E2E Network Slicing Architecture

Version 2.0

10 May 2022

This is a Whitepaper of the GSMA

Security Classification: Non-confidential

Copyright Notice

Disclaimer

Antitrust Notice

Table of Contents

1 Introduction

1.1 Overview

The purpose of this document is to emphasize necessity of cross-standardization collaboration to realize E2E Network Slicing Architecture designed from an End-to-End (E2E) perspective, spanning over different technical domains (e.g., device, access network, core network, transport network and network management system) and multiple vendors and to provide guidance on actions in each Standards Developing Organizations (SDOs) and open source projects.

1.2 Scope

The intention of this document is to show a guidance of the entire industry ecosystem; for operators, vendors and service providers to be able to consider common solutions of network slicing which have not been fully defined yet.

Network slicing is defined in 3GPP as a logical network that provides specific network capabilities and network characteristics and is composed of RAN, CN, and transport network. From operator perspective, it is important to provide customers with network service in non-3GPP domain such as SGi LAN/N6 network and fixed access network as well as in 3GPP domain. In this document, E2E network slicing is defined as a logical network spanning over both 3GPP and non-3GPP domains. It is essential to provide customized E2E network service complying with agreed Service Level Agreements (SLAs). Roaming feature is out of scope of this document.

The scope of this document is to provide the description of:

- Blueprint of E2E Network Slicing Architecture
    - Definition of network slicing in this whitepaper
    - Motivation to introduce a network slicing technology
    - Usage scenario and high level requirements
- Technology aspect
    - Indication of several aiming technology aspects to be achieved by network slicing
- SDOs and open source projects
    - Ongoing work and outlook
    - Gaps between existing specifications and expected network slicing which can realize the technology aspects above in each domain
    - Potential proposal of specifications to be considered

These descriptions are snapshot based on current status of these standardization activities, typically activities of the 3rd Generation Partnership Project (3GPP) Release 17.

1.3 Acronyms

Acronym	Description
3GPP	3rd Generation Partnership Project
5GC	5G Core
5GS	5G System
AGF	Access Gateway Function
AI/ML	Artificial Intelligence/Machine Learning
AMF	Access and Mobility Management Function
AN	Access Network
API	Application Programming Interface
B2B	Business to Business
B2B2X	Business to Business to everything
B2C	Business to Consumer
BNG	Broadband Network Gateway
BSS	Business Support System
CAC	Connection Admission Control
CN	Core Network
CNF	Containerized Network Function
COTS	Commercial off-the-Shelf
CPE	Customer Premises Equipment
CSC	Communication Service Customer
CSMF	Communication Service Management Function
CSP	Communication Service Provider
CU	Central Unit
CUPS	Control and User Plane Separation
DN	Deterministic Network
DNN	Data Network Name
E2E	End-to-End
FMIF	Fixed-Mobile Interworking Function
GST	Generic network Slice Template
HNF	Hybrid Network Function
IOC	Information Object Class
KPI	Key Performance Indicator
LSO	Lifecycle Services Orchestration
MNO	Mobile Network Operator
MOI	Management Object Instance
NaaS	Networks as a Service
NBI	Northbound Interface
NF	Network Function
NFV	Network Functions Virtualization

Acronym	Description
NFVI	Network Functions Virtualization Infrastructure
NFVO	Network Function Virtualization Orchestration
NOP	Network Operator
NRF	Network Repository Function
NRM	Network Resource Model
NSC	Network Slice Customer
NSI	Network Slice Instance
NSMF	Network Slice Management Function
NSP	Network Slice Provider
NSSAI	Network Slice Selection Assistance Information
NSSMF	Network Slice Subnet Management Function
OSS	Operation Support System
PCF	Policy Control Function
PDU	Protocol Data Unit
PNF	Physical Network Function
PoP	Point of Presence
PRB	Physical Resource Block
RAN	Radio Access Network
RCA	Root Cause Analysis
RIC	RAN Intelligent Controller
RRM	Radio Resource Management
SBA	Service Based Architecture
SBI	Southbound Interface
SBMA	Service Based Management Architecture
SDN	Software Defined Networking
SDO	Standards Developing Organization
SDTN	Software Defined Transport Network
SLA	Service Level Agreement
SLS	Service Level Specification
S-NSSAI	Single Network Slice Selection Assistance Information
SOP	Standard Operating Procedure
SR	Segment Routing
TN	Transport Network
TSG	Technical Specification Group
TSN	Time Sensitive Networking
TSR	Telecom Security Requirement
TT	Time Sensitive Networking Translator
UE	User Equipment
UPF	User Plane Function

Acronym	Description
URLLC	Ultra-Reliable and Low Latency Communications
URSP	UE Route Selection Policy
V2X	Vehicle to Everything
VAF	Virtualized Application Function
VIM	Virtual Infrastructure Manager
VNF	Virtualized Network Function
VTN	Virtual Transport Network
WIM	WAN Infrastructure Manager
WWC	Wireless and Wireline Convergence

1.4 References

Ref	Doc Number	Title
[1]	NGMN 5G White Paper (2015)	NGMN 5G White Paper (2015)
[2]	3GPP, TS23.501	System architecture for the 5G System (5GS)
[3]	GSMA	Network Slicing Use Case Requirements
[4]	3GPP, TS28.530	Management and orchestration; Concepts, use cases and requirements
[5]	GSMA, NG.116	Generic Network Slice Template
[6]	3GPP, TS28.541	Management and orchestration, 5G Network Resource Model (NRM), Stage 2 and Stage 3
[7]	3GPP, TR28.801	Study on management and orchestration of network slicing for next generation network
[8]	ETSI GS MEC 030 V2.1.1	Multi-access Edge Computing (MEC);V2X Information Service API
[9]	TIP White Pater	Creating Ecosystems for End-to-End Network Slicing
[10]	3GPP, TS38.300	NR and NG-RAN Overall Description; Stage 2
[11]	O-RAN White Paper	O-RAN: Towards an Open and Smart RAN
[12]	Broadband Forum, TR-470	5G Wireless Wireline Convergence Architecture
[13]	Broadband Forum, TR-456	AGF Functional Requirements
[14]	Broadband Forum, TR-124 Issue 6	Functional Requirements for Broadband Residential Gateway Devices
[15]	Broadband Forum, MR-521.1	Marketing Report on 5G Network Architecture - Overview
[16]	IEEE, Std 802.1AS	IEEE Standard for Local and Metropolitan Area Networks--Timing and Synchronization for Time-Sensitive Applications
[17]	Metro Ethernet Forum, MEF 22.3.1	MEF 22.3.1 - Amendment to MEF 22.3: Transport Services for Mobile Networks
[18]	Metro Ethernet Forum, MEF 84 (R1) Draft	Network Slicing
[19]	3GPP, TS28.532	Management and orchestration; Provisioning; Stage 2 and stage 3
[20]	3GPP, TS28.531	Management and orchestration; Provisioning; Stage 1
[21]	IETF, RFC 8453	Framework for Abstraction and Control of TE Networks (ACTN)
[22]	ETSI GR NFV-IFA 024	Network Functions Virtualisation (NFV) Release 3; Information Modeling; Report on External Touchpoints related to NFV Information Model

Ref	Doc Number	Title
[23]	ETSI GR NFV-EVE 012	Network Functions Virtualisation (NFV) Release 3;Evolution and Ecosystem; Report on Network Slicing Support with ETSI NFV Architecture Framework
[24]	ETSI GS NFV-IFA 014	Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Network Service Templates Specification
[25]	ETSI GS NFV-IFA 013	Network Functions Virtualisation (NFV) Release 3; Management and Orchestration; Os-Ma-nfvo reference point - Interface and Information Model Specification
[26]	ETSI GS NFV-IFA010	Network Functions Virtualisation (NFV) Release 4; Management and Orchestration; Functional requirements specification
[27]	ONAP Guilin E2E NS	ONAP Guilin White Paper E2E Network Slicing
[28]	OSM Release NINE, Release Notes	OSM Release NINE, Release Notes, December 2020
[29]	TM Forum, IG 1194	Focus on Services not Slices
[30]	Analysys Mason, Research strategy report	Network slicing: The future of connectivity in a 5G and fibre era
[31]	GSMA, NG.113	5GS Roaming Guidelines
[32]	3GPP, TS28.533	Management and orchestration; Management and orchestration architecture
[33]	3GPP, TS28.540	Management and orchestration, 5G Network Resource Model (NRM), Stage 1
[34]	3GPP, TS28.545	Management and orchestration; Fault Supervision (FS)
[35]	3GPP, TS28.550	Management and orchestration; Performance assurance
[36]	3GPP, TS28.552	Management and orchestration; 5G performance measurements
[37]	3GPP, TS28.554	Management and orchestration; 5G end to end Key Performance Indicators (KPI)

2 Blueprint of Network Slicing

This section covers a high level description of network slicing. It includes definition, motivation to introduce a network slicing technology, usage scenario and high level requirements.

2.1 Definition of technologies for Network Slicing development

Network slicing is a concept for running multiple logical customized networks on a shared common infrastructure complying with agreed SLAs for different vertical industry customers and requested functionalities. To achieve this goal, network slicing needs E2E Network Slicing Architecture to be designed from an E2E perspective, spanning over different technical domains (e.g., device, access network, core network, transport network and network management system) and multiple vendors.

Network slicing contains different technical domains, and many SDOs are working in parallel to provide a slicing solution under their area of competence. As a result, the technical content is fragmented. In order to form an E2E solution, significant work is required in terms of cross-SDO and open source project cooperation and coordination.

Network slicing was outlined as a vision for 5G capability empowering value creation in NGMN 5G White Paper [1]. Then, it becomes one of the key features which are specified in 3GPP to be supported by 5G System [2]. In 3GPP Release 15 specifications, the fundamental system architecture with dedicated identifier to recognize each slice has been specified in both core and RAN domains.

This whitepaper will illustrate that an overall network slicing architecture includes several technical domains from E2E perspective. Detail of roaming feature is out of scope of this document.

2.2 Motivations for Operators to introduce Network Slicing

For operators there are several motivations to introduce network slicing which is able to be deployed with 5G network. These are to provide tailored network service to meet SLA which each customer requests, with flexibility, agility and at low cost.

First of all, network slicing is a promising feature to permit business customers to acquire connectivity tailored to their specific business requirements by using common infrastructure. For example, some need high throughput, but others need massive connectivity, or low latency, or high reliability, and so on. In the 4G era, characteristic of common infrastructure, related to quality of communication services such as throughput and latency, are same. This is because basically common infrastructure and common mobile network are used as mobile services for consumers and enterprise at the same time. Parameters that can be configured to the network as SLA are limited, for instance, redundancy, availability factor and maintenance level. Therefore, it is difficult for operator to provide customized E2E network service complying with

agreed SLAs and to avoid harmful consequences of congestion and failure in common mobile network with consumer mobile services.

In the 5G era, network slicing is expected to be one of major services provided by 5G System. It is expected also for operators that providing tailored services has significant commercial potential. The enterprise opportunity in 5G era is predicted to be $300bn by 2025 [3]. Examples of parameter to be configured are assumed to be throughput, latency, maximum number of UEs, availability factor, limitation of utilization time and place and support or no support of mobility and voice capability.

To realize network slicing, network slice logically consists of dedicated or shared network functions (NFs) of 5G SA network and resources by utilizing emerging technologies such as virtualization so as to provide required network capability. The resources including compute resource, storage resource, bandwidth in transport network and radio resource can be assigned by the operator in a dedicated or shared manner depending on the requirements to meet. For instance, network functions can be assigned in a dedicated manner so as not to receive effect of congestion and failure in other network functions. As the result, a deployed network slice can be formed by a set of NF instances and the required resources with flexibility.

Network slicing deployment using virtualized NFs will benefit from network automation. Network slice customer will be able to order a creation of network slice via APIs provided by network slice provider. The network automation technology is expected to roll out network slices and to modify the network slices in an agile way.

Operators can utilize fully their deployed infrastructure by managing the resource usage of the logical networks. For instance, operators can assign the network slice instances to infrastructure which has spare resources and manage them.

Network automation and resource management can bring cost advantage on deployment and operation for operators due to reduction of workload. As a result, it is expected for operators to maximize their profit with same amount of cost.

2.3 Roles in Network Slicing

Multiple roles related to network slicing are specified in 3GPP TS 28.530 [4] and GSMA NG.116 [5]. According to GSMA NG.116, there are the following five roles;

- Communication Service Customer (CSC): Uses communication services, e.g. end user, tenant, vertical.
- Communication Service Provider (CSP): Provides communication services. Designs, builds and operates its communication services. The CSP provided communication service can be built with or without network slice.
- Network Operator (NOP): Provides network services. Designs, builds and operates its networks to offer such services.
- Network Slice Customer (NSC): The CSP or CSC who uses Network Slice as a Service.
- Network Slice Provider (NSP): The CSP or NOP who provides Network Slice as a Service.

An organization can play one or several roles simultaneously. According to 3GPP TS 28.530 [4], there are several types of communication service, which are Business to consumer (B2C), Business to business (B2B) and Business to business to everything (B2B2X). As an example, assuming the B2B2X case, operators provide network slice service as NSP to other CSPs, e.g. content providers and application providers. CSP can offer their own communication services to their own customer, CSC, by utilizing the provided network slice.

2.4 SLA/SLS attributes for Network Slicing

A set of attributes that can characterise a type of network slice/service are specified as Service Profile in 3GPP TS28.541 [6] and as Generic network Slice Template (GST) in GSMA NG.116 [5], in order to express CSC's Service Level Specification (SLS) which is a set of service level requirements associated with SLA to be satisfied by a network slice. The Service Profile is assumed to be utilized in 3GPP 5G System. From an E2E perspective, GST has been specified as a common set of attributes to be referred from 3GPP as well as transport domain. It is noted that alignment of attributes between Service Profile and GST is proceeding at the present time. Attribute details can be found in [5] and [6].

2.4.1 Service specific security

This section describes the potential security requirements that industry verticals may place in their SLAs and how those may be reflected in the network.

The GST attributes contain currently the following optional security related "isolation attribute". It can have the following values:

- No Isolation
- Physical Isolation
- Logical Isolation

The isolation attributes apply for the whole network slice and not to individual elements of the slice e.g. RAN, core network etc. The details are outlined in section 3.3.

A service might have further security requirements e.g. RAN security,

2.5 High level Requirements

Many use cases for network slicing, along with requirements for each use case, were mentioned in GSMA document [3] for instance.

In this document, technology aspects to satisfy these requirements are described in section 3. Target architecture and interface aspects are described in section 4.

3 Technology Aspects to achieve Network Slicing

This section covers technologies that are required in each of the different technical domains (e.g., device, access network, core network, transport network and network management) to realize the target network slicing. These technology aspects are categorized and introduced in this section.

3.1 SLA/SLS

A SLA is a commitment of provisioned network services between an operator and a consumer. The consumer declares communication service(s) requirements to the operator. These requirements are called SLS. Network performance attributes such as throughput, latency and reliability could be part of technical specification of SLS and specified in GSMA NG.116 [5]. SLS with some consumer may need to be change dynamically.

In order to guarantee a SLS with each consumer, network slice corresponding to each CSC has to reserve appropriate amount of resource (e.g. radio resource, compute resource) and to deploy network function such as UPF at right location, especially for low latency communication. SLS assurance mechanisms to tune slice behaviour dynamically is also required, especially in RAN domain.

GSMA North Americas Network Group's (NANG) Network Slicing Taskforce (NETSLIC) considered methods to identify key verticals and interpret the respective needs in order to map that information into attributes (existing and extended GSTs) and values (NESTs) during 2019-21. Based on these experiences, it was noted that the task is oftentimes challenging because the terminology of the mobile communications industry and standards setting organizations differs from the one that verticals tend to use. It is thus important to ensure the Network Slice Operators capture the practical requirements of the field correctly, understanding current and future key needs that are expressed in various non-standardized terms. More information on the process and selected examples of the NETSLIC can be seen in their Whitepaper "Network Slicing: North America's Perspective" that will be available at the GSMA North America's web page (https://www.gsma.com/northamerica/) in spring 2021.

GSMA North America brings forward the above mentioned work in North Americas Vertical Application Taskforce (NAVA) as of 2021 by exploring the new enablers of the networks based on the 3GPP development. The NAVA TF addresses inter-operable vertical applications development. It identifies suitable application deployment architectures specific to North America regional interests and fills the gaps on the related tasks within the current GSMA North American entities. The focus of the NAVA TF is on vertical applications relevant to the North America region in both the 4G and 5G domains.

3.2 Provisioning and Operation

There are some phases of lifecycle in network slice such as service design, provisioning, deployment, operation and removal. In order to provide consumers network slicing services, network service provider is required to map CSC requirements into typical value of GST attributes and to deploy resources and network function properly so as to satisfy CSC's SLAs. In operation phase, it is required that each of the network slices are monitored

whether its service quality meets CSC's SLA and that there is a mechanism to take actions to resolve service degradation automatically if service quality doesn't meet the SLA.

As network slicing is spanning over several network domains, E2E orchestration across these domains is a key capability to provide network slice service offerings to NSC from perspective of lifecycle management. E2E orchestration can control management function at each domain to provide lifecycle management service such as provisioning and can make services available faster to NSC by reducing offering time.

To utilize E2E orchestration effectively, automation is a key enabler to bring benefit, e.g. cost advantage, of removing manual interventions in lifecycle management, although there is automation in various phase such as provisioning, operation fault and performance management.

3.3 Isolation

Isolation is one of the key expectations of network slicing. A network slice instance (NSI) may be fully or partly, logically and/or physically, isolated from another network slice instance. There are the different types of isolation mentioned in GSMA NG.116 ver. 4.0. For the isolation at virtualization level, it is noted that only VM type of virtualization technologies are explored but not OS containers as the other type of virtualization technology. Isolation requirements on OS containers is work in progress as current native solutions don't offer much and additional support is required. ETSI NFV is working on SEC023 GS spec dedicated to this. Therefore, it is also noted that these isolation types may be modified in NG.116 in the future. It is recommended to refer to NG.116 for getting up-to-date information.

Network slicing allows the concurrent execution of multiple NSIs on top of a common infrastructure, satisfying their individual service Key Performance Indicators (KPIs) while guaranteeing their independence. The use of a single shared infrastructure for this makes isolation a key requirement in the support of network slicing. Isolation can be defined as the ability of a NSP to ensure that congestion, attacks and lifecycle-related events (e.g. scaling in/out) on one NSI does not negatively impact other existing NSIs.

Isolation in network slicing is a multi-faceted problem, with multiple dimensions that need to be carefully addressed. The dimensions include performance, management and security/privacy. *Isolation in terms of performance* represents the ability to ensure that service KPIs are always met on each NSI, regardless of the workloads or faults of other existing NSIs. *Isolation in terms of management* represents the ability to ensure that individual NSIs can be managed as separate networks, with the possibility of the NSC to retain control of the slice. Finally, *isolation in terms of security and privacy* represents the ability to ensure that any type of intentional attack occurring in one NSI must not have an impact on any other NSI. This means that each NSI shall have appropriate mechanisms preventing unauthorized entities to have read and write access to slice-specific configuration / management / accounting / user information, and be able to record any of these attempts.

The following subsections provide more details on these isolation dimensions.

3.3.1 Slice performance isolation

This isolation dimension is about segregating resources of a NSI from other NSI. To that end, the NSP shall leverage mechanisms providing means to (i) split the physical infrastructure into a set of partitioned resources; and (ii) allocate these resources into separate slices.

On the one hand, resource partitioning requires the definition separate resource quotas out of the infrastructure. The mechanisms allowing for this infrastructure partitioning are defined per resource domain, including radio resource domain (e.g. Radio Frequency RF carriers arranged into flexible time-frequency resource grids), transport network (i.e. switching and routing nodes, multi-technology connectivity links) and compute resource domain (i.e. providing virtualized execution environments where VM hosting access network (AN) and core network (CN) functionality can be deployed). For the radio resource domain, mechanisms based on spectrum planning and admission control can be considered for both single-cell and multi-cell scenarios. For the transport network, hard isolation (i.e. circuit switched connections) or soft isolation (i.e. statistical multiplexing) approaches can be applied, with some solutions in between. Finally, for the compute resource domain different technology options can be selected for virtualization. Table 1 illustrates some of them and their impact on the isolation degree achieved. Physical isolation probably is required depending on required isolation level in compute resource domain, too. For example, it may be necessary to isolate CPU cores or memory region for each virtualization technology.

Virtualization technology	Ring Level	Isolation provided by	Image size	Operation agility
VM	Level 0 or -1	Hypervisor	Large	Low
POD	Level 3, enforced at Level 0	Host OS	Medium	High
Unikernel	Level 0	Hypervisor	Small	Medium

Table 1: Comparative analysis of Virtualization technologies

On the other hand, the objective of resource allocation mechanisms is to dispatch the individual resource quotas to corresponding NSIs, in such a manner that the slice-specific KPIs are met. This includes Physical Resource Block (PRB) scheduling algorithms for radio resource quotas, encapsulation solutions for network resource quotas, and cloud orchestration engines for compute resource quotas.

3.3.2 Slice management isolation

This isolation dimension requires the NSP to have multi-tenancy support. Mechanisms for multi-tenancy support allows the NSP to share its management resources, e.g. Network Function Virtualization Orchestration (NFVO), Network Slice Management Function (NSMF), Network Slice Subnet Management Function (NSSMF), among the participant NSCs, in such a manner that individual NSCs are able to operate their NSIs with complete independence. To that end, the NSP shall provide different NSC with separate management spaces, each defining the (performance, configuration, lifecycle, fault) management capabilities that the

NSC can consume from the NSP. This depends on the capability exposure level agreed between both parties. It is noted that this is related with content in Section 3.7 – Industry Customer Management

3.3.3 Slice security isolation

Slice security isolation is a security feature, as it separates slice related data and activities of different customers. With the corresponding slice specific enforcements slice security isolation can prevent unauthorized access and modification to data, processes, services or functions. If the SLA includes slice security isolation.

Then the customer has the options of isolation type for physical isolation and for logical isolation, as addressed in section 3.3.

Currently NG.116 does not cover heterogeneous scenarios, where e.g. one network domain provides physical isolation and the other network domains logical isolation. Also inter-slice communication is not part of the current NG.116.

3.4 API Exposure

API exposure is important to interact with application functions, network slice subnets, cloud platform and so on. These could be operated by other network operator or public cloud provider. There are some helpful discussions of slice interaction between each role in network slicing [7]. For example, the discussion of Communication Service Management Function (CSMF), NSMF and NSSMF suggests possibility that a slice consists of some slice subnets. In other words, a slice provider could need collaboration with other network operator providing slice subnet(s) to guarantee E2E service level, e.g. RAN slice subnet and core slice subnet(s) are from different providers. There are also discussions of multi-operator edge computing scenarios. According to ETSI MEC specifications [8], service providers have to provide a robust network to customers by offering a service across all the territories including several network operator's coverage areas.

If a service provider chooses or creates some network slices to meet customer's demand, the challenge of network slice provider's API exposure is as follows:

- Standardization or openness for northbound API of NSMF which is an interface for service provider.
- Linkage enhancement to public cloud (Public cloud provider could become a slice provider or a service provider.)
- API exposure for design, creation, management, modification, deletion and utilization of a network slice
- Providing as much E2E network assets (i.e., RAN, transport, mobile core, cloud and so on) as possible, like a network slice subnet. If E2E is not possible, then the miss-usage of service APIs and data leakage need to be prevented.

3.5 Low latency communication

A feature to achieve low latency communication is included in 3GPP Release 16 or later. In RAN domain, optimization of physical layer such a subcarrier spacing, the maximum number of HARQ processes and so on are discussed. In transport domain, capability that controls network path for each slice identifier shall be required to achieve low latency communication. In addition to that, low latency communication could be achieved by deploying UPF and CU at the edge (i.e. nearby a base station) due to reduction of transmission delay time. Assuming that UPF and CU are virtualized, flexibility of where UPF and CU are deployed based on demand is important, while it is needed to take care of power consumption and footprint of hardware for UPF and CU due to limitation of the edge office's capacity.

3.6 Security

Network slicing involving logical segregation of physical network infrastructure into distinct logical units called slices each offering distinct characteristics and SLA's. However, when the underlying infrastructure is shared there are certain challenges that need to be addressed including:

- How to ensure resources from different slices not to impact each other
- Define a pre-emption rules when resources are scarce
- Define E2E security architecture
- Define security requirements of inter operator slices (e.g. in case of Edge slices) [9]

With the standardization of 3GPP Release 16 and many operators commercializing plans it is very important to evaluate slicing security impacts in the context of industry and other verticals as the stringent industry requirements make it necessary to apply an industry grade security and Telecom security requirements (TSR) as defined by both international and local regulations. The most important requirement to apply slicing in industry is that the data generated and owned by the industry partner do not leave the industry premises. Further it requires clear data governance for example demarking data boundary between industry control, management, service data and MNO. These security requirements also stipulate that the industry data is not available or leak to any outside organization including MNOs who may be the infrastructure provider for such solution.

This kind of data protection of the industry partner poses several challenges:

- Slice isolation on transport and signalling (application) layer in the core network, if no network slices spanning over all domains to provide E2E communication are used (RAN and CN) as the MNO may not want to reproduce a whole core network for the industrial partner
- Authorization not only on API service, slice and node level, but down to information element level.

Those challenges can be addressed by matching information from different layers and combining them to avoid accidental data leakage.

3.7 Industry Customer Management

Industry customer management is both the business and technical requirement to offer network slicing to industry and enterprise. As per network slicing promise the CSP's have to deliver network service (e.g. Deterministic Network (DN)) with distinct SLA characteristics to every slice and industry customer. In addition, there must exist a way to monitor slice in real time to ensure the industry customer can measure it and take any actions or pose requirements to the MNO to optimize it.

3.7.1.1 Business requirements

The business requirements from industry and enterprise require a distinct portal and provisioning mechanism through which an industry customer can both provision and monitor the slices within the scope of defined SLA's. The range of actions the industry customer can perform needs to be defined by the SLA.

The former is addressed using offering slicing through CSP's market place by integrating slicing management solutions with E2E orchestration and BSS solutions. However, the latter is complex primarily due to fact that there are certain slicing KPIs which are global in nature and cannot be adjusted/tuned as per business requirements e.g. Infrastructure utilization, resilience etc.

3.7.1.2 Technical requirements

The technical requirements for network slicing require the delivery of network service such as DN that can promise a defined and distinct SLA for each industry customer. Today with 3GPP Release 16 and key features like Time Sensitive Networking (TSN), 5G LAN, Campus networks it has been made possible to control and define a DN. However, to ensure that each provisioned slice delivers the required performance as resources scale requires features such as

- Real time slice monitoring
- Optimization and the possibility to take correcting actions
- Policy management

The slice monitoring requires that the business verticals are able to monitor slice and when a certain slice is not delivering SLA there are mechanisms that can advise customers the Root Cause Analysis (RCA) and remedial actions without the involvement from NOP. Not only the system should be capable to output the corrective action but it should be able to issue the corresponding policy to infrastructure to correct and enhance it.

3.7.1.3 Operational requirements

The business requires to adapt and follow each industry standard operating procedures (SOPs), that is official or usual way to be expected to do particular things in an organization, and operational practices to manage its network and slices, it requires the slicing to be both scalable and flexible as we apply it to different use cases and verticals.

It is vital for CSP's to automate and deliver the different slices by simple operation to ensure it can become a managed services provider for many different industries.

4 Network Slicing Architecture

This section covers an overall Network Slicing Architecture which illustrates technical domains in order to be referred from following sections.

As mentioned previously, network slicing is a concept for running multiple logical customized networks on a shared common infrastructure complying with agreed SLAs for different vertical industry customers (or tenants) and requested functionalities. To achieve this goal, network slicing needs to be designed from an E2E perspective, spanning over different technical domains (e.g. device, access network, core network, transport network and network management system). The creation of simplicity out of this complex environment requires applying the principles of abstraction and separation of concerns into the architecture design. The result is an architecture structured into different stratums, with segregated scope and different technology pace in each, namely infrastructure stratum, network and application function stratum, and O&M stratum. Figure 1 captures the logical structure of these stratums into the Network Slicing Architecture.

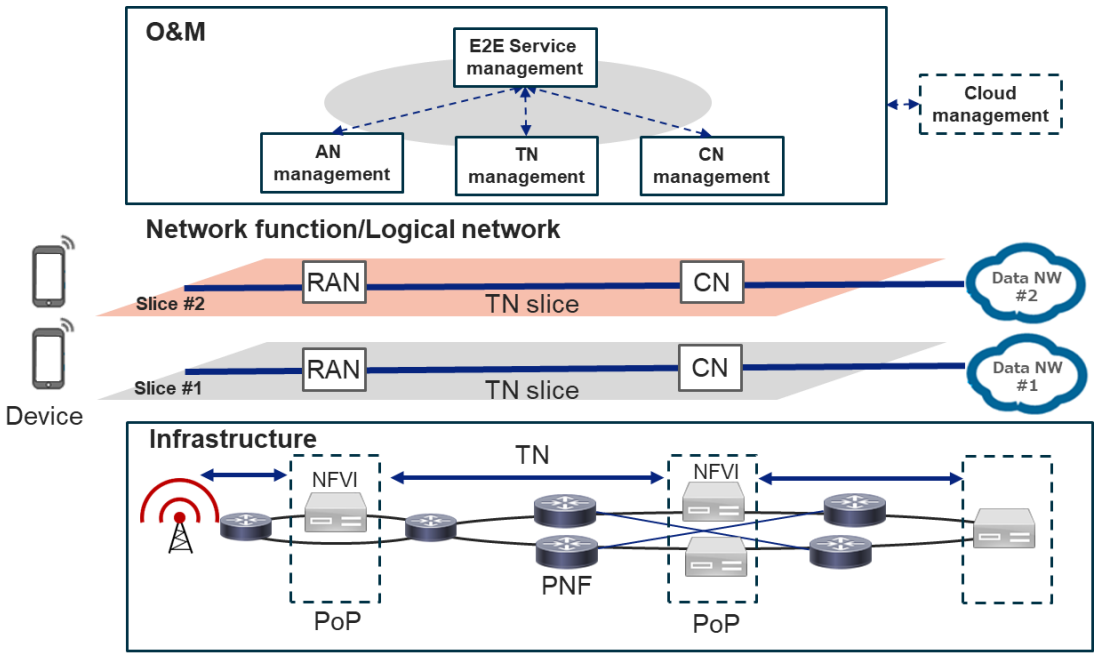

Figure 1: Overall Network Slicing Architecture example for mobile network

4.1 Design principles of infrastructure stratum

The infrastructure stratum compromises all the hardware and software resources building up the operator's substrate, including user equipment and a compute, storage and networking fabric. This fabric can be used to implement physical network nodes and/or to define a distributed cloud environment. i.e. a NFV Infrastructure (NFVI). While the former delivers bespoke Physical Network Functions (PNFs), the latter allows providing a multi-site

virtualized execution environment for VxF hosting, including Virtualized Network Functions (VNFs) and Virtualized Application Functions (VAFs). The lower side of Figure 1 shows the archetypal topology of an operator's 5G infrastructure. This infrastructure consists of a RAN with micro and macro cells attached via dedicated fibres to a multi-tier Transport Network (TN). This capillarity in TN design, based on distributing compute capacity across Points of Presence (PoPs) physically deployed at different aggregation levels, paves the way for the rapid decentralization. In this example, we consider a three-tier TN, with operator's managed PoPs classified into regional and central PoPs. Unlike central PoPs, which represent typical core cloud sites, regional PoPs take the role of edge computing nodes. These nodes provide virtualization capabilities closer to service delivery endpoints in order to reduce the delay budget, making them ideal to host functions and applications for URLLC services.

4.2 Design principles of network and application stratum

The network and application function stratum is formed of a collection of PNFs and VxFs. It provides the user, control and application plane functionality across the different network segments, including RAN, TN and CN. A network slice is defined as an E2E composition of PNFs and VxFs. The resource provisioning and allocation of the individual VxFs depend on the needs of the associated network slices.

Network slice is built and designed on the basis of requirements described in the GST. To fulfil these requirements a specific design has to be achieved in RAN, TN and CN. As a result, the network slice is the interlocking of RAN sub slice, TN sub slice and CN sub slice. Even though TN slicing is not well defined by SDOs, the TN is fully part of the network slice as it can be dedicated to a slice or shared between slices. The design of the TN sub slice is essential to fulfil E2E slice requirements such as E2E latency, E2E availability, isolation level and throughput for example. Such shared network nodes can also be potentially possible in the CN and requires special attention to avoid unauthorized access.

4.3 Design principles of O&M stratum

Finally, the O&M stratum conveys the OSS functionality that allows for the deployment and operation of individual network slices. This stratum aims at handling the operational complexity that network slicing may bring.

This complexity is due to the need to manage and orchestrate a wide variety of slices across all network segments, with an E2E perspective. The specificities of these segments, with different pace of each technology evolution and with solutions from different vendors, unveil unneglectable integration issues for operators. This is exacerbated as the number of slices running in parallel increases. To address these integration and scalability challenges, the operators are required to adopt novel architecture approaches on the O&M stratum. And Service Based Management Architecture (SBMA) is one of them. This architecture style means migrating from functional blocks exposing telecom-style point-to-point interfaces (e.g. Network Managers / Element Managers providing 3GPP Itf-N interfaces) to management services exposing APIs based on web-based technology. This change of paradigm facilitates a rapid evolution of management and orchestration capabilities in compliance with the innovation of the underlying network, by simply adding or updating APIs using libraries and other enablers (e.g. development tools, specification tools, code generators) which are

broadly available. This approach allows service innovation with minimal integration effort. Different SDOs have already captured the benefits of having a SBMA in their specifications. For example, 3GPP SA5 and ETSI ISG ZSM defined their architectural framework based on SBMA. Even ETSI ISG NFV, which originally chose an interface-centric approach for the design of the MANO framework, has now decided to migrate towards a SBMA from NFV Release 4 on.

Architecture stratum	Design principles
Infrastructure stratum	Coexistence of purpose-built and COTS hardware; in-network computing, clustered NFVI; multi-technology transport network
Network and application function stratum	Control and User Plane Separation (CUPS); cloud-native VxFs; RAN functional splitting; Integration of O-RAN ALLIANCE framework.
O&M stratum	SBMA; extensibility; MF stateliness; model-driven operation; reproducibility; management capability exposure;

Table 2: Design principles of stratums

5 Ongoing work and outlook

5.1 Radio Access Network

5.1.1 3GPP RAN

The 3GPP covers cellular telecommunications technologies, including radio access, core network and service capabilities, which provide a complete system description for mobile telecommunications. The three Technical Specification Groups (TSGs) in 3GPP are Radio Access Networks (RAN), Services & System Aspects (SA) and Core Network & Terminals (CT).

The 3GPP RAN is responsible for the definition of the functions, requirements and interfaces in the RAN domain. Regarding network slicing, the RAN slicing-awareness features were discussed in the 3GPP RAN. In Release 15 specification, the following key principles apply for support of network slicing in NG-RAN [10]:

- RAN awareness of slices in order to support a differentiated handling of traffic for different network slices which have been pre-configured.
- Selection of RAN part of the network slice by NSSAI provided by the UE or the 5GC.
- Resource management between slices, which is that NG-RAN supports policy enforcement between slices as per SLAs.
- Resource isolation between slices which may be achieved by means of radio resource management (RRM) policies and protection mechanisms that should avoid shortage of shared resources in one slice breaking the SLA for another slice.

The Release 15 specification enables RAN to identify network slice and the enhancement of some features are considered. However, these features RAN slice provides are implementation dependent. It may lead to service level inconsistency between vendors.

At this moment, the 3GPP RAN studies enhancement of RAN features such as mechanisms to enable UE fast access to the cell supporting the intended slice and mechanisms to support service continuity as Release 17 work. Specifications to be improved is assumed to be 3GPP TS 38.300.

5.1.2 O-RAN ALLIANCE

O-RAN ALLIANCE is a world-wide community of mobile network operators, vendors, and research & academic institutions operating in the RAN industry. O-RAN ALLIANCE's mission is to re-shape the RAN industry towards more intelligent, open, virtualized and fully interoperable mobile networks. The new O-RAN standards will enable a more competitive and vibrant RAN supplier ecosystem with faster innovation to improve user experience. O-RAN based mobile networks will at the same time improve the efficiency of RAN deployments as well as operations by the mobile operators.

To achieve this, O-RAN ALLIANCE is active in 3 main streams:

- The specification effort
  - new standards for open and intelligent RAN
- O-RAN Software Community
  - open software development for the RAN (in cooperation with Linux Foundation)
- Testing and integration effort
  - supporting O-RAN member companies in testing and integration of their O-RAN implementations. O-RAN focuses on technical aspects of the RAN and stays neutral in any political, governmental or other areas of any country or region. O-RAN does not get involved in any policy-related topics.

O-RAN ALLIANCE published specifications and reports relating to network slicing, available on the O-RAN ALLIANCE web site (https://www.o-ran.org/specifications) under the O-RAN ADOPTER LICENSE AGREEMENT.

In the technical report "O-RAN Study on O-RAN Slicing", the high level view of O-RAN slicing framework and architecture with the purpose of helping identify requirements for O-RAN defined interfaces and functions are provided.

In the technical specification "O-RAN Slicing Architecture", multiple O-RAN Slicing Use Cases are captured. O-RAN Slice Subnet Management and Provisioning Use Cases are the use cases and procedures which is necessary for O-RAN Slice Subnet Management feature that is in-line with 3GPP Slice Management framework.

The other O-RAN Slicing Use Case is the RAN Slice SLA Assurance Use Case. The network slicing should support the needs of the business through the specification of several service KPIs such as data rate, traffic capacity, number of users, latency, reliability and availability. These capabilities are specified based on a SLA between the mobile operator and the business customer, which has resulted in increased interest in mechanisms to ensure slice SLAs and prevent its possible violations.

O-RAN's open interfaces combined with its AI/ML based innovative architecture can enable such challenging RAN Slice SLA assurance mechanisms. The RAN Slice SLA Assurance use case is captured as one of network slicing use cases in "O-RAN Use Cases and Deployment Scenarios Whitepaper". In the use case, the non-Real-Time (RT) RAN Intelligent Controller (RIC) and Near-RT RIC can fine-tune RAN behaviour based on RAN specific slice SLA requirements to assure RAN slice SLAs dynamically

These two Use Cases are also captured in O-RAN Use Cases Analysis Report and Detailed Specification. Further enhancements of the specification to support O-RAN Slicing use cases are required.

As explained previously, O-RAN Slice Subnet Management and Provisioning specifications are to be defined.

In the white paper "O-RAN Use Cases and Deployment Scenario", there are some impacts on O1/O2/A1/E2 interfaces to support RAN Slice SLA Assurance Use Case. Utilizing the slice specific performance metrics received from E2 Nodes, Non-RT RIC monitors long-term trends and patterns regarding RAN slice subnets' performance, and trains AI/ML models to be deployed at Near-RT RIC. Non-RT RIC also guides Near-RT RIC using A1 policies, with possible inclusion of scope identifiers (e.g. S-NSSAI) and statements such as KPI targets. Near-RT RIC enables optimized RAN actions through execution of deployed AI/ML models or other slice control/slice SLA assurance xApps in near-real-time by considering both O1 configuration (e.g. static RRM policies) and received A1 policies, as well as received slice specific E2 measurements. O-RAN slicing architecture enables such challenging mechanisms to be implemented which could help pave the way for operators to realize the opportunities of network slicing in an efficient manner as well as potentially change the way network operators do their business.

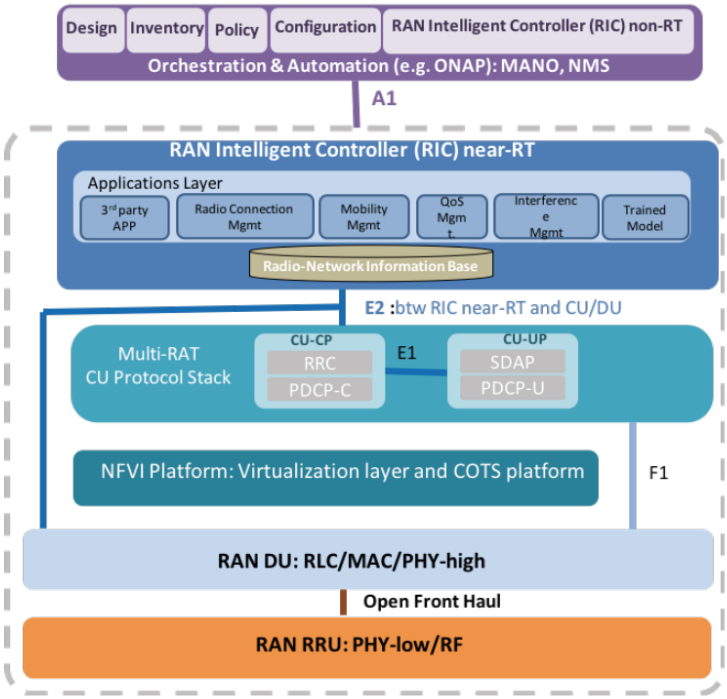

Figure 2: O-RAN ALLIANCE Reference Architecture [11]

5.2 Mobile Core Network

5.2.1 3GPP SA

The 3GPP SA is one of the TSGs in 3GPP and is responsible for the overall architecture and service capabilities of systems based on 3GPP specifications. When it comes to network slicing, the Release 15 specifications enables the following basic network slice features [2]:

- Identification and selection procedure of a network slice when the UE registers and connects
- S-NSSAI handling from subscription information aspects
- UE NSSAI configuration
- Operation Overview

In Release 16, 3GPP SA specified not only enhancements of network slicing such as access to specific network slices authorized and authenticated through additional user identifiers but also enhancements of Service-based Architecture (SBA). The SBA provides higher flexibility and better modularization of the 5G System for the easier definition of different network slices. As the results, it may realize full-scale virtualization.

The existing specification enables basic network slice features only for static operation with single vender, but dynamic control in accordance with SLA and multi-vendor are not addressed. For instance, support for the GST parameters enforcement is being discussed as Release 17 work items. Hence, implementation of the enforcement capability depends on vendors at this moment.

In Release 17, several parameters of the GST parameters such as maximum number of UEs and PDU sessions and maximum UL and DL data rate are studied in order to keep SLS. It is expected to also consider the rest of the GST parameters.

In terms of policy enforcement for security, the central point of security enforcement for the UE connecting to a specific slice lies in the AMF. The AMF can trigger extra secondary authentication, if required for this slice. General policy control information is provided by the Policy Control Function (PCF) to other control plane functions so that they can enforce them. For the core network, the NRF has the main responsibility for authorizing access to a specific network function. The NRF can optionally validate if a network function belonging to a specific slice is allowed to consume a service API of another network function.

Specifications to be considered are 3GPP TS 29.510, TS 23.501, TS 23.502 and TS 23.503.

5.2.2 Broadband Forum
- 5G Fixed Mobile Convergence architecture and requirements -

The Broadband Forum is specifying two mediation functions for the interworking of existing fixed access and the 5GC. These are the access gateway function (AGF) and the fixed-mobile interworking function. Logically they have a similar function to that of a gNodeB, but differ in that they attach to existing wireline networks. An AGF accepts an Ethernet or L2TP handoff from the access network and supports both legacy and 5G enabled CPE. The fixed-mobile interworking function (FMIF) accepts an IP handoff from deployed BNGs and only supports legacy CPE.

For the initial phase of the Broadband Forum work, legacy CPE support only extends as far as a single PDU session to a single slice/DNN. Support for 5G enabled CPE does not have this limitation. In both cases the resource model for the access is assumed to be a static contract (possibly sourced from a 3PP access provider) and the AGF or FMIF will adapt the 5G QoS onto the available access resources (including performing CAC where appropriate).

The FMIF specification is still under development. An overview of the BBF WWC architecture can be found in BBF TR-470 [12]. Issue 1 of the AGF specification can be found in BBF TR-456 [13]. And the BBF residential gateway specification as augmented for 5G operation is in publication as TR-124 issue 6 [14].

5.3 Transport Network

5.3.1 IETF

The IETF is a standardization body for internet and IP network. The IETF has recently been discussing the overall system for transport network between 3GPP defined user plane functions such as gNB, UPF and DN. In order to meet an agreed SLS, usage of transport network is important in the context of path control and SLS assurance. Identifier of 3GPP network slice is S-NSSAI, while on transport network level there can be DSCP, VLAN tag, MPLS label, and so on. The aims of the discussion in the IETF are how to use existing technology applicable to support 3GPP network slice as well as the management system with northbound interface. The management aspects are described in Section 5.4.2.

5.3.2 Broadband Forum
- 5G Transport architecture and requirements-

With the emergence of 5G, Network Operators are finding that they need to enhance their Transport Networks both to meet the increase demands and to deliver the new services that 5G makes possible. Broadband Forum is developing recommendations for the architecture of these new transport networks to enable operators to design, develop, and deploy the needed capabilities. See also the Marketing Report [15],

5.3.3 IEEE 802
- switched Ethernet networking and Time Sensitive Networking -

IEEE provides specifications for switched Ethernet networking and for TSN. According to 3GPP TS 23.501 [2], for supporting TSN time synchronization, the 5GS is integrated with the external network as a TSN bridge modelled as an IEEE Std 802.1AS [16] compliant entity. For TSN time synchronization, the entire E2E 5G system can be considered as an IEEE Std 802.1AS [16] "time-aware system". Only the TSN Translators (TTs) at the edges of the 5G system need to support the IEEE Std 802.1AS [16] operations. The TTs located at the edge of 5G system fulfil some functions related to IEEE Std 802.1AS [16], e.g. gPTP support, timestamping, rateRatio.

5.3.4 ITU-T SG15
- optical networking, and synchronization -

The international standards (ITU-T Recommendations) developed by Study Group 15 detail technical specifications giving shape to global communication infrastructure. The group's standards define technologies and architectures of optical transport networks enabling long-haul global information exchange; fibre- or copper-based access networks through which subscribers connect; and home networks connecting in-premises devices and interfacing with the outside world.

Network, system and equipment features covered by SG15 include routing, switching, interfaces, multiplexers, cross-connect, add/drop multiplexers, amplifiers, transceivers, repeaters, regenerators, multilayer network protection switching and restoration, operations, administration and maintenance (OAM), network synchronization for both frequency and precision time, transport resource management and control capabilities to enable increased transport network agility, resource optimization, and scalability (e.g. the application of software-defined networking (SDN) to transport networks).

5.3.5 MEF

MEF, Metro Ethernet Forum, is a global industry forum for network and cloud providers. MEF initially focused on carrier Ethernet but now covering much broader scope including underlay optical and IP transport and overlay SD-WAN, but also orchestration with MEF Lifecycle Services Orchestration (LSO). MEF provides foundational definitions and concepts for Metro Ethernet Services, service attributes and parameter requirements and as well as traffic classification, traffic profiles and related recommendations to deliver Carrier Ethernet Services but the developed framework is applicable to any transport connectivity services. Transport Services for mobile networks, i.e. Carrier Ethernet service definitions for interconnecting 5G RAN and Core equipment, are described in [17]. MEF is currently investigating network slicing as defined by 3GPP in the context of MEF LSO with 2 options: one where the slices use separate physical links and another where the slices share physical links.

MEF has started to identify some requirements regarding the fact that the slices must be instantiated prior to being used to offer a service, or that service providers must enforce isolation of the network slices instantiated on its infrastructure. Regarding management, the new requirements are that the service provider must confine the service and resource orchestration, control and management to the corresponding network slice but also the subscriber's orchestration, control and management activities on the network presented by a network service to the corresponding service agreement and utilized network slice.

Current MEF specification to consider is MEF 84 (R1) Draft [18].

5.4 O&M

The E2E service operation and management requires interconnections with E2E network & service management domain and controllers across different technological domains to produce an E2E view of the entire network slicing. Figure 3 depicts high level diagram of O&M domain. In the following subsections, the SDOs and open source projects related to the O&M domain is addressed.

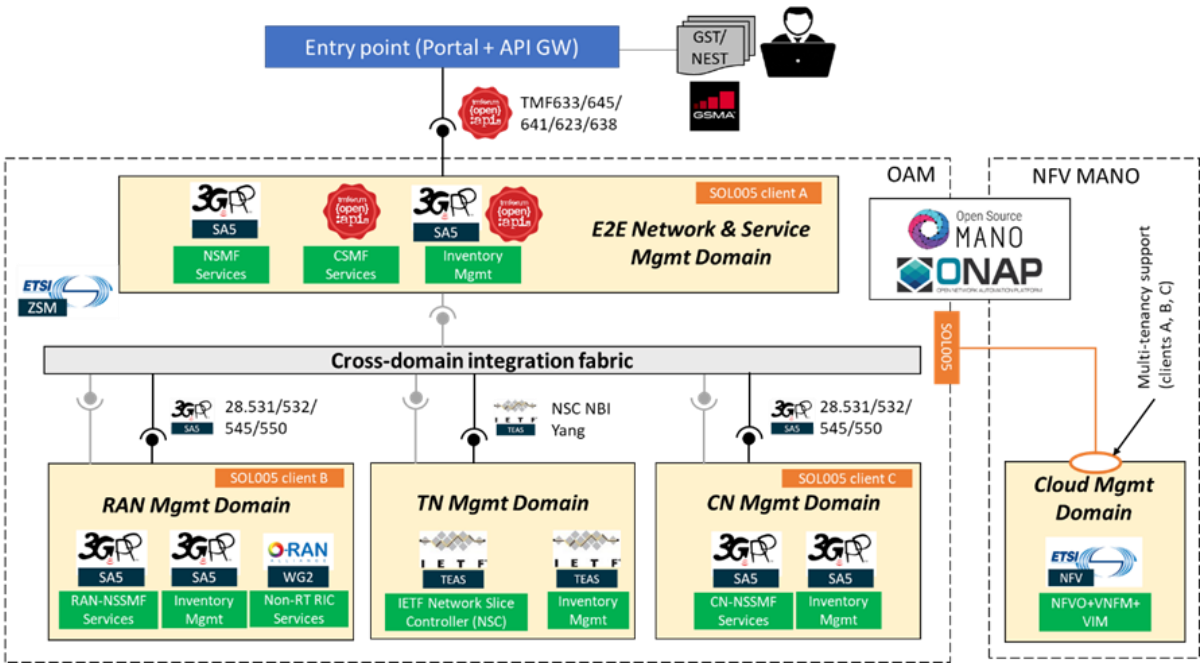

Figure 3: High level diagram of O&M domain

5.4.1 3GPP SA5

3GPP SA5 is responsible for all specification work pertinent to Telecom Management of the 3GPP network. This includes aspects such as operation, orchestration, assurance, fulfilment, automation and charging. With the advent of 5G technology from Release 15 on, 3GPP has defined the concept, architecture framework and management services in 3GPP TS 28.533, the generic management services 3GPP TS 28.532 [19], the provisioning service in 3GPP TS 28.530 [4], 3GPP TS 28.531 [20], as well as the 5G network resource model (NRM) in 3GPP TS 28.541 [6], fault supervision in 3GPP TS 28.545, performance assurance measurements in 3GPP TS 28.552 and KPI definitions in 3GPP TS 28.554.

3GPP SA5 defines the management and orchestration of network slicing, where the concept is defined in 3GPP SA2 (TS 23.501). The network slicing concept is about transforming a single mobile network to a mobile network that consist of “many” logical networks, where each logical network (an E2E network slice) has specific configuration to serve various service requirements. As an example, one (or more than one) communication service instances can be allocated to a specific logical network (network slice) with specific characteristics. The network slice instance (E2E network slice) are in 3GPP SA5 modelled to be the aggregation of the domain specific network slice subnets (e.g. RAN, Core and Transport). This allows the lifecycle of a network slice subnet instance to be managed

independently from the lifecycle of a network slice instance. Therefore, the management aspects of a network slice instance can be described by the following four phases:

- Preparation: This 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 a network slice instance.
- Commissioning: provisioning in this phase includes creation of the network slice instance. During network slice instance creation all needed resources are allocated and configured to satisfy the network slice requirements. The creation of a network slice instance can include creation and/or modification of the network slice instance constituents.
- Operation: This phase includes the activation, supervision, performance reporting, resource capacity planning, modification, and de-activation of a network slice instance. Provisioning in the operation phase involves activation, modification and de-activation of a network slice instance.
- Decommissioning: network slice instance provisioning in this phase includes decommissioning of non-shared constituents if required and removing the network slice instance specific configuration from the shared constituents. After this phase, the network slice instance is terminated.

Similarly, provisioning for a network slice subnet instance includes operations of network slice subnet instance creation, activation, de-activation, modification and termination.

The 3GPP SA5 management architecture adopts a service-oriented management architecture which is described as interaction between management service consumer and management service provider. For example, a management service consumer can request operations from management service providers on fault supervision service, performance management service, provisioning service and notification service, etc.

The NRM is an information model representing the management aspects of 3GPP 5G networks [6]. As captured in Figure 4:

- 5G NRM supports modelling of the following network resources: NG-RAN, 5GC and network slicing. This modelling is based on the definition of Information Object Classes (IOCs), based on which different IOC instances (a.k.a. Management Object Instances (MOIs), can be created. For example, different NetworkSlice instances can be defined from a NetworkSlice IOC.
- 5G NRM provides Stage1, Stage 2 and Stage 3 definitions for network resources. The Stage 1 (requirements-level) provides conceptual and use case definitions for individual network resources, and derives their requirements. The Stage 2 (information service-level) provides the technology-independent specification of a network resource. Finally, the Stage 3 (solution set-level) provides the mapping of stage 2 definitions into one or more technology-specific solution sets.

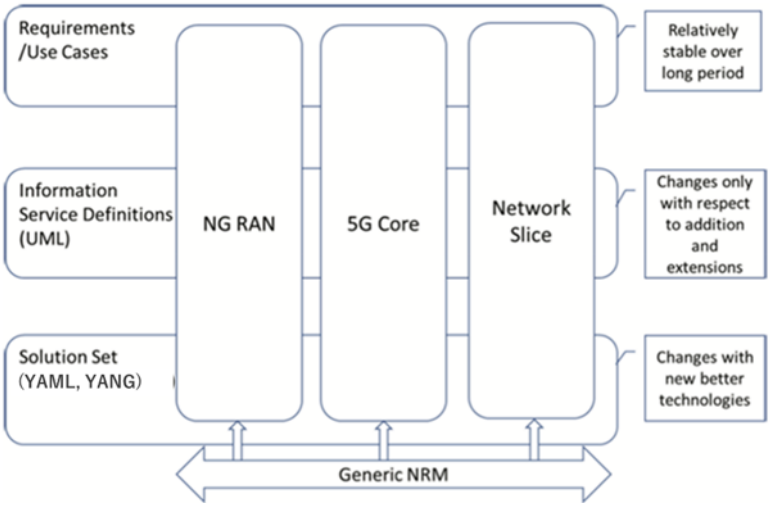

Figure 4: 3GPP NRM for 5G network

A network slice defined by 3GPP SA5 spans across RAN, Transport and Core domain specific network slice subnets (E2E network slice).

The 3GPP SA5 resource model includes the modelling of the TN end-points. It does not, however, include the modelling for the 5G transport network itself, nor the modelling of other access networks other than RAN.

In 3GPP SA5 the RAN resource partitioning is modelled by RRMPolicyRatio in NR NRM [6]. Furthermore, there are indication of whether a service defined by a ServiceProfile can share resources or not in a network slice instance. This indicates cooperation with other bodies, e.g. ETSI, as mentioned above, is needed. However, the problem is that many other bodies define management function and interfaces regarding what and how they could allocate resources. Yet, an E2E view is lacking since transport and NFVI is not part of 3GPP. It is expected to specify management framework for SLA compliance and that is ongoing in SA5 with regards to RAN and Core. In addition, if resources are handled by customers directly, further discussion will be needed.

Specifications to be considered are listed in Annex A.1.

5.4.2 IETF

The IETF is developing a framework to fulfill the requirements of transport part of the network slice. This framework is updated and maintained by the Network Slice Design Team (NSDT), a task force formed in the Traffic Engineering and Architecture Signaling (TEAS) Working Group. The mission of the NSDT is to utilize existing and under-development IETF technologies (including IETF traffic engineered technologies, IETF traffic engineering architectures and IETF network & service delivery models) to deploy and operate IETF network slices, each representing a transport network slice subnet.

For the provisioning of individual IETF network slices, a new management entity is defined: the IETF Network Slice Controller. Conceptually equivalent to the 3GPP referred TN Domain Manager, the Network Slice Controller is the entity responsible for IETF network slice lifecycle

management activities. For the interaction with upper/lower management systems, the Network Slice Controller defines two interfaces:

- Network Slice Controller Northbound Interface (NBI)

  It allows the Network Slice Controller to exposing built-in capabilities to higher level operation systems, e.g. 3GPP SA5 management system. It is a unified, technology-agnostic interface. Over this NBI, slice characteristics and other requirements can be communicated to the Network Slice Controller, and the operation state of an IETF network slice may be requested.

- Network Slice Controller Southbound Interface (SBI)

  It allows the Network Slice Controller to interact with underlying domain-specific network controllers, e.g. IP/MPLS controller, optical/DWDM controller, microwave controller. Unlike NBI, the SBI represent technology-specific interfaces that leverage many of the existing network models (e.g. ONF Transport-API, ONF TR-532, etc.)

Figure 5 captures the Network Slice Controller interfaces and their interaction with other management elements/systems. Figure 6 illustrates an example of the interaction between 3GPP management system and the Network Slice Controller.

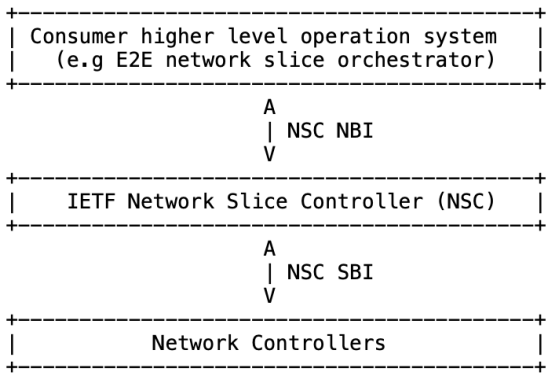

Figure 5: IETF Network Slice Controller interfaces

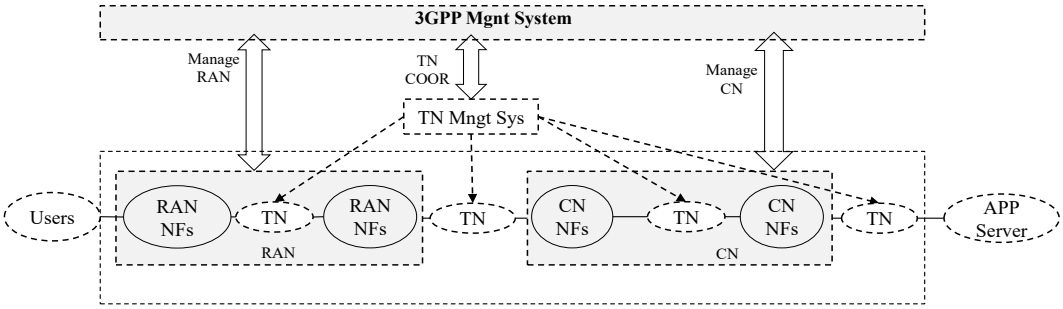

Figure 6: An example of deployment scenario for management [4]

However, a solution on how to map S-NSSAI such that the manager of each domain identifies the corresponding transport network of the network slice has not been seen and may not been created, since the IETF network slice is not specifically focusing on the 3GPP network slice, but should be technology agnostic in the context of IETF technology. In particular, Transport Network Management System (TN Domain Manager) in Figure 6, which is expected to have a capability to manage the transport domain of the network slice, in the context of SLA/SLS assurance, and an interface with 3GPP Management System, it does not seem to be defined.

The NSDT focuses on discussion related network slice technology and employs existing IETF technology for the necessary interfaces, while practically IETF TEAS WG is working to develop several I-Ds for the framework and northbound interface. The on-going work streams and I-Ds are shown in the Annex A.2.

Building upon the agreed IETF network slice and NSC concepts, the NSDT contributors are exploring solutions for transport network slicing in multiple work streams: (i) IETF network slice definition and terminology; (ii) architectural framework for IETF network slices; (iii) on modelling the Network Slice Controller NBI, e.g. YANG data model for the management and control of transport slice, for performance monitoring telemetry and so on. It is noted that RFC 8453 [21], which provides a framework for Abstraction and Control of TE Networks (ACTN) to support virtual network services and connectivity services, can be used as a base framework in case of TE-enabled underlying technology.

There are many contributions being proposed in IETF that suggest the reuse of existing mechanisms, among which are

- ACTN – Abstraction and Control of TE networks (draft-king-teas-applicability-actn-slicing)
- the definition of Virtual Networks (draft-ietf-teas-actn-vn-yang) and
- the possibility to bind L2/L3 services to them (draft-ietf-teas-te-service-mapping-yang) and
- Enhanced VPNs (draft-ietf-teas-enhanced-vpn).

5.4.3 ETSI ISG ZSM

ETSI ISG ZSM is a standardization group to provide definition of the E2E architecture and solutions to enable agile, efficient and qualitative management and automation of emerging and future networks and services. The goal is to have all operational processes and tasks (e.g., delivery, deployment, configuration, assurance, and optimization) executed automatically, ideally with 100% automation. ZSM provides an open framework, which is designed for closed-loop automation and data-driven operations, and optimized for AI/ML algorithms. The ZSM framework is versatile and built on service-bases principles offering scalability, modularity, extensibility and flexibility.

ETSI ISG ZSM is working to specify solutions and management interfaces for the orchestration and automation of network slicing across multiple management domains. The framework enables zero-touch management and orchestration of network slicing based on service level specification of a network slice or service from E2E perspective. エラー! 参照元が見つかりません。7 depicts a ZSM deployment example for network slicing management and automation.

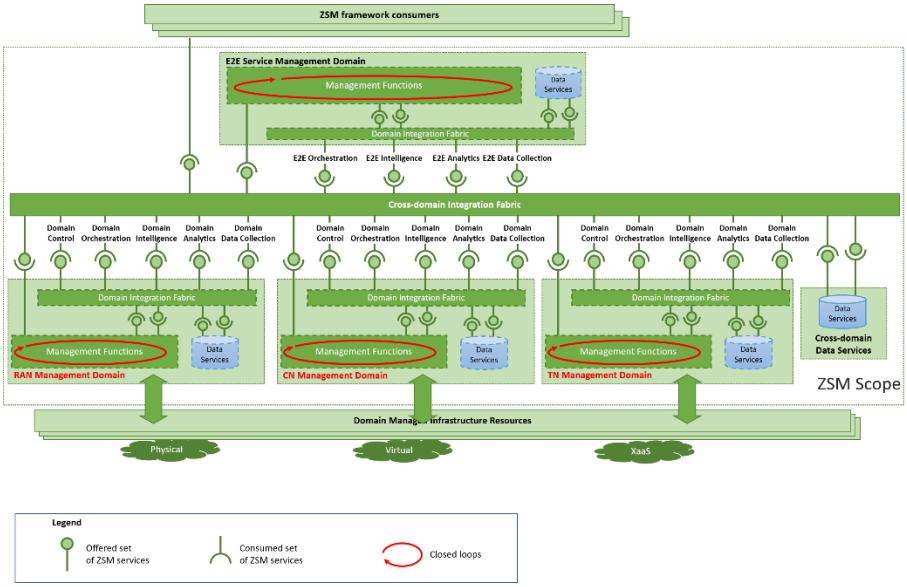

Figure 7: ZSM deployment example for network slicing management and automation

The network slicing lifecycle is managed using different processes across domains, which can be divided into fulfilment processes and assurance processes.

ETSI ISG ZSM leverages interfaces provided by TM Forum, 3GPP SA5, IETF, ETSI ISG NFV, and supports provisioning, performance and fault management of a network slice (in network slice as a service model) or E2E communication service based on network slice (in network slice as NOP internal mode).

In addition, ETSI ISG ZSM works on generic enablers (GS ZSM 009-1) and solutions (GS ZSM 009-2) for closed-loop as well as on advanced topics for next generation closed-loop operations (GR ZSM 009-3). It also studies security aspects (GR ZSM 010) related to the ZSM framework and solutions to identify potential security threats and mitigation options that should be considered by the ZSM specifications to ensure that the automated processes are secured and deliver the intended business outcomes.

ETSI ISG ZSM has started working on generic aspects related to intent-driven autonomous networks as well as on enablers for Artificial Intelligence-based Network and Service Automation.

3GPP specifies requirements and solutions for network slice, RAN and Core network slice subnets management and orchestration. Specification of requirements for transport network slice subnet for backhaul, front haul, and transport between 5GC and data network, is out the scope of 3GPP.

In addition, communication service management and its relationship to network slice management are not specified in 3GPP, or any other SDOs.

ETSI ISG ZSM may fill gaps to define requirements for transport network slice subnets and coordinate with 3GPP and IETF for the feasibility of the requirements. Further ETSI ISG ZSM may fill gaps for translating communication service requirements to network slice or

network slice subnet requirements. And ONAP has used ETSI ISG ZSM's architecture in the implementation of Transport slicing.

Specifications to be considered are "Reference Architecture" (GS ZSM002) and "End to end management and orchestration of network slicing" (GS ZSM003).

5.4.4 ETSI ISG NFV

From a resource management viewpoint, the external touchpoint of the NFV model with the 3GPP NRM is captured in the ETSI GR NFV-IFA 024 [22]. The 3GPP network slice (subnet) has a touchpoint with the NFV Network Service (NS). The composition mechanisms used in the design of individual NFV Network Services lead to a wide case-based variety in their complexity, ranging from single-VNF network services to composite network services (i.e. a network service including nested network services).

The initial 2017 study done by ETSI ISG NFV on gaps and concerns relating to network slicing was captured in ETSI GR EVE-012 [23]. After the study, the published NFV Release 3 normative specifications included support for the network slicing feature, the list of impacted specifications and the summary of the impacts is available on the ETSI wiki: FEAT05 (Network Slicing).

The updates in the specifications are:

- Updated NS descriptor (NSD) in normative ETSI GS NFV-IFA 014 [24], to include NS priority support;
- Updated Os-Ma-nfvo reference point specification in normative ETSI GS NFV-IFA 013 [25], to include the LCM coordination and enhance LCM notifications with information related to resource shortage errors, including resource pre-emption by a higher priority NS;
- Updated functional requirements in normative ETSI GS NFV-IFA010 [26] to include the functional requirements on NFVO to handle the enhancements for slicing;
- Updated external NFV information model, to include the 3GPP slicing touchpoint in ETSI GR NFV-IFA 024 [22].

5.4.5 ONAP

ONAP (Open Network Automation Platform) provides a comprehensive platform for real-time, policy-driven orchestration and automation of physical, virtual and cloud-native network function. The project has been hosted by Linux Foundation since February 2017. The initial version of ONAP was a merge of the Open ECOMP developed by AT&T and the Open-Orchestrator project in which the major contributors were: China Mobile, Huawei and ZTE.

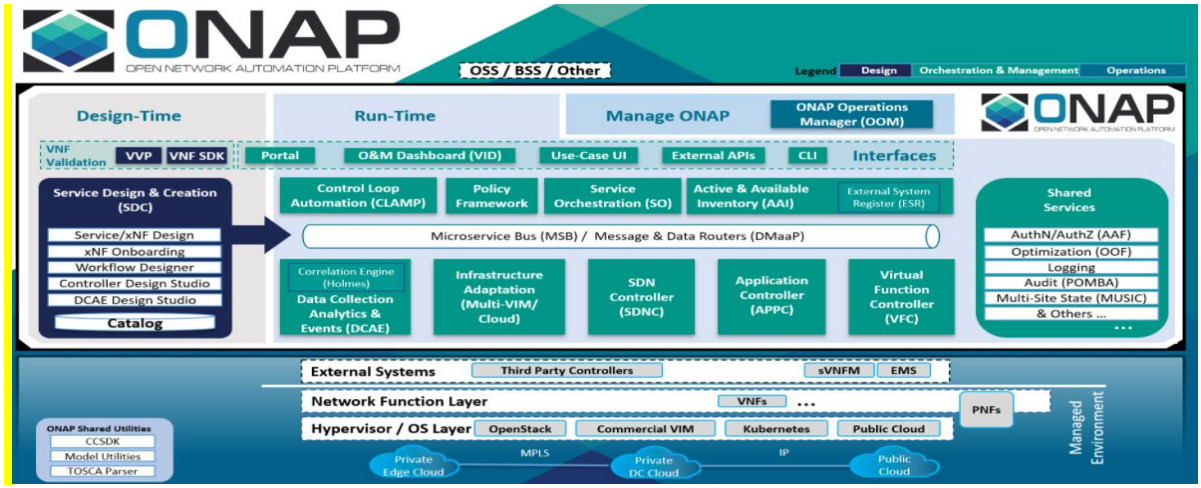

Figure 8: ONAP Architecture (https://www.onap.org/architecture)

ONAP platform aims to support full lifecycle management of VNF, CNF and PNF along with design, onboarding and modeling, orchestration and configuration management. Currently, it hosts a wide range of use cases for Communication Service Providers (CSP) such as, Virtual CPE (vCPE), Optimization and SON, 3GPP & O-RAN alignment, E2E Network Slicing, etc. Network Slicing belongs to the most important technologies in 5G and the following sections will provide more in-sights into the current ONAP status.

The Network Slicing in ONAP is designed to support end-to-end 5G Network Slicing across RAN, CN and TN in alignment with relevant standards (i.e. 3GPP, IETF, ETSI, TM Forum). It envisions to implement multi-vendor, interoperable ONAP platform including Network Slice Management Functions such as CSMF, NSMF and NSSMF. This would enable easier interoperability of slice management functions to be realized within ONAP with 3rd party slice management functions, as well as northbound and southbound systems. It also intends to support modelling and designing of CST, NST, NSST along with orchestration, lifecycle management, assurance and policy driven closed-loop.

ONAP community has started Network Slicing development in the Frankfurt release with basic functionality of CSMF and NSMF. The release of Guilin [27] provided a Network Slicing use case with RAN, Core and Transport in alignment with 3GPP NRM (Network Resource Model) TS 28.540 / TS 28.541 as well as ETSI ISG NFV, IETF and TM Forum APIs. Currently ONAP supports E2E Network Slice design, E2E Network Slice creation and activation, deactivation, and termination.

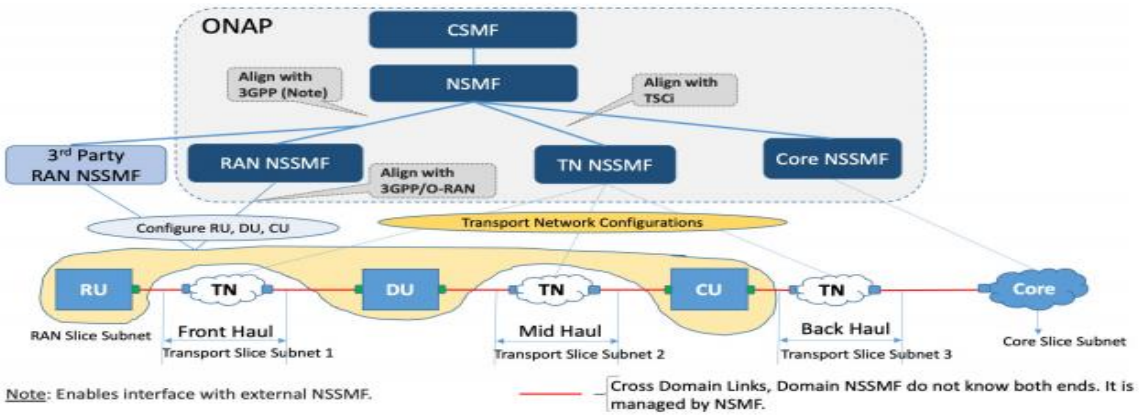

Figure 9: Network slicing implementation in light of Guilin ONAP

It supports Network Slicing Design template like Communication Service Template, Network Slice Template, and Network Slice Subnet template for RAN, Core and TN domains. Current support for Slice template design is minimal and needs to create manually. ONAP supports defining of RAN with parameters like set of cells, bandwidth, latency, priority, etc., Core with a specified “capacity” and other parameters and Transport with QoS, bandwidth, resiliency and other parameters.

ONAP supports creation of Communication Service Instance, Network Slice Instance, and Network Slice Subnet Instances. CSMF is realized as part of Workflow, which allows operator to create communication service request based on CST with minimal parameters for now. And NSMF supports creation of service profile & NSI in addition to slice template and slice instance selection as per slice requirements and it requests NSSI creation to appropriate NSSMF. Since the Guilin release, ONAP also supports NSSI creation using RAN, Core, and Transport NSSMF which has Proof-of-Concept level implementation for now. In E2E demo scenario the RAN NSSI is assumed to be already created. Transport NSSMF is currently able to create new NSSI Instances only, reusing and modification of existing TN NSSIs is not supported. Core NSSMF has been implemented using CNF approach, with HELM charts containing dummy Network Functions with only one configuration parameter (S-NSSAI) and that will be improved in future releases.

Crucial functionality of Monitoring, Automated Closed Loops with support of Intelligent Slicing is in future consideration. In the Guilin release there is isolated Closed Loop provided as a base work for further enhancements.

In the Honolulu release, enhancements in Service and Slice Profile modeling, RAN slice sub-net modeling enhancements which can support front-haul and mid-haul are in plan. E2E Slice creation related enhancements like E2E Slice allocation, NSMF level enhancements like NST selection optimization, improvements in interactions between NSMF & NSSMFs, NSI/NSSI selection based on account capacity, resource occupancy levels are under progress. NSSMF level improvements like RAN Slice Profile decomposition, instantiation of RAN NFs and initial configuration are also in stretch goal. Enhancement in configuration of Core Slice Subnet along with reuse of existing TN NSSI are in future consideration. Slice Monitoring & reporting supports with closed loop and AI/ML at E2E slice and slice subnet level for RAN & Core are in plan.

The alignment between ONAP and O-RAN started in the Frankfurt release with initial implementation of Non-RT-RIC.

In the ONAP Guilin release, the A1 Controller was replaced with a reference implementation of A1 Policy Management Service and A1 Policy Adapter up streamed by O-RAN software community. The Non-RT-RIC supports released versions of A1-Policy protocol, multiple near-RT-RICs, TLS based REST for SBI and NBI, unified REST & DMaaP NBI for A1 Policy Management.

The Non-RT-RIC is evolving as per the released specification by O-RAN ALLIANCE and in coming releases it plans to incorporate many other features like alignment with a new version of A1 Policy (e.g. A1 feedbacks), REST-based configuration of Non-RT RIC, Security management etc. Also, for other aspects of A1 Application Protocol – A1 Enrichment Information and A1 AI/ML – a discussion is ongoing in ONAP community and is open for later implementation whenever specification is available. An important concept of Non-RT RIC is a rAPP – a modular application designed to run on Non-RT RIC and to provide value added services. Currently there is no specification of rAPPs and ONAP community is open to add support for Non-RT RIC applications in future. O1 interface was introduced to support observability and configuration of managed elements. In the ONAP Guilin release, support for O1 compatible VES ingestion was started and is planned for the Honolulu release.

ONAP community is working to add O-RAN slicing requirements and support for mapping of Slice Profile to each Near-RT RIC-level configuration, support for A1 interface based closed loop control and O1 interface based configure/re-configure of O-RAN resources is planned in the Honolulu release.

5.4.6 ETSI OSM

Open Source MANO (OSM) is an ETSI-hosted project to develop a community-driven, production-quality Management and Orchestration (MANO) stack for the Telco Cloud. The scope of OSM project covers both design-time and run-time aspects related to Network-as-a- Service (NaaS) offering in telco service provider environments, including NFV network services and network slices.

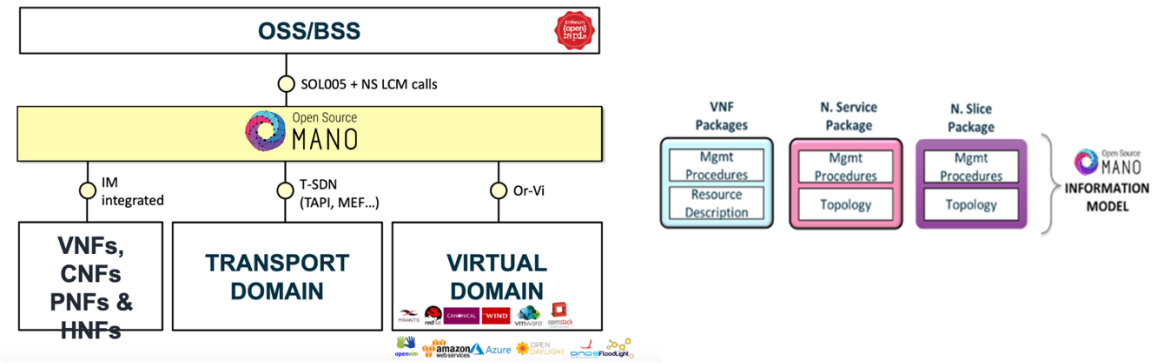

Figure 10: OSM Reference Architecture

The OSM Reference Architecture is presented in Figure 10. As it can be seen, the scope of OSM includes the ability to orchestrate E2E network services and slices across multiple sites and different domains – virtual, transport and physical domains. For the interaction with external systems, OSM is imbued with:

- A unified NorthBound Interface (NBI). OSM uses the NBI to expose management capabilities to other network and service management systems (e.g. 3GPP management system), collectively represented as OSS/BSS in Figure 10. The OSM's NBI provides a superset of ETSI ISG NFV SOL005 APIs together with the ability to handle network slices from a resource management viewpoint. This viewpoint allows modelling a network slice as a composition of individual network slice subnets, each deployed as an exclusive or shared NFV network service.
- A number of SouthBound Interfaces (SBI), allowing for the communication with underlying assets: commercial NFV infrastructures (NFVI+VIM), Software Defined Transport Networks (SDTN) and NFs. These SBIs include plugins towards virtual and transport domains (i.e. VIM, WIM and SDN controller plugins) as well as configuration interfaces (e.g. NETCONF/YANG, Ansible, SSH+script, Except, etc.) towards individual NFs. These NFs can be of different types - VNFs, Containerized Network Functions (CNFs), PNFs, and Hybrid Network Functions (HNFs) -, and are modelled in an infrastructure-agnostic manner using a OSM's Information Model, which is openly available and in alignment with ETSI ISG NFV's SOL006.

Focusing on network slicing management, it is worth mentioning that OSM allows different modes to control and manage the lifecycle of network slice instances. In the **full E2E management mode**, OSM takes the full control over individual instances, managing them through their entire lifecycle, including day-1 and day-2 operations. In the **stand-alone management mode**, a 3rd party standalone slice manager takes the role of managing slices via the OSM exposed SOL005 APIs, with OSM acting as NFVO. These two different management modes reflect aspects of the OSM capability exposure.

The OSM community keeps six-month cycles of releases, with OSM Release NINE being the current OSM version. Figure 11 shows the OSM Release NINE architecture. [28] provides a detailed description of the gold nuggets in this new Release. Among them, two features deserve to be mentioned, due to their impact on OSM network slicing management capabilities: the definition of the Placement optimization module (PLA), and the quotas management functionality.

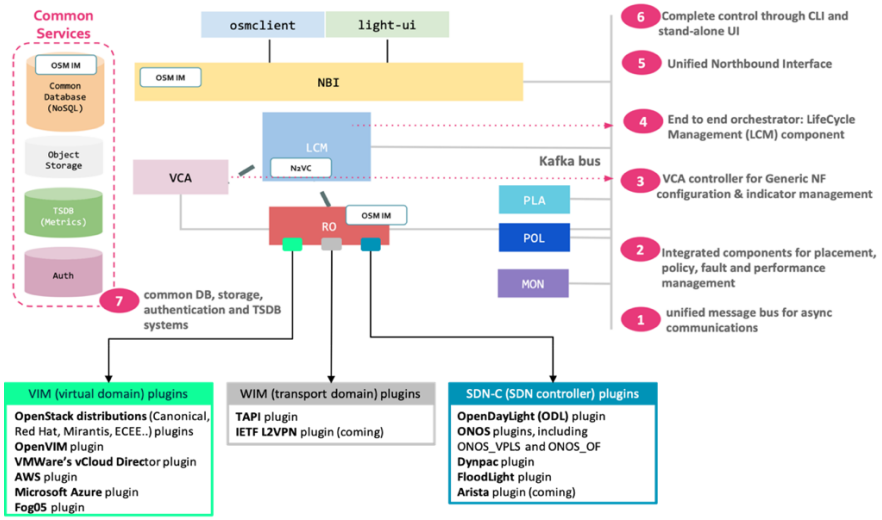

Figure 11: OSM Release NINE architecture

On the one hand, the definition of the PLA enriches OSM network slicing management capabilities at instantiation time. This module helps the OSM user to find an optimal deployment of network slice, distributing the individual VNFs/CNFs/HNFs over the set of available VIMs. This distribution should be done according to optimization criteria, based on user-provided models of (i) compute and networking cost, and/or (ii) latency and jitter metrics of inter-VIM connectivity. Optimal placement of VNFs/CNFs/HNFs over the VIMs is done by matching network slice specific requirements to infrastructure availability and metrics, while considering cost of compute and networking. Figure 12 shows an example on how the PLA module can be apply for a particular network slice setup.

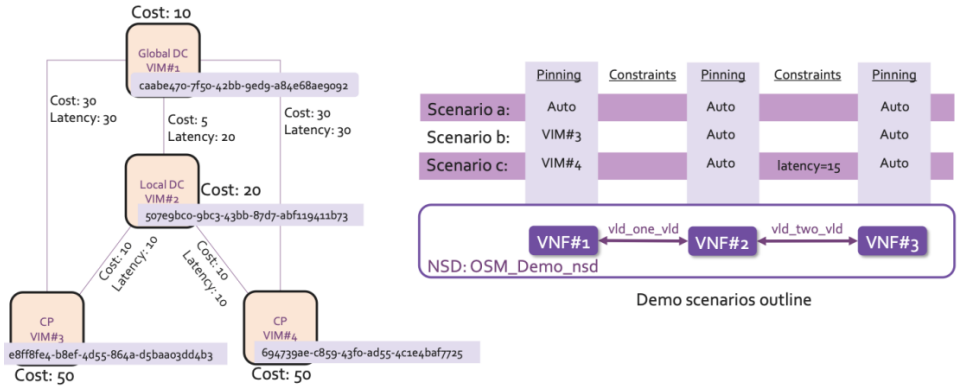

Figure 12: Scenario for PLA-based allocation. This scenario is part of the demo presented in the OSM Hackfest 8 (OSM-MR#8).

On the other hand, the quotas management functionality allows setting limits for the infrastructure (number of VIM, WIM, SDN controllers, Kubernetes clusters), packages (VNFDs, NSDs) and deployed instances (network service and slice instances). Once the limit is reached, any attempt to create a new item will be rejected to prevent overloading the system. This feature enhances security and provides a more granular control of OSM usage, which is key for multi-tenancy support. OSM can support multi-tenancy environment by

provide separate projects for different OSM clients. The project information contains the quotas assigned to the corresponding tenant. These quotas can be changed on-demand, by simply re-configuring the project settings at run-time.

At the date of the publication of this document, the OSM community has identified some areas where further coordination efforts might be really beneficial to speed up deployments of network slicing use cases:

- On the one hand, there is an opportunity to improve the specifications for scenarios of *Shared Network Slice Subnets*. The mechanisms to avoid potential side effects of reconfigurations of a given *Shared Network Slice Subnet* motivated by different Network Slices are ambiguous and have room for improvement in many cases. Thus, the requirements for its constituent NFs and/or the configuration mechanisms might help to ease the practical adoption of subnets of this kind.
- On the other hand, for a slice deployed across different domains (e.g. RAN, transport and telco cloud) with stringent latency requirements, the criteria to split the *latency budget* among the domains is mostly heuristic and driven by experience. However, while this static approach is perfectly valid at early stages or with a limited number of slices, more dynamic mechanisms to determine such a split among domains in an automatic fashion would ease massive and on-demand provisioning of dynamic long-distance slices.

The ongoing work in the OSM are listed in Annex A.3.

5.4.7 TM Forum

Network slicing management needs to standardize both the SBI for provisioning and managing the slices as well as NBI to enable the BSS/OSS and CFS to integrate with the slicing manager to provision the slices in the hybrid network. This domain is complex as the CSP's have many systems including PNF, VNF and now CNF which all need to be integrated and manage for the successful management of slices

The TM Forum is the industry initiative which includes both the operator and vendor communities to define a collection of architecture, interface and API models to manage the BSS/OSS side provisioning and management of slices. It further abstracts the underlying infrastructure thereby explains both the business, operational processes and objectives which should be achieved by CSP's introduction of slicing and digital transformation.

The TM Forum "Catalyst" program brings the community together to introduce and validate the slicing features and bring awareness for market adoption by promoting industry best practices for example, automation of a key process, interworking between a number of proprietary systems or latest programmes like The Aviator, which is enabling multi-vertical innovation through 5G slicing

The most important role of TM Forum standards like Open Digital Architecture and Open API is to define both the interfaces like TMF631 to 635 and API specs which are abstracted at a reasonable degree. It means we can achieve vision of "Open" systems and reduce the

amount of time to solve integration issues as these standards can support interworking even when the underlying implementation details are different.

As slicing relies strongly on BSS/OSS systems and these systems are architected using TM Forum API's so TM Forum serves as an important industry initiative to commercialize the slicing and 5G services in the broader context.

Regarding gaps and concerns, the TM Forum has already developed the technical report, IG1194 "Focus on Services Not Slices v1.0.1" [29]. The purpose of this technical report is to share some concerns and suggestions about network slicing adoption as the technical and standards work starts to reach maturity.

Specifications to be considered are the followings;

- IG1130E TAM Impacts by SDN/NFV for Network Slice Orchestration R18.5.1
- TM Forum Open Digital Architecture
- TM Forum Open API's
- Slicing in the context of AI driven business assurance for 5G (https://www.tmforum.org/ai-driven-business-assurance-5g/)

5.5 Device

5.5.1 UE Route Selection Policy

The existing 3GPP specification enables device steering of traffic to PDU session on a network slice using UE Route Selection Policy (URSP). URSP is provided by PCF (H-PCF in case of roaming to V-PCF) via the AMF to the device. The device uses URSP to determine which PDU session be chosen for particular traffic based on a URSP rule, which consists of rule precedence, traffic descriptor and route selection descriptor; in 3GPP Release 16 also Route Selection Validation Criteria. The traffic descriptor has many possible entries such as DNN, IP address, OS ID plus APP ID and so on in order to determine whether a URSP rule matches.

The route selection descriptor may contain e.g. SSC Mode Selection, Network Slice Selection, DNN Selection, and PDU Session Type Selection, but not all of these entries must be present. For example, it is possible to just include the DNN Selection and to leave the selection of the network slice to the network (determined during PDU session establishment). If using DNN as traffic descriptor, then DNN selection must not be included in Route Selection Validation Criteria, however, if using OS ID and App ID, then both DNN Selection and Network Slice Selection may be present in route selection descriptor.

It is impossible and not wanted for applications in device to be aware of network slice. The traffic of many applications use the same PDU session to the Internet DNN, and hence the IP address which an application uses is not unique. There is also no standard for APP ID definition and procedure. However, the URSP rules are generated by the PCF under control of the network and hence there is need for alignment between network and devices (OEM, OS and modem) on traffic descriptor for URSP rules in order to steer particular traffic to a PDU session, and to decide which network slice to be used for a PDU session. This traffic on a PDU session can be from a single application or from a group of applications. Note that

there is an upper limit of eight network slices that be used by a device, but the number of applications on a device can be much higher.

Specifications related to URSP are 3GPP TS 23.501, TS 23.502, TS 23.503 and corresponding stage 3.

5.5.2 GSMA Terminal Steering Group

GSMA Terminal Steering Group (TSG) facilitates operator and vendor alignment to drive device related matters and manages/coordinates capability requirements activities. With the fast growing requirement on network slicing, work item TSG Network Slicing was set up in the TSG. In the work item, gaps between current 3GPP specification and device implementation and clear requirements for device relating network slicing are being discussed.

The TSG Network Slicing is a task forces and is discussing on APP ID as part of the traffic descriptor in URSP (see section 5.5.11). Its key issues are APP ID definition, implementation requirements of APP ID based URSP in device and authentication and authorization mechanism of network slicing.

6 On a phased-based rollout for E2E network slicing

Network slicing rollout may follow in a phased approach, as long as standardization work and commercial products get mature.

Static slicing could be available along *2021*, but will not help that much. The allocation of specific radio and core network resources for private networks is something that is ongoing.

Dynamic slicing may take a bit longer. The specificities of the different network segments (RAN, transport, core), with different Technology Readiness Level (TRL) and pace of technology evolution each, may take dynamic slicing follow an evolutionary approach:

- *CN slicing* will be the first step -- 5GC was designed to support network slicing from the very beginning, i.e., 3GPP Rel-15. As the 5GC comes cloud-native and with a microservice architecture, dynamic slicing will be easier and earlier available (*late 2021, early 2022*)
- *RAN slicing* may be the next natural step (*2022* onwards), once slicing-aware RRM policy management and associated models get consolidated. To that end, further discussion in convenient SDOs (3GPP and O-RAN) is needed.
- *TN slicing* will allow for programmable service-tailored connectivity throughout the end-to-end data path, across all network segments (fronthaul, midhaul, backhaul) and technology domains (IP/MPLS, optical, microwave). The existing heterogeneity (in terms of resources and topology) on the transport underlay makes TN slicing a challenge, and naturally the last part to be consolidated (*2023* onwards). This requires the completion of SDN-C standards and a wider adoption of SDN technology across the different domains – current SDN implementations only focus on certain domains.

N**etwork Slice as a Service (NSaaS)** is the ultimate goal, though still far from being a reality in commercial networks (*2024* onwards). NSaaS represents a service delivery model that allows the operators to provision customized network slices to individual customers, and eventually enable these customers to gain access to some network slice management capabilities. It is up to the operator to decide on which specific management capabilities are made available to each customer, typically exposed through customer-facing APIs (e.g. TM Forum APIs).

- A baseline NSaaS will require at least (i) the standardization of dynamic slicing mechanisms, so individual network slices can be deployed and operated in an E2E manner; and (ii) the definition of mechanisms allowing for network slice capability exposure to individual B2B customers, and the integration of corresponding customer-facing network APIs into the operator's service platform. GSMA is considering this in the next phase of Operator Platform Group (OPG) work.
- An advanced NSaaS will require full automation in the entire orchestration pipeline, with focus on the assurance phase. To that end, the operation based on closed loops automation is a must.

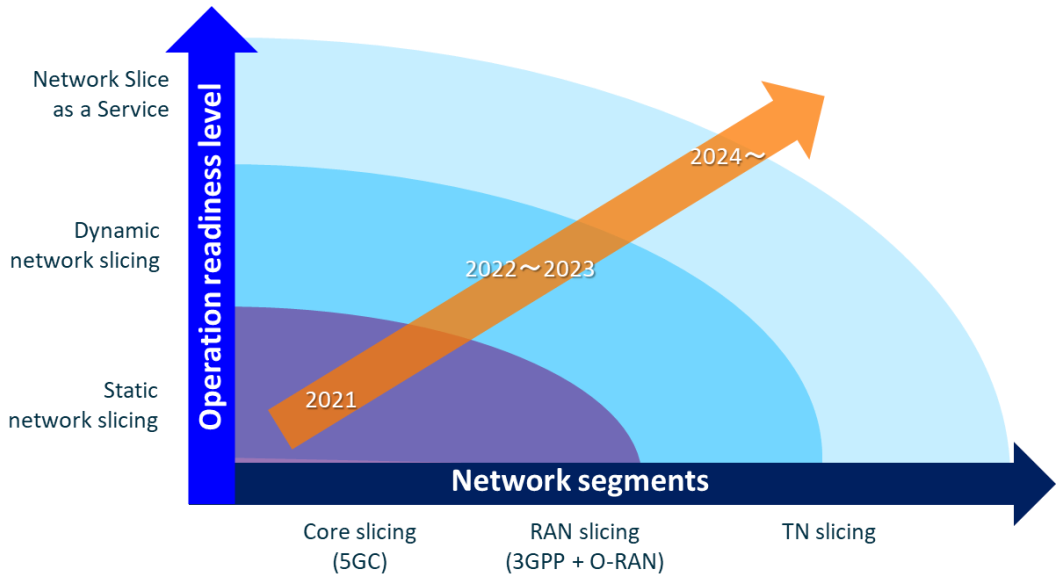

Figure 13: Phased-based rollout approach

Facing a go-to-market strategy, operators do not need to wait to have E2E slicing with fully-fledged network capabilities to include slicing related service offerings in their portfolio. Facing a go-to-market strategy, the following service offerings can be identified. [30]

Phase 1: Capacity-based network slices	Pre-configured by provider Slices built with particular application types in mind, e.g. eMBB traffic slices. Most related to capacity rather than advanced network capabilities (for example, low latency) Few, static slices (coarse-grained design) Key drivers related to CSP's total cost of ownership (TCO), efficiency and automation
Phase 2: Service-specific network slices	High degree of virtualization Increasingly provisioned on an on-demand basis. Larger number of slices with wider variety of characteristics than capacity-based slices, as scalability issues at OSS solutions gets solved. Configured independently from the physical slices, but still call on these resources Allocated by provider's engine
Phase 3: Tenant-driven network slices	VNFs and applications made available in a marketplace. (B2B/industry) Customers can flexibly define an end-to-end slice following a 'pick-and-choose' approach, selecting components from the marketplace and compose them as needed. Slices may be ephemeral, i.e. short-lived slices More-critical functions require links to specifical physical resources (such as ultra-high reliability); higher value

Table 3: Phase-based rollout strategy

7 Conclusions

It has begun to deploy 5G core network which has a new capability of network slicing which is specified in 3GPP specifications such as [2], GSMA documents such as [5] and so on. Network slicing can be utilized mainly in core network domain at the moment. On the other hand, the standardization work related to network slicing remains and is ongoing, in order to realize network slicing for providing E2E network services and to satisfy requirements from vertical industries and demands from operators.

For instance, there is an ongoing work, which are not addressed in previous sections.

- **Roaming**: Every operators deploying 5G System with network slicing capability will support 5G System roaming for network slices. For technical aspect, the guidelines are covered in GSMA NG.113 [31]. The commercial requirements, charging models and agreements can be found in GSMA documents, which are under development. Successful completion of all network, device and billing aspects is required to support network slice roaming.

This GSMA white paper indicates the several aiming technology aspects and the snapshot of ongoing work in SDOs and open source projects toward E2E Network Slicing. GSMA will keep working to validate network-slicing-related specifications from these SDOs from E2E

perspective, to coordinate with these SDOs and industry activities such as open source projects as next steps.

Finally, it is important to continue the interaction with SDOs and open source projects and to coordinate their specifications especially from inter-domain perspective, in order to realize E2E Network Slicing.

Annex A Reference List

A.1 O&M, 3GPP SA5

Specification	Title
3GPP, TS28.530 [4]	Management and orchestration; Concepts, use cases and requirements
3GPP, TS28.531 [20]	Management and orchestration; Provisioning; Stage 1
3GPP, TS28.532 [19]	Management and orchestration; Provisioning; Stage 2 and stage 3
3GPP, TS28.533 [32]	Management and orchestration; Management and orchestration architecture
3GPP, TS28.540 [33]	Management and orchestration, 5G Network Resource Model (NRM), Stage 1
3GPP, TS28.545 [34]	Management and orchestration; Fault Supervision (FS)
3GPP, TS28.550 [35]	Management and orchestration; Performance assurance
3GPP, TS28.552 [36]	Management and orchestration; 5G performance measurements
3GPP, TS28.541 [6]	Management and orchestration, 5G Network Resource Model (NRM), Stage 2 and Stage 3
3GPP, TS28.554 [37]	Management and orchestration; 5G end to end Key Performance Indicators (KPI)

Table 4: Specifications to be considered in 3GPP SA5

A.2 O&M, IETF

Work stream	Internet-Drafts	Description
IETF network slice Definition and Terminology	draft-ietf-teas-ietf-network-slice-definition	This document provides a definition of the term "IETF Network Slice" for use within the IETF and specifically as a reference for other IETF documents that describe or use aspects of network slices. It is noted that this document will be revised from 5G slice specific terminology to IETF technology agnostic terminology and will be eventually merged into the following framework document. See Note 1
Architectural framework for IETF network slices	draft-ietf-teas-ietf-network-slice-framework	The document discusses the general framework for setting up special-purpose transport connections using existing IETF technologies. These connections are called transport slices for the purposes of this memo. See Note 1
On modelling the NSC NBI	draft-contreras-teas-slice-nbi	This document analyses different use cases deriving the needs of potential customers of network slice realized with IETF techniques in order to identify the functionality required on the Network Slice Controller NBI to be exposed towards such IETF network slice customers. See Note 2
	draft-liu-teas-transport-network-slice-yang	This document describes a provider-centric YANG data model for the management and control of individual IETF network slices. This model can be used by the Network Slice

		Controller operator to provision transport slices towards customers using the NBI. See Note 2
	draft-wd-teas-ietf-network-slice-nbi-yang	This document describes a customer-centric YANG data model for the management and control of individual IETF network slices. This model can be used by the Network Slice Controller customers to request, configure and manage the components of individual IETF network slices. See Note 2
	draft-ietf-teas-actn-vn-yang	This document provides a YANG data model generally applicable to any mode of Virtual Network operation. This data model is prospected to be used as NBI API. See Note 1
	draft-ietf-teas-te-service-mapping-yang	This document provides a YANG data model to map customer service models (e.g., the L3VPN Service Model (L3SM)) to TE models (e.g., the TE Tunnel or the Virtual Network model). See Note 1
	draft-ietf-teas-actn-pm-telemetry-autonomics	This document provides YANG data models that describe performance monitoring telemetry and scaling intent mechanism for TE tunnels and Virtual Networks. See Note 1
Existing technologies applicable to network slicing	*RFC8453*	This document defines a framework for the abstraction and control of traffic engineered networks. It also defines the concept of virtual network, which provides connectivity between end points in a multi domain environment.
	draft-king-teas-applicability-actn-slicing	This document outlines the applicability of ACTN to network slicing in a Traffic Engineering (TE) network that utilizes IETF technology. It also identifies the features of network slicing not currently within the scope of ACTN, and indicates where ACTN might be extended. See Note 2
	draft-ietf-teas-actn-vn-yang	This document provides a YANG data model generally applicable to any mode of Virtual Network (VN) operation. See Note 1
	draft-ietf-teas-te-service-mapping-yang	This document provides a YANG data model to map customer service models (e.g., the L3VPN Service Model) to Traffic Engineering (TE) models (e.g., the TE Tunnel or the Virtual Network (VN) model). This model is referred to as TE Service Mapping Model and is applicable generically to the operator's need for seamless control and management of their VPN services with TE tunnel support. See Note 1
	draft-ietf-teas-enhanced-vpn	This document describes the framework for Enhanced Virtual Private Network (VPN+) service. VPN+ will be used to form the underpinning of network slicing, but could also be of use in its own right providing enhanced connectivity services between customer sites. See Note 1

Table 5: NSDT work streams and related Internet-Drafts

Note 1: These drafts are IETF work in progress. They are subject to change or removal at any time. See https://www.ietf.org/standards/ids/ for details.

Note 2: Opinions expressed in these drafts are the authors'. The drafts do not necessarily have any standing in the IETF. They are subject to change or removal at any time. See https://www.ietf.org/standards/ids/ for details.

A.3 ETSI OSM

Ongoing work	URL
OSM's User Guide - Use of Network Slices	https://osm.etsi.org/docs/user-guide/05-osm-usage.html#using-network-slices
OSM's IM	https://osm.etsi.org/docs/user-guide/11-osm-im.html
YANG definitions of OSM's IM	https://osm.etsi.org/gitweb/?p=osm/IM.git%3Ba=tree
OSM's Northbound Interface	https://osm.etsi.org/docs/user-guide/12-osm-nbi.html
Open OSM NBI in Swagger-UI	https://forge.etsi.org/swagger/ui/?url=https%3A%2F%2Fosm.etsi.org%2Fgitweb%2F%3Fp%3Dosm%2FSOL005.git%3Ba%3Dblob_plain%3Bf%3Dosm-openapi.yaml%3Bhb%3DHEAD
OSM Scope and Functionality	http://osm-download.etsi.org/ftp/Documentation/201902-osm-scope-white-paper/#!02-osm-scope-and-functionality.md
Network slicing in practice with OSM: A complete Network Slice with Magma and OSM	https://osm.etsi.org/gitlab/vnf-onboarding/osm-packages/tree/master/magma

Table 6: Ongoing work in ETSI OSM

Annex B Document Management

B.1 Document History

Version	Date	Brief Description of Change	Approval Authority	Editor / Company
1.0	April 2021	Initial version of the White Paper	Technology Group	Masaharu Hattori / KDDI Javier Sendin / GSMA
2.0	May 2022	Updates	Networks Group	Masaharu Hattori / KDDI Javier Sendin / GSMA

B.2 Other Information

Type	Description
Document Owner	GSMA, Network Group Vertical Taskforce
Editor / Company	Masaharu Hattori / KDDI Javier Sendin / GSMA
Reviewed by	GSMA, Network Group Vertical Taskforce