Deliverable 2.3

Control Plane System Definition, APIs and Interfaces

Editor:	George Agapiou, OTE
Deliverable nature:	Report (R)
Dissemination level: (Confidentiality)	Public (PU)
Contractual delivery date:	31/12/2017
Actual delivery date:	13/04/2018
Suggested readers:	Control Plane Providers, Management Plane Providers, Vertical Industries
Version:	1.0
Total number of pages:	77
Keywords:	Control Plane, Network Slice, 5G

Abstract

The main objective of this document is to provide the architecture of the control plane of the SliceNet project. It starts with the definition of a slice and relevant projects and standardization bodies/interest groups that deal with the control plane architecture of 5G networks. It then details the main components of the SliceNet Control Plane, the P&P and the QoE along with their APIs. Also the Intra domain and Inter domain slicing have been analysed and workflows have been detailed. Workflows of the main components of the SliceNet CP such as P&P and QoE have been detailed and workflows of P&P and QoE have been described with regards to the SliceNet use cases. Data programmability and control of the resources by the CP are also detailed.

Disclaimer

The research leading to these results has received funding from the European Union Horizon 2020 Programme under grant agreement number 761913.

Impressum

[Full project title] End-to-End Cognitive Network Slicing and Slice Management Framework in Virtualised Multi-Domain, Multi-Tenant 5G Networks

[Short project title] SLICENET

[Number and title of work-package] WP2- SLICENET System Definition

[Number and title of task] T2.3 - Control Plane System Definition, APIs and Interfaces

[Document title] Control Plane System Definition, APIs and Interfaces

[Editor: Name, company] George Agapiou, OTE S.A.

[Work-package leader: Name, company] Pedro Neves, Altice Labs

Copyright notice

List of authors

Company	Author
ALTICE LABS SA, Portugal	José Cabaça, Mário Rui Costa, Pedro Neves, Rui Calé
CREATIVE SYSTEMS ENGINEERING (C.S.E) MONOPROSOPI EPE, Greece	John Vavourakis, Konstantinos Koutsopoulos
Dell EMC INFORMATION SYSTEMS INTERNATIONAL	Zdravko Bozakov, Thuy Truong
EURECOM	Chia-Yu Chang, Navid Nikaein, Xenofon Vasilakos
EURESCOM-EUROPEAN INSTITUTE FOR RESEARCH AND STRATEGIC STUDIES IN TELECOMMUNICATIONS GMBH	Anastasius Gavras
IBM ISRAEL - SCIENCE AND TECHNOLOGY LTD	Dean H Lorenz, Katherine Barabash, Valleriya Perelman
NEXTWORKS	Giacomo Bernini, Pietro G. Giardina
ORANGE ROMANIA SA	Marius Iordache, Vlad Sorici
ORANGE SA	Elie El Hayek
HELLENIC TELECOMMUNICATIONS ORGANIZATION S.A. - OTE AE	Christina Lessi, George Agapiou
ERICSSON TELECOMUNICAZIONI	Carmine Galotto, Ciriaco Angelo, Giuseppe Celozzi, Raffaele De Santis
UNIVERSITAT POLITECNICA DE CATALUNYA	Albert Pagès, Fernando Agraz, Salvatore Spadaro
UNIVERSITY OF THE WEST OF SCOTLAND	Jose M. Alcaraz Calero, Hector Marco, Qi Wang, Zeeshan Pervez
REDZINC SERVICES LIMITED	Ricardo Figueiredo

Table of Contents

List of figures

List of tables

Abbreviations

Abbreviation	Full form
5G	Fifth Generation (mobile/cellular networks)
5G CN	5G Core Network
5G PPP	5G Infrastructure Public Private Partnership
5G-S-TMSI	5G-S-Temporary Mobile Subscription Identifier
AMF	Access and Mobility management Function
AN	Access Network
AUSF	Authentication Server Function
CN	Core Network
CP	Control Plane
DN	Data Network
DU	Distributed Unit
E2E	End-to-end
eNBs	Enhanced Node-Bs
FG	Forwarding Graph
gNBs	NR gigabit Node-Bs
KPI	Key Performance Indicator
NFV	Network Functions Virtualization
NR	New radio
NRF	NF Repository Function
NSSAI	Network Slice Selection Assistance Information
P&P	Plug & Play
PNF	Physical Network Function
QoE	Quality of Experience
QoS	Quality of Service
RAN	Radio Access Network
RRU	Remote Radio Unit
RTC	Real-Time Control
SBA	Service Based Architecture
SBI	Southbound Interface
SDN	Software Defined Networks
SliceNet	End-to-End Cognitive Network Slicing and Slice Management Framework in Virtualised Multi-Domain, Multi-Tenant 5G Networks
SMF	Session Management Function
S-NSSAI	Single NSSAI
SON	Self-Organizing Networks
UDM	Unified Data Management
UE	User Equipment
UP	User Plane
UPF	User Plane Functions
VAT	Value-added tax
VNF	Virtualised Network Function

1 Introduction

Future networks need to provide technical and service requirements in terms of performance, functional and operational requirements which are set by the operators and the end users and/or verticals. These requirements need to be conveyed throughout the network in an E2E architecture. One of the key requirements of the 5G networks will be to support a variety of vertical industries such as those proposed by SliceNet, thus smart grid, e-health and smart city. These verticals derive different use cases which impose very strict requirements than today services do. It is well understood that these requirements can be satisfied after significant improvements in the architecture is done.

This deliverable considers the introduction of components/enablers in the Control Plane (CP), such as Plug & Play (P&P), which is the main enabler for E2E slice runtime customization and QoE optimizer used for monitoring and optimizing the network metrics to keep the QoE in a fixed level according to the vertical SLA. The control plane is based on two level of abstractions for the provision of slice monitoring and performance improvement. The first level of abstraction is provided by the southbound interfaces of the CP that provide information for the underlying infrastructure, while the second level of abstraction is provided by the P&P and QoE optimizer enablers. The above enablers are using Software Defined Network (SDN) and Network Functions Virtualization (NFV) technologies for leveraging network functions and services in the slices.

The document is structured as follows:

- Section 2 presents a background study on 5G research projects, interest groups and standardization bodies that work on the control plane issues and architectures.
- Section 3 outlines the key concepts for the network slicing concept and its logical model as well as a high level description of the control plane architecture.
- Section 4 presents the two enablers of the SliceNet CP which are described in detail along with the interfaces, adapters and detail analysis of the P&P and QoE optimizer. Workflow diagrams that provide the configuration of a slice in terms of KPI collection, slice VNF configuration and slice VNF deployment are provided. Workflow diagrams, for the optimization of a slice, are provided by the QoE optimizer. Intra-domain multi-tenant and Inter-domain slicing is detailed and topologies are provided. The document continues with a high level description of the APIs for the SliceNet CP.
- Section 5 presents a high level description of the SliceNet CP interfaces and APIs.
- Section 6 identifies the high level workflows of the SliceNet CP and the mapping to the use cases.
- Section 7 describes the data plane programmability for QoS-aware slicing.
- Section 8 concludes the document.

2 Background

This section provides overview information relevant to SliceNet on work related to the Control Plane in standardization bodies and other projects.

The 5G technology implementations, architecture and concepts are accepted to be driven by the use-case and industry verticals, as it should provide a layer of control and data plane and that should be exposed to any application and functions, for programmability and resource usage.

The requirements for the control plane in the 5G architecture are referred to self-configuration and self-management for the functions, there is required a logically separation between the data plane and control plane, that are impacting the architecture, decoupling the modification of the data plane due to control plane modifications.

The control plane interacts, on the vertical axis, via a North Bound Interface (NBI) with the software applications and via a South Bound Interface (SBI) with the data plane. On the horizontal axis, the control plane, ensures that the different services are using the proper resources through the orchestration.

The control plane in SliceNet is not only related to a pool of resources underlying the user plane with a set of SDN controllers, but it is more related to an abstract implementation of a network with full capabilities to implement optimization, inter & intra-domain service and slice composition, P&P control with APIs support.

The service-driven 5G network architecture aims to flexibly and efficiently meet diversified mobile service requirements. With SDN and NFV supporting the underlying physical infrastructure, 5G comprehensively cloudifies access, transport, and core networks [1].

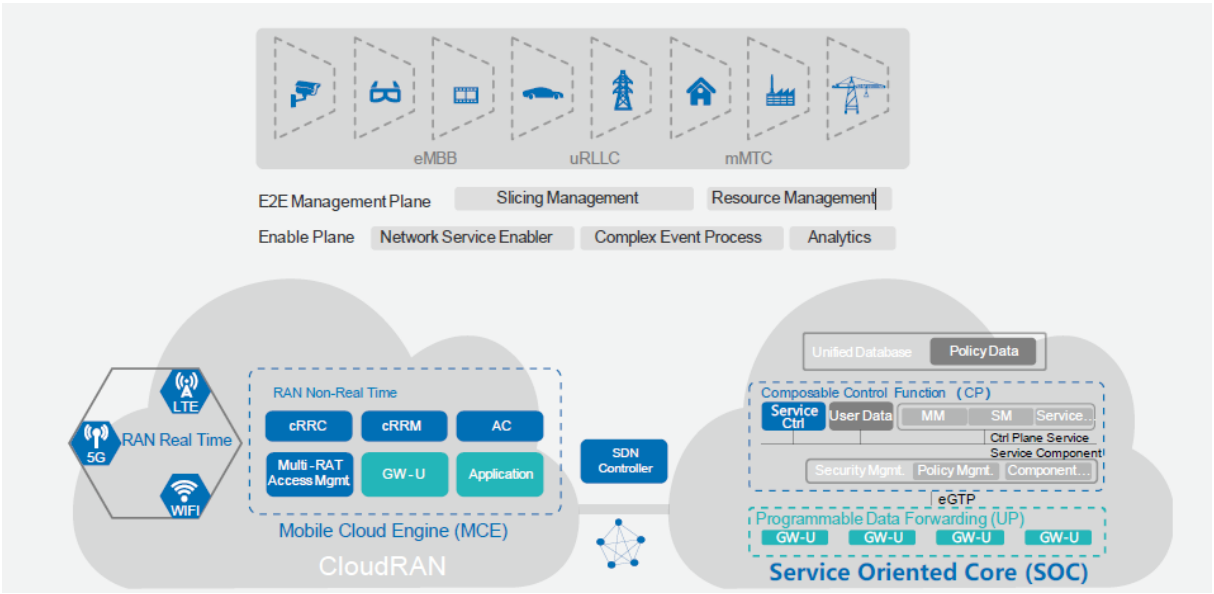

Figure 1: Cloud adoption diversified 5G services [1]

CloudRAN consists of sites and mobile cloud engines. This facility coordinates multiple services, operating on different standards, in various site types for RAN real time resources that require a number of computing resources. Multi-connectivity is introduced to allow on-demand network deployment for RAN non-real time resources. Networks implement policy control using dynamic policy, semi-static user, and static network data stored in the unified database on the core network function orchestration to ensure that networks can select corresponding CP or UP functions according to different service requirements [1].

2.1 Mobile Cloud Engine (MCE)

Mobile Cloud Engine is the logical entity of central **control and management** for CloudRAN, incorporating RAN non-real time functions, Wi-Fi AC, distributed gateway, service-related application distribution entity, and Cache [1].

2.2 Proposed framework

The perspectives of this proposal are described as separate planes. Although separately defined, the planes are not completely independent: key items in each are related to items in the other planes. However, the planes are sufficiently independent to simplify reasoning about the complete system requirements. The interworking between planes is manifested by groups of interfaces (i.e. reference points) that would be used for exchange of information and/or controls between separate (sub) systems sharing boundaries. The proposed separation in distinct planes is [2]:

- Service layer
- Management & Orchestration layer
- Control layer
- Data layer

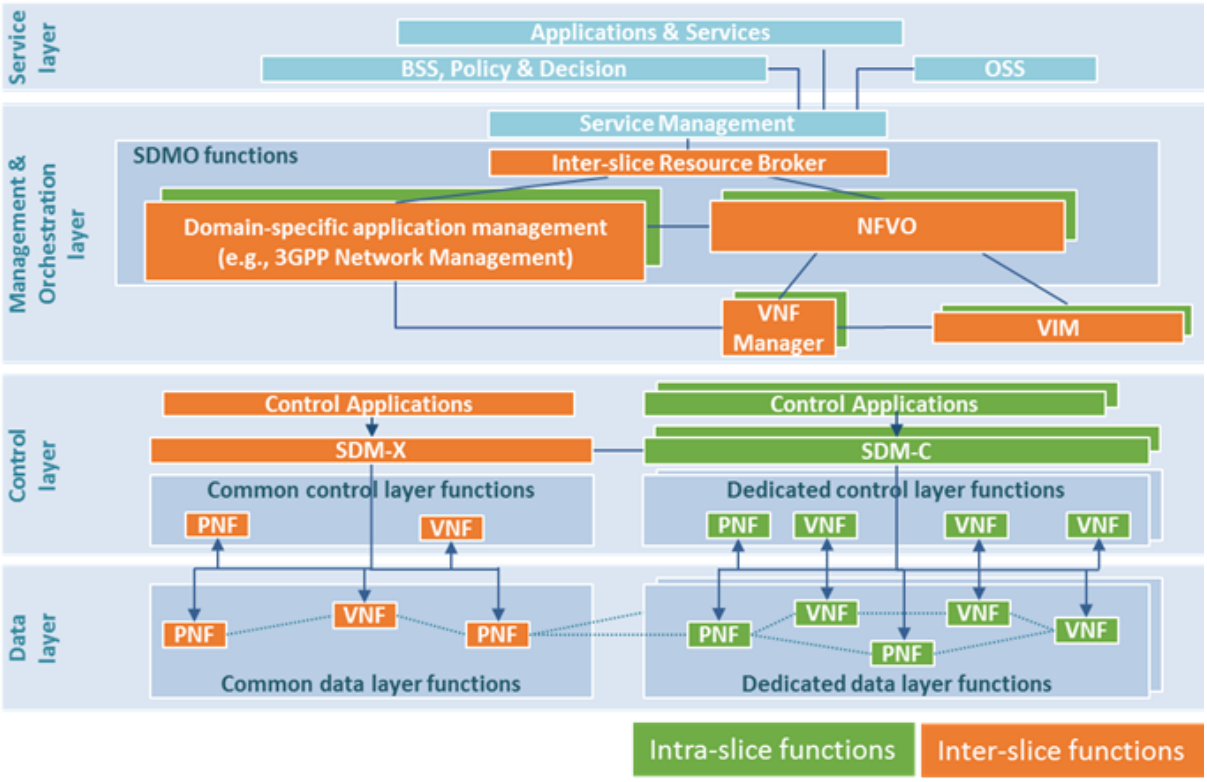

Figure 2: 5G architecture proposed by the 5G-PPP architecture group [2]

The 5G System architecture shall leverage service-based interactions between CP network functions where identified. Some key principles and concept are to [2]:

- Separate the UP functions from the CP functions, allowing independent scalability, evolution and flexible deployments, e.g. centralized location or distributed (remote) location.
- Modularize the function design, e.g. to enable flexible and efficient network slicing.
- Wherever applicable, define procedures (i.e. the set of interactions between network functions) as services, so that their re-use is possible.

- Enable each Network Function to interact with other NF directly if required. The architecture does not preclude the use of an intermediate function to help route Control Plane messages (e.g. like a DRA).
- Minimize dependencies between the Access Network (AN) and the Core Network (CN). The architecture is defined with a converged core network with a common AN - CN interface which integrates different 3GPP access and non-3GPP access.
- Support a unified authentication framework.
- Support "stateless" NFs, where the "compute" resource is decoupled from the "storage" resource.
- Support capability exposure.
- Support concurrent access to local and centralized services. To support low latency services and access to local data networks, UP functions can be deployed close to the Access Network.
- Support roaming with both Home routed traffic as well as Local breakout traffic in the visited PLMN.
- Involve different administrative domains and thus will be built on network slicing based on a per service basis, therefore building an E2E slice structure.

2.3 Network Function Service Consumer – Network Function Service Producer Interactions

The interaction between two NFs – Consumer and Producer – within this NF service framework follows two mechanisms [3]:

- **Request-response:** A CP NF Service Producer (NF_B) is requested by another CP NF Service Consumer (NF_A) to provide a certain NF service, which either performs an action or provides information or both. NF_B provides NF service based on the request by NF_A. In order to fulfil the request, NF_B may in turn consume NF services from other NFs. In request – response mechanism, communication is one to one between two NFs (consumer and producer) and a one – time response from producer to a request from consumer is expected within a certain timeframe [2].

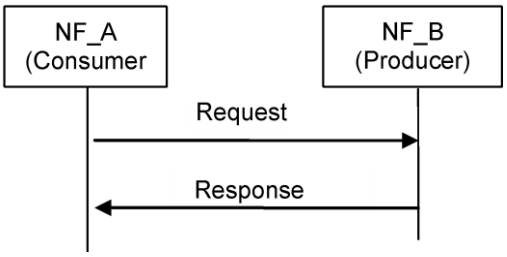

Figure 3: Request – response Network Function Service illustration [3]

- **Subscribe-Notify:** A CP NF Service Consumer (NF_A) subscribes to NF Service offered by another CP NF Service Producer (NF_B). Multiple CP NFs may subscribe to the same CP NF Service. The subscription request shall include the notification endpoint of the NF Service Consumer to which the event notification from the NF Service Producer should be sent to. In addition, the subscription request may include notification request for periodic updates or notification triggered through certain events (e.g. the information requested gets changed, reaches certain threshold, etc...). The subscription for notification can be done through one of the following ways:
    - A separate request or response exchange between the NF Service Consumer and the NF Service Producer, or

- The subscription for notification is included as part of another NF Service operation of the same NF Service, or
- Registration of a notification endpoint for each type of notification the NF Consumer is interested to receive, as a NF Service Parameter with the NF Repository Function (NRF) during the NF and NF Service Registration procedure.

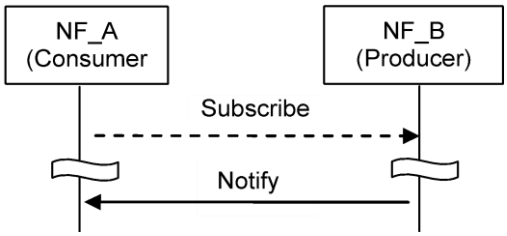

Figure 4: Subscribe – Notify Network Service Illustration [3]

CP NF_A may also subscribe to NF Service offered by CP NF_B on behalf of CP NF_C, i.e., it requests the NF Service Producer to send the event notification to another consumer/ consumers. In this case, NF_A includes the notification endpoint of the NF_C in the subscription request.

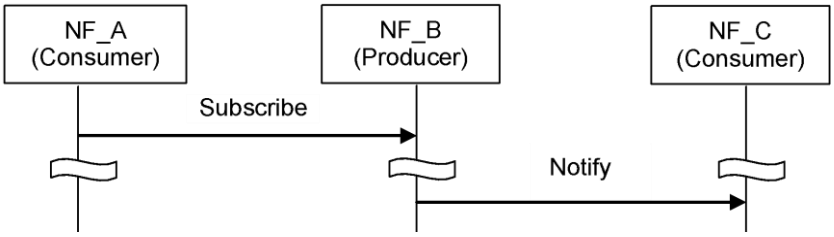

Figure 5: Subscribe – Notify Service discovery Illustration [3]

2.4 Control Plane Protocol Stacks between Access and Core Network in 5G

The CP interface between 5G – AN and the 5G Core supports:

- The connection of multiple different kinds of 5G – AN (e.g. 3GPP RAN, N3IWF for Untrusted access to 5GC) to the 5GC via a unique CP protocol: A single NGAP protocol is used for both the 3GPP access and non-3GPPaccess.
- There is a unique N2 termination point in the Access and Mobility Management Function (AMF) per access for a given UE regardless of the number of PDU Sessions of the UE.
- The decoupling between AMF and other functions such as SMF that may need to control the services supported by 5G-AN (e.g. control of the UP resources in the 5G-AN for a PDU session). For this purpose, NGAP may support information that the AMF is just responsible to replay between the 5G-AN and the SMF [3].

2.5 Relevant projects of 5G-PPP Phase I

Relevant projects that some aspects can be exploited by SliceNet are presented below in Table 1.

Table 1: Relevant 5G-PPP projects

Related project	Main topic	Potential exploitable items
5Gnorma	Control layer: • accommodates the two main controllers, SDM-X and SDM-C, control applications, and distributed control NFs • SDM-X and SDM-C translate decisions of the control applications into commands to VNFs	The project describes the mapping of services to the network functions. This can be adapted by SliceNet to improve slice isolation and performance

	and PNFs • SDM-X and SDM-C as well as other control applications can be executed as VNFs or PNFs themselves	
METIS II	A common control and user plane framework providing the means for an efficient support of the broad versatility of services expected for 5G as well as a future-proof and cost-efficient implementation of the 5G integration.	The separation of the CU and DU can be exploitable by SliceNet for the improvement of the traffic programmability
5G-Xhaul	A software-defined cognitive control plane, able to forecast traffic demand in time and space, and the ability to reconfigure network components.	Some issues can be exploited by SliceNet to improve aspects of the QoE optimizer
COHERENT	Based on the developed control framework, COHERENT will develop advanced traffic steering and resource allocation techniques, novel joint radio and front-haul resource allocation schemes in heterogeneous mobile networks, and investigate radio access network (RAN) virtualization techniques allowing network slicing, and over-the-top applications and services delivery at the RAN level.	Concepts of the global controller and the lower layer abstractions used by the project can be found useful.
5GEx	Open platform enabling cross-domain orchestration of services over these multiple domains, with a set of open source software tools and extensions	Aspects of the cross-domain can be found useful for the multi-domain slicing

3 Control Plane Architecture

3.1 Slice concept

An end-to-end (E2E) service might comprise different domains, each one having different technologies. The E2E slice will consist of sub-slices that belong to one or more domains. The functions such as P&P control, QoE optimisation, service function chaining are handled by the control plane. The slice is an instance that will implement and run the services requested by the SliceNet verticals independently of each other with a distinct set of resources. Therefore, slicing is an enabler to support the SliceNet verticals on a single infrastructure while maintaining and satisfying the QoS guarantees and SLA agreements with the verticals.

The slice concept in the SliceNet project is fully compliant with and further develops the network slice definitions by NGMN and 3GPP. As shown in Figure 6, in line with the recursive nature of a Network Slice Instance (NSI) and Network Slice Subnet Instance (NSSI) defined in NGMN and 3GPP, a SliceNet slice (instance) has the following recursion levels:

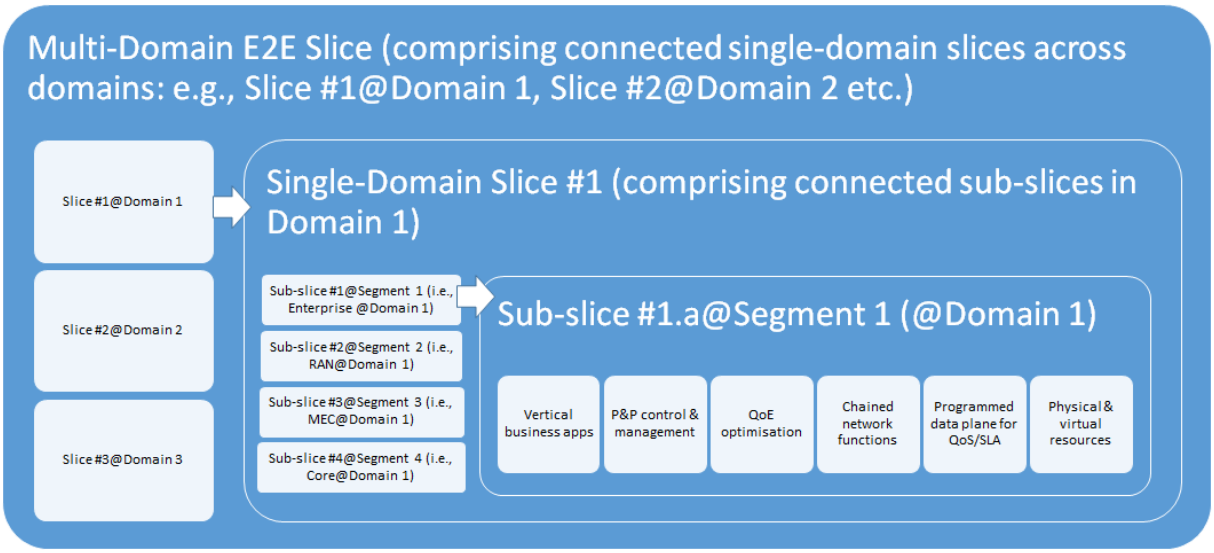

Figure 6: E2E network slice concept in SliceNet

As it is seen in the above figure, each slice is identified by a single id for a specific administrative domain and for each network segment, (ex. RAN, MEC, etc.)

- At the top level, for an E2E slice across multiple domains, it comprises the ordered, structured and connected slices from individual involved single domains, i.e., Slice #1@Domain 1, Slice #2@Domain 2, Slice #3@Domain 3 etc. These single-domain slices are the NSIs of the E2E multi-domain NSI, and they are essentially controlled and managed by the corresponding domains respectively.
- At the intermediate level, within an individual domain, a (single-domain) slice, a single slice consisted of many slices, called sub-slices, e.g., Slice #1@Domain 1, consists of the various sub-slices offered by the involved network segments, including Enterprise network, Radio Access Network (RAN), Mobile Edge Computing (MEC) or Edge network and Core network. The corresponding example sub-slices are Sub-slice #1.a@Enterprise segment, Sub-slice #1.b@RAN segment, Sub-slice #1.c@Edge segment, and Sub-slice #1.d@Core segment, all in Domain 1. These segment-specific sub-slices are the NSIs of a single-domain, and they are essentially controlled and managed by the control and management functions in this domain. It is noted that the Enterprise network segment is typically controlled and managed by the corresponding enterprise/vertical that owns this Enterprise network.

- At the bottom level, a segment-specific sub-slice (instance) may contain the following or a subset of the following:

1. Structured and connected network functions through service function chaining, typically based on a predefined network slice template/blueprint, as being defined in NGMN and 3GPP.
2. The physical and virtual resources to run these network functions, as defined in the NGMN network slice model.
3. Configured data plane through data plane programmability to enable user/tenant traffic engineering to meet the QoS/SLA requirements of the slice-based service.
4. QoE optimisation functions, typically deployed on demand by the network operator to optimise the run-time performance of a specific slice.
5. P&P control and management functions, requested on demand by the vertical to achieve particular P&P functionalities regarding a specific slice for the vertical.
6. Vertical's Apps, required to run for the business level application in the use case, typically deployed by the vertical in the Enterprise network segment.

It is noted that only the first two points have been addressed by NGMN and 3GPP definitions as the state of the art to create a slice instance, whilst the remaining ones (3 – 6) are innovations proposed in SliceNet to achieve advanced slicing for verticals.

3.2 Logical Slicing Model

In response to the slice concept as envisioned in Figure 6, a logical model for slicing in SliceNet is proposed in Figure 7. The slicing model highlights the different levels of innovative and advanced slicing.

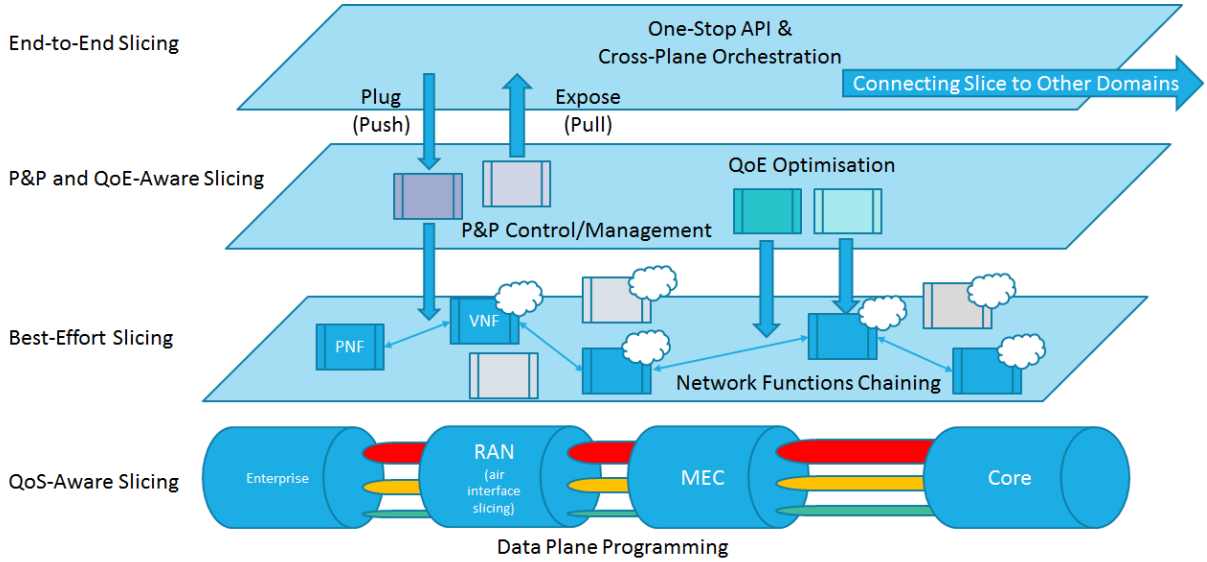

Figure 7: Logical Slicing Model in SliceNet

- Best-effort slicing, which is the state-of-the-art slicing, follows a network function forwarding graph to achieve specific network services for a use case slice. The network functions can be Virtual Network Functions (VNFs) or Physical Network Functions (PNFs).
- QoS-aware slicing, which attempts to realise network slicing with guaranteed network-layer QoS (e.g., in terms of bandwidth, delay/latency jitter, …), to be achieved by programming the data plane across the various network segments, and introducing QoS-enabling mechanisms such as QoS control points and traffic engineering functions in the network. This is in line with the vision from 3GPP to "Provide slice-as-a-service with guaranteed QoS".

- QoE-aware slicing, which further advances QoS-aware slicing by addressing the perceived quality of the user (QoE) receiving the slice-based service in a use case. It is well-known that there is a non-linear, complicated relationship between network-layer QoS and application-level QoE, and thus QoS-aware slicing may not be good enough for some QoE-critical use cases such as eHealth applications and thus should be addressed.
- P&P enabled slicing, which allows a vertical user to request on-demand control functions to be enforced regarding a slice for the vertical, depending on the levels of exposure of run-time network control capabilities from the operator. Conceptually, it can be a push from the vertical to insert a P&P control function or a pull e.g., in terms of monitoring functions exposed by the operator.
- Finally, the E2E slicing, which targets to create a multi-domain slice across various administrative domains. It is noted that SliceNet does not assume uniform deployment of the SliceNet architecture among different administrative domains, and thus the cross-domain slicing is achieved through a negotiation-based approach among the serving domain that receives the slice-based service request from a vertical via the One-Stop API and the distributed candidate partner domains.

3.3 High level control plane Architecture

The SliceNet CP sits on top of an heterogeneous physical and virtualized infrastructure where slice instances are deployed as a composition of different types of NFs, that could be implemented as MEC applications, 4G/5G Control and User Plane network functions, and VNFs deployed and instantiated through the SliceNet management and orchestration tools. In this context, the main goal of the SliceNet CP is to enable the enforcement of specific and dedicated per-slice runtime configurations and adaptations aiming at fulfilling the required QoS and QoE performances expressed by verticals and slice consumers at large at slice template and SLA levels. This per-slice runtime operation is conceived to be applied once the slice instance, as a combination of network domain specific sub-slices, is deployed and instantiated by the SliceNet management and orchestration tools.

The SliceNet CP is designed around two main principles: i) a Service Based Architecture (SBA) approach, ii) and a double level of control and operation APIs abstraction. Concerning the former, SliceNet defines its Control Plane functionalities, components and building blocks as fine granular and decoupled, aiming to improve automation and agile slice operation via loose-coupling control services and logics. The SliceNet CP SBA approach follows the NGMN principles defined for the 5G network architecture [4] and targets lightweight interfaces across control plane services and functions thus enabling rapid interface development and high-level of resource utilization.

On the other hand, the SliceNet CP is conceived to implement two levels of abstraction of slice control and operation APIs (including information models and control logics). The **first level** aims to **abstract specific technology implementations and peculiarities**, and it is the main relevant abstraction capability for the SliceNet platform itself. The **second level** is applied by the P&P functionalities and aims at **further abstracting the technology and implementation agnostic APIs** towards more high-level and vertical oriented logics and models that allow verticals and slice consumers to avoid dealing with the SliceNet abstraction level details and APIs granularity.

Figure 8, shows a high level view of the SliceNet Control Plane and positions it in the context of the overall SliceNet logical architecture presented in deliverable D2.2. As defined above, the SliceNet CP (represented by the green box in the picture) lays on top of the intra-domain SliceNet physical and virtualized infrastructure, which in general is a multi-Point-of-Presence (PoP) infrastructure spanning across multiple network domains covering RAN, MEC and Core segments of 4G/5G networks and belonging to the same administrative entity. As shown in Figure 8, the physical and virtualized infrastructure hosts the E2E slice instances as a combination of NFs, e.g. the yellow and purple MEC Applications and VNFs belonging to two different slices, deployed by the SliceNet NFV/MEC Management and Orchestration (MANO) tools. The different network domains and segments have

different purposes and scopes: while the MEC/Edge and Core ones are intended to be considered as NFV and MEC PoPs for deploying virtualized VNFs and MEC Applications (possibly exploiting micro services), the RAN, backhaul and WAN segments are intended to provide both radio and wired slice network connectivity fulfilling vertical customized QoE and QoS requirements, and guaranteeing cross-slice isolation following SDN principles where applicable and supported by the specific technology implementations. In particular, the backhaul segments can be considered as SDN enabled networks interconnecting MEC/Edge and NFV core PoPs, while the WAN segments as SDN enabled networks providing inter-administrative domain connection.

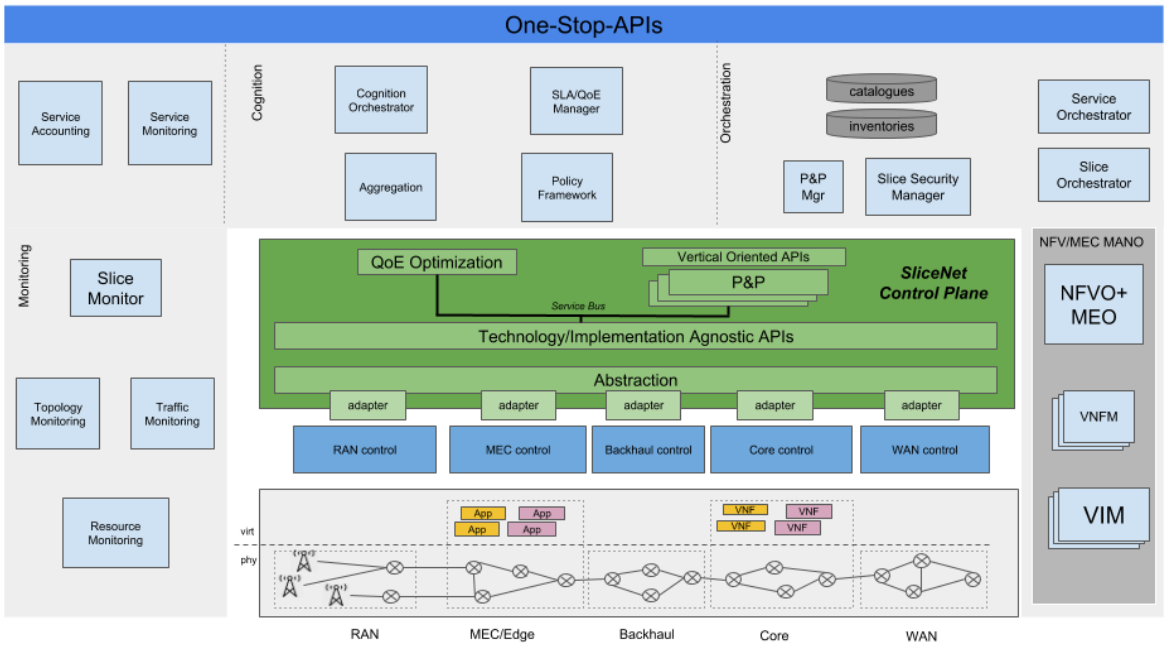

Figure 8: SliceNet Control Plane high level view

The **first level of abstraction** provided by the SliceNet CP, as anticipated above, aims to abstract the several technology and implementation options for controlling the different network segments in the SliceNet physical and virtualized infrastructure. Each of this SliceNet infrastructure segment is defined as an infrastructure pillar, and each pillar is assumed to be controlled (in terms of runtime network resources and functions control logics) according to different technologies and implementations possibly provided by different vendors. With reference to Figure 8, each blue box on top of the network (virt/phy) represents per-pillar control functionalities able to cope with specific implementation and technologies. It is important to highlight that the same pillar in a given administrative domain may consist of several instances of the same or different implementations activated in parallel, thus allowing having different RAN technologies, as well as multi-vendor backhaul and WAN pillars with different control technologies and implementations on top. This is depicted in Figure 9 where two controllers are depicted per pillar.

For this reason, the SliceNet CP is built at its SBIs with a set of control adapters that support the interaction with specific control technologies and implementations. These adapters are the enablers for common SliceNet CP information model and control logics that translate, as depicted in Figure 8, into technology and implementation agnostic slice control APIs exposed towards QoE optimization and P&P functions, and in general to slice management and orchestration tools, following the SBA approach. Every controller is associated with one adapter and the adapters are instantiated in parallel to allow the integration of each controller with the SliceNet CP.

The **second level of abstraction** in the the SliceNet CP is provided, as shown in Figure 8, by the P&P control functions. The SliceNet P&P control is the main enabler for E2E slice runtime customization,

as it intends to offer verticals and slice consumers direct access and control to their slice instances. The P&P aims at exposing dedicated per-slice and vertical oriented runtime control and management APIs towards slice consumers to allow specialization of slice instances according to heterogeneous specific vertical services and use case requirements, as detailed in section 4.1. In particular, one of the main goals is to provide a second level of abstraction on top of the slice technology and implementation agnostic APIs shown in Figure 8: this further abstraction is conceived to hide the complexity (in terms of logics, information models, control and management workflows) of the SliceNet control, management and orchestration tools and components, aiming to offer verticals APIs and runtime control logics closer to the verticals space and produce a vertical view of each slice instance, where possible still leveraging on open and standard SDN and NFV APIs. The P&P control can be considered as a dedicated set of functions available for each slice instance, as depicted in Figure 8 where multiple P&P boxes appear. Each P&P sits on top of the slice technology and implementation agnostic APIs offered by the first abstraction in the SliceNet CP, and follows the SBA approach for interacting with the underlying SliceNet control functions and components. The QoE Optimization functionalities also sit on top of the slice technology and implementation agnostic APIs utilising the aggregated operations offered by the Slice Control Context. In this regard, NSI/NSSI-level (re-)configuration operations may be triggered without the need to know the fine details of the specific technologies involved in each segment that constitute the E2E slice. The specific control and configuration operations are responsibility of the underlying slicing functions (as seen in Figure 9, exposing their capabilities in the form of agnostic APIs which can then be invoked by external modules/services, such as the QoE Optimization module, which also follows the SBA approach to leverage the capacities of other SliceNet control functions.

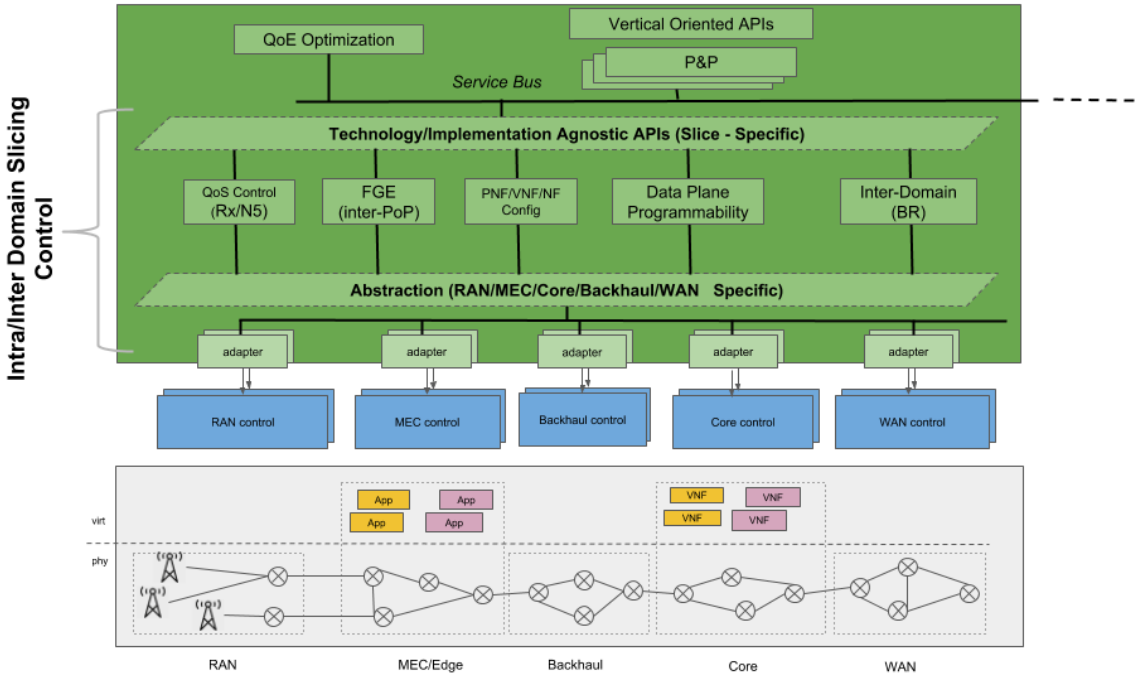

Figure 9: SliceNet Control Plane SBA approach and slice control functions with DP programming

In this SBA enabled SliceNet CP context, each control plane service (or function, e.g. each P&P) is an atomized capability, with high-cohesion, loose-coupling, and independent management from other control plane services. This allows each control plane service to be deployed on-demand, updated and where needed decommissioned independently and with minimal impact to others. The SBA approach envisages that each control plane service produces certain outputs based on specific inputs, as expected by its "service" capability and required by the SliceNet CP and verticals (especially in the case of P&P) specific needs.

The SliceNet SBA principle is further adopted within the intra- and inter-domain slicing control functionalities and services (i.e. those implementing the first level of slice control technology and implementation abstraction), as shown in Figure 9. Here, the capabilities exposed by each underlying infrastructure pillar, together with its set of dedicated control functions and related SliceNet adapters, are integrated and packed in the context of a set of SliceNet control services, which sit on top of the control adapters exploiting common and technology independent control primitives. The arrangement of infrastructure pillars functionalities into control services formulates the slice technology and implementation agnostic APIs which represent the set of slice configuration endpoints that can be accessed by other SliceNet platform components (from P&P and QoE optimization to management and orchestration tools), thus providing the control context of a slice. The concept of the "Slice Control Context" requires the integration of a workflow execution engine that is responsible for segregating every single control operation exposed via the Agnostic APIs into a number of steps. Each step of an operation workflow involves runtime discovery of the appropriate service endpoints to be invoked, dynamic preparation of the payload to be used and endpoint invocation. The result of each step might be also subject to be processed in the next one. The slice control context is exposed via the Agnostic APIs and aggregates all the control operations that are supported in the context of a slice. Interaction with the Agnostic API is done on the basis of usage of slice and sub-slice identifiers. In particular, each SliceNet control service shown in Figure 9 aims to provide specific slice configuration capabilities, following an SBA approach, in terms of:

- Inter-PoP forward graph configuration, mostly focusing on inter PoP network configuration (i.e. cross and inter pillar) to be combined with intra-PoP forwarding graphs that are normally offered by NFV and MEC MANO functions (via specialized VIM capabilities) in the context of NS support. Forward Graph Enabler (FGE) intents to account for the role of WIM Manager and remain compliant with ETSI MANO specifications. In this way the component is expected to be forward compliant with future NFVO implementations that support WIM Management.
- PNF, VNF and 4G/5G (either user or control) NF configuration
- Inter-domain connectivity, focusing on user traffic classification and forwarding beyond 4G/5G Core pillars (i.e. after PGW)
- UE Session QoS Control (e.g. PCF/PCRF interaction)
- Data Plane Programmability for quality based user traffic management that takes advantage of packet processing and acceleration techniques

The SBA in the Intra/Inter Domain Slicing control depicted follows the same principles and same service bus approach described above for QoE optimization and P&P. It allows for dynamic expansion of the SliceNet CP services, adapters and pillars (in terms of new control technologies and implementations to be on-boarded). Indeed, new control adapters can be on-boarded in the SliceNet CP and registered to be used with updated or newly added pillar implementations, following the SBA service registration and discovery features. The SliceNet control services can therefore detect new adapters instances allocated by SliceNet management functions and configure/use them according to their specific logic. In turn, the slice control context exposed at the level of the slice technology and implementation agnostic APIs is updated accordingly extending (or reducing) the view of each slice instance.

The same applies to new available SliceNet control service that might be on-boarded within the control plane to fulfil specific requirements or needs. These control adapters and services have to be considered anyway as a slow-pace dynamicity, with respect to runtime slice control operations, that follows SliceNet management workflows and procedures.

4 Control Plane description of components

This section describes the SliceNet CP components such as Plug & Play (P&P) and QoE with high level workflow procedures, it then details the components decomposition when intra and inter plane slicing are implemented and finally it details the CP high level APIs and interfaces.

4.1 P&P control

SliceNet aims at providing truly customized runtime control, management and operation of E2E slice instances with the main objective to offer vertical-tailored services. The P&P control can be considered as one of the key enablers of slice customization, and aims to offer a novel combination of tailored control functions, APIs and tools to enable verticals to even plug their own control logics to specialize their slice instances according to their needs. The ultimate goal is to provide an innovative control environment, dedicated per slice, which offers to the verticals, and in general to slice consumers, significantly enhanced degree of flexibility for deploying services to the end users.

The P&P control functions are conceived to be activated for the runtime operation of slice instances, and therefore are not be considered as part of the slice provisioning control features and tools in SliceNet. The main goal of the P&P is to expose per-slice instance dedicated and vertical customized control features and capabilities.

Due to the heterogeneous roles and actors envisaged in SliceNet, as described in deliverable D2.2, the P&P control, at least in its concept, has to be considered as agnostic of the specific provider-to-consumer interaction. This way it can be applied to any “slice provider-to-vertical” or “slice provider-to-slice provider” case in support of either single-domain or multi-domain E2E slices. This way, P&P control functions can be exposed to verticals for their customized slices operation, and to other customers in general (like other slice providers) in the context of E2E slices spanning across multiple administrative domains.

As a general approach, the SliceNet P&P is built on top of the slice control exposure principles proposed by Nokia within an Internet-Draft (I-D) part of the IETF Common Operation and Management of network Slicing (COMS) initiative, which is currently tackling network slicing investigation and management architecture definition. In particular, this I-D [15] envisages that depending on the specific slice provider – slice consumer relationship (at different levels, from business-to-business, to SLA, to confidentiality, etc.), the slice provider can offer various and heterogeneous levels of control to the slice consumers, thus translating into different abstractions of slices resources and access to different slice provider management entities.

In this context, the SliceNet P&P control paradigm relies on open and standard SDN and NFV APIs, aiming to enable a higher degree of extensibility of the SliceNet framework offering the possibility to third parties in general to plug their own control logics and functions.

4.1.1 P&P architecture and functionalities

The P&P functionalities are considered mostly as part of the SliceNet CP, and thus are included among all those features and components that allow the runtime operation, configuration and dynamic re-configuration of slice instances, in terms of network functions composition and forwarding graph, QoS and QoE optimization, programmable control of VNFs and PNFs). Therefore, as the whole SliceNet CP, the P&P control logically lays on top of the heterogeneous infrastructure composed by the integration of 5G RAN, MEC and Core network segments, including where applicable the vertical enterprise segments, providing those per-slice customization functions needed to accommodate vertical´s requirements.

Figure 8 shows the SliceNet CP architecture, and in this context, the P&P control has to be considered as the composition of those specific and customized control functions plugged and available for each

slice instance and which allow verticals, or slice consumers in general, to apply their own control logics.

The main goal of the P&P control is therefore to offer verticals with an isolated control environment, specific per slice instance that can be activated on-demand when new E2E slice instances are provisioned. The idea is that each P&P control instance can have access to a limited set of slice control and management primitives, strictly depending on the P&P requirements specified by the vertical and included in the slice catalogue maintained in the SliceNet Cross-Plane orchestration layer (see Figure 10), offering a specific vertical-tailored level of slice control exposure. Moreover, to guarantee isolation, each P&P control instance insists and has access (either direct or indirect) to those physical and virtualized resources and network functions (e.g. for configuration purposes) owned and used by the given slice instance for which it is activated.

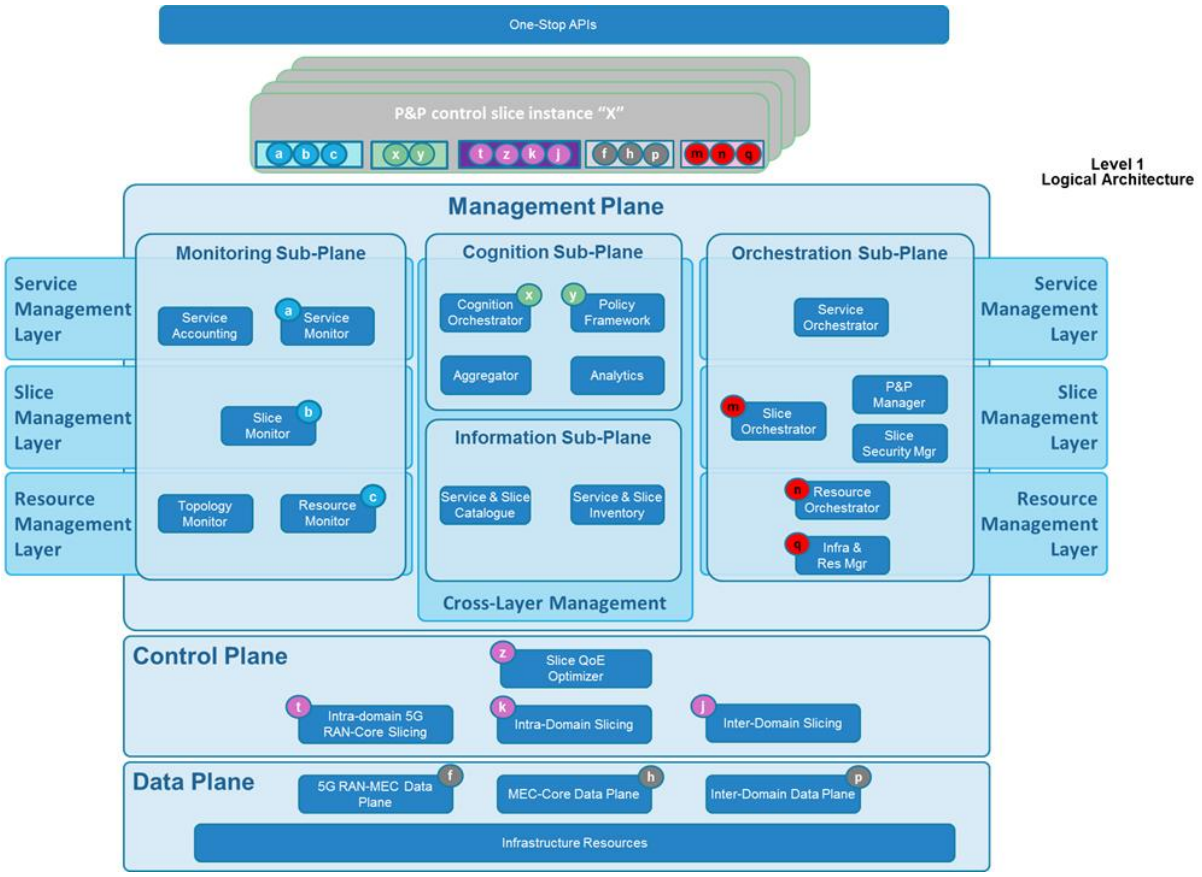

Figure 10: P&P positioning in the SliceNet logical architecture

Moreover, when E2E slices span multiple administrative domains, and thus are composed by sub-slices provided by different providers, the ultimate P&P control instance offered to the vertical should be the composition and combination of per sub-slice P&P control instances managed and offered by each provider in the chain. More details on multi-domain and multi-provider P&P control aspects are provided in section 4.1.4.

Figure 10 starts from the logical architecture approach shown in D2.2, and provides a more coherent view of how the P&P control instances (i.e. one logical dedicated per slice instance) leverage on and consume other SliceNet control and management plane APIs and features. In Figure 10 each P&P control box refers to the specific and slice dedicated logical instance that is responsible to offer a specialized and customized E2E slice view to the vertical, along with dedicated runtime control operations exposed.

Each P&P control instance can have access, through dedicated plugins (shown in Figure 10 with the coloured circles), to SliceNet control and management plane components and features to allow each

slice instance to be customizable at different levels and layers. Figure 10 presents a one-to-one mapping between APIs and primitives on control and management components, and related plugins/drivers on the P&P control side. In particular, all the possible APIs and primitives to which the P&P control instance can have access are shown in the picture, from monitoring, to control and orchestration planes. For each specific P&P control instance, this access is regulated by the P&P management features residing in SliceNet management plane, according to slice template capabilities and requirements combined with vertical (or generic slice consumer) needs.

Following the SliceNet CP principles defined in section 3.3, the P&P implements the second level of abstraction oriented to expose to the verticals a simplified view of E2E slice instances, that on the one hand can be aligned and compliant with each vertical logic and needs, and on the other hides and further abstracts the slice technology agnostic APIs according to the agreements with the slice provider in terms of control exposure. Moreover, the P&P follows the SBA approach described in section 3.3, and therefore each P&P control instance can be considered as a dedicated per-slice control plane service exploiting the slice technology agnostic APIs and providing atomized per-slice control exposure capabilities loosely coupled to other control plane services (i.e. either other P&P control instances dedicated to other slices or other SliceNet CP services) with deployed on-demand with independent management workflows.

In this context, Figure 11 shows a more detailed view of the P&P integration within the overall SliceNet framework, focusing on three main categories of P&P envisioned interactions (enabled by different types and classes of P&P plugins depicted with different colours in the picture): monitoring, slice control plane, management and orchestration. In particular Figure 10 highlights the different types of interactions and accesses that a given P&P instance may have. For sake of readability of the picture, a single P&P control logical instance is represented; following the SBA approach, each P&P instance accesses the SliceNet CP service bus and thus discovers the available technology agnostic APIs offered by the SliceNet CP services. In summary, this approach translates into three main categories of control and management features and APIs that can be consumed by a given P&P instance through dedicated plugin types:

- *service, slice and resource monitoring services*: these allow the P&P to retrieve KPIs, metrics and status information at different granularities (i.e. resource, slice and service) directly from the SliceNet monitoring vertical components via dedicated plugins
- *SliceNet control plane services*: these allow the P&P to enforce runtime slices and sub-slices configurations leveraging on the slice technology agnostic APIs exposed the SliceNet CP services and following the SBA approach. These include, among others: i) enforcement of vertical customized QoS/QoE policies, ii) configuration of slice Network Functions (NFs, e.g. deployed as PNFs or VNFs in a given slice instance)
- *Slice and NFV/MEC orchestration services*: these allow the P&P to access slice lifecycle management and NFV/MEC orchestration primitives directly from the SliceNet management and orchestration vertical. This enables an high level of control (and management) exposure to the vertical, offering the possibility to trigger deployment, scaling or termination of customized vertical MEC Applications or VNFs, possibly at specified locations.

As described in detail in the next subsection, on top of these specific P&P plugins an abstraction layer allows to specialize a generic and common slice information model into a specific and vertical customized slice view. In this case, the specialized vertical view of an E2E slice instance is offered through a set of vertical oriented APIs (implementing the second level of abstraction in the SliceNet CP) exposing a well-defined set of vertical friendly control operations.

4.1.1.1 P&P functional decomposition

As depicted in Figure 11, each P&P instance is an independent logical run-time component which has access to multiple SliceNet control and management logics and APIs by means of dedicated drivers

and plugins. The restriction and selection of the tailored subset of these primitives is performed by P&P management features during the activation phase. The P&P management within the SliceNet management plane is responsible for each P&P control instance lifecycle management (activation, plug of specific tailored drivers, configuration of proper abstraction features, deactivation).

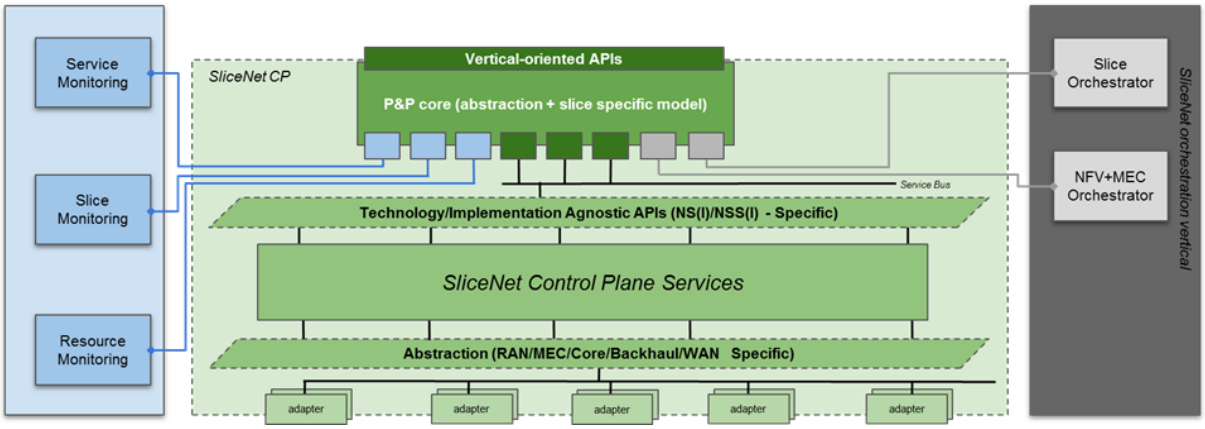

Figure 11: SliceNet P&P main interactions with other SliceNet components

In general, the P&P deals with control and management primitives and APIs which follow the SDN and NFV principles. In particular, each P&P instance should be able to plug specific drivers and plugins (while abstracting their logics and complexity) for SDN controllers, NFV MANO tools and even programmable PNFs and VNFs) running in the programmable E2E infrastructure (spanning from RAN to MEC and core networks).

Due to the heterogeneous nature and needs of vertical actors and slice consumers at large, the SliceNet P&P is conceived to provide tailored exposure of control and management primitives. A slice consumer can be either a vertical actor, as ultimate consumer of an E2E slice instance, or another slice provider in the case of slice instances spanning multiple providers. The SliceNet P&P paradigm applies to both cases following the same principles, while the customization of control and management features exposed to consumers may strongly differ from case to case, especially in the provider-to-provider case where more restrictions and limited control may apply.

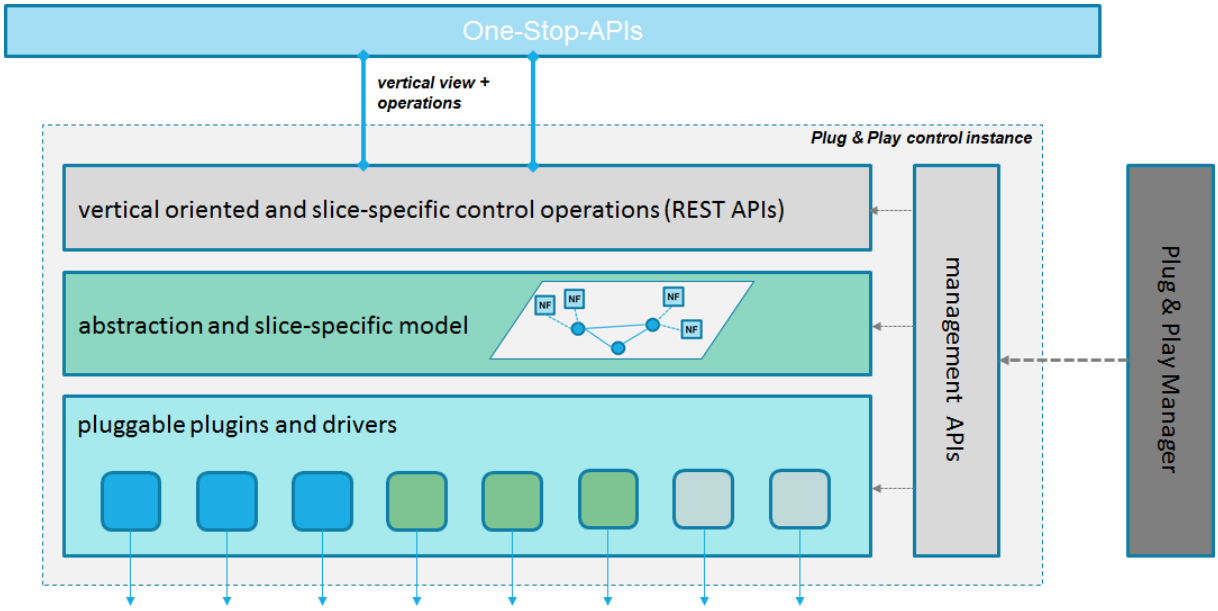

Figure 12: SliceNet P&P control logical architecture decomposition

Figure 12 presents the functional decomposition of a P&P control instance, which follows a layered approach built by three main high level components and targeting an high degree of flexibility and

customization in creating customized views of slice instances for the verticals. In particular, the SliceNet P&P is built around a generic, common and technology agnostic slice information model to be specialized in the context of each E2E slice (thus for each P&P instance) according to vertical needs and control exposure level. In particular, this generic slice model has to be considered as a kind of topology-like abstract view of a slice, where a set of vertices and edges can be arranged in the form of graphs. The specialization of the generic slice model, providing the customized view of the slice instance for a given vertical, is based on the provision of a specific slice-context to each vertice and edge, together with allowed control operations for each of them. Therefore, as an example, a specific customization of the generic slice model may result into a topology graph including interconnected PoPs with NFs (e.g. MEC Applications or VNFs) associated to each PoP, along with a set of operations associated to each entity in the slice-specific model (e.g. reconfigure a VNF, or deploy a new MEC Application in a specific PoP). The detailed specification of the model (as well as of the whole P&P) is undergoing in WP4, where some specific reference work is considered for this generic slice model. In particular, the technology independent information model proposed in a IETF COMS I-D [16] is considered as a relevant candidate starting point to model properties, attributes and operations on each slice entity, following a generic approach on top of the data model for network topologies [17].

With reference to Figure 12, the three main components of the P&P control instance can be briefly described as follows:

- *Pluggable plugins and drivers*: this layer includes the set of plugins and drivers that provide the required adaptation between the P&P generic slice model and the monitoring, control plane and orchestration SliceNet framework primitives. In the P&P, these plugins and drivers are considered as pluggable modules providing access to specific SliceNet control and management logics. In particular, for the SliceNet CP they follow the SBA approach and provide an implementation of the slice technology agnostic APIs. In order to enable flexibility, each P&P driver exposes its capabilities to the abstraction layer to enable its dynamic pluggability and usage by the P&P logics.
- *Abstraction and slice-specific model*: this layer provides the specialization of the generic slice information model into the customized vertical view of a given E2E slice instance. The slice-specific model produced implements an abstracted and vertical friendly view enabling the second level of abstraction within the SliceNet CP. These abstraction features allow also to map control operations coming from the vertical oriented APIs into specific actions to be enforced through the pluggable plugins and drivers.
- *Vertical oriented APIs*: this layer implements the set of control and management APIs that are exposed to the vertical. It basically offers the vertical tailored control operations over the slice-specific model and view, following the slice control exposure level agreed by slice consumer and slice provider and described as part of the slice template or descriptor. These vertical oriented and slice-specific APIs are exposed to the vertical with the mediation of the One-Stop-APIs layer, which embeds these P&P northbound APIs for each given E2E slice providing all the required authentication and user management logics required to expose a secure access to third parties.

The three high level P&P layers described above are coordinated and properly configured on a per-instance basis (i.e. per-slice instance basis) from dedicated management APIs that cover vertically the whole P&P as depicted in Figure 12. The P&P management APIs allows to:

i) plug and activate the proper set of plugins and drivers, according to the control logics to be exposed to the vertical for the given slice instance;
ii) specialize the generic slice model into the customized vertical view of the slice, thus enforcing the specific slice context to each vertice and edge in the view;

iii) enable slice-specific control operations in the form of dedicated APIs on top the specialized slice model, thus configuring the vertical oriented APIs to expose the required control logics towards the vertical.

These management APIs are exploited by the P&P Manager component (residing in the SliceNet management and orchestration) when after the E2E slice deployment and provisioning, the associated P&P control instance has to be activated. Indeed, the P&P Manager is the lifecycle manager of the whole set of P&P control instances. This means that it takes also care of dynamic adaptations of P&P instances at runtime, e.g. by enabling the plugging of new plugins and drivers whenever required, e.g. to fulfil slice upgrades in terms of level of control exposure upon request (validated and accepted by the slice provider) from the vertical.

4.1.2 Levels of slice control exposure

As said, the SliceNet P&P control enables a generic slice consumer to access a set of customized slice instance control and management APIs and primitives. Not all slice instances will be allowed to have the same level of control exposure towards the consumer, especially in the case of slice provider-to-slice provider interactions for the provisioning of E2E slices, where limited control could be exposed from one provider to another due to confidentiality reasons and to restrict access to specific slice resources and control primitives.

According to the agreements, relationships and trust among the different actors involved (i.e. vertical-to-slice provider and slice provider-to-slice provider) different flavours and levels of P&P control exposure are defined in SliceNet:

- very basic monitoring only option, where the slice provider offers only means to monitor slice KPIs (e.g. in terms of performance, resource availability, etc.), while slice configuration and customization is chosen from a catalogue of pre-designed slice templates
- limited control option, where the slice consumer can have also access to a limited set of SDN and NFV control and configuration primitives to customize the slice runtime operation
- extended control option, where the slice consumer can also access the slice instance lifecycle management, thus opening and offering the full operation of the slice

As a further decomposition of the above control exposure approaches, Table 2 provides a full list of potential options. These options for the levels of control exposure should be reflected into slice capabilities and requirements in the slice templates and descriptors, stored in the slice catalogue and advertised (or at least made available) through the One-Stop-API. It is worth to mention that the control exposure levels listed in Table 2 have to be considered as preliminary as they are under detailed definition and analysis in the context of WP4 P&P dedicated activities.

Table 2: SliceNet P&P levels of slice control exposure

Control Exposure Level	Main offered control capabilities
Level 0	• Monitoring only of slices NFs • Consumer can collect KPIs for the NFs
Level 1	• Level 0 control capabilities • Access to limited set of NFs configuration and reconfiguration options (e.g. catalogue based)
Level 2	• Level 0 + Level 1 control capabilities • Possibility to onboard and request activation/instantiation of proprietary NFs • Full access to NFs configuration and reconfiguration (proprietary and non-proprietary)
Level 3	• Level 0 + Level 1 + Level 2 control capabilities

	• Access to pre-defined SDN APIs to control chain/composition of NFs for QoS/QoE purposes
Level 4	• Level 0 + Level 1 + Level 2 + Level 3 control capabilities • Access to NFs lifecycle management (instantiate, terminate, scale, start, stop)
Level 5	• Level 0 + Level 1 + Level 2 + Level 3 + Level 4 control capabilities • Access to slice/sub-slice lifecycle management (instantiate, terminate, scale, start, stop, collection of KPIs)
Level 6	• Level 0 + Level 1 + Level 2 + Level 3 + Level 4 + Level 5 control capabilities • Limited access to slice management platform
Level 7	• Level 0 + Level 1 + Level 2 + Level 3 + Level 4 + Level 5 + Level 6 control capabilities • Full access to slice management platform

4.1.3 P&P in multi-domain slices

From a technical perspective, the SliceNet P&P control is not limited to customize the slice control at the vertical space only. Indeed, in addition to the P&P control exposed to the verticals, the multi-domain interactions across providers for building and composing E2E slices will be also exploiting P&P functions. Assuming that E2E slices will always be offered to verticals by a single provider (which in turn can trade and negotiate part of the slice resources and network functions with other providers), a multi-provider (i.e. multi-domain) slice instance can be considered as fragmented into several single provider sub slices (see Figure 13). When provisioned, an E2E slice instance exposes a set of P&P control functions and primitives to the vertical according to its needs and requirements. This P&P is the result of the combination and abstraction of the P&P control functions that each provider contributing to the E2E slice is offering and exposing.

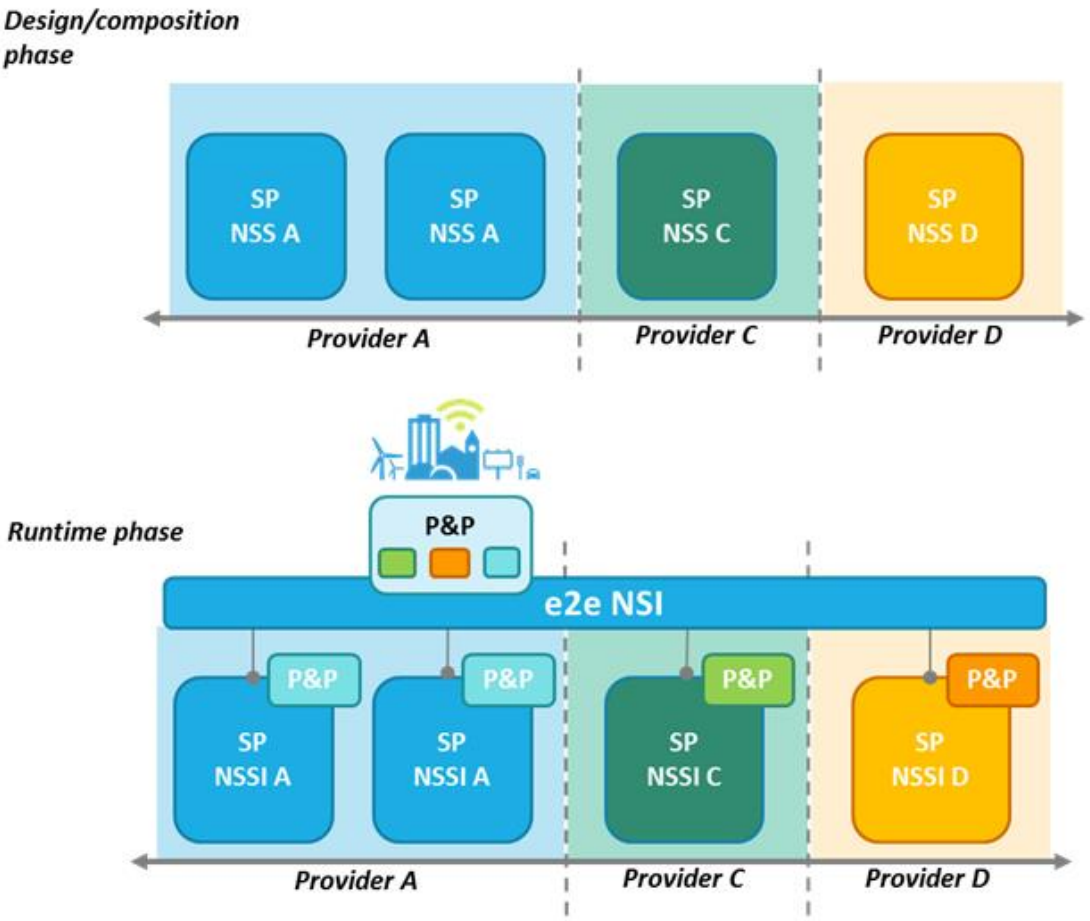

Figure 13: SliceNet P&P approach in multi-provider scenarios

4.1.4 High Level workflow diagrams

The SliceNet P&P is conceived to provide verticals and slice consumers at large runtime access to their slice instances, offering them the possibility to directly control the operation of slices according to the specific level of exposure agreed with the slice provider. The P&P, as it has presented and designed above, is highly flexible by design, allowing dynamic plugging of southbound drivers, customization of slice view and model to fit vertical needs, and exposure of specialized control operations.

This means that the behaviour of each P&P instance, in terms of control and management workflows, is customized and specialized for each specific vertical use cases. Indeed, each P&P instance is designed to adapt to the vertical needs on the one hand, and to the slice provider policies in terms of control exposure to be offered. However, a reduced set of common internal P&P workflows are presented in this section to describe how P&P works at an high level in three specific cases of control operations exposed: i) slice metrics and KPIs collection, ii) slice VNF configuration, iii) slice VNF deployment.

4.1.4.1 P&P slice metrics collection

This P&P workflow is related to one of the most conservative control exposure level, that basically refers to allowing a vertical or slice consumer to collect slice related metrics and KPIs, without offering the possibility to apply runtime modifications at slice configuration, forwarding graph or network functions composition level.

This slice metrics collection internal P&P workflow, see Figure 14, can be summarized with the following step-wise description, assuming that an E2E slice instance is already provisioned and consumed by a given vertical:

1. the vertical accesses its E2E slice instance and requests through the One-Stop-API, which embeds and exposes the P&P vertical oriented APIs, the collection of slice level KPIs and metric according to the agreed contract with the slice provider, e.g. with explicit queries
2. the request from the vertical is processed by the P&P northbound vertical oriented API layer, which maps it to the correspondent specialized slice view set up by the P&P for this slice instance. In the case of slice level metrics and KPIs collection, the target is to retrieve from the SliceNet monitoring vertical aggregated information at the slice level.
3. the metrics collection request is further mapped from the specialized slice view to the proper southbound plugin, which is selected as implementation of the specific models and logics to access the Slice Monitoring components and functions within the SliceNet monitoring vertical
4. the Slice Monitor southbound plugin retrieves the required slice metrics and KPIs from the Slice Monitoring component(s) within the SliceNet framework, thus collecting the aggregated slice level information
5. the collected metrics and KPIs are mapped back to the abstracted and customized slice view implemented by the P&P and are then provided to the vertical by means of the One-Stop-APIs

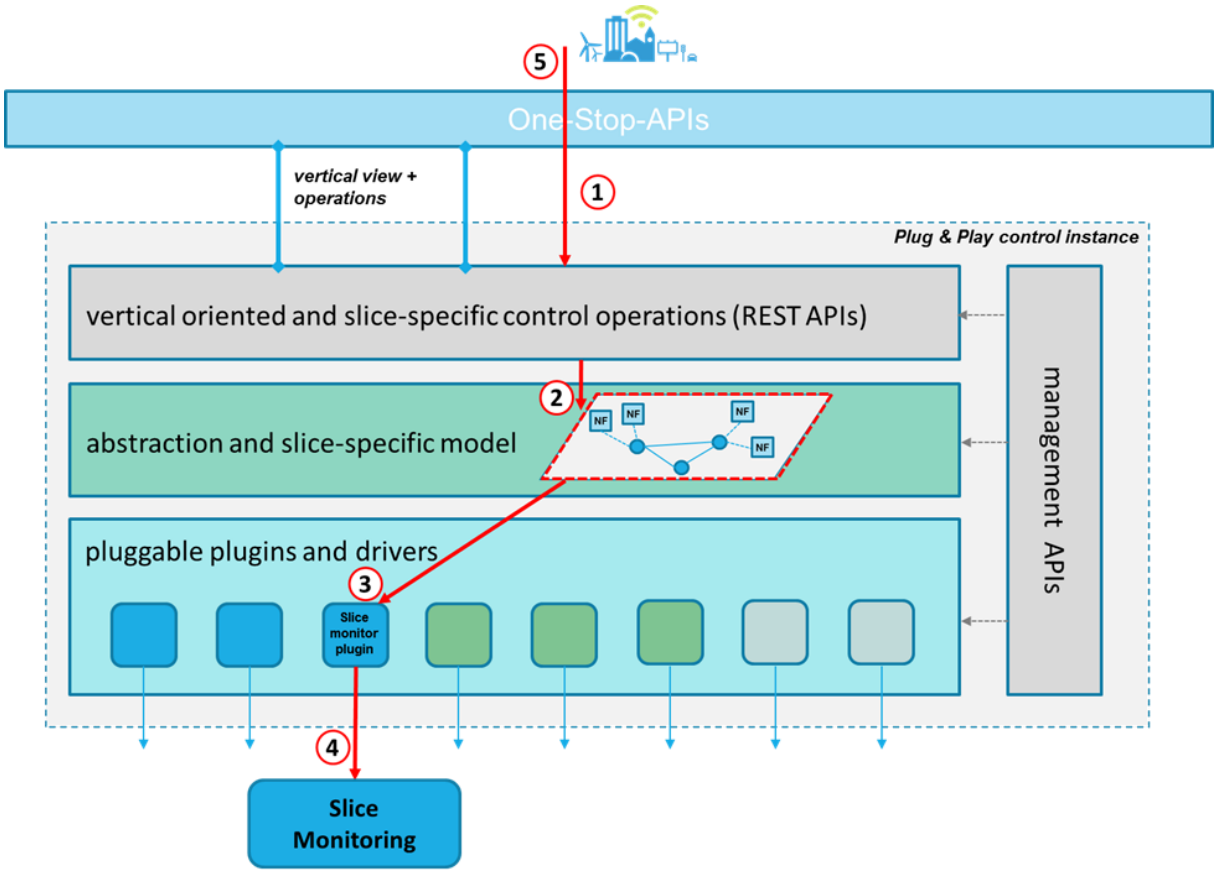

Figure 14: SliceNet P&P internal workflow for slice metrics collection

4.1.4.2 P&P slice VNF configuration

This P&P workflow considers a more open control exposure level with respect to the previous one, and describes the case where the slice provider allows the vertical to directly access some of the

network functions (e.g. one or more VNF instances or MEC applications) composing the E2E slice instance to apply some configurations, e.g. according to vertical-specific logics and requirements.

This slice VNF configuration internal P&P workflow is presented below with a step-wise description, still assuming that an E2E slice instance is already provisioned and consumed by a given vertical. The workflow is shown in Figure 15.

1. the vertical accesses its E2E slice instance and requests through the One-Stop-API the configuration of a VNF instance deployed in its E2E slice, including the whole set of attributes and parameters to be enforced in the VNF
2. the request from the vertical is processed and mapped by the P&P northbound vertical oriented API layer to the correspondent specialized slice view. Here the proper slice instance view entity, i.e. the one modelling the VNF to be re-configured, is identified and selected
3. the VNF configuration request is then mapped from the abstract and customized slice view to the proper southbound plugin, which is selected as implementation of the specific interface to access the related SliceNet control services providing the required control operation
4. the VNF config southbound plugin enforces the required configuration through the SliceNet CP technology and implementation agnostic APIs, which are offered by a dedicated control service taking care to translate the request towards the actual VNF instance
5. the result of the VNF config control operation is reported back to the vertical with any additional information required according to the specific configuration request

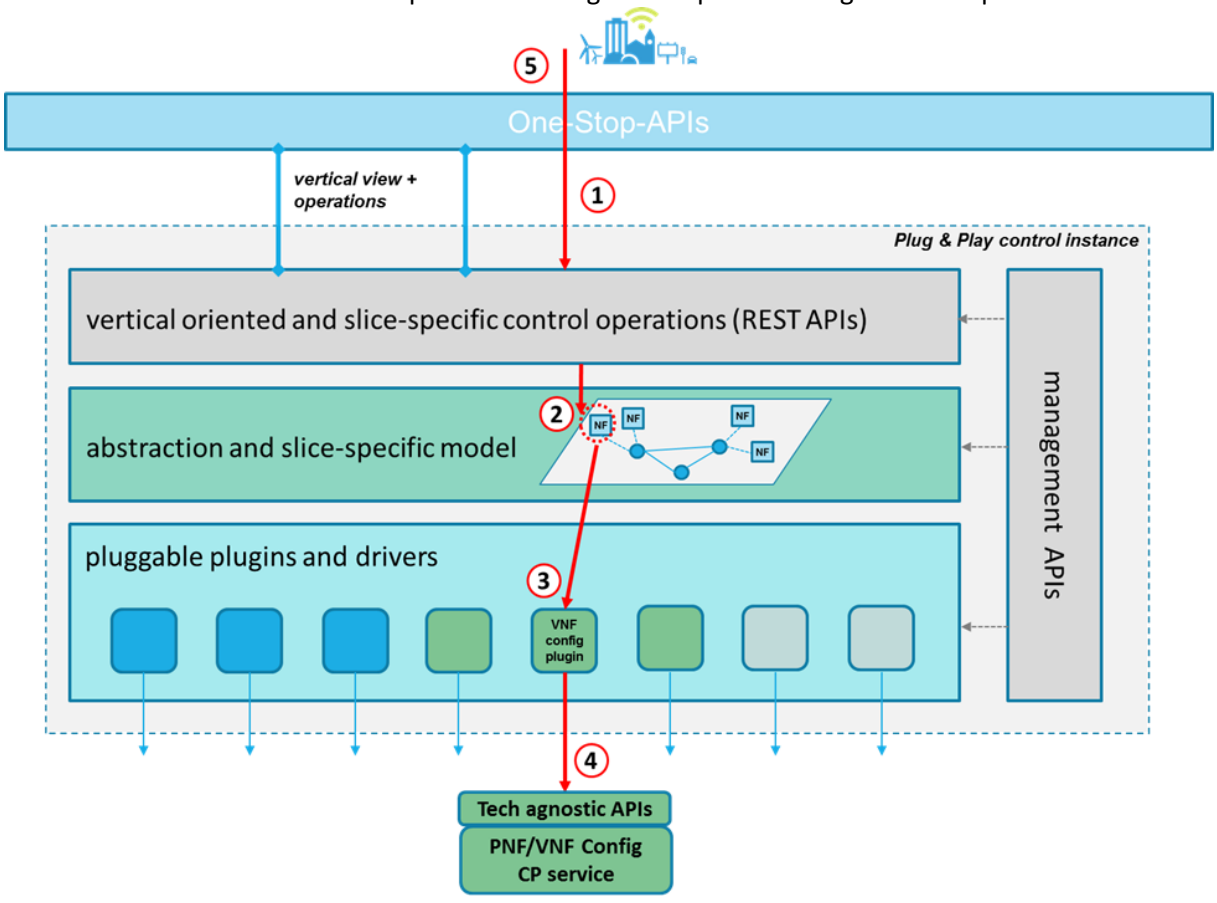

Figure 15: SliceNet P&P internal workflow for slice VNF configuration

4.1.4.3 P&P slice VNF deployment

The option of exposing a reduced set of slice network functions lifecycle management operations is also considered within the SliceNet P&P. This workflow describes the case where the slice provider

allows the vertical to request specific lifecycle management actions for some network functions (e.g. one or more vertical custom VNF instances or MEC applications), e.g. to trigger the deployment of new VNF instances to be composed in the E2E slice.

This slice VNF deployment internal P&P workflow is briefly summarized below with a step-wise description, considering as a prerequisite that an E2E slice instance is already provisioned and consumed by a given vertical:

1. the vertical accesses its E2E slice instance and requests through the One-Stop-API the deployment of a new specific VNF (e.g. a customized VNF provided by the vertical itself) in its E2E slice, possibly including additional information and attributes, i.e. how many instances to scale and where in terms of location
2. the request from the vertical is processed and mapped by the P&P northbound vertical oriented API layer to the correspondent specialized slice view. Here the proper slice instance view entity, i.e. the one(s) modelling the PoPs or locations where the new VNF instance(s) have to be deployed
3. the VNF deployment request is then mapped from the abstract slice view entity to the proper southbound plugin, which is selected as implementation of the specific interface to access the related SliceNet management and orchestration component offering the required VNF lifecycle operation. In this case, depending on the specific control exposure agreed with the slice provider, as well as on the interface approach and mechanisms towards orchestration components (currently under discussion in WP4), the candidates plugin could be:
    a. Slice orchestrator plugin,
    b. NFV/MEC MANO plugin (the one considered in this workflow).
4. the NFV/MEC MANO southbound plugin translate the request coming from the abstraction layer and enforces the required VNF deployment operation through the NFV/MEC orchestration tools in the SliceNet management layer
5. the result of the VNF deployment operation is reported back to the vertical with any additional information required, e.g. for later accessing the new VNF instance for controlling it

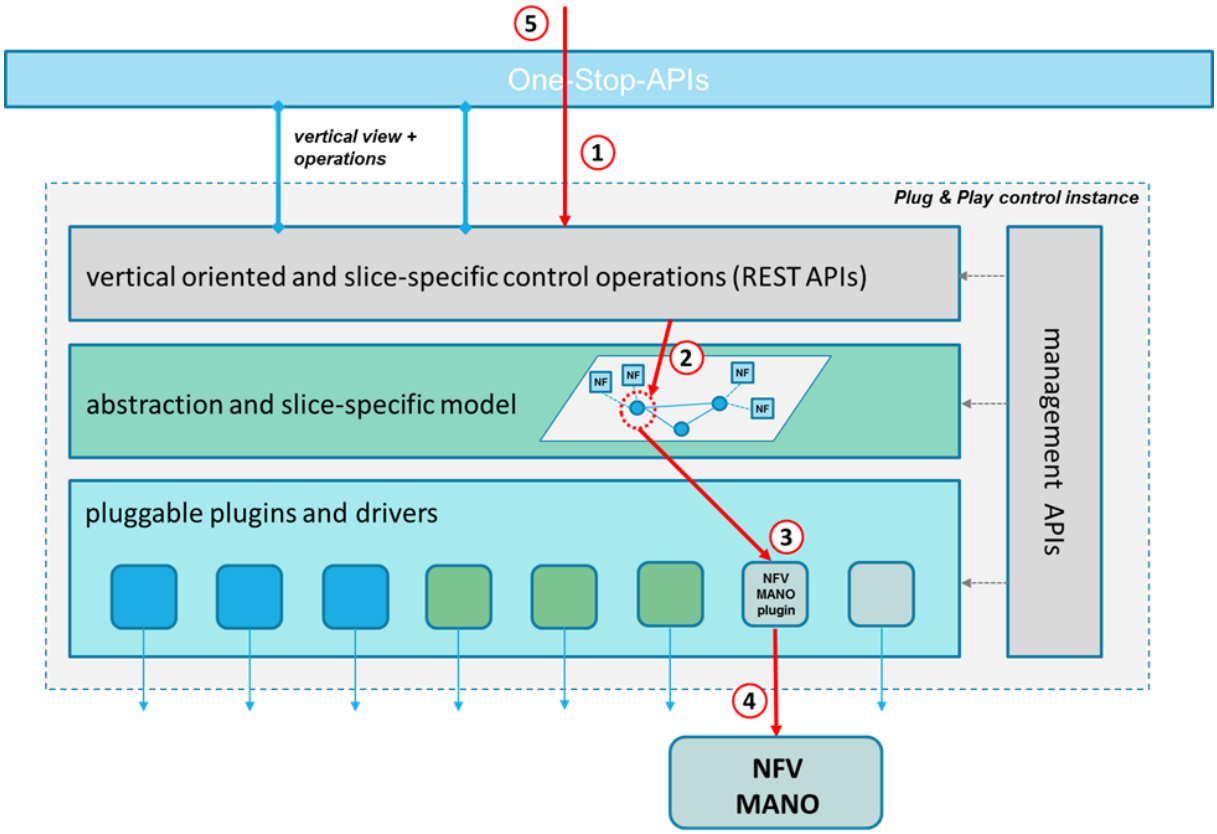

Figure 16: SliceNet P&P internal workflow for slice VNF scaling

4.2 Slice QoE Optimizer

One of the key features of SliceNet architecture is the QoE/QoS-driven slice provisioning, tailored to the specific needs of the vertical. In this regard, besides the (re-)configuration of the physical/virtual resources needed in both provisioning and runtime phases of the slice instances, one of the missions of the SliceNet CP is the maintenance of QoE levels for every active slice instance in order to deliver an optimized service towards the vertical user. Within the control plane, the Slice QoE optimizer module is the responsible of such task.

Thus, the Slice QoE optimizer is conceived as a per-slice optimisation framework in which QoE/QoS performance metrics are monitored to measure the delivered quality, analyses and determines the most optimal actions needed to be taken to re-establish desired quality levels, and enforces them through vertical-informed actuators, which, at their turn, interact with the slicing functionalities of SliceNet CP to trigger the necessary actions at the underlying infrastructure level (physical and/or virtual).

From a design perspective, the Slice QoE optimizer module leverages mainly on two aspects: cognition/machine learning (ML) and policy-based configuration of resources/functions. On the one hand, ML techniques are employed with the scope of infer current QoE levels for the slice given the observed QoS metrics at the slice level, predict future QoE levels so as to proactively act against poor quality situations and determine suitable actions to be taken for QoE levels maintenance. For these goals, the appliance of ML techniques is performed at two levels, spanning both management and control functions of the SliceNet architecture. While long-term analysis and prediction is performed at the management plane, more specifically at the Cognition sub-plane, real-time analysis and actions are performed at the control plane, i.e. the Slice QoE Optimizer module. This enables for fast reaction against dynamic conditions of the physical/virtual infrastructure, without impacting on the overall learning procedures at the Cognition sub-plane side.

As for the policy-based configuration approach, this is also performed in a cross-plane level. The outcomes of the long-term ML mechanisms taking place at the management plane are translated into concrete system policies that dictate how resources and functions should be (re-)configured and instantiated, both in a system-wide fashion, for common infrastructure resources/function, and specialized policies tailored to current active NSIs, which will mandate the potential re-configurations only within the particular instance. The generated policies are then pulled towards the Slice QoE Optimizer module, which its mission in this regard is to select the most appropriate policies to be applied when QoE level violations are experimented. Note that the Slice QoE Optimizer does not infer or construct any policy from the monitoring data received, since the goal of such module is to be able to react fast against the observed degradations/failures so all learning is performed at the management level of the SliceNet architecture.

Given these design principles, the following sub-section specifies in more depth the functions of the module and their relationships.

4.2.1 Slice QoE Optimizer functions decomposition

The Slice QoE Optimizer module is considered as a dedicated control function within the SliceNet CP, following the function splitting of 4G/5G architectures, which separate control functions into common and dedicated ones. From this perspective, SliceNet common control functions are the ones that are transversal for all NSI configuration operations, namely both intra- and inter-domain slicing functions, while the dedicated control functions are the ones that are specifically tailored for every NSI, such as the P&P control and the Slice QoE Optimizer. Thus, for every deployed NSI, an instance of the Slice QoE Optimizer is created and deployed, which then will take care of the QoE optimization for that specific NSI during its runtime. Figure 17 depicts the internal functional structure for a Slice QoE Optimizer instance, highlighting the main relationships within the module, with other control plane functions and with management plane functions.

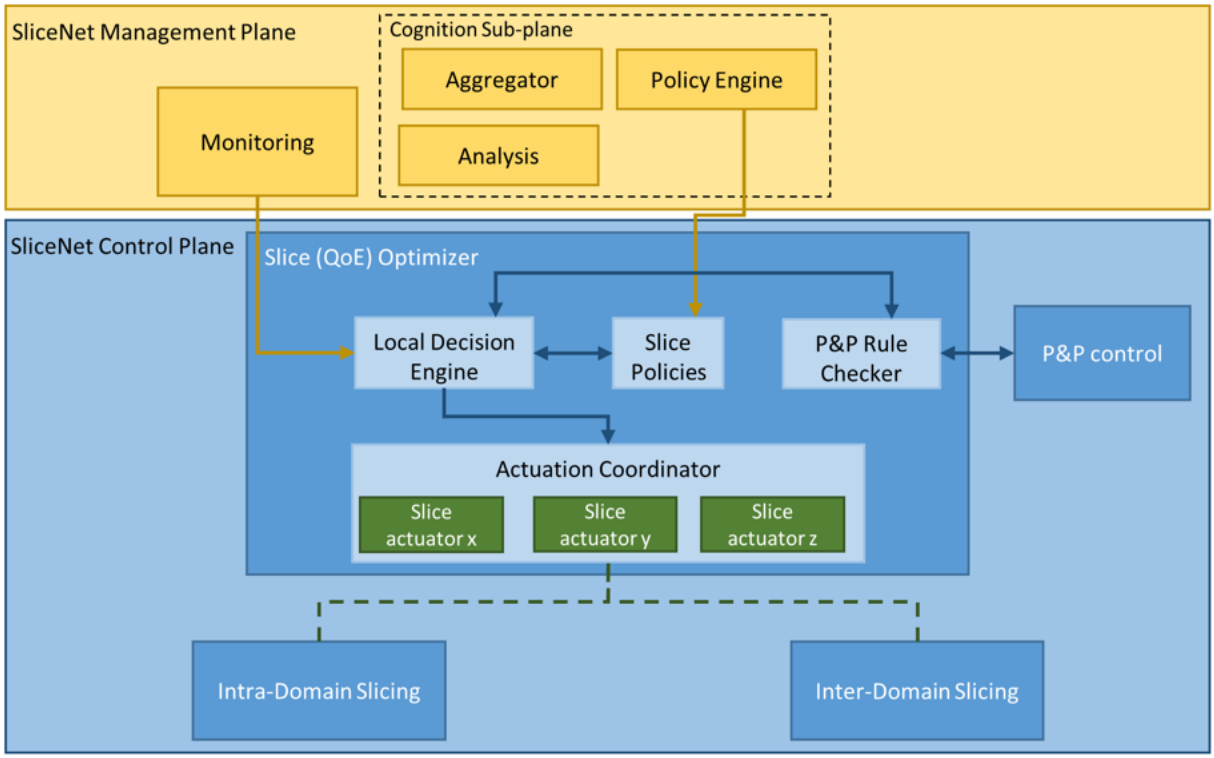

Figure 17: SliceNet Slice QoE Optimizer logic architecture decomposition

The core of the Slice QoE Optimizer is the Local Decision Engine (LDE). In the LDE, already trained ML models are executed. Such models are fed with one or several slice KPIs (e.g. E2E latency, bandwidth, packet loss ratio, etc.), gathered at the monitoring sub-plane at the management plane, more

specifically, the slice monitoring sub-module. With such information, the models running at the LDE infere QoE values for the NSI. These values are then employed to trigger corrective actions against quality degradations and/or failures in the NSI. In this regard, the LDE may act in a reactive way, in which once QoE levels are violated, the whole actuation framework is triggered, or in a proactive way, in which if bad QoE situations given current and past measurements are predicted, the actuation framework is triggered to avoid such future events in a timely manner.

In regards of the actions to be triggered, as mentioned before, a policy-based approach is followed. The Slice Policies sub-module acts as a local repository where potential policies to be applied at the NSI are stored. Such policies are pushed from the cognition sub-plane, more specifically from the Policy Engine, determined thanks to the long-term analysis of monitoring data performed at such sub-plane. Once corrective actions are needed, the LDE checks the policies stored at the Slice Policies sub-module, which will have the most updated policies concerning the NSI. Then, it is the responsibility of the LDE to analyse and decide the most suitable policies in accordance with the type of anomaly detected/predicted.

The generic optimization goal may be re-configured on demand from the vertical user to reflect current preferences in regards of quality maintenance, for instance, due to changes on the offered service on top of the NSI or the introduction of new ones. Such interaction is achieved through the P&P control, which, as mentioned in previous subsections, is employed to expose control capabilities of the deployed NSI towards the vertical user. In this regard, the interaction between the P&P control and the Slice QoE Optimizer is performed through the P&P Rule Checker sub-module. In essence, it checks for the allowed actions that are enabled through the P&P control instance of the particular NSI. Then, if needed and if they affect the QoE optimization process, are reported to the LDE, which will take them into account during the runtime and the decisions on how to solve unsatisfactory QoE situations.

To finalize the functional description, the Slice QoE Optimizer enforces the chosen policies through a set of slice-level actuators. From a functional perspective, these actuators are the responsible to interact with the capabilities exposed through control and data plane programmability APIs for the different segments. The interaction with these APIs, that is, the desired configuration orders are performed via the slicing functions of the SliceNet CP, namely, Intra-Domain Slicing and Inter-Domain Slicing. In this sense, technology specific configuration details are (mostly) hidden from the Slice QoE Optimizer thanks to the technology/implementation agnostic abstraction layer, providing a view of the underlying infrastructure as a set of APIs to enforce QoE/QoS guarantees decided during the optimization process. Given that multiple actuations may be needed to achieve the desired NSI re-configuration, both in several NSSIs and segments (RAN, MEC, core, ...) and infrastructure levels (physical, virtual and functional), the different sets of actuators are coordinated through an Actuation Coordinator, which, once contacted by the LDE with the set of actions to be performed, is in charge to timely contact the different slicing functionalities in the correct order and from the different infrastructure perspectives (physical, virtual or functional) to achieve a holistic re-configuration of the NSI.

4.2.2 High level workflow diagrams

After having described the main functional blocks of the Slice QoE Optimizer, their mission and the relationship with each other, this section describes the high level workflows for the main operations to be achieved through the module.

4.2.2.1 Runtime QoE optimization

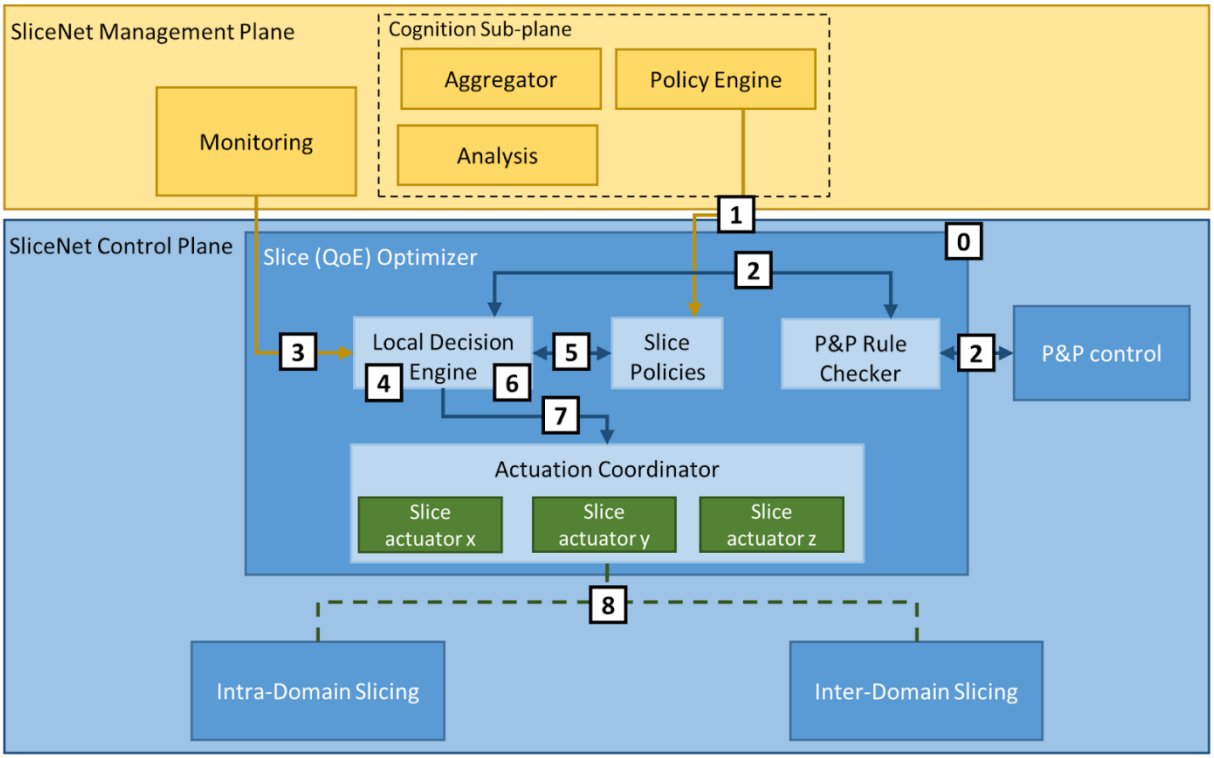

Figure 18: Runtime QoE optimization workflow

Table 3: Runtime QoE optimization steps

Step	Description
0	After the NSI initial provisioning and configuration, an instance of the Slice QoE Optimizer is deployed with particular NSI details.
1	During runtime, potential policies to be applied at the NSI for QoE optimization are periodically copied/disseminated from the Cognition sub-plane to the Slice Policies repository.
2	Allowed P&P actions/operations are checked and reported to the LDE.
3	Monitoring information coming from the Monitoring sub-plane is gathered by the LDE.
4	The LDE determines QoE levels and checks/predicts for QoE violations.
5	If such events happen, potential actuation policies are gathered from the Slice Policies repository.
6	The LDE determines among all available active policies the most suitable to be applied according to the QoE violations at hand.
7	After deciding the most proper policies, the LDE contacts the different sets of actuators (common and dedicated) to enforce them.
8	The Actuation Coordinator contacts the different slicing functions of SliceNet CP to trigger the concrete (re-)configuration actions through the set of exposed APIs (slice actuators).

4.3 Intra-domain multi-tenant slicing and service composition

4.3.1 Logical Intra-domain slicing functions decomposition

The “intra-domain multi-tenant slicing” functionality is thought to handle tasks for the enforcement of network functions (PNFs/VNFs) configuration rules and policies governing the run time operations of RAN, Core and MEC network segments within the same administrative domain.

The internal architecture is following the SBA principles; SW components are connected to a common bus, they offer and/or consumes services communicating by an agreed Service Based Interface (SBI).

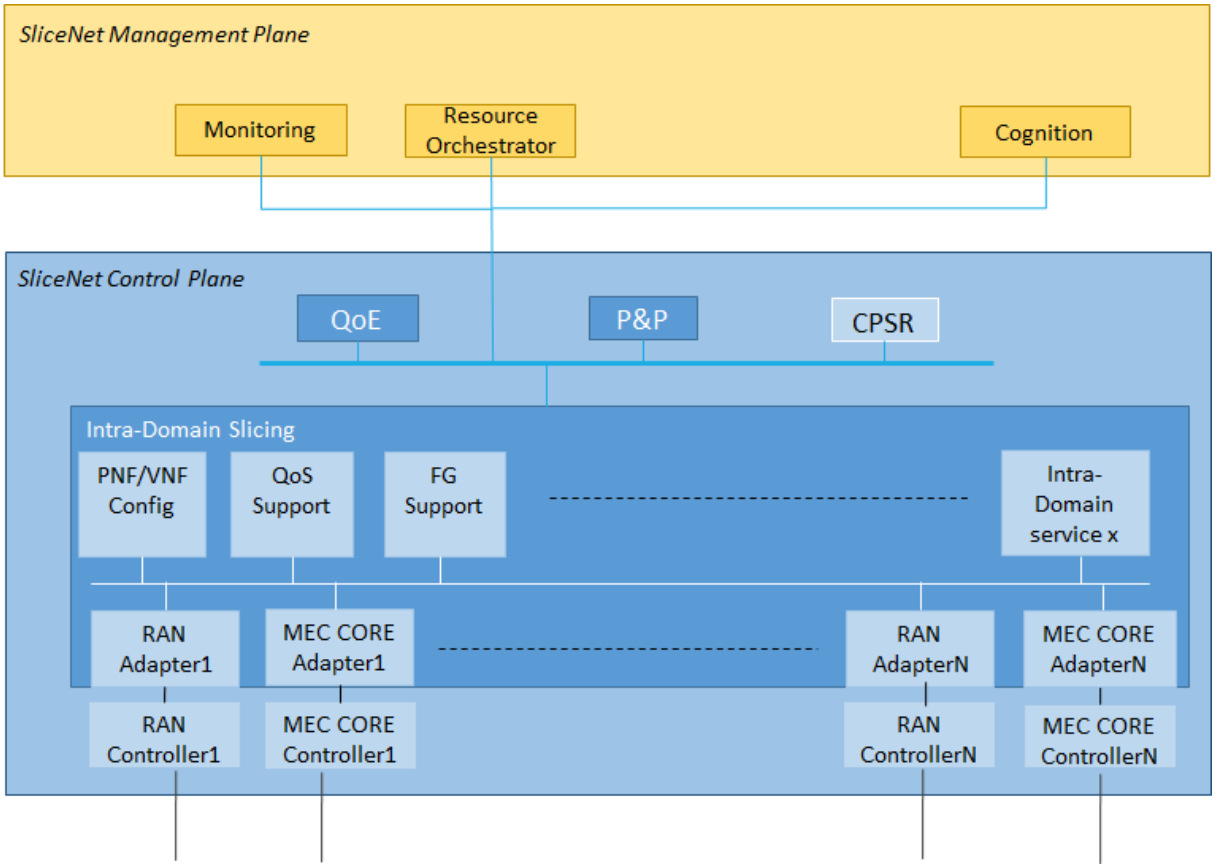

Figure 19: Intra-Domain Slicing, internal architecture

Functional decomposition:

- *PNFs/VNFs configuration support:*

  It provides a SliceNet centralized access to PNFs/VNFs application configuration (both initial and run time configuration) by offering a generic API abstracting the Network Element specific configuration interfaces.

- *Forwarding Graph control support:*

  It provides SliceNet centralized interface for completion of Forwarding Graph if needed (ex: for inter PoP connection or Service Function chaining)

  In order to deploy or update the concerned Forwarding Graph, it might interwork with the Resource Orchestrator (if the MANO has such capabilities) or with the appropriate underlying network Adapters e.g. SDN-Controllers.

- *Service QoS control support:*

It provides SliceNet centralized access to dynamic QoS setting per UE or UE-session.

- *RAN Adapters, MEC CORE adapters:*

  They are plugins that translates the high level abstracted request to the specific commands foreseen by the RAN and MEC-CORE controllers Northbound Interfaces.

  These plugins are specific for interfacing the underlying RAN and MEC-CORE 4G or 5G networks meaning they are vendors / technology dependant.

4.3.2 High level workflow diagrams

4.3.2.1 Configuration support

This workflow shows how the use case of PNFs/VNFs Configuration is handled by the "Intra-Domain Slicing" internal functional components.

Below the first PNFs/VNFs Configuration workflow is detailed but similar workflow applies in case of PNFs/VNFs Re-Configuration required by QoE or P&P during Slice run-time phases.

The first Slice PNFs/VNFs Configuration is required by the Resource Orchestrator after Slice instantiation and before Slice activation phases.

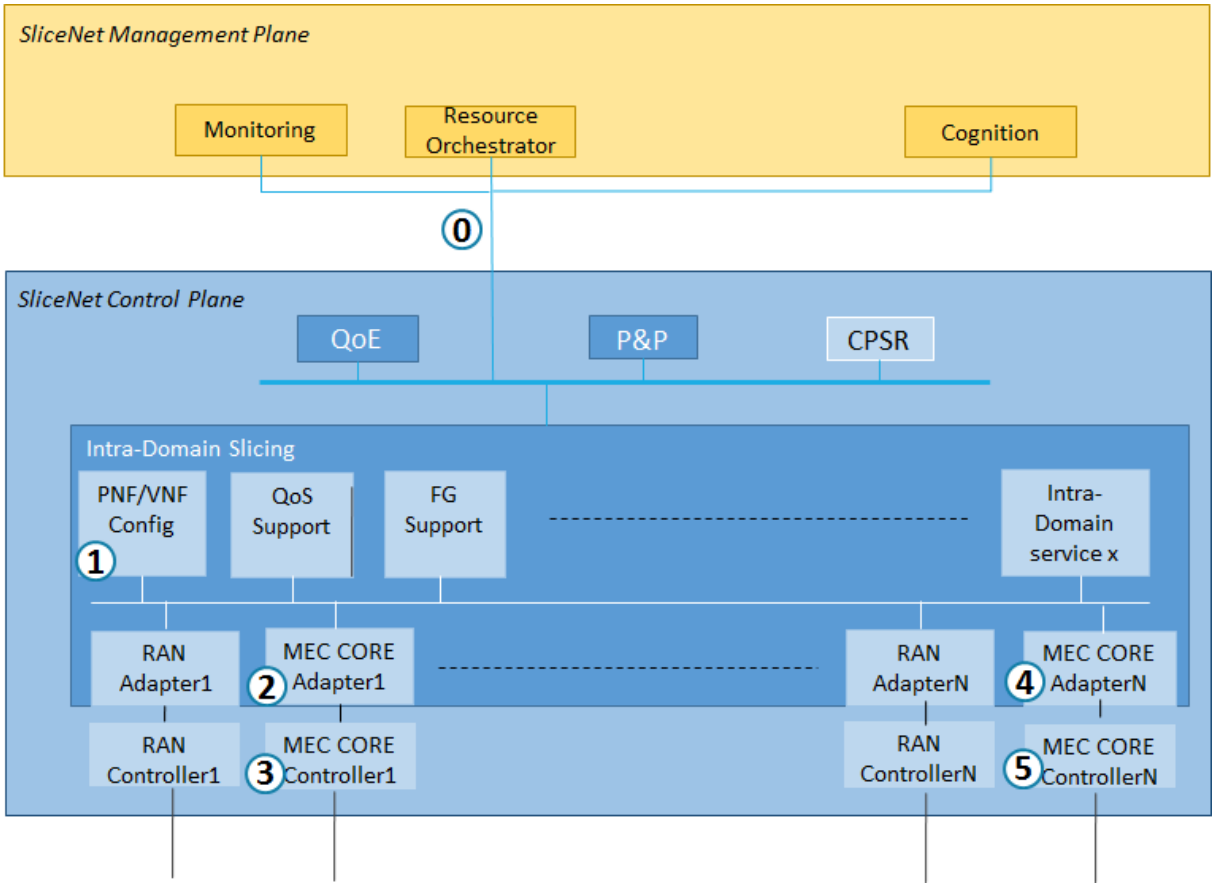

Figure 20: Intra-Domain Slicing, Configuration support

Table 4: Intra-Domain Slicing, Configuration support steps

Step	Description
0	Resource Orchestrator orders a VNF first configuration to SliceNet CP

1	PNF/VNF Config orchestrates the configuration operation and selects the concerned MEC CORE Adapter.
In case of MEC CORE type 1	
2	MEC CORE Adapter1 translates the high level abstracted request to the specific commands foreseen by the controller NBI.
3	MEC Core Controller enforces the PNF/VNF configuration towards the underlying MEC CORE Network.
In case of MEC CORE type N	
4	MEC CORE AdapterN translates the high level abstracted request to the specific commands foreseen by the controller NBI.
5	MEC Core Controller enforces the PNF/VNF configuration towards the underlying MEC CORE Network.

4.3.2.2 QoS control support

This workflow shows how the use case of QoS customization at run time is handled by the "Intra-Domain Slicing" internal functional components.

Below the QoS customization workflow is detailed in case it is required by QoE but similar workflow applies if the request comes from P&P or any other authorized SliceNet functional component.

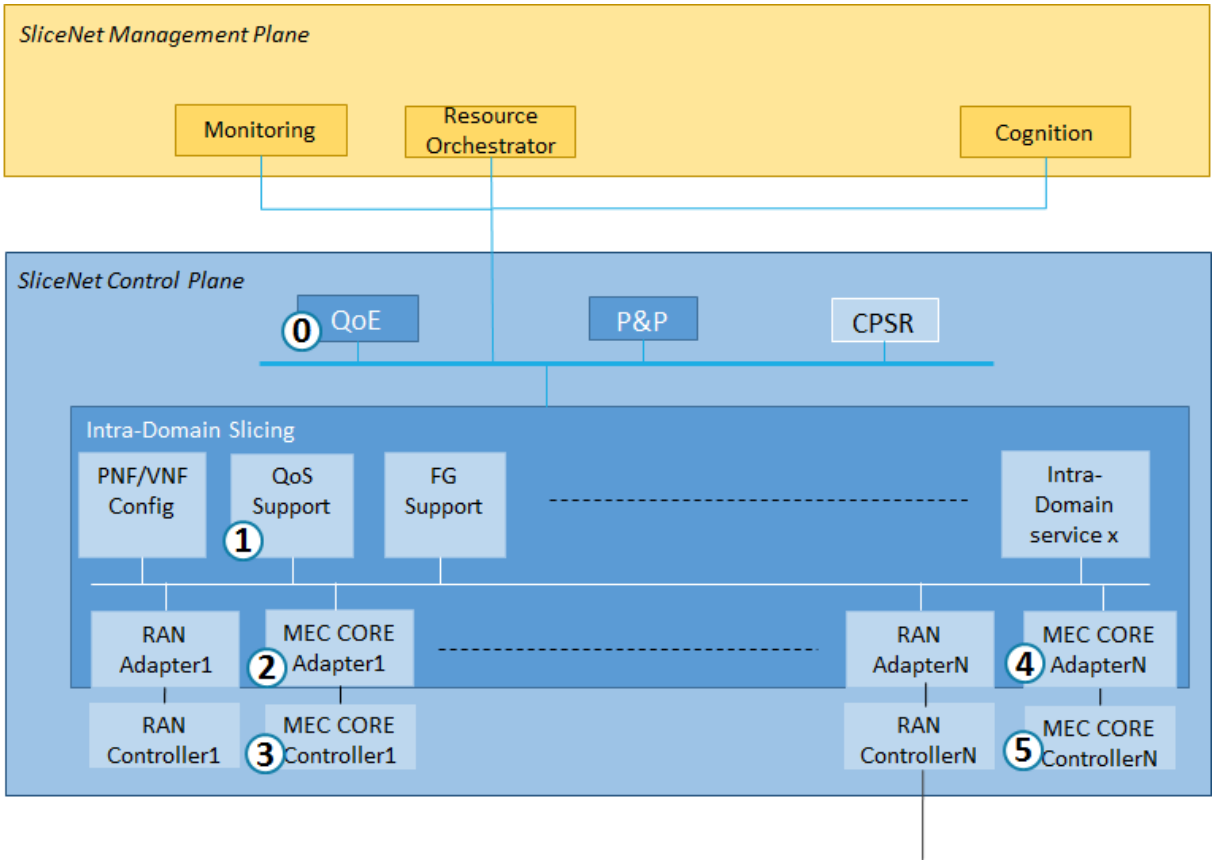

Figure 21: Intra-Domain Slicing, QoS support

Table 5: Intra-Domain Slicing, QoS support steps

Step	Description
0	QoE orders a QoS setting
1	QoS Support orchestrates the QoS operation and selects the concerned MEC CORE Adapter.

In case of MEC CORE type 1	
2	MEC CORE Adapter1 translates the high level abstracted request to the specific commands foreseen by the controller NBI.
3	MEC CORE Controller enforces the QoS setting towards the underlying MEC-CORE Network.
In case of MEC CORE type N	
4	MEC CORE AdapterN translates the high level abstracted request to the specific commands foreseen by the controller NBI.
5	MEC CORE Controller enforces the QoS setting towards the underlying MEC-CORE Network.

4.3.3 Forwarding Graph control support

This workflow shows how the use case for Forwarding Graph updating at run time is handled by the "Intra-Domain Slicing" internal functional components.

Below the FG updating workflow is detailed in case it is required by QoE or P&P at run time but similar workflow applies if the request comes from Resource Orchestrator or any other authorized SliceNet functional component.

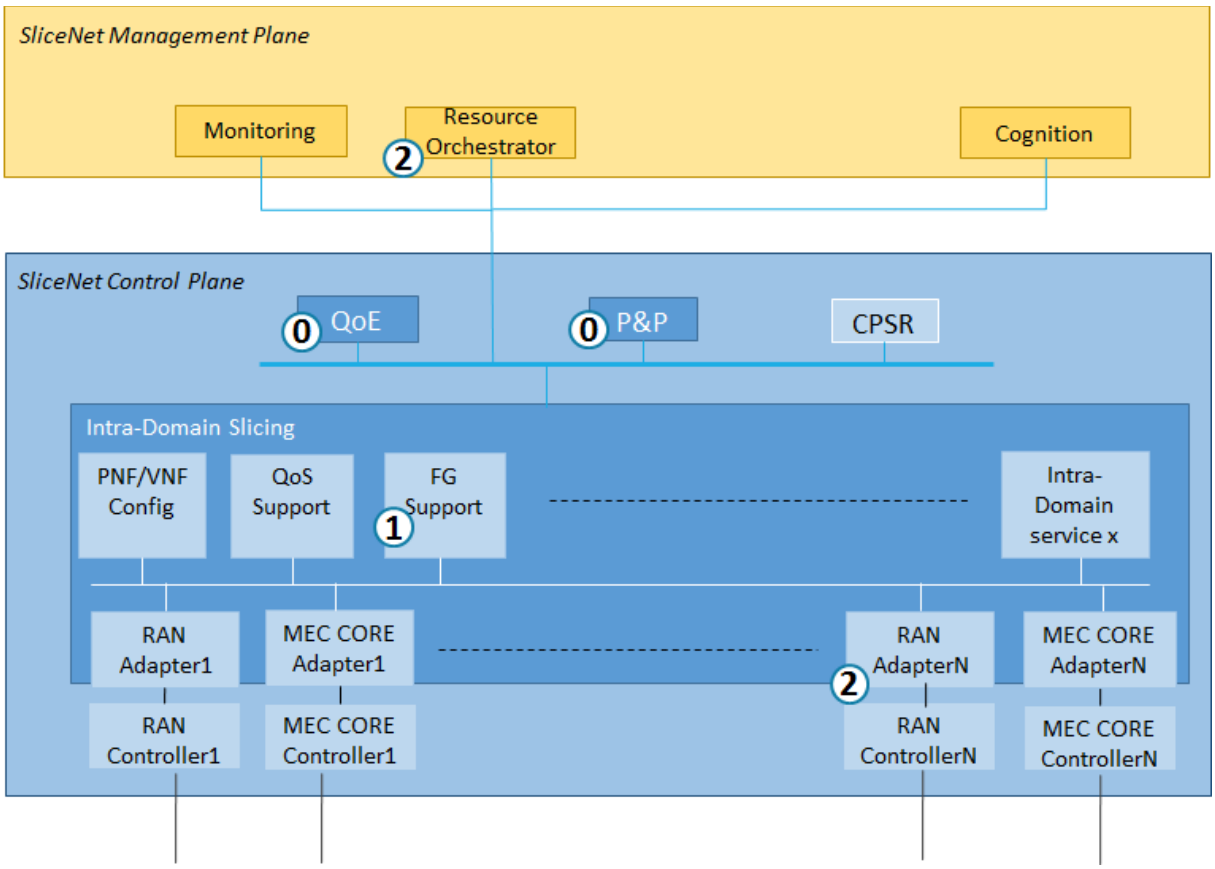

Figure 22: Intra-Domain Slicing, FG support

Table 6: Intra-Domain Slicing, FG support steps

Step	Description
0	QoE or P&P orders updating of a Forwarding Graph
1	Forwarding Graph Support has the control logic to translate the requested action into orchestrated steps towards Resource Orchestrator (if the connected MANO is capable of deploying FGs in multi PoP environment and/or to realize Service Function Chaining) or towards the underlying network controller by identifying the relevant Forwarding Graph

	templates
2	Resource Orchestrator or Network Controller receives the Forwarding Graph templates and enforce relevant configuration updating in the underlying infrastructure.

4.4 Intra-domain 5G-RAN core slicing

4.4.1 Logical Intra-domain 5G-RAN Core slicing functions decomposition

3GPP Next Generation Radio Access Network (NG-RAN) consists of New Radio (NR) gigabit Node-Bs (gNBs) and/or enhanced LTE Node-Bs (eNBs), providing the UP and the CP protocol terminations for the radio interfaces towards the user equipment (UE). GNBs and eNBs may be interconnected via an Xn interface or an Xx interface when involving LTE only eNBs. In addition, gNBs and eNBs are connected to the 5G CN (a.k.a. 5GC) via NG interfaces. More specifically, they are connected to the Access and Mobility Management Function (AMF) via the NG-C or N2 interfaces, and to the User Plane Functions (UPF) via the NG-U or N3 interfaces. Furthermore, a gNB is disaggregated into three modules, namely, (i) the Centralized Unit (CU), (ii) the Distributed Unit (DU) and (iii) the Remote Radio Unit (RRU). Therefore, a gNB may consist of a CU-CP and CU-UP, and one or more DUs connected to the CU via F1-C and F1-U interfaces for the CP and UP, respectively, with a flexible function split between CUs and DUs.

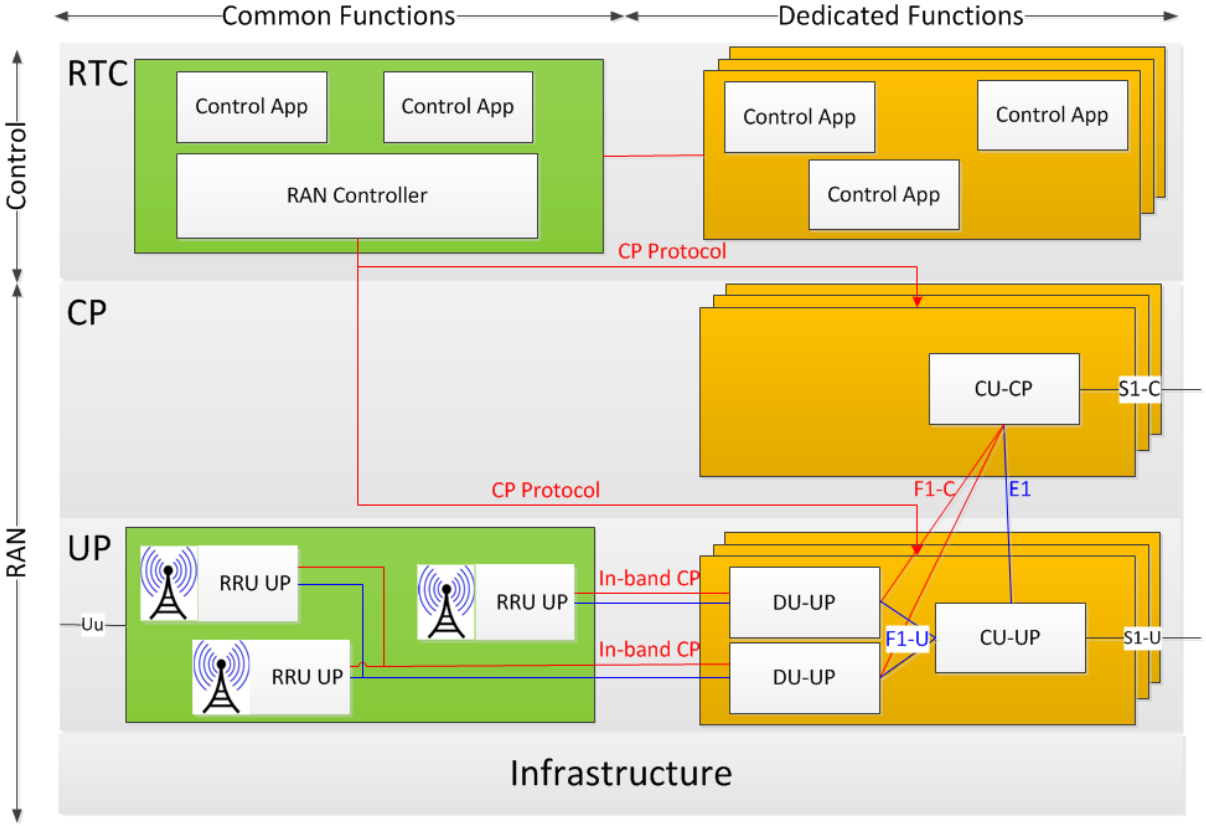

Figure 23: Intra-domain 5G RAN slicing function decomposition

Figure 23, above illustrates the intra-domain 5G RAN slicing function decomposition for both the UP and CP, as well as their interconnection with the Real-Time Control (RTC) plane. In Figure 23 we can observe three types of eNB/gNB function decomposition:

1. functional split between CP and UP;
2. functional split between CU and DU with an one-to-many (1:n) relationship;
3. functional split between DU and RRU, again with an one-to-many (1:m) relationship.

While the CU-CP manages the RAN CP, the CU/DU/RRU runtime reconfigurability and programmability is supported by the RAN controller, as well as the corresponding *Common* App and *Dedicated App* control applications in the RTC plane. Note that the RRU reconfigurability/reprogrammability is supported through an in-band mechanism by DU. Control applications are running on top of the RAN controller and can be either common such as a QoE optimizer or dedicated such as a handover logic that is only applied to a specific slice. At this point, we acknowledge the challenge of applying both a common and a dedicated application (or function, as discussed in the next paragraph), e.g. a common one parallel to a dedicated one (per slice) for UE handovers. In this case, there is a need to resolve conflicts between possibly "incompatible" common and dedicated control actions, for instance by defining priorities: a common handover policy function may prevail over the dedicated ones or alternatively can let them define some of the handover decision parameters.

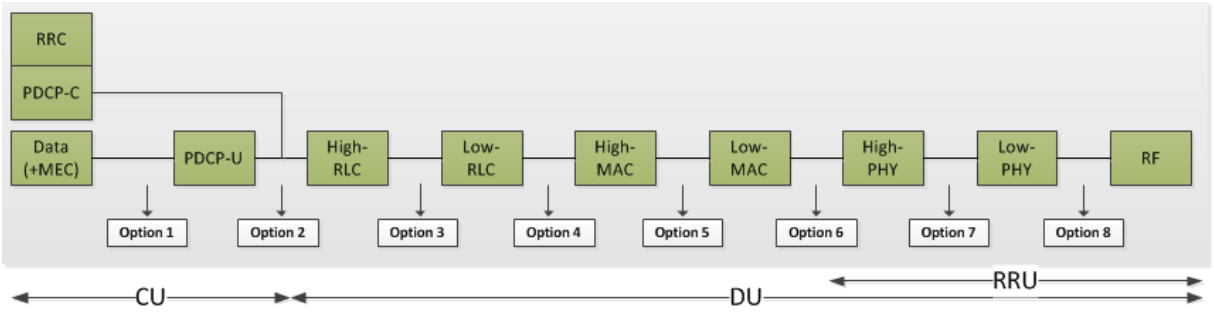

Figure 24: Functional split in 3-tier RAN

Regarding functional splitting, this is fixed between CU and DU and dynamic between DU and RRU, as shown in Figure 24 above. In particular, CU-CP includes RRC and PDCP-C functions, CU-UP performs GTP and PDCP-U functions, DU includes RLC, MAC and (optionally) low and high PHY functions depending on the split option with RRU.

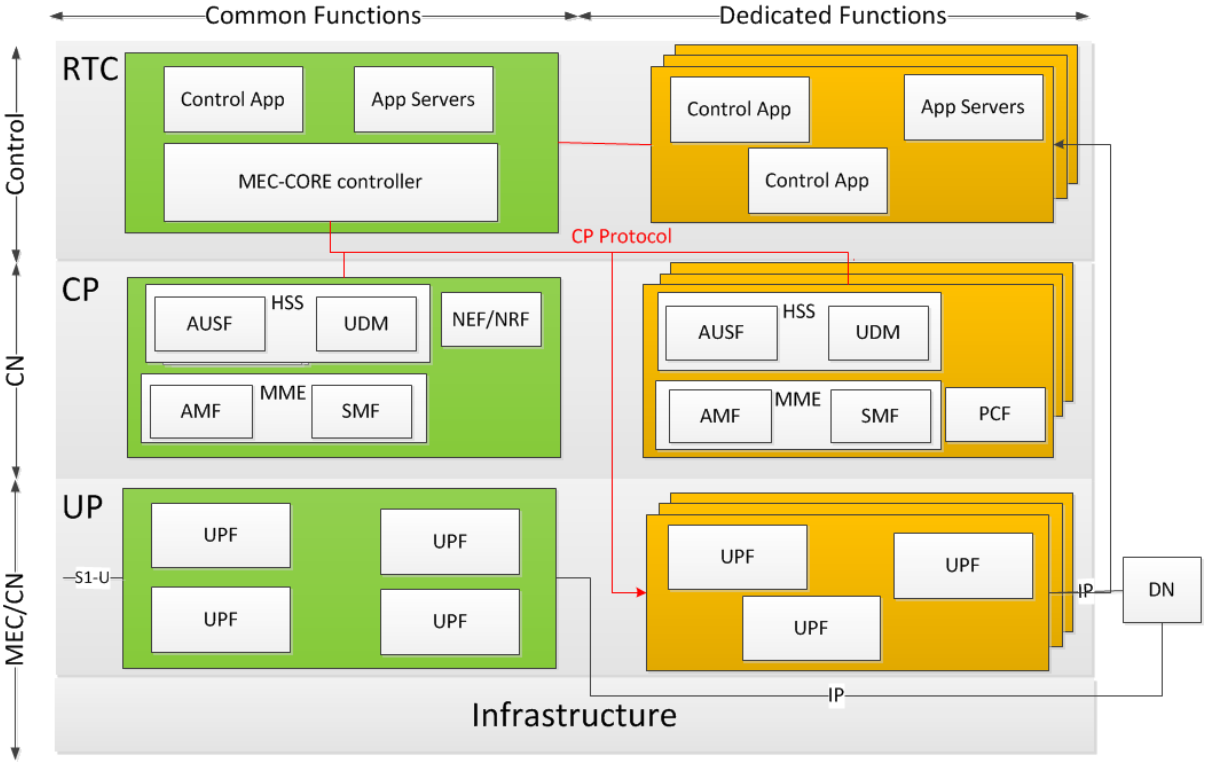

Figure 25: Intra-domain 5G MEC-CORE slicing function decomposition

Similar to the RAN case, Figure 25 above illustrates the intra-domain 5G MEC-CORE slicing function decomposition in UP and CP, as well as their interconnection with the RTC plane. As it can be seen in Figure 25, both CP and UP functions can either be *common* or *dedicated* depending on the slice description. The *MEC-CORE controller* sits on the top and is in charge of the data plane programmability in the segment among the RAN, edge, and Core networks on a per slice basis. This allows establishing the default and dedicated data path to the local app server located at the edge of network or to the default Data Network (DN). The UPF is in charge of managing user data and it is under the control of the Session Management Function (SMF). The AMF is in charge of managing the user mobility and access bearers through eNB/gNB. It is also in charge of authenticating the user through interaction with the Authentication Server Function (AUSF) and the Unified Data Management (UDM) component. The NF repository Function (NRF) is responsible for function discovery and communication. Note that AMF derives from MME, SMF from MME and -optionally- from SPGW-C. Also, the UPF derives from SP-GW, either includes both SP-GW-C+SP-GW-U or just SP-GW-U.

4.4.2 High level workflow diagrams

A UE is typically pre-configured with a given Network Slice Selection Assistance Information (NSSAI). This configuration can be done by the manufacturer-retail in collaboration with the service provider. This NSSAI provides a list of Single NSSAI (S-NSSAI), one per network slice and up to 8.

For each Slice request, the UE connects to the network as follows:

1. The desired NSSAI is provided to the RAN, which is also passed to the RTC, allowing the RAN to select the associated AMF (if any); otherwise the default AMF will handle the request.
2. The selected AMF (MME) checks with the UDM and AUSF (HSS in 4G) if this UE is allowed to use this NSSAI based on its current subscription.
    a. If it is allowed, then the AMF checks whether it can serve that NSSAI and all included S-NSSAI.
    b. If it is *not* allowed, then the AMF asks the NSSF to provide a proper AMF and passes the UE & NSSAI information to this new AMF
3. The associated AMF checks with the NRF the available SMF (4G MME and potentially part of SP-GW) instances to serve the different S-NSSAI
4. The associated AMF (MME) selects the returns a 5G-S-Temporary Mobile Subscription Identifier (5G-S-TMSI) to the UE to be included in any connection request for the RAN to route the network signalling for any application traffic to the corresponding UPF.

The network and the UE are now ready to receive some application traffic that will be carried through the selected slice thanks to the assigned SMF instances that will set the UPF (SPGW in 4G), accordingly.

4.5 Inter-domain slicing and service composition

4.5.1 Logical Inter-domain slicing functions decomposition

In a slice that involves multiple administrative domains, the slice can be divided into multiple intra-domain slices connected on the network level, with this element being the only one that actually connects the two domains.

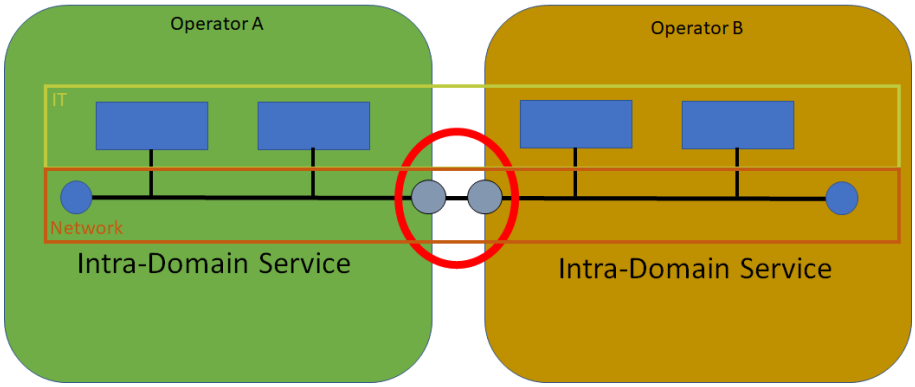

Figure 26: Inter-Domain Slice decomposition

To achieve this, an inter-domain slice must work like an extended intra-domain slice, with the extra configuration of the cross-domain connectivity. So for an inter-domain slice to exist, a local intra-domain slice must also exist.

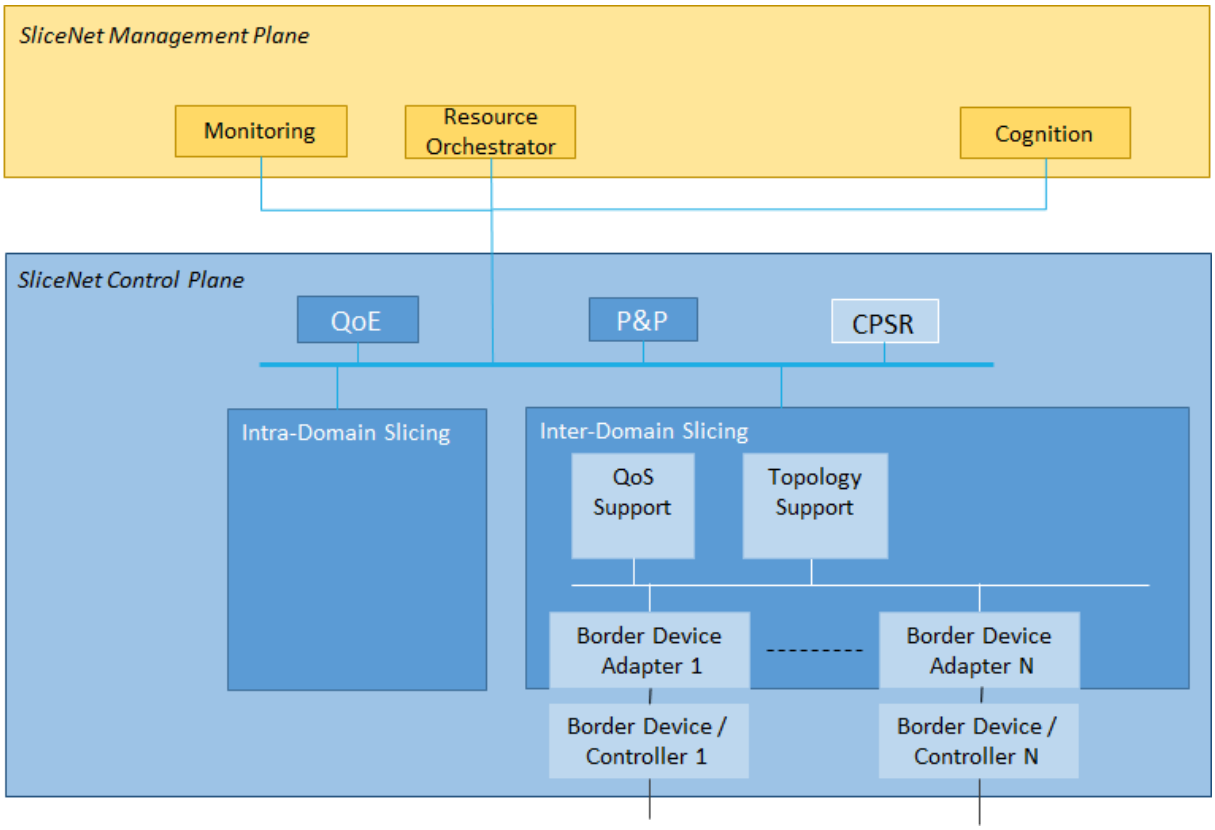

Figure 27: Inter-Domain Slicing internal architecture

The internal architecture for inter-domain slicing follows the same principles as the intra-domain slicing.

The main difference in the inter-domain is that it uses pre-established business relations between network operators and, as such, requires it to be less independent of the management plane, providing feedback for every and any operation being executed in order to maintain E2E slicing across domains. This communication and orchestration must be done on the management plane for security concerns, so that neighbour operators do not have direct access to sensitive border devices.

Functional decomposition:

- *Topology support:*

It provides a SliceNet centralized access to topology configuring via network policies for inter-domain virtual networks.

- *Service QoS control support:*

  It provides SliceNet centralized access to dynamic QoS setting.

- *Border Device Adapters:*

  These are plugins that translate the high level abstracted requests to the specific commands expected by either border devices or controllers of border devices in their Northbound Interfaces.

4.5.2 High level workflow diagrams

4.5.2.1 Virtual Topology Configuration

This workflow shows how the system can deploy cross-domain virtual topologies.

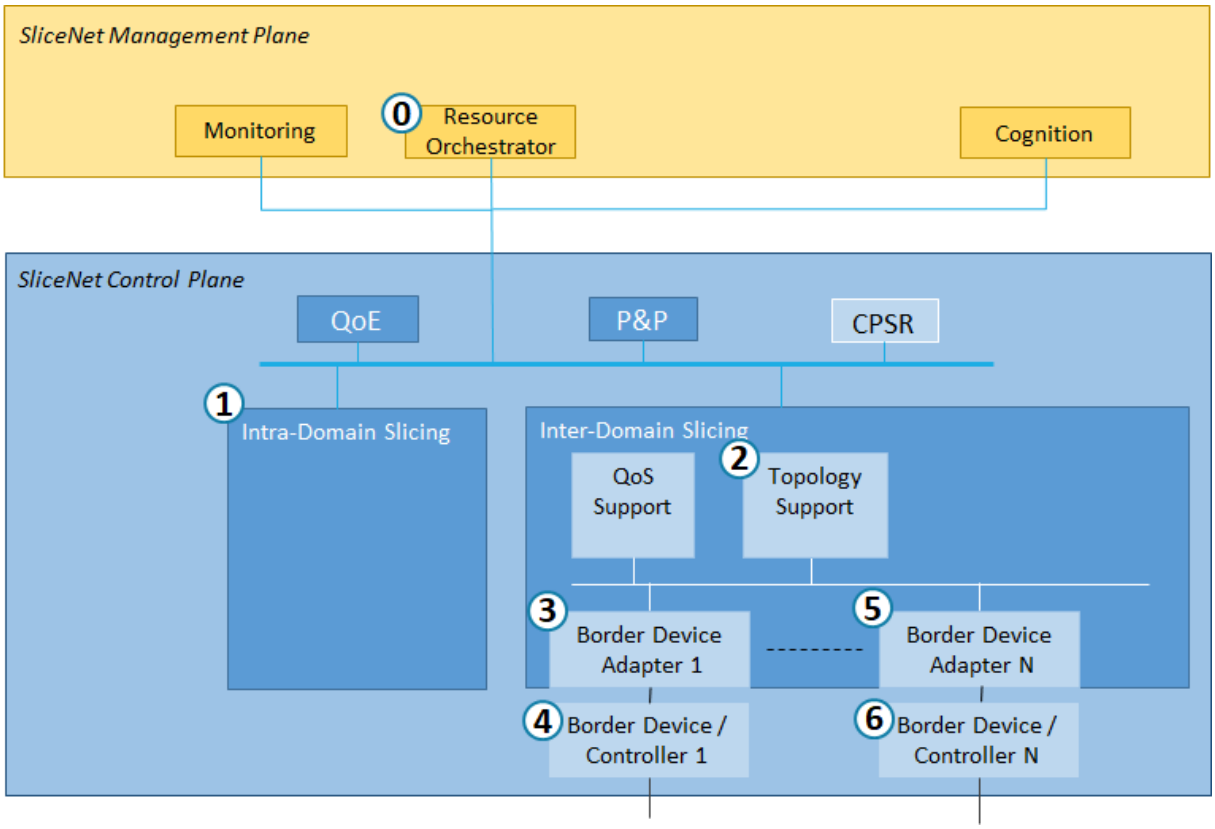

Figure 28: Inter-Domain Slicing virtual topology configuration

Table 7: Inter-Domain Slicing virtual topology configuration

Step	Description
0	Resource Orchestrator orders an inter-domain network topology to SliceNet CP
1	Intra-Domain Slicing orchestrates the internal Network Topology
2	Topology Support orchestrates the operation selecting the appropriate Adapter.
In case of Border Device type 1	
3	Border Device Adapter 1 translates the high level abstracted request to the specific commands foreseen by the device or network controller NBI.

4	Border Device or Controller enforces the policy settings towards the underlying Network.
In case of Border Device type N	
5	Border Device Adapter N translates the high level abstracted request to the specific commands foreseen by the device or network controller NBI.
6	Border Device or Controller enforces the policy settings towards the underlying Network.

4.5.2.2 QoS Support

In this workflow, the support for QoS settings for cross-domain inbound and outbound traffic are shown.

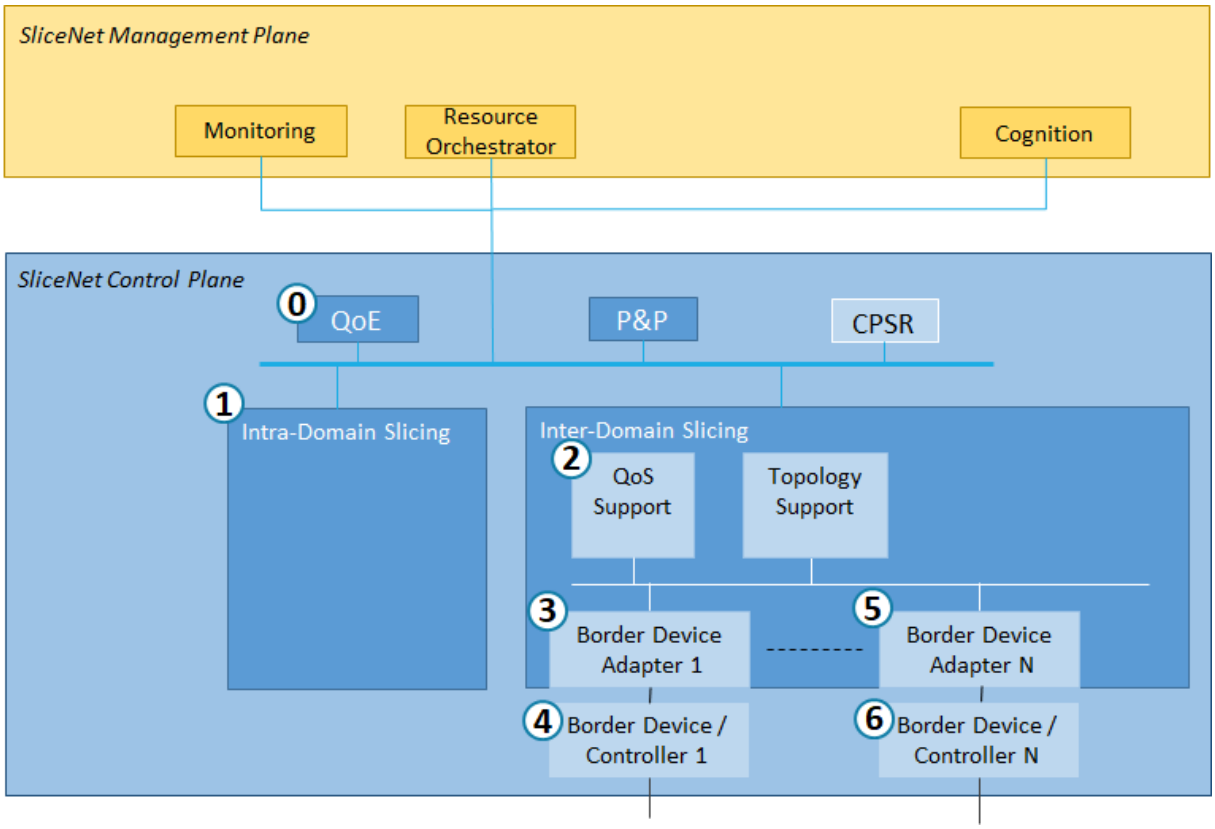

Figure 29: Inter-Domain Slicing QoS configuration

Table 8: Inter-Domain Slicing QoS configuration steps

Step	Description
0	QoE orders a QoS setting
1	QoE Operations are executed in the intra-domain part of the slice
2	QoS Support orchestrates the QoS operation and selects the appropriate Adapter
In case of Border Device type 1	
3	Border Device Adapter 1 translates the high level abstracted request to the specific commands foreseen by the device or network controller NBI.
4	Border Device or Controller enforces the QoS setting towards the underlying Network.
In case of Border Device type N	
5	Border Device Adapter N translates the high level abstracted request to the specific commands foreseen by the device or network controller NBI.
6	Border Device or Controller enforces the QoS setting towards the underlying Network.

5 SliceNet CP High Level APIs and interfaces

This section presents a high level description of the SliceNet CP interfaces and APIs, taking as reference the high level architecture split presented in section 3.3. The main goal is to highlight the available operations and the exchanged information across the different level of abstractions within the SliceNet CP. Indeed, following the abstraction principles described in section 3.3, three different macro-sets of APIs and interfaces are reported:

i) P&P APIs,
ii) Technology Agnostic APIs,
iii) Infrastructure Pillar Abstraction APIs.

These three sets of interfaces aim to define operations exposed at the different levels of the SliceNet CP.

Low level definition of these operations and translation into concrete APIs is already undergoing in WP4 and will be reported in related deliverables.

5.1 P&P APIs

The P&P APIs are conceived to provide the second level of abstraction in the SliceNet CP paradigm, thus exposing towards slice consumers and verticals a customized and dedicated per-slice view for their runtime control. The main goal, as detailed in section 4.1, is to enable specialization of slice instances according to verticals' requirements and slice provider's policies in terms of control exposure, by offering a limited and regulated set of control and management APIs to slice consumers.

The following tables summarize the high level description of the P&P interfaces exposed towards slice consumers and verticals, aiming to list the high level operations that could be exposed by P&P control instances. According to the customized view exposed to the specific slice consumer or vertical, not all these operations may be actually offered by the P&P.

Moreover, it is worth to mention that these operations are under a detailed definition process in WP4 in the related P&P activities, and actually they may map in multiple low level APIs. Also, further operations may be defined in the context of WP4 to fulfil P&P and vertical requirements at large.

Operation	**Collect Slice Monitoring Metrics**
Callers	Slice consumer / Vertical
Description	This operation allows verticals and slice consumers at large to monitor their slice related metrics and KPIs, where applicable at multiple granularities (i.e. at resource, network function or application, slice, service levels) according to the exposure allowed by the slice provider. This operation enables verticals and slice consumers to apply their own analytics processes and possibly leverage other P&P for slice runtime control and adaptation. Minimum set of required inputs to be passed for this operation includes the identification of the slice instance and the entities for which retrieve monitoring information. Filtering of specific metrics and KPIs may also be possible.

Operation	**Reconfigure Slice Performances**
Callers	Slice consumer / Vertical
Description	This operation offers the possibility to verticals and slice consumers of requesting upgrades or downgrades in the performances of their slices. In this case, performance may have different declinations, from QoS attributes and parameters at either slice or resource/network function levels, as well as expression of more declarative QoE Different options may be available and exposed to verticals and slice consumers according to the level of exposure agreed with the slice provider. Indeed, slice provider may offer a limited set of QoS levels for the vertical (or slice consumer) to choose and select at runtime as a mean for dynamic slice adaptation. This may be at the level of whole slice, as well as specific for network functions (e.g. VNFs or MEC applications) and SDN network resources and applications. On the other hand, the vertical (or slice consumer) may also exploit this interface for informing or expressing declarative QoE requirements. Minimum set of required attributes to be specified by verticals and slice consumers for this operation include the identification of the slice instance and the expression of QoS/QoE to be upgraded or downgraded according to the options described above.

Operation	**Deploy new Customized (Network) Function**
Callers	Slice consumer / Vertical
Description	This operation allows the vertical (or the slice consumer) to request specific lifecycle management actions for some network functions (e.g. one or more vertical custom VNF or MEC application), in the specific case to trigger the deployment of one or more new instances to be composed in the E2E slice. According to the customized vertical slice view exposed by the slice provider, different options may be available to specialize the deployment. For example, the vertical may be allowed to request the deployment of a new VNF or MEC application indicating location, forwarding graph and QoS related constraints. Minimum set of required attributes to be specified by verticals and slice consumers for this operation include the identification of the slice instance and the Network Function type to be deployed (corresponding to an already catalogued VNF or MEC application). Additional constraints may be included as described above.

Operation	**Configure Customized (Network) Function**
Callers	Slice consumer / Vertical
Description	This operation allows the vertical to directly access a limited set of the network functions instances (e.g. one or more VNF instances or MEC applications) composing the E2E slice offering the possibility to apply specific configurations according to vertical logics and requirements. Following the SliceNet CP abstraction principles, this operation aims to hide any technology and implementation detail for the network function configuration logic, leveraging on the first level of abstraction provided by the underlying Technology Agnostic APIs (and the related SliceNet CP control services). Minimum set of required attributes to be specified by verticals and slice consumers for this operation include the identification of the slice instance and the VNF or MEC application instances to be reconfigured, along with the set of configuration parameters to be enforced. If applicable the P&P may restrict and provide a limited set of options for configuration attributes to be validated and enforced in the given Network Function instances.

Operation	**Scale Customized (Network) Function**
Callers	Slice consumer / Vertical
Description	This operation offers the possibility to have access to further lifecycle management actions for customized vertical Network Functions (i.e. VNFs and MEC applications mostly), in the specific case to trigger the scale of a well defined set of instances types composing an E2E slice. Different options may be exposed to verticals and slice consumers to specify scaling constraints, like explicit number of instances to be added or removed (optionally with location constraint also), as well as based on more NFV-like Network Function specific deployment flavors that may involve additional constraints for QoS and forwarding graph. Minimum set of required attributes to be specified by verticals and slice consumers for this operation include the identification of the slice instance and the VNF or MEC application instances to be scaled, augmented with optional constraints as described above.

Operation	**Scale Slice to Different Flavor or Profile**
Callers	Slice consumer / Vertical
Description	This operation allows verticals and slice consumers to dynamically upgrade and downgrade the overall slice to a new profile. This has to be considered as a slice-level operation involving the combination of heterogeneous slice aspects and building blocks, from individual vertical custom Network Functions, to QoS and forwarding graph. Therefore, it leverages on the whole set of Technology Agnostic APIs exposed by the SliceNet CP control services for runtime slice control and adaptations, in terms of forwarding graph, QoS control, VNF/PNF/NF configuration, etc. Minimum set of required attributes to be specified for this operation include the identification of the slice instance and the expression of the new slice profile or deployment flavor for upgrading or downgrading the slice.

5.2 Technology agnostic APIs

The technology agnostic APIs provide the first level of **slice control abstraction** in the SliceNet CP approach. In particular, exploiting the common information model and control logics provided by the southbound SliceNet CP plugins and adapters, the technology agnostic APIs hide the specific control technology and implementation details exposed by each infrastructure pillar, and provide a slice view to the control operations offered.

In practice, the technology agnostic APIs aggregate the whole set of interfaces and operations exposed by the SliceNet CP services and provide a slice context towards other SliceNet components.

The following tables provide an high level view of the SliceNet CP technology agnostic exposed interfaces, in terms of operations offered to apply specific slice aware control logics as described in section 3.3. The tables are grouped following the SliceNet CP control services identified in section 3.3, and have to be considered as a preliminary set of slice control operations offered by the SliceNet CP that are already under a more detailed definition and consolidation process in the context of WP4 activities. Following the service based approach, it is possible that an operation offered by a given SliceNet CP control service (and described in a table below) could be invoked by another SliceNet CP control service. Moreover, since the whole set of technology agnostic APIs are conceived to offer slice-level control logics, each of the operation described below has to be considered as per slice or subslice (i.e. an explicit slice context or identification is required).

SliceNet CP control services are processing the operation requests according to internal workflows and the various execution steps may involve a combination of service invocations that are offered through the Infrastructure Pillar's abstractions. The combination of execution steps depends on the slice/subslice composition with respect to the actual allocation of infrastructure resources to each particular slice/subslice. For example a QoS tuning request may result in a forwarding rule update request to be directed to the backhaul abstraction, whereas in other cases the same request may result in an access network allocation request. The overall decision is subject to a combination of metrics and conditions detected in the actual Data Plane resources.

5.2.1 QoS Control Interface

Operation	**Configure Session QoS Parameter**
Callers	P&P, Inter domain control service, Slice Orchestration, QoE Optimizer
Description	This operation is used to configure a QoS parameter for the user plane sessions that are served by a slice/subslice within an administrative domain. The indicated quality level should be ensured for all user plane communications that are served by the infrastructure resources under the control of the specific domain's CP. Additionally, the requested level should be also provided across the entire domain. For this purpose the QoS service should propagate, by adapting and following the appropriate information model and semantics, the required quality level to all the services involved in the slice support across pillars.

Operation	**Configure Session Priority**
Callers	P&P, Inter domain control service, Slice Orchestration, QoE Optimizer
Description	The purpose of this operation is to configure the priority constraints pertaining to user plane differentiation among active sessions. Priority based decisions might be practiced at a number of control points within the infrastructure. These points should be discoverable and reachable by the QoS control service via the Service Based Architecture so that priority constraints should be appropriately propagated.

5.2.2 FGE Interface

Operation	**Provision Forwarding Graph**
Callers	Slice Orchestration, Resource Orchestration
Description	This operation allows to enforce a full forwarding graph to properly interconnect the whole set of network functions (VNFs, PNFs, NFs, MEC Apps, etc) in the slice, considered at this FGE level as generic "nodes" in a graph. Whenever required, e.g. if the SliceNet resource orchestration components (NFV and MEC orchestrators) cannot enforce the slice forwarding graph themselves through dedicated interaction with VIMs and NFV infrastructures, the FGE can offer this operation to provision the slice forwarding graph by interacting with one or more infrastructure pillar adapters at the SliceNet CP southbound to enforce the proper forwarding rules.

Operation	**Add nodes to forwarding graph**
Callers	P&P, Slice Orchestration, Resource Orchestration (e.g. NFVO), Cognitive Management
Description	This operation has to be used to apply changes to a slice forwarding graph, i.e. by offering the possibility to add a new node (i.e. a new network function) in the graph. For example, this may be required in the context of an overall slice scale out operation, where one or more new network functions are deployed in the context of the E2E slice (e.g. VNFs, MEC Apps, etc.)

Operation	**Remove nodes from forwarding graph**
Callers	P&P, Slice Orchestration, Resource Orchestration (e.g. NFVO), Cognitive Management
Description	This operation is counterpart of the above one for adding new nodes in the forwarding graph. It allows to remove one or more nodes form the graph. For example, this may be required in the context of an overall slice scale in operation, where one or more new network functions have to be decommissioned in the context of the E2E slice (e.g. VNFs, MEC Apps, etc.)

Operation	**Provision link in forwarding graph**
Callers	P&P, Slice Orchestration, Resource Orchestration (e.g. NFVO)
Description	Similarly to the two previous operations, this one is conceived to apply changes to a slice forwarding graph, i.e. by offering the possibility to provision a new interconnection (i.e. a new virtual/logic link) in the graph among two or more nodes. This may be required when a slice scale out operation requires (following the NFV terminology) the migration to a new deployment flavor where existing nodes (i.e. network functions) have to be interconnected following a different pattern.

Operation	**Remove link in forwarding graph**
Callers	P&P, Slice Orchestration, Resource Orchestration (e.g. NFVO)
Description	This operation allows to dynamically update a slice forwarding graph by removing an existing interconnection among two or more nodes in the graph. This may be required upon a slice scale in operation where the rollback to a lower slice profile or deployment flavor implies the deletion of one or more virtual/logical links among slice network functions.

Operation	**Configure routing scheme**
Callers	P&P, Slice Orchestration, Resource Orchestration (e.g. NFVO), Cognitive Management
Description	This operation allows to enforce a specific routing scheme within the slice forwarding graph, e.g. unicast, multicast and broadcast. This operation can also be used to update the routing scheme within existing forwarding graphs. Where applicable, this operation may be supported by the SliceNet CP as part of others described in this table for the FGE component (e.g. “provision forwarding graph”)

Operation	**Delete forwarding graph**
Callers	Slice Orchestration, Resource Orchestration (e.g. NFVO)
Description	This operation is the counterpart of the “Provision forwarding graph” one and it is conceived to be used to delete a whole slice forwarding graph, thus coordinating the interactions with the involved infrastructure pillar adapters to decommission the related forwarding rules.

5.2.3 VNF/PNF/NF Configuration Interface

Operation	**Add VNF/PNF/NF Configuration**
Callers	P&P, Slice Orchestration, QoE Optimizer

Description	This operation is intended to offer the callers with a common primitive for applying a control configuration to one or more VNF/PNF/NFs running in the context of a slice. The operation has to be considered as technology and implementation agnostic, thus not referring to any particular control protocol or logic to apply the given configuration. The SliceNet CP is in charge to derive the proper infrastructure pillar adapters to be exploited and augment the attributes received as inputs with further information collected from the SliceNet platform. Indeed, depending on the specific VNF/PNF/NF involved in the configuration action, different logics may be required to be applied (e.g. a configuration of a vertical customized MEC application vs. a configuration of a new policy in the 5G PCF NF). Minimum set of required information to be passed includes the identification of the VNF/PNF/NF and the set of configuration attributes (e.g. in the form of key/value tuples).

Operation	**Update VNF/PNF/NF Configuration**
Callers	P&P, Slice Orchestration, Cognitive management
Description	This operation allows to dynamically modify an existing and previously enforced configuration in one or more VNF/PNF/NFs running in the context of a slice. It follows the same technology agnostic principles of the "Add VNF/PNF/NF Configuration" operation and requires a stateful approach for the management of VNF/PNF/NF configurations within the SliceNet CP and its control services. Minimum set of required information to be passed includes the identification of the VNF/PNF/NF and the configuration to be updated together with the new set of configuration attributes (e.g. in the form of key/value tuples).

Operation	**Rollback VNF/PNF/NF configuration**
Callers	P&P, Slice Orchestration, Cognitive management
Description	This operation allows to delete an existing and previously enforced configuration in one or more VNF/PNF/NFs running in the context of a slice. As the other two above operations, it follows the technology agnostic principles, and it has to be used when during the lifecycle of a slice instance, a given (or a set of) VNF/PNF/NF configuration(s) has to be rollbacked. This operation may be invoked as a result of the execution of the cognitive loops in response of some specific slice behavior or status (e.g. deletion of a given policy in a 5G PCF NF), as well as during the decommissioning of a slice (or subslice).

	Minimum set of information required as input to perform this operation includes the identification of the VNF/PNF/NF and the configuration to be rollbacked.

5.2.4 Data Plane Programmability Interface

Operation	**Set flow priority**
Callers	P&P, Inter domain control service, Slice Orchestration, Cognitive management, QoS service
Description	The operation is expected to prioritise certain flows with respect to the way data plane functions manipulate the user plane flows. Applicability of the rule is valid as long as the flow is active.

Operation	**Set session priority**
Callers	P&P, Inter domain control service, Slice Orchestration, Cognitive management, QoS service
Description	The operation is expected to prioritise certain sessions with respect to the way data plane accommodates user sessions. Applicability of the rule is valid as long as the session is active. This operation is expected to affect all the flows activated in the context of a user session.

Operation	**Set flow acceleration**
Callers	P&P, Inter domain control service, Slice Orchestration, Cognitive management, QoS service
Description	The operation is expected to accelerate the processing of certain flows within the forwarding graphs. Applicability of the rule is valid as long as the flow is active.

Operation	**Set session acceleration**
Callers	P&P, Inter domain control service, Slice Orchestration, Cognitive management, QoS service

Description	The operation is expected to accelerate the processing of all the flows activated in the context of certain sessions within the forwarding graphs. Applicability of the rule is valid as long as the session is active.

5.2.5 Inter Domain Interface

Operation	**Interconnect Slice to Peering Domain(s)**
Callers	Slice Orchestration
Description	This operation allows to enforce the interconnection of single-domain slices (e.g. per-administrative-domain slice subnets) to E2E slice instances. It is therefore the operation exposed by the SliceNet CP to be invoked when an E2E slice spans across multiple administrative domains, specifically to properly configure border devices of the administrative domain under the ownership of the SliceNet CP. This can translate into applying proper slice traffic tagging or encapsulation depending on the given E2E slice requirements, where applicable following SDN principles. Minimum set of information required as input to perform this operation includes the identification of the set of peering domains to interconnect to (e.g. as a list of remote endpoints) and the traffic tag/encapsulation to be adopted. The resolution of proper border devices to be properly configured is performed leveraging on topology related functions embedded in the SliceNet CP, as described in section 4.4. As an option, this operation may offer the possibility to directly enforce specific QoS policies at border devices in support of dedicated per-slice SLAs agreed across providers.

Operation	**Modify Slice Interconnection to Peering Domain(s)**
Callers	Slice Orchestration, Cognitive management
Description	This operation allows the modification, addition or removal of interconnection parameters such as traffic tagging, encapsulation and QoS of an interconnected slice. The tuning of QoS parameters in the context of multi-domain E2E slices, in particular, allows application and enforcement of policies when it is required to dynamically update (e.g. upon trigger from cognitive functions) per-slice QoS rules and policies at the border of administrative domains with the aim of fulfil SLA agreements with peering domains. Minimum set of information required as input to perform this operation includes the interconnection parameters to modify, such as updated QoS rules and policies towards specific peering domains (identified with a list of remote endpoints). As

	for the above operation, the resolution of proper border devices where to apply the new parameters is performed leveraging on topology related functions provided in the SliceNet CP.

Operation	**Remove Slice Interconnection to Peering Domain(s)**
Callers	Slice Orchestration
Description	This operation allows the removal of an established slice interconnection to a peering point, clearing any parameters set on installation or modification operations. Minimum set of information required as input to perform this operation includes the slice interconnection to terminate.

5.3 Infrastructure Pillar Abstraction APIs

Pillar abstraction interfaces and APIs intent to provide a homogenised layer over the main segments of an underlying network. The purpose of the abstraction supports the possibility of using different technologies per domain and also within the same domain. At this level the overall slice/subslice context is expanded into more technology/segment related contexts able to support overall slice/subslice provision. Each of the technological pillars of the infrastructure is supporting a common set of operations and semantics identified for the particular segment. The specialisation is applied thereafter by software artifacts instantiated in the form of adapters that hide both the implementation details and the actual technology used.

5.3.1 Access, Edge, and Core Network Control Interface

Operation	**Collect Cell/Slice/User Monitoring Metrics**
Callers	SliceNet CP control services, Slice Orchestrator
Description	This operation enables to retrieve the monitoring information with the desired level of granularity form the underlying 5G infrastructure.

Operation	**ReConfigure Cell/Slice/UE**
Callers	SliceNet CP control services, Slice Orchestrator
Description	This operation enables to reconfigure a cell/slice/UE given their respective

	identifier.

Operation	**Allocate access resources to slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	This operation enables the allocation of (radio) access resources to a slice. The pool of resources are to be used when a UE is attached to the slice and it is utilizing the related slice services. The operation is iteratively applied across all the instances of the Access Network based also on any geographical constraints to allow for complete or partial coverage.

Operation	**Reconfigure basic RAN and Core services**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation intends to configure (R)AN with respect to basic core services (e.g. MME or AMF/SMF) that will be required to resolve the details that are required for the management of user session and their allocation to subslices.

Operation	**Apply QoS constraints**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation will allow the indication of quantitative parameters to be enforced for adequate servicing of user sessions in the access and core networks. These parameters are expected to be considered in the process of access network resource planning (e.g. utilisation of radio channels/bearers, reservation, scheduling, packet marking, etc.)

Operation	**Reconfigure MEC Breakout**
Callers	SliceNet CP control services, Slice Orchestrator

Description	The operation is used in case MEC Applications require traffic redirection of data plane sessions from access network to a service hosted by MEC. E.g a PDU session may have to be terminated at a PDU Session Anchor in the same PoP with the RAN functions potentially by bypassing all the relay processing that occurs in typical PDU sessions.

5.3.2 MEC Control Interface

Operation	**Allocate MEC Resources**
Callers	SliceNet CP control services, Slice Orchestrator
Description	Whenever MEC applications have to be deployed this operation is used to request allocation of computing and networking resources that are required to host the intended applications. The operation indicates the edges where the resources have to be allocated and the performance attributes that are required.

Operation	**Instantiate MEC Applications**
Callers	SliceNet CP control services, Slice Orchestrator
Description	Previously allocated MEC resources are utilised for the provision of MEC applications instances.

Operation	**Reconfigure MEC Applications**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation aims at applying specific configuration set to active MEC applications.

5.3.3 Backhaul Control Interface

Operation	**Provision SDN intent**
Callers	SliceNet CP control services, Resource Orchestration (e.g. NFVO)

Description	This operation offers dedicated SDN primitives following an intent-based approach, thus it simplifies the interaction between the SliceNet CP control services (and possibly Resource Orchestration components, where applicable) and the SDN controllers towards a declarative control operations. In practice, an SDN intent may be expressed as a connectivity service between two or more endpoints with given QoS requirements (e.g. a VPN, a MPLS tunnel, etc.) and with a given routing scheme, as well as an high level firewall rule to block specific user traffic at a border of a network domain. This operation therefore exposes what to do on top of the SDN controllers, without specifying how to do it. In this case, following the SliceNet CP abstraction principles, the underlying Backhaul/Transport SDN controller are abstracted as black boxes for slice related QoS-aware resource allocation with isolation in support of the SliceNet CP control services logics.

Operation	**Remove SDN intent**
Callers	SliceNet CP control services, Resource Orchestration (e.g. NFVO)
Description	This operation is the counterpart of the above “Provision SDN intent”. It follows the same declarative approach and it has to be used to enforce the decommissioning of a given SDN intent by means of the underlying Backhaul/Transport SDN controller. The decommissioning logic is mandated to the SDN controllers, aiming to reduce conflict at the SliceNet CP control services level and increase the success of the requests.

Operation	**Provision QoS-enabled SDN Forwarding Rules**
Callers	SliceNet CP control services, Cognitive/QoE control services
Description	This operation allows to directly enforce QoS-enabled SDN forwarding rules in the data plane and thus apply dedicated per-flow SDN rules (e.g. mirror, forward, drop, etc. with QoS attributes). It has to be used whenever there is the need to apply per-network-device (either physical, logical or virtual) SDN forwarding rules, as alternative of the intent based approach implemented by the two above operations. In particular, this operation, and the related primitives/APIs can be exploited by several SliceNet CP control services, e.g. the FGE in the context of forwarding graph adaptations (i.e. add/remove nodes and links). Following the SliceNet abstraction principles, this operation is conceived to hide the specific technology implementation details of the underlying SDN controller(s), and expose a common model for QoS-enabled SDN forwarding rules that can be later translated by the specific infrastructure pillar adapters into custom and

	specific logics and protocols.

Operation	**Update QoS-enabled SDN Forwarding Rules**
Callers	SliceNet CP control services, Cognitive/QoE control services
Description	This operation allows to dynamically update and modify existing QoS-enabled SDN forwarding rules. It follows the same abstraction principles described for the operation above, and it is mostly conceived to enable dynamic updates of per-flow SDN rules (e.g. for QoS upgrade/downgrade) according to outcomes produced by Cognitive/QoE control functions and components in the analysis of heterogeneous slice monitoring metrics.

Operation	**Remove QoS-enabled SDN Forwarding Rules**
Callers	SliceNet CP control services, Cognitive/QoE control services
Description	This operation enables the decommission of one or more SDN forwarding rules previously provisioned. The operation follows the SliceNet CP abstraction principles described for the above SDN forwarding rules operation and is an alternative of the “Remove SDN Intent” operation, being it applicable on a per-flow and per-network-device basis.

Operation	**SDN Topology Exposure**
Callers	SliceNet CP control services, Resource Orchestration (e.g. NFVO)
Description	This operation is conceived to expose a common logical view of the backhaul/transport network segments topology to other SliceNet CP control services (and if applicable to Resource Orchestration ones). Following the SliceNet CP abstraction principles, the SDN topology view offered through this operation has to be considered as independent from specific SDN controller(s) technologies and models, and could be for example based on a network graph approach.

5.3.4 Core Control Interface

Operation	**Add Core Function Instance to Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	For an existing NF that has been provisioned, the operation registers it as a an NF that can be used in the slice procedures for CP or DP. This in the case of 4G core will be affecting MME whereas in 5G it will have an impact of AMF/SMF

Operation	**Configure UE Set for Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation is used to register the UE identifiers (e.g. SIM, IMSI) of the equipment units that have to be attached to a slice.

Operation	**Configure Traffic Classification per Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation intends to configure the classification to be applied to user data traffic within a slice. The classification will be used for the proper routing of user data.

Operation	**Apply UE Addressing scheme for Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation is used to define how IP addresses will be allocated to UEs that are attached to a Slice (static IPs, etc.)

Operation	**Enable broadcasting support for a Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	This operation indicates the support of distribution of broadcasted information to and from slice endpoints.

Operation	**Enable direct point to point communications support for a Slice**
Callers	SliceNet CP control services, Slice Orchestrator
Description	This operation will allow direct communication among endpoints in a slice avoiding as much as possible excessive relaying an encapsulation.

Operation	**Get Sessions per Slice**
Callers	SliceNet CP control services, Slice Orchestrator, Cognitive Services
Description	The operation is used to retrieve information about active sessions in the context of a slice.

Operation	**Get Flows per Slice Session**
Callers	SliceNet CP control services, Slice Orchestrator, Cognitive Services
Description	The operation is used to retrieve information about active flows in the context of slice sessions.

Operation	**Apply QoS via AF Rules**
Callers	SliceNet CP control services, Slice Orchestrator, Cognitive Services

Description	The operation is used to inject policies and rules that influence user data management

Operation	**Create MEC Breakout (UPF level)**
Callers	SliceNet CP control services, Slice Orchestrator
Description	The operation is used in case MEC Applications require direct forwarding of data plane sessions from core network.

5.3.5 WAN Control Interface

Operation	**Provision Border QoS-enabled SDN Forwarding Rules**
Callers	Inter Domain SliceNet CP control services, Cognitive/QoE control services
Description	This operation allows to directly enforce QoS-enabled SDN forwarding rules at the border devices of the administrative domain under the ownership of the SliceNet CP. It is therefore to be used to apply dedicated per-flow and per-network-device SDN rules with QoS attributes at domain borders. In particular, this operation, and the related primitives/APIs can be exploited by the Inter Domain SliceNet CP control services in the context of interconnection of per-administrative-domain slices (or slice subnets/subslices) to E2E cross-provider slices. Following the SliceNet abstraction principles, this operation is conceived to hide the specific technology implementation details of the underlying SDN approach , and expose a common model for QoS-enabled inter domain SDN forwarding rules.

Operation	**Update Border QoS-enabled SDN Forwarding Rules**
Callers	Inter Domain SliceNet CP control services, Cognitive management
Description	This operation allows to dynamically update and modify existing QoS-enabled SDN forwarding rules at domain borders. It follows the same abstraction principles described for the operation above, and it is mostly conceived to enable dynamic updates of per-flow inter domain SDN rules (e.g. for QoS upgrade/downgrade) according to decision made at the Cognitive/QoE control layers and components in the analysis of heterogeneous slice monitoring metrics, aiming to keep and satisfy agreed cross-provider SLAs and performances.

Operation	**Remove Border QoS-enabled SDN Forwarding Rules**
Callers	Inter Domain SliceNet CP control services, Cognitive/QoE control services
Description	This operation enables the decommission of one or more SDN forwarding rules previously provisioned at domain borders. The operation follows the SliceNet CP abstraction principles described for the above inter domain SDN forwarding rules operations and it is intended to be offered to other SliceNet CP services (e.g. the Inter Domain ones) and Cognitive/QoE control functions, and applied on a per-flow and per-network-device basis

Operation	**Inter-domain SDN Topology Exposure**
Callers	Inter Domain SliceNet CP control services, Resource Orchestration (e.g. NFVO)
Description	This operation allows to expose a common logical view of to the Inter Domain SliceNet CP control services (and if applicable to Resource Orchestration ones). Following the SliceNet CP abstraction principles, this Inter Domain SDN topology can be considered as independent from specific SDN technologies and models, and could be for example based on a network graph approach.

6 Use cases and High Level Workflows of SliceNet Control Plane

This section identifies the high level workflows of the SliceNet CP and the mapping to the use cases. The lifecycle of a slice is used to manage a network service with various states (created, provisioned, configured, stopped, etc.) When some action is applied to a network service (e.g. configuration of a service), many activities need to be applied on the components of this network service. Therefore, a workflow is used to execute a number of tasks in the correct order and consequently each lifecycle can generate many activities on workflows. The following subsections present the workflows that take place between the components of the control plane, P&P and QoE with regards to the SliceNet use cases. The difference between the lifecycle and workflow is shown in the following Figure 30.

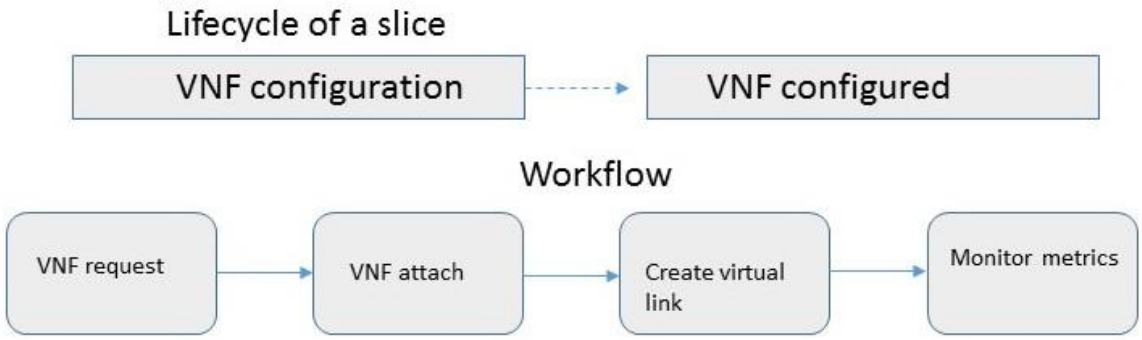

Figure 30: Difference of Lifecycle and Workflow, (ex. For a VNF function)

6.1 Smart Grid Use Case – P&P and QoE related high Level Workflows

The Smart-Grid (SG) Use Case (UC) is oriented towards protecting the power grid from unexpected faults. Its ultimate goal is to accelerate the fault detection and protection of the power grid and therefore minimize the impact on the power consumers. Intelligent Electric Devices (IEDs) are placed in strategic points of the power grid in order to sense and actuate on it. Towards this goal, the IEDs must communicate with each other with minimum latency. This will allow the IED that detected the failure in the power grid to advertise the neighbour IEDs, enabling these to “open” the electric circuits and therefore protect their customers. Simultaneously, this procedure will also enable the affected IEDs to be supplied with power through an IED from another path.

6.1.1 P&P workflows and mapping to Smart Grid UC

Frequently, the Vertical needs to add and remove IEDs from the power grid. This is a manual procedure made by the Vertical – going to the field, adding the required infrastructure to the power grid and installing the IEDs. However, besides adding the IEDs, these have to be recognized by the communications network in order to enable its communication with the neighbour IEDs. This means that the Network Slice needs to be configured with the new IED communication endpoints and add it to the multicast group. This procedure requires interactions through the P&P. An example is as follows:

1. The P&P control instance for the SG Network Slice exposes the capability of reconfiguring the Network Slice and indicating the new IEDs information;
2. The Vertical accesses the Network Slice instance (e.g. through a Self-Care Portal) and configures the new IED in the Network Slice (providing all the required attributes and parameters related with the IED);
3. The request from the Vertical is processed and mapped by the P&P northbound vertical oriented API layer to the correspondent specialized slice view;

4. The P&P selects the appropriate southbound plugin and enforces the required configuration (e.g. add the IED to the multicast group) through the implementation agnostic APIs;
5. The result of the operation is reported back to the vertical.

Other P&P interactions might exist for this UC and are still being discussed and defined within the project (e.g. monitoring of the Network Slice instance delivered to the Vertical exploring the SG UC).

6.1.2 QoE workflows and mapping to the Smart Grid UC

Providing an ultra-low latency and highly reliable service is mandatory for the SG UC. The provided Network Slice must therefore keep these requirements always under "surveillance" in order to fulfil the agreed SLA with the Vertical. Therefore, the slice provider should be permanently sensing and optimizing its network resources to satisfy the customer / Vertical needs. Towards this aim, optimization procedures, driven in SliceNet by the "QoE Optimizer", are required. An example is as follows:

1. The Monitoring sub-plane monitors and feeds the slice KPIs to the QoE Optimizer (**Local Decision Engine**);
2. The QoE Optimizer runs trained ML models to infer QoE values (e.g. latency) for the network slice;
3. The QoE Optimizer predicts that the QoE level (e.g. latency) for the Network Slice instance is in risk;
4. As an action/mitigation plan, the QoE Optimizer applies the most suitable policy(ies) to achieve the desired level of QoE.

Other QoE optimization procedures might exist for this UC and are still being discussed and defined within the project.

6.2 E-Health use Case P&P and QoE related high level Workflows

This section describes the workflows that are taken place during a slice life cycle between P&P and the e-Health use case. The common and dedicated functions of the P&P component and the functions of the new core of 5G are identified. To support different type of QoE and maintain the required QoE level during runtime, two scenarios for P&P workflow mapping to eHealth are provided: reconfiguring existing VNFs and customising CDN slicing during runtime.

6.2.1 P&P Workflows for reconfiguring VNFs

At the hospitals, experts examine a video showing conditions of the patient being treated at the Ambulance. The interface provides the users different options for QoE regarding the video resolutions and frame rates, for example, standard DR 1080p with 8Mbps, High DR 1080p with 10Mbps, etc. Assuming during the slice instantiation, 8Mbps DR is provided for video streaming at the hospital, and now the user wants to change to 10Mbps.

1. Slice Control Framework checks the exposure level of P&P control to see if it has access to the VNFs that are composed for the Video Streaming service. If not, it returns without doing anything or sends a warning to the User.
2. P&P has access to the VNFs so it contacts relevant common and dedicated actuators to perform actions to increase the QoE.
3. Common/Dedicated actuators reconfigure the VNFs, and reassign new KPIs for them.
4. Common/Dedicated actuators contact the MANO to enforce the resource reallocation to scale up relevant VNFs.

Other similar scenarios with the same workflow to reconfigure VNFs are re-selecting the audio/video codec to achieve a new required QoE level and supporting video streaming on different devices at the hospital.

6.2.2 P&P Workflow to customise CDN slicing during runtime

The paramedics might request certain videos instructing to perform certain actions. Due to different locations of Ambulances and the mobility of the Ambulance, SliceNet supports slice customisation and optimisation during runtime to deploy new VMs closer to the location of the Ambulance to host a great different number of videos by means of caches, encoders, transcoders and streamers. The P&P workflow for this scenario is as follows.

1. Slice Control Framework checks the exposure level of P&P control to see if it has access to the chained VNFs that are composed for the CDN service. If not, it returns without doing anything or sends a warning to the User.
2. P&P has access to the chained VNFs on this CDN service, directly contacts the MANO to orchestrate the resources to run those VMs on specific NFVI.
3. All relevant repositories are updated accordingly, e.g. VNF instances, NFVI resource mapping, etc.

6.2.3 QoE workflows and mapping to the e-Health UC

In this eHealth UC, it is important to maintain the QoE level during the lifecycle of the slice instance, and thus we create different workflows and map to the UC: QoE level drop (in general, and in case of the ambulance mobility).

6.2.3.1 QoE level drop – general workflows.

1. The Monitoring sub-plane at the MP constantly monitors and feeds the slice KPIs to the Local Decision Engine.
2. The LDE runs trained ML models with new input (slice KPIs fed by the monitoring sub-plane) to infer QoE values for the NSI.
3. The LDE checks that the level of QoE is poorer than the desired level.
4. The Slice Policies Repository is checked and the most suitable policies are selected to achieve the desired level of QoS.
5. Depending on the actions in the selected policies, the dedicated and/or common actuators are contacted to increase the QoE, by scaling up the slice data path (increase dedicated capacity) or by increasing the slice resources.
6. Results of the decisions taken are reported back to the Cognition sub-plane to further refine the policies.

6.2.3.2 QoE level drop due to mobility during the runtime of Face Recognition VNF.

1. The Monitoring sub-plane at the MP constantly monitors and feeds the slice KPIs to the Local Decision Engine, including the KPI of Face Recognition VNF.
2. The LDE runs trained ML models with new input (slice KPIs fed by the monitoring sub-plane) to infer QoE values for the NSI.
3. Due to the mobility of the ambulance, it is moving out of one domain and entering another domain, the LDE checks that the level of QoE is poorer than the desired level and Face Recognition KPI is lower than required (the latency is higher than required).
4. The Slice Policies Repository is checked and the most suitable policies are selected to achieve the desired level of QoS. The policies require new deployment of this Face Recognition VNF at the MEC in new domain to reduce latency.
5. The dedicated and/or common actuators are contacted to deploy Face Recognition VNF in the new domain, then terminate the VNF in the previous domain.
6. Results of the decisions taken are reported back to the Cognition sub-plane to further refine the policies.

6.3 Smart City Use Case – P&P and QoE related high Level Workflows

6.3.1 P&P slice metrics collection

1. The local city administration application may require through One-Stop-API, information about Smart lightning network slice.
2. The relevant information could be maximum simultaneous authenticated users, maximum throughput, number of disconnections

6.3.2 P&P slice VNF configuration

1. The operators of the lightning infrastructure want to push for a software upgrade of the lightning controllers.
2. The throughput needed for this operation is more than the maximum data rate configured for the usual data traffic of the slice.
3. In this situation, the vertical will inform the P&P through One-Stop-API that the VNFs need to be updated to allow this increased data rate from the devices.
4. Once the upgrade is finalized, the vertical will inform the P&P to restore the default configuration for the slice

6.3.3 P&P slice VNF deployment

1. In the Smart City slice, there is a need to introduce a new User Plane Function that will act as firewall in the data path allowing any some type of traffic and acting also as an antivirus inspector to protect the lighting controllers from malware attacks.
2. This user plane function will be required by the vertical using the P&P that will further introduce it in the data path of the Smart City slice using the orchestration tools in SliceNet management layer

6.3.4 Smart City use case and QoE high level workflows

The QoE workflows which are described below refer to some network metrics monitoring which are significant for this use case. These metrics are packet loss, TCP SYN retries and unusual high throughput.

6.3.4.1 QoE monitoring of Packet loss

1. Monitoring information coming from the monitoring sub-plane indicates a high packet loss on the Smart Lightning slice
2. Local decision engines checks that the level of packet loss is higher than the desired level
3. Local slice policy repository is checked and the time interval when the desired level of QoS (could be that only during night the packet loss should be low)
4. Both dedicated and common actuator are involved to increase the reliability of the communication either by scaling up the slice data path (increase dedicated capacity) or by increasing the QoS marking of the packets
5. Results of the decisions taken are reported back to the Cognition sub-plane to further refine the policies

6.3.4.2 QoE monitoring of TCP SYN retries

1. Monitoring information coming from the monitoring sub-plane indicates a lot of TCP SYN retries on the Smart Lightning slice
2. Local decision engines checks what is the threshold for TCP SYN retries/ minute

3. Local slice policy repository is checked if the device should have access or not to the IP address of the destination of the SYN and what is the current policy implemented for the user data flows.
4. If the current data flow does not allow access to the destination or the bandwidth allocated is too low, implement a change in data flow policies
5. Results of the decisions taken are reported back to the Cognition sub-plane to further refine the policies

6.3.4.3 QoE monitoring of Unusual high throughput

1. Monitoring information coming from the monitoring sub-plane indicates the throughput used by a device
2. Local decision engines identifies that the device has reached maximum throughput
3. Local slice policy repository is checked to see if the high throughput can be caused by a good event, like a software upgrade from a trusted server or a bad event like a malware attack
4. Current maximum bit rate allowed can be increased or decreased based on the type of event and policy
5. Results of the decisions taken are reported back to the Cognition sub-plane to further refine the policies

7 Data Plane Programmability & Resource Control for QoS-Aware Slicing

This section describes the data plane programmability for QoS-aware slicing, as envisioned in Figure 8 In addition, resource control for slicing is discussed.

7.1 Data Plane Programmability and QoS Support for the Non-RAN Segments

Network traffic flows along the different network segments in the E2E 5G architecture, including at least Enterprise/Vertical Network to RAN, RAN to Edge/MEC, MEC to Core Network and Core Network to Inter-Domain Network. Across these network segments, the traffic traverses some Commercial Off-The-Shelf (COTS) computers, compute nodes and switches allocated in different key network locations in the 5G architectures, such as MEC and Core locations. In this section, we focus on the data plane in the non-RAN segments. In SliceNet, as highlighted before, data plane programmability will be explored to achieve QoS-aware slicing, advancing the state-of-the-art best effort slicing.

The control of the network traffic to be achieved through programming the data plane is key in this vision and would expose the following key requirements to the envisioned 5G architecture:

- Support for traffic isolation to enable the definition of isolated tenant networks.
- Support of the slicing of the physical network resources to warranty specific QoS requirements in terms of network-level metrics such as bandwidth, latency/delay and jitter within the isolated tenant network.
- Support for the deployment of virtualized network control functions (VNCF) within the isolated tenants to provide required control capabilities.
- Support of a control function within the isolated tenant network to provide mobility management of the 5G users that are sharing the same tenant network.
- Support for the slicing of the resources of the tenant network to warranty specific QoS requirements within the different 5G users sharing the same tenant network.

The usage of virtualization and COTS network elements requires the extension of the traditional hardware-based data paths only to more mixed scenarios composed by both hardware- and software-based data paths, where virtual machines (VMs) can be interconnected efficiently within the software-data path and at same time with the physical ports of the computers. This combination of both hardware- and software-based data paths leads to the inclusion of control and monitoring points along the path from source to destination in the E2E data plane.

In terms of the physical data path, it is envisioned that each of the network elements should provide at least two different physical network interfaces to be used as the data path, which may be connected respectively to different hardware forwarding devices to inter-connect the network segments involved in the connectivity of the network elements. In this sense, some high-end hardware-based network cards already provide basic network traffic telemetry and control functions embedded, allowing the configuration of such existing control functions. Furthermore, there are other approaches to enable the programmability of the hardware-based data path to allow the implementation of network control and telemetry function in hardware. They are usually based on programmable hardware data paths provided by Field Programmable Gate Arrays (FPGA) network cards such as NetFPGA [7] and NetCOPE [8].

After the network traffic flows through the hardware network cards, it is received in memory, generally using the PCI interface to allocate the data packets into the DRAM memory using the available Direct Memory Access (DMA) channels. At this point, two main approaches can be explored to allow the flowing of the data packets along the host computer to reach the appropriate VM. On

one hand, a hardware-based approach can be used to directly connect the hardware network card into the VM, using the Single Root I/O Virtualization (SR-IOV) [9] and the Virtual Machine Device Queues (VMDq) [10] technologies. On the other hand, a software-based approach can be employed to allow the usage of a software-based data path to control and monitor the communications passing through the computer.

Each of these approaches has advantages and disadvantages. The hardware-based approach would provide better performance results in terms of latency although it imposes the implementation of all the network control functions to achieve the 5G requirements in hardware, which would lead to costly hardware that needs to deal with advanced capabilities such as VXLAN/GRE-based tenant encapsulation, GTP-based mobility management, OpenFlow programmability, 5G telemetry, among others. It would also reduce the number of control points since this approach would not allow the control of traffic in the host computer. The software-based approach would yield inferior performance but would allow the inclusion of additional points to enforce network control functions and policies, which would foster both scalability of policies and flexibility. Traditionally, this software-based data path has been implemented using the Linux/Windows network stack. However, recently with the inclusion of high speed transmission rates such as 10GbE (Gigabit Ethernet) and 100GbE, novel approaches to deal with performance scalability have been provided based on accelerated software-based data paths such as Data Plane Development Kit (DPDK) [10], Open Datapath [12], etc. The advantages of this approach are that it is fully programmable in software and provides very significant performance improvements, e.g., DPDK has achieved 100GbE at the line rate with 64-bytes packet size. This approach would also allow the inclusion of a control and telemetry point where specific network functions such as tenant encapsulation can be enforced. This approach would require the usage of a software switch to allow the connectivity between different software-based data paths to allow an efficient communication between VMs. It includes an additional control and telemetry point along the data path. The different VMs allocated in the computer can have an efficient inter-VM zero copy sharing mechanism to achieve effective and fast communications between VMs. This technology has different names according to the hypervisor being used, for example, *virtio* is the backend used in KVM. This backend finally delivers the network packets into the VNCF virtualized network card. At this point, both telemetry and control inside of the VNCF can be implemented similarly to those software-based approaches already indicated.

In summary, two network control and telemetry points are available in the hardware-based approach: one in the hardware card and the other one in the virtualized data path available inside of the VNCF. In contrast, three different network control and telemetry points are available in the software-based approach: one in the host machine, one in the software switch and one inside of the data path of the VNCF. All the software-based data paths are programmable due to their own nature and the hardware-based data path can also be programmable when network cards are used. Additionally, a hybrid approach could complement the software-based approach with an additional hardware-based point inside of the network card to allow hardware acceleration of key functions.

Figure 31, shows an overview of the different approaches foreseen based on the vision of introducing programmable points to both hardware and software spaces from RAN to MEC and to Core, focusing on the non-RAN segments. The round labels indicate the locations where the programmable data path can provide network control functions.

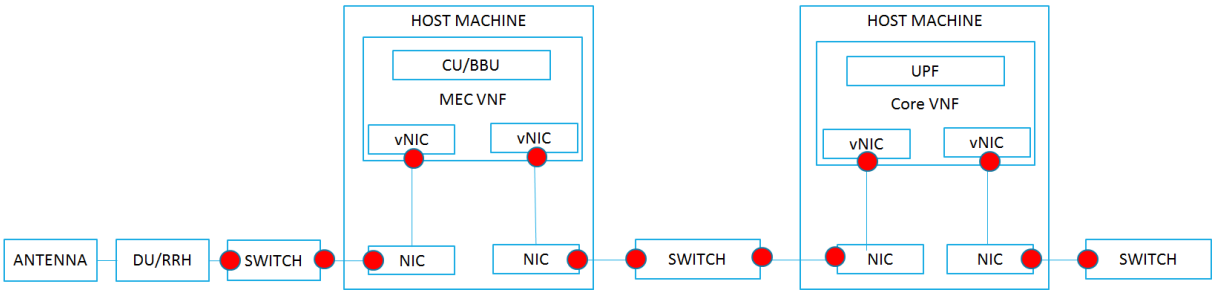

Figure 31: Different data plane programming approaches

This architecture allows the identification of different architectural points in the network where to apply the required network control and telemetry functions required to provide network slicing along the network. In each of these points, traffic engineering functions such as flow classification, redirection, mirroring, and dropping, flow policing and shaping, queueing and de-queue scheduling, encapsulation and de-encapsulation should be implemented therein to provide network slicing with guaranteed QoS (e.g., bandwidth, delay and jitter) as required in the SLAs.

7.2 Data Plane Programmability and QoS Support for the RAN Segment

QoS in data plane for any particular user within a slice, is initially managed through the attachment procedure based on the NSSAI. The RAN controller provides the runtime UE monitoring, control, and coordination to support QoS, hence allowing adapting UP based on spatio-temporal traffic variability and network dynamics among the others. For this purpose, each RAN entity/module (i.e., CU, DU, and RRU) can be decomposed into two parts: (i) the control logic, which makes the decisions for the radio link, and (ii) the control action that is responsible for applying the decisions. For example, the control logic of the MAC makes scheduling decisions like resource block allocation, modulation and coding scheme, while the action logic applies such decisions to user logical channels. Similarly, part of the logic of the RRC protocol decides on UE handovers, while the actual handover operation requires RRC to perform the corresponding action. Based on this taxonomy, the separation of the RAN CP from the UP can be taken a step further by detaching the control logic from the action and by consolidating all the control operations in a *logically centralized RAN controller*. This allows BSs to focus on performing all the UP related action functions such as applying scheduling decisions, performing handovers, applying DRX commands, (de)activating component carriers in carrier aggregation, etc.

To control and manage the eNB/gNB UP actions, the RAN API is introduced to provide a set of functions that constitute the southbound APIs. These functions are operated by an agent that resides at eNB/gNB to allow the CP to interact with the UP in five ways:

1. get and set configurations like the UL/DL bandwidth of a cell;
2. request and obtain statistics like transmission queue sizes of UEs and signal-to-interference and noise ratio (SINR) measurements of cells;
3. issue commands to apply control decisions (e.g., calls for applying MAC scheduling decisions, performing handovers, activating secondary component carriers);
4. obtain event notifications like UE attachments and random access attempts;
5. perform a dynamic placement of control functions to the master controller or the agent (e.g., centralized scheduling at the master controller or local scheduling at the agent-side).

These APIs can be invoked in the RTC plane or directly at eNB/gNB UP if control for some operation has been delegated to it.

Table 9: RAN APIs in support of QoS -based UP programmability

API	Target	Direction	Example	Applications
Configuration (synchronous)	eNB, UE, Slice	Agent → RTC RTC → Agent	QoS support UL/DL cell bandwidth, reconfigure Data Radio Bearer (DRB), Measurements	Monitoring, Reconfiguration, Self-Organizing Networks (SON)
Statistic, Measurement, Metering (asynchronous)	List of eNB, UE, Slice	Agent → RTC	Channel Quality Indicator (CQI) measurements, SINR measurements, Reference Signal Receive Power (RSRP) / Reference Signal Receive Quality (RSRQ) / UL/DL performance	Monitoring, Optimization, SON
Commands (synchronous)	Agent	RTC → Agent	QoS Support Scheduling decisions, Admission control Handover (HO) initiation	Realtime Control, SON
Event Trigger	Master	Agent → RTC	TTI, UE attach/detach, Scheduling request, Slice created/destroyed	Monitoring, Control actions
Control delegation	Agent	RTC → Agent	Update DL/UL scheduling, Update HO algorithm	Programmability, Multi-service

Table 9, above provides a list of some exemplary RAN-specific API calls to support QoS-based UP programmability. Note that the eNB/gNB agent is in charge of retrieving the cell and user related information from the underlying eNB such as cell bandwidth, slice-specific radio resource partitioning and allocation through the API calls, and can trigger events when a state changes such as user attachment and transmission time interval (TTI or equally frame and sub-frame). In addition, such API calls may be related to the NFs, resources, UEs, etc. belonging to a particular slice. Table 9 lists the different types of network applications that can be developed, ranging from monitoring for better decision making (e.g., adaptive video optimization) to control and programmability for better adaptability and flexibility to services (e.g., by controlling resource allocation, adjusting the handover logic, changing functional splits, updating precoding matrix, or even disabling/enabling ciphering/deciphering, etc.). Thus, RAN data plane support for QoS is enabled via an API allowing a control app (common or dedicated) to reconfigure and/or reprogram (a) the data radio bearer (DRB) associated to a user within the slice, and (b) the resources available for the user within the slice.

To support the QoS among different slices, the radio resources are first partitioned among different slices based on the enforced/default RRM policy to support the requested SLA/QoS. In addition, the resource partitioning also abstracts the physical resources among slices to maximize the multiplexing gain when allocating resources. Each slice has a dedicated scheduler that allocates resources for UEs belonging to its slice according to the applied scheduling algorithm (e.g. Proportional Fair - PF, Round Robin – RR, Priority-based, Delay-based). It can be implied that this scheduling is performed in two levels, namely intra-slice and inter-slice, to decouple how UEs are served and how the resources are granted and mapped to the physical channels.

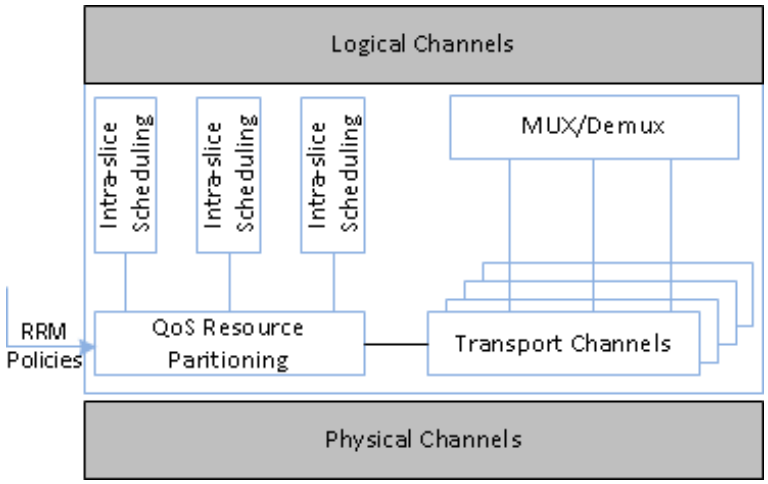

Figure 32: Slice-Aware MAC scheduling architecture

7.3 Resource Control for Slicing

In this section, we focus on the resource control in the data plane. In SliceNet, a NSI/NSSI decomposed to a set of network functions (PNFs and/or VNFs) where each NF will eventually be allocated/assigned with a set of resources. The resources can be physical/virtual bandwidth on a network link, a network forwarding element (switch, router), processing capacity of servers, processing capacity of network elements, RAN elements/functionalities, and can be owned and managed by different infrastructure providers in intra- and inter-domain, and they can be either isolated or shared by different NSI/NSSI. These features make resource isolation one of the key enablers for network slicing to guarantee the required performance (QoS/QoE), security and privacy.

In order to fully achieve the requirements, coordination and cooperation between components in the DP, CP and MP are required. Firstly, the Resource Orchestrator has the information or description of what is required for each VNFs regarding the resources, this information is sent by the Slice Orchestrator who is requesting the resources for all VNFs in a slice instance during the slice instantiation phase. The Resource Orchestrator will finalize the request and passes it to the Infrastructure & Resource Manager. The Infrastructure & Resource Manager, with the ability to dynamically allocate virtualised/non-virtualised resources for each VNF, is responsible for the VNF's lifecycle management, e.g., instantiate, scale (increase/reduce capacity of the VNF), update/upgrade (reconfiguration for new VNF software update/upgrade), and terminate VNFs (release the associated resources and return them to the pool). There are some good example solutions for managing the virtual infrastructure including OpenStack, vSphere from VMware, Kubernetes, etc.

Further, any KPIs associated with a VNF during the NSSI/NSI design or instantiation and on boarding of a NSI/NSSI to guarantee a certain of SLA/QoE level. The KPIs of the VNF will be monitored by the Resource Monitor during its lifecycle (lifecycle of VNF instance). This information may be reported to the Slice/Service Monitor in MP if relevant and be used for scaling/optimising operations, e.g. scale up/scale down by adding/removing CPUs, scale out/scale in by adding/removing VMs. For example, in case of the performance of a NSI/NSSI is being downgraded, or a threat is detected, the Resource Orchestrator will send request to the Infrastructure & Resource Manager to enforce the reallocation/mitigation of resources for the affected VNFs in that NSI/NSSI. This type of monitoring and slice operations can also come from the P&P Control when the control exposure level allows, e.g. P&P control can monitor the slice NF or collecting KPIs for the NFs only if the control exposure level is at level 0, while full access to slice management platform when P&P control has access level 7.

To control the infrastructure resources, assign appropriate KPIs and monitor the resources, the resource components in MP and/or P&P Control in CP (depending on the access exposure levels) will communicate directly with the NFV infrastructure (NVFI) in the DP. More specifically, the Virtualized

Infrastructure Manager (VIM), which is a part of the Infrastructure & Resource Manager is responsible for controlling and managing the NFVI compute, storage and network resources within one infrastructure domain, while also collecting the performance measurements and events of the resources. Finally, the resource control, depending on the used virtualization solution, will communicate with the corresponding hypervisor who is actually takes care of CPU scheduling and memory partitioning for each virtualized function.

A number of mature virtualization solutions are available and have seen widespread deployment in production environments during the last decade. In addition to commercial offerings like VMware's ESXi and Microsoft Hyper-V, a several open source solutions such as KVM, Xen, LXC, and Docker are available. While all virtualization solutions essentially provide a similar core functionality – abstracting the resources of a physical host to appear as multiple isolated logical counterparts – the underlying technologies and associated performance/isolation trade-offs differ significantly. Today the two prevalent virtualization approaches can be classified as full virtualization (aka. hypervisor-based virtualization) and OS-level virtualization (aka. container-based virtualization). Full virtualization solutions employ a hypervisor layer allowing the execution of highly isolated VMs. Each VM executes in its own kernel, with the hypervisor providing isolated access to the host's physical resources. In contrast OS-level virtualization provides virtual environments, known as containers, which share the host's kernel. Each container is assigned a dedicated process and network space and is allocated a share of the available system resources to provide a level of isolation. The associated mechanisms - namespaces and cgroups - are part of the mainline Linux kernel. Container solutions for managing and configuring these kernel features include Docker, LXC, LXD.

In general, full virtualization approaches offer a higher level of isolation between guests. The downside is that the guest Kernel and the host hypervisor layer incur a performance overhead. This can be partially mitigated by using pass-through approaches that grand the guest OS direct access to some hardware components. In contrast, container-based virtualization, which is a lightweight alternative to hypervisor-based virtualization, does not isolate resources as well as hypervisors but it offers higher density of virtualized instances on the same resources and thus offers superior performance and faster deployment [13][14].

8 Conclusions

The current traditional networks face a lot of challenges for managing, configuring, deploying and managing network services. An E2E slicing architecture that utilizes the existing SDN and NFV enablers and builds additional components and methods in all layers will provide an added value to future network architectures. In addition, network slicing can play the role of a key mechanism to provide flexibility and adaptability in the management of network resources. The full lifecycle phases of a slice when it is adapted to the operation of verticals can lead to better performance of the network services. It is also expected that multi-domain coordination and management for the E2E creation, configuration and monitoring of slices can ease the traffic management across different domains. The advancement of the control plane with the inclusion of additional components such as the P&P and QoE optimizer can lead to the provision of keeping an SLA between the operator and the vertical to an agreed and fixed value. These components provide additional abstractions to the upper layers for the easier management and orchestration of the services requested by the verticals.

This report has provided the SliceNet architectural approach with the new components of P&P and QoE in the CP plane and has described the component decomposition which is implemented when the intra and domain slicing are implemented which are significant means for providing isolation during the lifecycle of a service. It has also been concluded that the E2E slicing and the DP can offer traffic differentiation and service isolation for better performance of the service instances requested by the vertical.

In this document the slicing concept was presented and described so that it covers issues of multi-tenancy and slice isolation. The high level workflows which are discussed for the SliceNet use cases can result to the provision of services with stability and higher performance. The SliceNet architecture can provide the means of offering a network that can be flexible and manageable.

References

[1] 5G Network Architecture A High-Level Perspective; url: http://www.huawei.com/minisite/hwmbbf16/insights/5G-Nework-Architecture-Whitepaper-en.pdf

[2] 5G PPP Architecture Working Group; url: https://5g-ppp.eu/wp-content/uploads/2018/01/5G-PPP-5G-Architecture-White-Paper-Jan-2018-v2.0.pdf

[3] 3GPP TS 23.501, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; System Architecture for the 5G System; Stage 2 (Release 15); url: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144

[4] https://www.ngmn.org/fileadmin/user_upload/170428_Liaison_Statement_from_NGMN_SBA_to_3GPP_v1_0.pdf

[5] 3GPP TS 23.502, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Procedures for the 5G System; Stage 2 (Release 15); url: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145

[6] TS 33.501, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Security Architecture and Procedures for 5G System (Release 15); url: https://portal.3gpp.org/ngppapp/CreateTdoc.aspx?mode=view&contributionUid=S3-173524

[7] NetFPGA, https://netfpga.org/site/#/

[8] NetCOPE, https://www.netcope.com/en

[9] SR-IOV, https://docs.microsoft.com/en-us/windows-hardware/drivers/network/single-root-i-o-virtualization--sr-iov-

[10] VMDq, https://www.intel.co.uk/content/www/uk/en/ethernet-products/converged-network-adapters/io-acceleration-technology-vmdq.html.

[11] DPDK, https://dpdk.org/

[12] Open Datapath, https://www.opennetworking.org/projects/open-datapath/

[13] Hypervisor Performance Analysis for Real-Time Workloads.

[14] Hypervisors vs. Lightweight Virtualization: A Performance Comparison.

[15] https://tools.ietf.org/html/draft-flinck-slicing-management-00

[16] https://tools.ietf.org/html/draft-qiang-coms-netslicing-information-model-02

[17] https://tools.ietf.org/html/draft-ietf-i2rs-yang-network-topo-20

