-rest-specification-r16-5-1/. [Accessed June 2019].

[85] TM Forum, “TMF641 Service Ordering API REST Specification R15.5.1,” 2019. [Online]. Available: https://www.tmforum.org/resources/standard/tmf641-service-ordering-api-rest-specification-r15-5-1/. [Accessed June 2019].

[86] Y. T. Lee, “Information modeling: from Design to Implementation,” 2008. [Online]. Available: https://pdfs.semanticscholar.org/d2f4/334f3acd1cd4084a65381c141b073b15e168.pdf. [Accessed April 2019].

[87] C. Wilson, “Telefonica's Elizondo: OSM Offers Best Information Model for Service Orchestration,” 2018. [Online]. Available: https://www.lightreading.com/nfv/nfv-

mano/telefonicas-elizondo-osm-offers-best-information-model-for-service-orchestration/d/d-id/746347.

[88] 3GPP, "TS 38.401 V15.5.0 Technical Specification Group Radio Access Network; NG-RAN; Architecture description (Release 15)," 2019.

[89] 3GPP, "TS 28.622 V16.0.0 Technical Specification; Generic Network Resource Model (NRM); Integration Reference Point (IRP); Information Service (IS) (Release 16)," 2019.

[90] 3GPP, "TS 28.708 V15.2.0 Evolved Packet Core (EPC); Network Resource Model (NRM); Integration Reference Point (IRP); Information Service (IS) (Release 15)," 2018.

[91] ETSI NFV ISG, "GR NFV-IFA 015 V2.4.1 - NFV Release 2; Management and Orchestration; Report on NFV Information Model," 2018.

[92] ETSI NFV ISG, "GS NFV-IFA 011 V2.5.1 - NFV Release 2; Management and Orchestration; VNF Descriptor and Packaging Specification," 2018.

[93] ETSI NFV ISG, "IFA 014, Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Network Service Templates Specification", v3.1.1," 2018.

[94] ETSI NFV ISG, "ETSI GR NFV-IFA 017 V2.1.1 NFV Release 2; Information Modeling; UML Modeling Guidelines," 2017.

[95] Open Source MANO, "OSM Information Model," 2019. [Online]. Available: https://osm.etsi.org/wikipub/index.php/OSM_Information_Model. [Accessed April 2019].

[96] 5G-TRANSFORMER, "H2020 5G PPP project 5GTRANSFORMER (5G Mobile Transport Platform for Verticals)," 2019. [Online]. Available: http://5g-transformer.eu/. [Accessed April 2019].

[97] 5G-TRANSFORMER, "Deliverable D3.1, Definition of vertical service descriptors and SO NBI," 2018.

[98] ETSI MEC ISG, "Multi-access Edge Computing (MEC)," 2019. [Online]. Available: https://www.etsi.org/technologies/multi-access-edge-computing. [Accessed 2019 2019].

[99] 5GTANGO, "H2020 5G PPP project 5GTANGO (5G Development and Validation Platform for Global Industry-Specific Network Services and Apps),," 2019. [Online]. Available: https://5gtango.eu/. [Accessed April 2019].

[100] SONATA, "H2020 5G PPP project SONATA (Service Programming and Orchestration for Virtualized Software Networks), 2015," 2018. [Online]. Available: www.sonata-nfv.eu/. [Accessed April 2019].

[101] 5GTANGO, "Deliverable 2.2, Architecture Design," 2017.

[102] 5GTANGO, "Deliverable 5.1, Service Platform Operational First Release," 2017.

[103] B. Chatras, U. S. T. Kwong and N. Bihannic, "NFV enabling network slicing for 5G," in *20th Conference on Innovations in Clouds, Internet and Networks (ICIN)*, 2017.

[104] P. S. Khodashenas, H. Khalili, D. Guija and M. S. Siddiqui, "TALENT: Towards Integration of Satellite and Terrestrial Networks," in *EUCNC 2019*, Valencia, 2019.

Annex A Relevant 5G-PPP Phase 2 Projects

This Annex provides complementary information about related Phase 2 5G-PPP projects, all of which are active in fields relevant to 5G-VINNI including network slicing, network slice orchestration and 5G-satellite integration. It should be noted that this is not supposed to be an exhaustive list of relevant 5G-PPP projects[1], either from an architectural perspective or by providing potentially reusable components.

- 5G-TRANSFORMER is especially relevant in relation to network slicing, by dealing with management operations at multiple abstraction layers in the service stack, including application layer (vertical services), network slice layer (network slice, network slice subnet, and networks services) and resource layer (VNFs/PNFs and underlying physical and virtual resources).
- 5GTANGO is focused on the flexible programmability of 5G networks, covering the three well-defined stages of the lifecycle of a service: (i) service development, (ii) service testing, verification & validation, and (iii) service operation. 5GTANGO has been responsible for the evolution of the SONATA orchestrator, including network slicing related capabilities.
- SaT5G pursues several research pillars, one of which related to this deliverable being the development of the enablers for an integrated 5G-SatCom virtual and physical resource orchestration and service management. To this end, SaT5G pursues the development of an orchestration solution able to perform lifecycle management of SatCom and 5G core virtualized functions, aiming to demonstrate the creation of end-to-end SatCom-5G slices.

A.1 5G-TRANSFORMER

5G-TRANSFORMER [96] is a 5G PPP Phase 2 project, whose purpose is to transform today's mobile transport network into an enriched service platform able to support a myriad of use cases from various vertical industries (e.g. eHealth, media, automotive) in an effective manner, while making an efficient resource usage. To this end, 5G-TRANSFORMER relies on the capabilities brought by network slicing. The result is a slicing-ready, SDN/NFV-based mobile transport and computing platform with a twofold technical approach: (i) provide vertical industries with customized network slices to meet their specific service requirements, while allowing per-slice management of virtualized resources; and (ii) aggregate and federate transport networking and computing fabric, from the edge all the way to the core and cloud, to deploy and operate network slices throughout a federate virtualized infrastructure.

Figure A-1 provides an overview of the 5G-TRANSFORMER system architecture, where three main subsystems have been defined: Vertical Slicer (5GT-VS), Service orchestrator (5GT-SO), and mobile transport and computing platform (5GT-MTP). The **5GT-VS** is the common entry point for verticals into the system, allowing them to gain access to the service portfolio, as well as consume the system exposed service capabilities. For this end, a high-level interface at the 5GT-VS northbound is defined. The 5GT-VS dynamically maps the incoming service requests onto NSIs (and their constituent NSSIs) according to their specificities, and manages their lifecycle. Finally, it translates the NSI/NSSI requirements into network service information, including selected network services along with their instantiation parameters, and sends this information to the 5GT-SO. The **5GT-SO** takes this information to create/reuse instances of network services and operate them throughout their lifecycle. This subsystem embeds complete NFVO functionality, including network service orchestration (NFVO-SO) and resource orchestration (NFVO-RO). Finally, the **5GT-MTP** manages the virtual resources on top of which network service instances (and their VNF/PNF instances) run,

[1] Information on other potentially relevant Phase 2 5G-PPP projects can be found at https://5g-ppp.eu/5g-ppp-phase-2-projects/

building up an ecosystem that serves the 5GT-SO, enforcing 5GT-SO operations at the infrastructure level. As seen from Figure A-1, 5GT-MTP responsibilities include at least the lifecycle management of virtual resources and of the underlying mobile transport, computing and storage infrastructure (VIM/WIM), as well as VNF management at the virtualized resource level (VNFM).

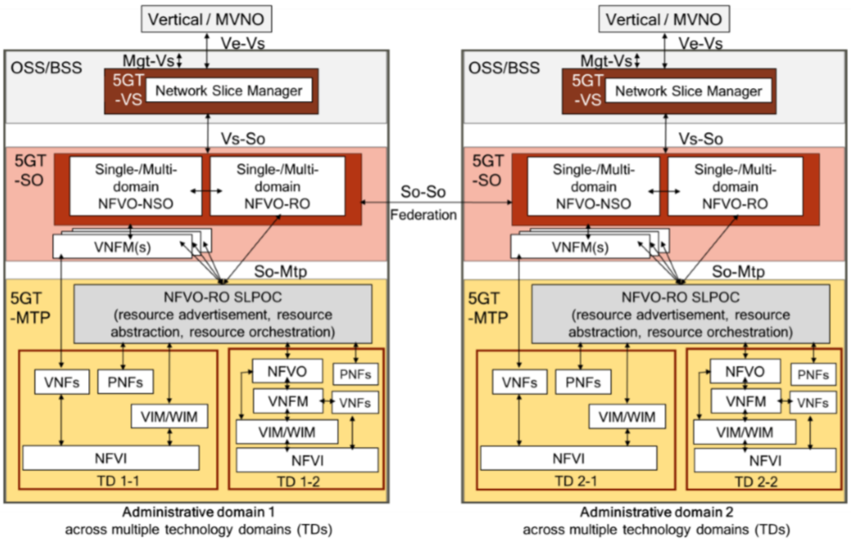

Figure A-1: 5G-TRANSFORMER system architecture [97]

5G-TRANSFORMER relevant features include the information models for network slice description and vertical service description, as well as the operation across federated administrative domains. The system architecture is also aligned with the 3GPP and ETSI NFV vision on network slicing (see Section 2.2). The tentative mapping between this architecture and the NFV-assisted 3GPP slicing framework is shown in Figure A-2 (see also Figure 2-5).

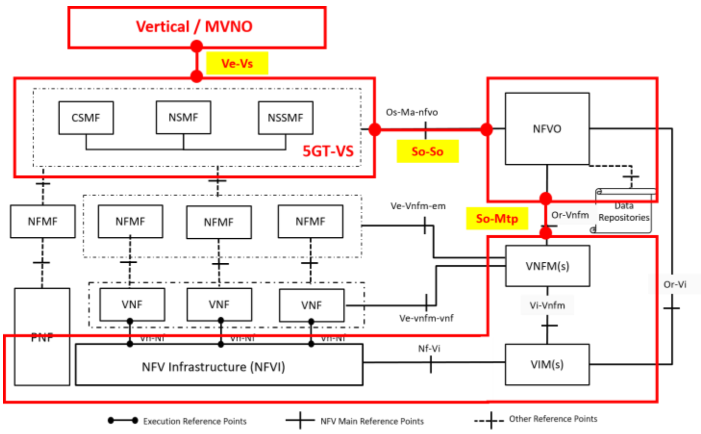

Figure A-2: Mapping 5G-TRANSFORMER architecture into 3GPP/NFV network slicing framework

The major innovative features that 5G-TRANSFORMER project brings for network slicing are:

1. *The definition of an information model for network slice description*. 5G-TRANSFORMER states that a network slice can be described as a (set of) forwarding graph(s) composed of a set of VNFs/PNFs and virtual applications connected with virtual links, where some of them have the specific location constraints and/or the characteristics of multi-edge access (MEC) applications [98]. Aligned with the ETSI NFV vision that a network service is the resource-facing view of a network slice (subnet), 5G-TRANSFORMER claims that a network slice (subnet) could be modelled with an extended NSD.
2. *The definition of an information model for vertical service description with network slicing support*. Logically placed just above NSD model, this model defines two new templates for the description of vertical services running into provided network slices. On one side, there

are Vertical Service Blueprints (VSBs), a baseline template that provides a high-level description of a given vertical service, including service components, topology, and SLA. On the other side, there are Vertical Service Descriptors (VSDs), obtained after providing the missing information in VSBs. To instantiate a vertical service, a vertical first selects a blueprint, then it provides the missing information in the VSB to prepare a VSD. The 5GT-VS then maps the VSD to the extended NSD, describing the network slice for this vertical service, and sends this NSD along with instantiation information to the 5GT-SO.

3. *Network slice operation across federated administrative domains*. 5G-TRANSFORMER allows 5GT-SO to extend NFVO functionality beyond a single domain for the cases where a network slice is composed of network services provided by different administrative domains. For this end, the So-So reference point (see Figure A-1) is defined. This interface allows 5GT-SO to exchange management data and operations with other 5GT-SOs when needed, including communication between NFVO-NSOs (network service federation) and NFVO-ROs (virtual resource federation). Although both options are theoretically doable in 5G-TRANSFORMER, current Proof of Concepts (PoC) are inclined for federation across NFVO-NSOs.

A.2 5G-TANGO

5GTANGO [99] has defined an integrated vendor-independent platform to design, test and operate NFV-based network services. Based on the concepts and architectural components developed in 5G-PPP phase-1 project SONATA [100], the 5GTANGO platform provides an extended multi-modal NFV DevOps approach between multiple actors (i.e. service developers, telecom operators and vertical industries), bringing increased operational efficiency and facilitating the implementation and validation of new services, while accelerating the adoption of NFV technologies.

As seen from Figure A-3, 5GTANGO platform distinguishes three main architectural subsystems: Service Development Kit (SDK), Verification & Validation (V&V) platform, and Service Platform (SP).

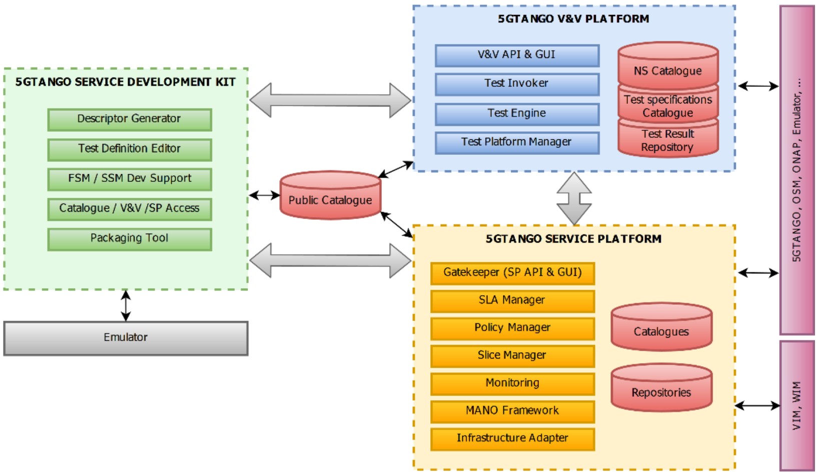

Figure A-3: Overall 5GTANGO architecture [101]

The **SDK** provides a set of tools, empowering the service developer to rapidly validate, verify and test NFV services. Although many tools can be used on their own, the SDK enables a developer workflow supporting the creation of an isolated workspace and project environment, the generation and validation of descriptors, the packaging and on-boarding, as well as the testing and emulation in a local development environment. The **V&V** Platform provides a sandbox environment available to service developers to conduct testing activities over designed VNFs and network services at design time, ensuring they meet a core set of requirements before launching them in operational networks. These activities consist of executing a sequence of tests over submitted VNFs and Network Services, using appropriate tooling (e.g. probes, sinks, traffic generators). In case of failure, developers may need to perform required amendments in the design of the VNF or network service, so interactions between SDK and V&V occur. Finally, the **SP** provides an operational environment for the deployment, management, and orchestration of VNFs and Network Services, as described in Section 2.4.2.

The three aforementioned subsystems are designed following the well-established architectural principles of software engineering (e.g. separation of concerns, single responsibility, and reuse), leading to a loosely coupled platform where each subsystem only deals with issues of a given service lifecycle stage (e.g. SDK focuses on development, V&V platform focuses testing and V&V, and SP focuses on service operation), and thus can be deployed and used independently of the rest of subsystems. Internally, those subsystems use a microservices approach, benefiting from similar architectural principles.

Network Slicing is a major feature that 5GTANGO brings, with applicability and relevance to both SDK and SP. Taking as reference the ideas defined in 3GPP (Section 2.2) and the vision adopted by ETSI NFV (Section 2.3), 5GTANGO introduces the following capabilities for network slicing support:

1. *The definition of a preliminary model for a Network Slice Template (NST)*. Following 3GPP terminology, the NST is a template that allows the design of a Network Slice, enabling the deployment and operation of Network Slice Instances (NSIs) in the SP. 5GTANGO provides a first approach on how a NST should look like, and specifies the minimum information it should have, including at least: (i) Network Slice identifier (Id), (ii) the set of Network Services (NSs) that shall be deployed to build different Network Slice Subnets, and the relationship among them, and (iii) expected QoS behavior. For (i) and (iii), the S-NSSAI and 5QI parameters defined by 3GPP in [53] are used. In (ii), references to NSDs are used to identify the NSs to be deployed, while information on their service access points are needed to enable their interconnection with VLDs. An example of a vanilla NST defined in 5GTANGO can be found in [102].
2. *The architectural definition of a Network Slice Manager*. This component aims at managing NSIs throughout their lifecycle, leveraging the capabilities provided by NFV. To this end, 5GTANGO extends, implement and validates the architecture proposed in [103], where ETSI and 3GPP SDOs combine their existing solutions to offer Network Slicing with NFV support. Based on this approach, the Network Slice Manager is a SP functional block (see Figure A-3) that interfaces with MANO framework to enforce Network Slice lifecycle management actions, consuming the REST APIs the latter exposes via Os-Ma-nfvo reference point.
3. T*he implementation of the Network Slice Manager logic*. This component partially performs NSMF/NSSMF functionality as defined by the 3GPP. Indeed, it perform lifecycle management operations over NSIs, but not performance/fault/configuration management activities. The lifecycle phases that the Network Slice Manager addresses include NST on-boarding, NSI instantiation and NSI termination (other NSI operations such as scaling may be supported in the future). Figure A-4 illustrates this lifecycle.

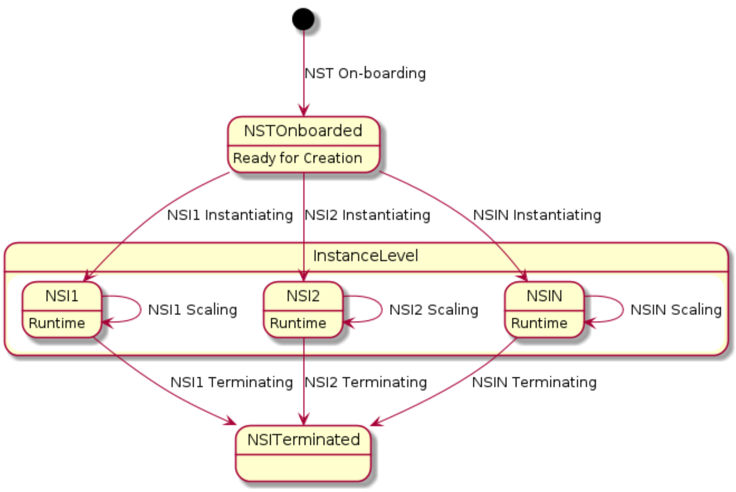

Figure A-4: 5GTANGO NSI Lifecycle Management [101]

The Network Slice Manager developed in 5GTANGO has been integrated as part of the OSM Release FIVE, making OSM one of the first open source initiatives providing management and orchestration functionality with Network Slicing support with ETSI compliance. As seen in Figure A-5, this functional block is implemented with the NetSlicer module, which consumes management services provided by the rest of modules using OSM NBI.

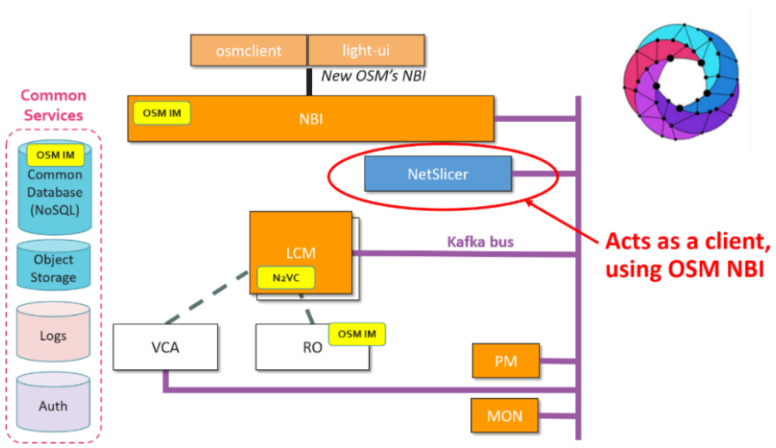

Figure A-5: Integration of 5GTANGO Slice Manager into OSM Release 5

A.3 SaT5G

SaT5G [75] is a 5G PPP Phase 2 project, whose vision is to develop cost effective "plug and play" SatCom (Satellite Communication) solutions for 5G to enable telecom operators and service providers to accelerate 5G deployment in all geographies and at the same time create new and growing market opportunities for SatCom industry stakeholders.

One of the technical challenges that needs to be addressed for the realisation of cost effective "plug and play" SatCom solutions for 5G, and one of the SaT5G research pillars, is the development of the enablers for an integrated 5G-SatCom virtual and physical resource orchestration and service management. To this end, SaT5G pursues the design and implementation of an ETSI-compliant orchestration solution able to perform lifecycle manage of SatCom virtualized functions as well as 5G core functions, aiming to demonstrate the creation of end-to-end satcom-5G slices. SaT5G management and orchestration solution will feature a northbound interface enabling interworking with BSS/OSS of both satellite and terrestrial network operators.

The main component of network slicing is the orchestrator that enables full lifecycle management of slices and associated resources, including SatCom specific NFV management as well as 5G Core management functions. The SaT5G network slicing approach covers end-to-end service provisioning, management, and operation workflows on a per-slice basis. It enables services control and orchestration for different verticals, which are provided over multiple terrestrial and satellite platforms. Deployed architecture will be capable of:

- Creating customisable and dynamic SLA-based multi-domain end-to-end 5G slices with flexible and dynamically deployable satellite resources.
- Dynamic provisioning and instantiation of network slices.
- Flexible multi-domain management and operation of virtualised satellite and terrestrial functions.
- Seamless service provisioning at the network level through automated processes.

Lifecycle management of virtual functions (instantiation, modification, termination) without effective management and orchestration systems is not possible. The management system is a coordination layer on top of all domains, responsible for automated cross-domain lifecycle management / coordination as well as interfacing with the operators (satellite and terrestrial). On one hand, 3GPP SA5 focuses on the specifications, requirements, architecture for provisioning and management of the network and its services. Following 3GPP SA5 specifications, an end-to-end management solution covers the general NF management concept which includes: physical and virtual NF lifecycle management; configuration; fault management and performance management. Merging 3GPP next generation architecture with ETSI MANO framework is a very challenging task, especially if a wide picture that includes both satellite and terrestrial communication technologies is considered. SaT5G targets to tackle this challenge by introducing a holistic management and orchestration system able to coordinate all terrestrial and satellite elements.

In this context, within the SaT5G project, an end to end satellite-terrestrial resource orchestration solution is developed, termed as “TALENT: Terrestrial Satellite Resource Coordinator” [104]. TALENT is a coordination solution which can support end to end services composed of satellite, radio access, cloud and mobile edge computing resources. TALENT is not vendor locked and is NFVO (e.g. OSM, ONAP) and VIM (e.g. OpenStack) agnostic. TALENT covers end to end service management over cloud and edge computational resources. TALENT provides a single and easy to use point of interaction for all stakeholders involved in the ecosystem, i.e. terrestrial and satellite operators, as well as different 5G verticals. TALENT is in line with 3GPP and ETSI architectures, extending it towards satellite systems. Figure A-6 illustrates the high-level design of the TALENT solution developed within the SaT5G project.

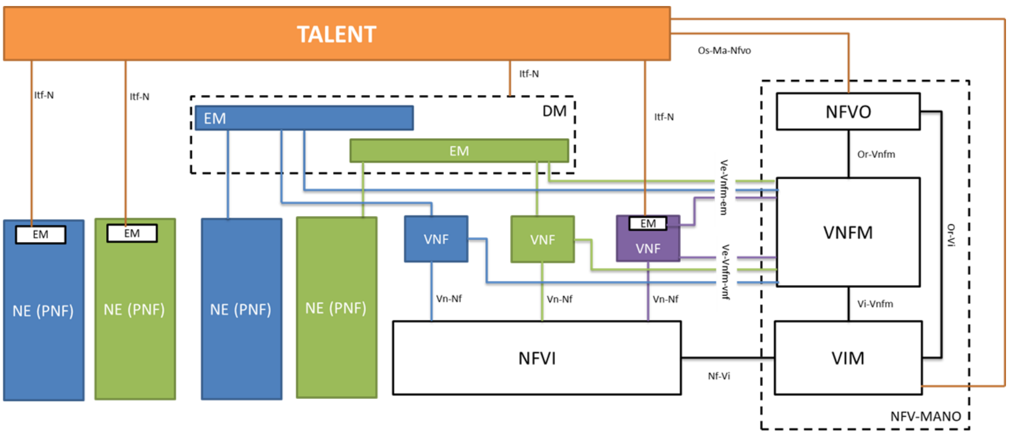

Figure A-6: “TALENT: Terrestrial Satellite Resource Coordinator” [104]

