ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Technical Specification

(07/2020)

ML5G-DEL2

Machine learning based end-to-end network slice management and orchestration

Technical Specification ITU-T ML5G-DEL2

Machine learning based end-to-end network slice management and orchestration

Summary

This Technical Specification contains the initial draft of Recommendation ITU-T Y.3182. It proposes the framework and requirements for machine learning based end-to-end network slice management and orchestration in multi-domain environments. FG ML5G approved this Technical Specification at its 9th and final meeting, held on 2-3 June 2020 (ML5G-I-247). FG ML5G is of the opinion that this technical specification should be further elaborated into an ITU-T Recommendation.

Keywords

FG ML5G Machine learning based end-to-end network slice management and orchestration machine learning, multi-domain, network slice management, network slice orchestration, network slice, use cases.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Editor: Slawomir Stanczak
Fraunhofer HHI Germany
E-mail:
slawomir.stanczak@hhi.fraunhofer.de

© ITU 2026

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

Technical Specification ITU-T FG ML5G-DEL2

Machine learning based end-to-end network slice management and orchestration

1 Scope

This Technical Specification provides the framework and requirements for machine learning based end-to-end network slice management and orchestration in multi-domain environments. It addresses the following subjects:

- Overview of machine learning based multi-domain end-to-end network slice management and orchestration;
- Use cases of machine learning based multi-domain end-to-end network slice management and orchestration;
- Functional requirements of machine learning based multi-domain end-to-end network slice management and orchestration;
- Framework of machine learning based multi-domain end-to-end network slice management and orchestration;
- Examples of test scenarios or results from experimentation.

2 References

[ITU-T X.1277] Recommendation ITU-T X.1277 (2018), *Universal authentication framework.*

[ITU-T X.1601] Recommendation ITU-T X.1601 (2015), *Security framework for cloud computing.*

[ITU-T Y.3103] Recommendation ITU-T Y.3103 (2018), *Business role-based models in IMT-2020.*

[5GPPP] 5GPPP Network Management & Quality of Service Working Group, *Cognitive Network Management for 5G*, white paper, March 2017.

[AIOps Platforms] Andrew Lerner, *AIOps Platforms*, Gartner Blog, August 2017. [online]: https://blogs.gartner.com/andrew-lerner/2017/08/09/aiops-platforms/

[IEC 61850-x] IEC 61850 series on Communication networks and systems for power utility automation.

[SliceNet Project] SliceNet Project Publications webpage: https://slicenet.eu/publications/

3 Definitions

3.1 Terms defined elsewhere

This Technical Specification uses the following terms defined elsewhere:

3.1.1 control plane (CP) [b-ITU-T Y.2011]: The set of functions that controls the operation of entities in the stratum or layer under consideration, plus the functions required to support this control.

3.1.2 data plane (DP) [b-ITU-T Y.2011]: The set of functions used to transfer data in the stratum or layer under consideration.

NOTE – In the ITU-T International Mobile Telecommunications-2020 (IMT-2020) related standard documents, "User plane" is used preferentially rather than "Data plane".

3.1.3 user plane (UP) [b-ITU-T Y.1714]: This refers to the set of traffic forwarding components through which traffic flows.

NOTE – "User plane" is referred to as "transport plane" in other ITU-T Recommendations.

3.1.4 management [b-ITU-T Y.3100]: In the context of IMT-2020, the processes aiming at fulfilment, assurance, and billing of services, network functions, and resources in both physical and virtual infrastructure including compute, storage, and network resources.

3.1.5 orchestration [b-ITU-T Y.3100]: In the context of IMT-2020, the processes aiming at the automated arrangement, coordination, instantiation and use of network functions and resources for both physical and virtual infrastructures by optimization criteria.

3.1.6 network slice (NS) [b-ITU-T Y.3100]: A logical network that provides specific network capabilities and network characteristics.

NOTE 1 – Network slices enable the creation of customized networks to provide flexible solutions for different market scenarios which have diverse requirements, with respect to functionalities, performance and resource allocation.

NOTE 2 – A network slice may have the ability to expose its capabilities.

NOTE 3 – The behaviour of a network slice is realized via network slice instance(s).

3.1.7 network slice instance (NSI) [b-ITU-T Y.3100]: An instance of network slice, which is created based on a network slice blueprint.

NOTE 1 – A network slice instance is composed of a set of managed run-time network functions, and physical/logical/virtual resources to run these network functions, forming a complete instantiated logical network to meet certain network characteristics required by the service instance(s).

NOTE 2 – A network slice instance may also be shared across multiple service instances provided by the network operator. A network slice instance may be composed of none, one or more sub-network slice instances which may be shared with another network slice instance.

3.1.8 role [b-ITU-T Y.3502]: A set of activities that serves a common purpose.

3.1.9 network slice service user (NSsu) [b-ITU-T Y.3100]: The NS service user uses the service(s) provided by the NS instance(s).

3.1.10 network slice service provider (NSsp) [ITU-T Y.3103]: The NS service provider is the user of the NS instance(s), and is responsible for providing services to its NS service users via the NS instance(s).

3.1.11 network slice provider (NSp) [ITU-T Y.3103]: The NS provider is the owner of the NS instance(s) and provides the NS instance(s).

3.1.12 network infrastructure provider (NIp) [ITU-T Y.3103]: The network infrastructure provider is the owner, the provider and the manager of the network infrastructure.

3.1.13 network slice management and orchestration provider (NSmop) [ITU-T Y.3103]: The network slice management and orchestration provider is responsible for orchestrating NS(s) and managing the lifecycle of NS instance(s) based on NS blueprint(s) (a term defined in [b-ITU-T Y.3100] (see definition in clause 3.1.2)).

NOTE – The NS blueprints can be provided by third parties.

3.1.14 network sub-slice provider (NSSp) [ITU-T Y.3103]: The network sub-slice provider is the owner and provider of the network sub-slice instance(s).

3.1.15 network sub-slice management and orchestration provider (NSSmop) [ITU-T Y.3103]: The network sub-slice management and orchestration provider is responsible for orchestrating network sub-slice(s) and managing the lifecycle of the network sub-slice instance(s).

3.1.16 network infrastructure management provider (NImp) [ITU-T Y.3103]: The network infrastructure management provider integrates infrastructures of multiple infrastructure providers to offer the combined resources to the NS management and orchestration provider.

3.2 Terms defined in this Technical Specification

None.

4 Abbreviations and acronyms

This Technical Specification uses the following abbreviations and acronyms:

ABAC	Attribute-Based Access Control
AI	Artificial Intelligence
AIOP	Artificial Intelligence for IT Operation
API	Application Programming Interface
CP	Control Plane
CSP	Customer Services Provider
DP	Data Plane
E2E	end-to-end
eMBB	enhanced Mobile Broadband
FCAPS	Fault, Configuration, Accounting, Performance, Security
IED	Intelligent Electronic Device
IMT-2020	International Mobile Telecommunications-2020
KPI	Key Performance Indicator
MAPE	Monitor Analyse Plan Execute
ML	Machine Learning
NIp	Network Infrastructure provider
NS	Network Slice
NSI	Network Slice Instance
NSp	Network Slice provider
NSS	Network Sub-Slice
NSmop	Network Slice management and orchestration provider
NSSmop	Network Sub-Slice management and orchestration provider
NSSp	Network Sub-Slice provider
NSsp	Network Slice service provider
NSsu	Network Slice service user
OSA	One Stop shop Access
PNF	Physical Network Function
QoE	Quality of Experience
QoS	Quality of Service

R-GOOSE	Routable GOOSE
RAN	Radio Access Network
SIM	Subscriber Identity Module
SLA	Service Level Agreement
UE	User equipment
UP	User Plane
uRLLC	ultra-Reliable Low Latency Communications
vFW	virtual Firewall
vIDS	virtual Intrusion Detection System
vIPS	virtual Intrusion Prevention System
VNF	Virtual Network Function

5 Conventions

None.

6 Requirements and challenges analysis of machine learning (ML) based end-to-end (E2E) multi-domain network slice management and orchestration

6.1 Service deployment requirements for E2E multi-domain vertical services

Network operators and service providers are investigating network slicing as a key enabler for delivering services on top of 5G networks to a wide plethora of vertical industries. The heterogeneity of these vertical actors (e.g., eHealth, automotive, smart cities, industry 4.0, energy and smart grid, etc.) poses a very different collection of requirements for their deployment, from infrastructure resources, network performance and service levels, and at the various phases of the lifecycle of the vertical services.

In particular, the following requirements posed by these vertical services to a network slice management and orchestration (NSmop) framework can be highlighted:

- First of all, from the perspective of a layered architecture in the framework, capable of managing and controlling services in a consistent and coordinated way, slices and resources, a vertical service will be built on the network slice with the required performance assurance, which in turn will require the underlying resources to be properly if not optimally orchestrated. Across the different levels involved, the framework shall combine resource orchestration, network slice orchestration and vertical service orchestration.
- At the subscription phase of the requested network slice based vertical services, the framework shall accept or reject the requests to create network slices e.g., depending on the availability of the required resources, and match a new network slice request with specific and diverging quality of service (QoS) and service level agreements (SLAs) requirements to a suitable pre-defined network slice template/blueprint.
- At the provisioning phase of the accepted vertical services, the framework shall provision (including deployment and configuration) the required network functions (either virtualized or physical) and allocate the required resources to run them along the E2E network infrastructure in order to create the network slice. Achieving E2E network slice based vertical services often requires a tight integration and combination of technologies and resources that span across several network and administrative domains. In practice, to achieve a truly E2E delivery of network slices, network operators and service providers shall consider the control

and orchestration of heterogeneous resources (and services) deployed in different network domains to fulfil vertical service requirements in terms of performance, SLAs and geographical constraints.

- At the runtime of the provisioned vertical services, the framework should assure that the QoS/SLA requirements are continuously satisfied, by integrating per-service performance monitoring and data analytics functionalities. With increasing requirements on quality of experience (QoE) for the end users of the vertical services, it is important that the framework is able to evaluate the QoE in an automatic and objective manner especially for those services where feedback from the verticals cannot be acquired and processed by the framework. In the framework, a given set of monitored QoS metrics at runtime may be mapped to real-time or even predicted QoE for both reactive and proactive QoE management approaches. For instance, if the QoE is predicted to be degraded, a corresponding corrective actuation shall be triggered by the framework to maintain the committed QoE for the end users.
- At the runtime of the vertical services, the framework should assure that the security requirements are continuously satisfied by integrating per-service performance monitoring and data analytics functionalities. It is expected from the framework to have an automated system for security threats monitoring, detection and response.

Therefore, coordinating the subscription, provisioning and operation of such complex network slice based vertical services requires a novel approach for E2E multi-domain service orchestration that is able to meet the above requirements. In addition, this novel approach for E2E multi-domain network slice management and orchestration can benefit from machine learning (ML) techniques to improve and enhance the runtime operation of the delivered 5G vertical services and network slices, with the main target of guaranteeing QoE and QoS levels agreed with verticals. ML can indeed be integrated as part of the multi-domain orchestration workflows and operations to support a full automation of the runtime optimization (e.g., including QoS-QoE mapping, fault corrections and performance degradation predictions, etc.) of provisioned network slices by leveraging advanced data analytics.

6.2 Vertical use case requirements for advanced network slicing management and orchestration

In addition to the above deployment requirements of vertical services, the variety of vertical services themselves has diverging QoE/QoS requirements that must be met. ITU has defined three typical classes of 5G network slices in terms of such requirements, including eMBB (enhanced mobile broadband), URLLC (ultra reliable low latency communications) and mMTC (massive machine type communications). The following set of vertical services shows some examples:

- An eHealth service requires eMBB network slices to stream live ultra-high-definition video from an ambulance on the move to a specialist in the hospital for early diagnosis of life-threatening diseases such as stroke. To further accelerate the diagnosis, mobile/multi-access edge computing (MEC) is required to intercept and process the video streaming and apply an ML-enabled stroke diagnosis virtual network function (VNF), which is able to detect stroke by analysing the face imagery of the patient and should be deployed on demand and added to the existing network slice instance (NSI), implying runtime network slice customisation. Requirements are provided in Appendix I.3.
- A smart grid service requires URLLC network slices to signal critical messaging between 5G-enabled intelligent electronic devices (IEDs) with a latency constraint of 5ms to trigger self-healing procedures on detecting and preventing a potential outage risk. To increase the reliability of the network slice for this critical mission, failure-tolerant network slicing is needed, e.g., by automatically switching to another 5G radio access network (RAN) whilst keeping the ongoing network slice instance uninterrupted. Requirements are provided in Appendix I.2.

- A Smart City service requires mMTC network slices to manage a huge quantity of smart assets such as lighting poles and waste bins to manage the energy and waste effectively, with a very high scalability requirement of up to 1 million such devices per square km. With such a high density of device-to-device communications, the quality of information is likely to be compromised e.g., resulting from noisy neighbours at either physical or virtualisation level.

To implement the above demanding vertical services in the 5G era, advanced network slicing has to be employed and automated QoE/QoS-aware management and orchestration of network slicing is needed to address these diverse and challenging requirements efficiently. The current existing mechanisms are not sufficient to address these challenges.

6.3 Functional requirements of ML-based network slicing management and orchestration

The provisioning of E2E network slices with proper QoE guarantees is seen as one of the key enablers for 5G and future networks. However, it poses several challenges in the network slice management that need to be addressed for efficient end-to-end services delivery, including estimating QoE key performance indicators (KPIs) from monitored metrics and reconfiguration operations (actuations) to support and maintain the desired quality levels. Cognitive network slice management that leverages ML techniques is entailed to proactively maintain the network in the required state to assure end-to-end QoE, as perceived by the vertical customers.

This document envisions an intelligent cost-effective network management, control and orchestration framework that can cope with the challenges of multi-domain network slicing, while minimizing human intervention towards full automation of slice lifecycle management and runtime operation. Some of the main requirements include:

- **E2E QoE** – in order to support service-level end-to-end QoE, a collection of complex tasks is required; these tasks must translate the vertical end-to-end use case requirements into network-level SLAs and concrete KPIs, *estimate* and *predict* QoE KPIs from network-level QoS metrics, and optimize the network resources to support the needs of multiple network slices.
- **Flexibility** – 5G and future networks will support many configurations, which will lead to a multitude of system states. It is no longer possible to prescribe what needs to be done for every conceivable system state, as there are too many options and heterogeneous cases. The management and control mechanism shall be able to flexibly handle states it has never encountered before and for which there are no specific rules (e.g., by applying rules from similar states); thus, it must be able to *generalize* rules and *comprehend* their intent.
- **Scale** – the deployment scale and density of 5G and future networks make it impractical (even from cost considerations alone) to require human input in the control loop. The control shall be *autonomous* and *automated* with humans only prescribing desired behaviour (e.g., as policies).
- **Dynamicity** – 5G and future networks will support a combination of heterogeneous types of workloads stemming from a variety of vertical services. These workloads can come and go and may even change dynamically as needed by the verticals. As a result, the derived requirement from the network may change often and these changes may be significant. The management mechanism cannot focus only on the common case; it shall constantly *adapt* to and *anticipate* changes.
- **Abstraction** – the 5G and future networks will be managed through several layers of abstraction. There are multiple information owners providing multiple data sources, each with its own semantics; moreover, depending on a specific role, only partial information may be available from some layers. For example, a network slice provider (NSp) or network infrastructure provider (Nip) that implement a network sub-slice (NSS) might provide only partial network information to the network service slice provider that manages the end-to-end

network slice. Thus, the management and control, per role, shall be able to combine multiple information sources, *interpret* their meaning, and fill in the gaps ("*guess*") when needed.

Cognitive network management addresses these challenges by utilizing ML to understand the network behaviour and proactively steer that behaviour towards its desired state (see [5GPPP] for a more detailed discussion). The proposed approach advocates a declarative rather than an imperative approach to network management, where cognition is used to learn the best actions to achieve declared goals, moreover, the goals are abstracted in a user-centric manner to utilize the full potential of network slicing and fulfil the "Verticals-in-the-loop" vision.

6.4 Advantages of deploying machine learning based multi-domain end-to-end network slice management and orchestration

ML application for network management automation becomes especially relevant for dynamic network slicing where the time to adapt the automation functions might be very limited. Additionally, in such a dynamic scenario, where network slices are deployed and terminated within short intervals it is highly likely that multiple network slices have similar requirements, for example, when requested by different tenants or when deployed over different temporal and spatial scopes. The application of AI will also enable required re-configurations to be orchestrated quickly and efficiently even when sufficient human resources, experience and special skills are not available.

In this context, ML can help introduce some cognitive capabilities in order to provide the means to analyse the changing environments and adjust the network slice function behaviour accordingly. One of the main application areas of ML is the prediction of future events based on past data. For network slicing this would be crucial from both the business and operational aspects of a network operator. Network service providers want to maximize the resource utilization by multiplexing different network slices but still avoid any SLA violations. Accurate assumptions about the user behaviour would allow a network operator to analyse a new network slice request and to decide whether to accept or reject it during the network slice design phase. During the operational phase, short term load predictions could also help to proactively manage the small-scale variations in different network slices load and corresponding resource demands for better SLA assurance. Apart from resource demand prediction another area of interest is the anomaly detection and diagnosis where ML-based prediction can help to proactively manage any resource outages.

Another area where ML can help network slicing management automation is to extract the insights from the previous network management actions and help improve the management process over time. During the network slice design phase, the insights from previously deployed network slices can, for example, be used to match a new network slice request with similar deployed or terminated network slices to better assess the resource requirements as well as to guide the network slice tenant of the new network slice if the assumptions about the required network slice are found to be inaccurate. Similarly, the experience from previously deployed network slices can also be used to optimize the network operator's internal design processes.

AI can also be helpful when predicting QoE from QoS metrics. Actually, in the context of slicing, measuring perceived QoE is a challenging problem for verticals because it is costly and complex due to the human involvement in the process. The challenging issue of subjective measurement is to predict it from the objective measurements; in other words predict QoE from a given set of QoS parameters. QoE shall be understood to be a multi-dimensional concept which ranges over different aspects of quality and how users perceive it. From the users' point of view, the quality of a service depends as well on the physical characteristics of the service and on the users' socio-economic and cultural background. The influence of a given factor on the way the quality of service is assessed by a user may be quite different depending on the actual situation. Hence, by forecasting the QoE degradation, the network can alert the orchestrator in advance which allows it to take remedial actions and to correct the problem before it occurs.

6.5 Challenges for machine learning based multi-domain E2E network slice management and orchestration

The multi-domain end-to-end network slice management and orchestration is a complex task that poses a set of technical challenges that have to be translated into functionalities and features offered by the orchestration logic. Moreover, as the proposed orchestration approach is augmented with machine learning and data analysis techniques, additional challenges are raised as a consequence of the required combination and integration of such techniques with the multi-domain orchestration logic. Some of these challenges can be summarized as follows:

6.5.1 Coordination of multiple layers of orchestration logic

The multi-domain orchestration approach takes care to manage the lifecycle of specific managed entities at different logical layers, which are respectively vertical services, end-to-end network slices and the single domain network slices composed of technology and domain specific resources. This means that dedicated orchestration functions are required to manage the lifecycle of these different managed entities which identify three main management domains and scopes:

- vertical services and end-to-end network slice orchestration domain,
- network slice orchestration domain,
- resource orchestration domain.

Cross-layer coordination is required for delivering to verticals tailored end-to-end services provisioned as a combination of single domain network slices.

6.5.2 Collection of heterogeneous performance metrics to feed machine learning based data analysis

As a key requirement to have ML techniques applicable and beneficial for an optimized runtime operation of vertical service and end-to-end network slices, the multi-domain orchestration framework needs to coordinate the collection of performance metrics and measurements to feed the data analysis. This means that the monitoring system, responsible for collecting and storing performance measurements for the running services and network slices has to be tightly coupled with the orchestration logic at all layers (i.e., service, network slice, resource). Ad-hoc performance metrics for each vertical service and network slice must be defined and supported by the orchestration framework, and their collection and exposure must be integrated with the service and network slice provisioning workflows. In practice, a full automation of multi-domain vertical services and network slices provisioning, monitoring and data exposure is required to enable ML based data analysis.

6.5.3 Integration of machine learning decisions and optimization actions

Another key aspect of the combination of ML techniques with multi-domain orchestration is the integration of ML decisions and optimization actions into the runtime operation of vertical services and network slices. Therefore, the multi-domain network slice management and orchestration framework is required to offer dedicated application programming interfaces (APIs) and operations for accepting the ML based decisions and applying the required optimization actions over end-to-end network slices. More than that, the orchestration framework is required to take care of the resolution of potential conflicts related to runtime optimizations affecting vertical services and network slices. Different sources of inputs for runtime modification operations may exist (e.g., including administrative actions) and need to be regulated. As a consequence, the orchestration framework needs to be aware of the status of services, network slices and resources to properly apply such resolution logic.

6.5.4 Validation of machine learning based decisions

ML models may trigger optimization actions. Once these runtime optimization actions are applied, the orchestration framework should take care to coordinate the validation process to ensure that the optimization provided the desired effect on the vertical services and network slices. Further

integration with services, network slices and resource monitoring mechanisms may be required for this purpose. Such evaluation can provide feedback to the ML techniques to refine their analysis strategies on one hand and can even trigger the orchestration framework to roll back workflows whenever the actions did not provide the desired effect.

6.5.5 Coordination of multiple layers of machine learning based decisions

Similarly to orchestration logic, ML techniques can be applied at all layers of the managed entities, namely vertical services, network slices and domain resources. This means that the cross-layer orchestration framework also needs to coordinate different levels of decisions coming from ML techniques applied over heterogeneous sets of data that are collected on top of different logical layers (i.e., service, network slice, resource).

7 Framework of ML-based E2E multi-domain network slice management and orchestration

Generally speaking, a telecommunication network/service management platform can be represented, in terms of its logical architecture, with monitoring, information, cognition and orchestration sub-planes. Taking into consideration the business roles from the network slicing perspective, represented in Figure 6-1 of [ITU-T Y.3103], the following figures show the players that need to implement the identified sub-planes aiming to achieve ML multi-domain end-to-end network slice management and orchestration.

The model described in Figure 7-1 can be further developed to support hierarchical network slicing (via network slices and network sub-slices) according to Figure 7-2.

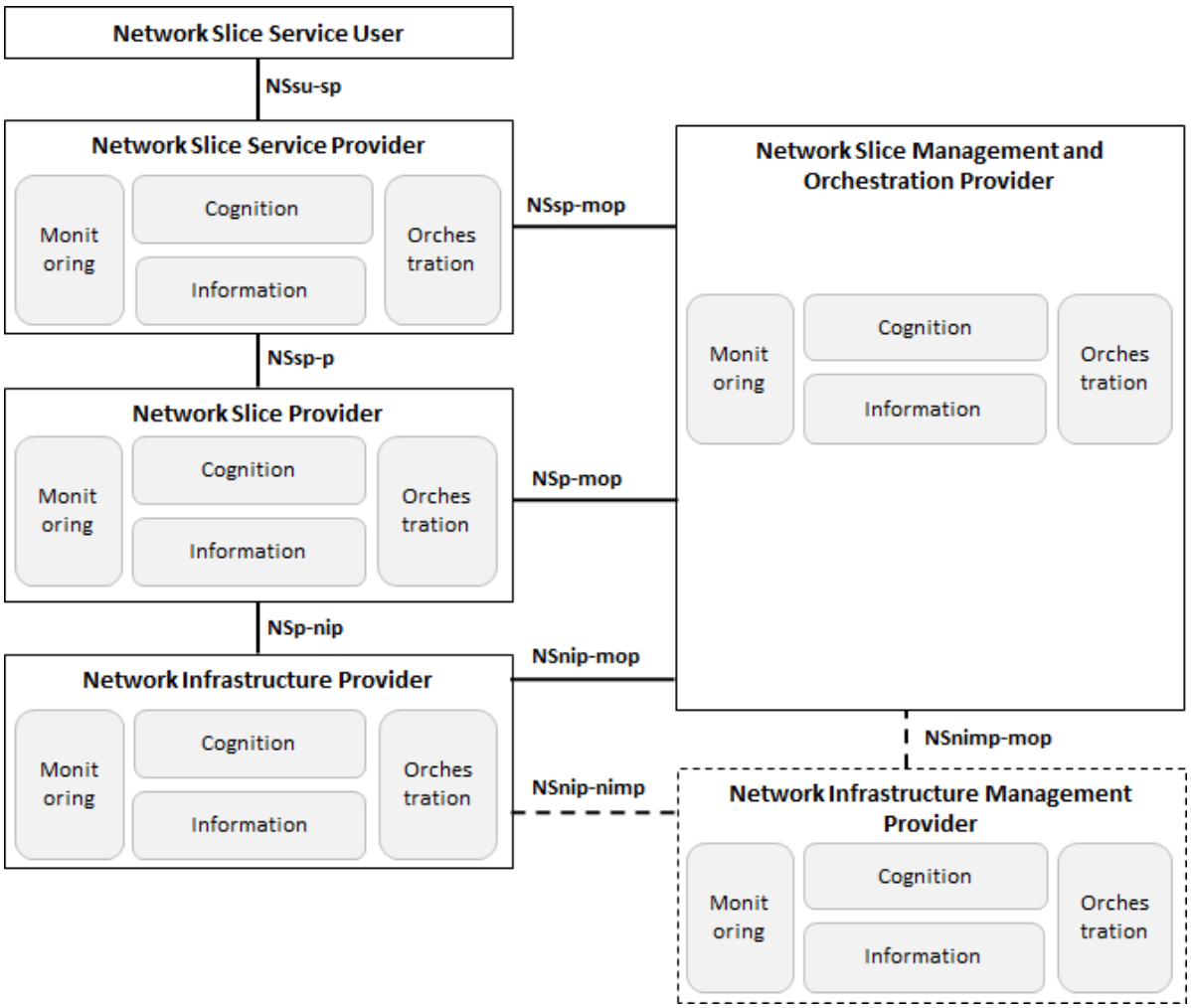


Figure 7-1 – Business roles from network slicing perspective with ML multi-domain end to end network slice management and orchestration

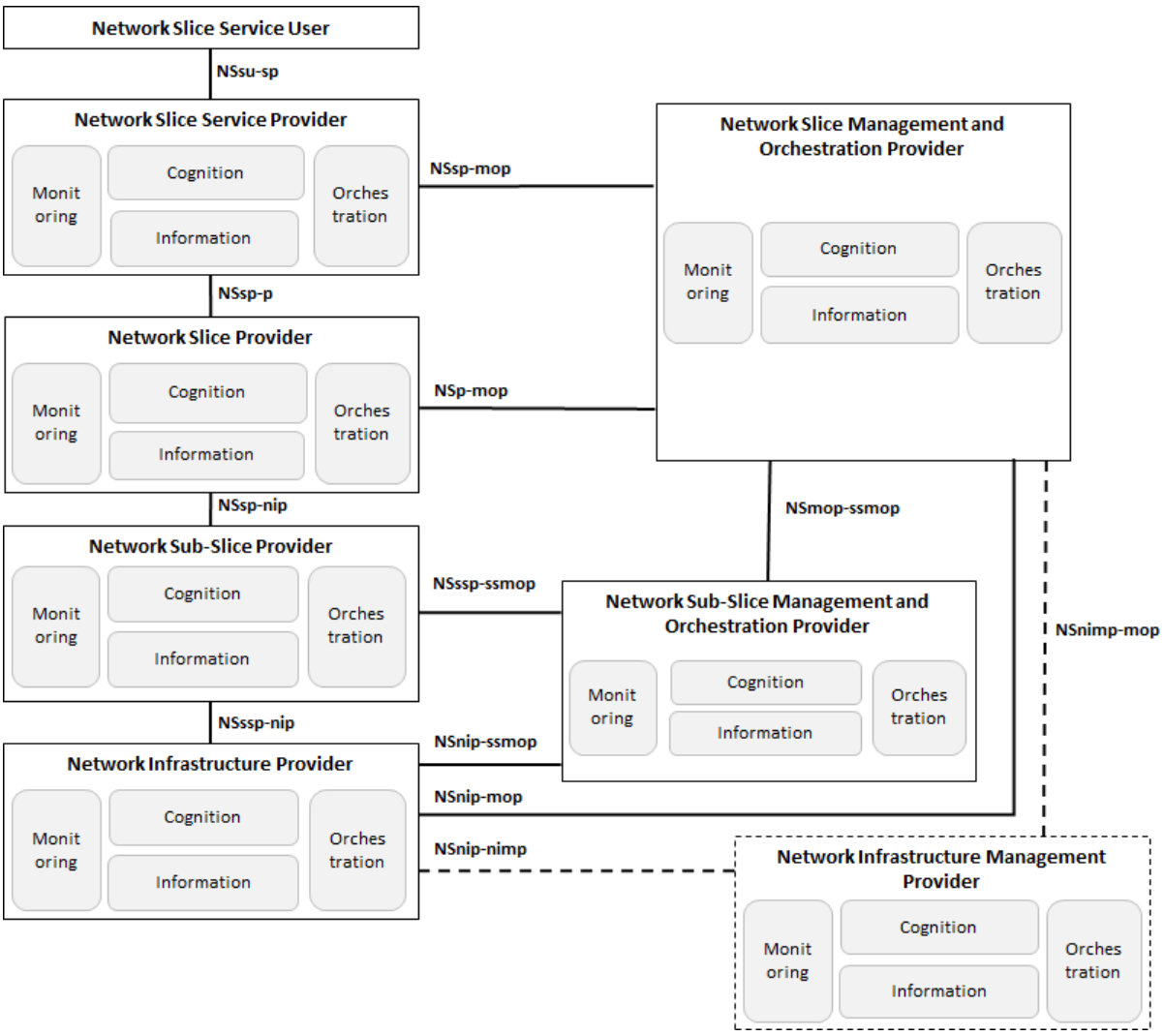


Figure 7-2 – Business roles from the hierarchical network slicing perspective with ML multi-domain end to end network slice management and orchestration

The monitoring, information, cognition and orchestration sub-planes aim together to go beyond the classical FCAPS (fault, configuration, accounting, performance, security) management functions and the silos of OSSs. These sub-planes will implement a closed-control loop approach as represented in the Figure 7-3.

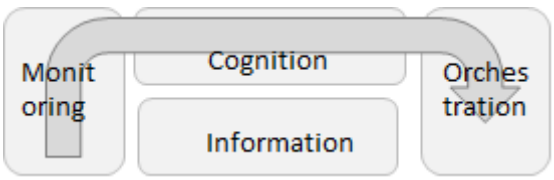


Figure 7-3 – Closed-control loop

The several players will implement this approach adapted to their specific realities, e.g., the monitoring sub-plane in the NIp is monitoring their network resources (physical network functions (PNFs) and/or VNFs) in terms of alarms and performance whereas the NSp is monitoring the requirements that belong to a given network slice like latency, bandwidth, etc.

What is relevant for the closed-control loop to function properly is that the several sub-planes can interact in order to ensure high availability and security of the network infrastructure and respective network slices. This will also contribute to minimising the human effort in maintenance and troubleshooting tasks, thus significantly reducing the operational expenditure (OPEX).

The **monitoring sub-plane** provides a cross-layer platform for collecting metrics and counters from virtual and physical network elements at the network infrastructure provider level and also collecting metrics and counters from the network slice requirements at the network slice provider and network slice service provider (NSsp) levels. The same concept applies to the network slice management and orchestration provider (NSmop) and to the network infrastructure management provider where monitoring sub-plane functions as a repository of events related to management and orchestration tasks. The main goal of the monitoring sub-plane is to integrate and combine heterogeneous sources of metrics and counters information in a common way, where applicable providing preliminary aggregation of data, at the network infrastructure and at the network slice level, exposing the collected and pre-processed data towards the cognition plane for further aggregation and analysis purposes. The monitoring sub-plane ingests all its raw data and aggregated metrics into the information plane, which functions as a "monitoring" database.

The **information sub-plane** provides a cross-layer platform to store metrics and counters, at the network infrastructure and at the network slice level, acting as a monitoring data lake repository and also storing all catalogues and inventories available. While the catalogues maintain the information related to the platform capabilities and offerings in terms of network infrastructure network elements (either VNFs and/or PNFs) descriptors and templates and network slice service descriptors and templates, the inventories keep track of all provisioned instances of network infrastructure network elements and network slice as a service offered.

The **orchestration sub-plane** provides a set of coordination functions required to onboard, provision and maintain network infrastructure network elements (VNFs and/or PNFs) and network slices. It provides functionalities to make the whole management and orchestration plane work in a coherent way. The orchestration plane also interacts with the information plane, which basically provides a heterogeneous set of catalogues and inventories.
The **cognitive sub-plane** uses AI/ML techniques to ensure the operational aspects for services, network slices and the underlying network infrastructure resources. The intelligence it provides is distributed among its inner modules. In this plane, data training is performed in order to achieve a specific model that will be later applied. Integration of AI/ML techniques into end-to-end network slice management and orchestration is essential to achieve an autonomous closed-control loop network.

8 Cognitive management and security management of ML-based E2E multi-domain network slice management and orchestration

8.1 Overview of the cognitive sub-plane

The cognitive sub-plane embraces the monitor analyse plan execute (MAPE)-K approach for automated and autonomic management, MAPE-K is a loop of **m**onitor-**a**nalyse-**p**lan-**e**xecute governed by a **k**nowledgebase that encapsulates policies, rules, algorithms, etc.

The monitoring sub-plane, separates the acquisition of monitoring data from the processing of that data and transforming it into network slice QoE metrics. The Analysis step uses the acquired knowledge to assess the network slice QoE and possible impact on corrective actions. This is done by both inferring learned cognitive models and by applying more traditional automated management methods. The planning and execution steps (termed *Actuation*) are governed through a policy framework.

The cognitive sub-plane employs a data-driven network operations methodology, also known as artificial intelligence for IT operations (AIOPs) [AIOps Platforms]. Network analysis applications react to collected operations data (both raw and processed) and generate new metrics and signals (e.g., QoE metrics and QoE-aware insights) that in turn trigger network operation actions. With this methodology, most components interact only with the data store, acting as consumers and producers. This approach minimizes the direct interfaces, provides flexibility, and facilitates easier integration of cognitive tools. It also allows existing techniques to be used with little change, as the outputs of the cognitive tasks can be treated as advanced sensor metrics.

Multiple data sources are logically merged to provide all the required information for QoE management. control and data sensor outputs are collected and persisted to support traditional monitoring through parsing, transformation, and aggregation. However, this data is also used for ML model training and for extracting QoS metrics. Additionally, feedback from the vertical is combined to allow the data processing application to assume the role of QoE sensors, learning and estimating the vertical perspective. The data source may also be fed from external data sources (both raw and processed) that are not created within the controlled system. Leveraging external data sources enables the training of ML models by means of historical data of the infrastructure or other deployed services. This allows for QoE network slice management under practical limitations, where some information must be curated to hide sensitive data or anonymize. For example, a network slice provider (NSp) may not be willing to provide some of its raw network metrics but may share processed alerts. As another example, data from multiple network slices may be merged and provided as an external source, allowing insights from one network slice to be applied to another.

The data-oriented approach, described above, is a continuation of the Monitoring sub-plane to support QoE sensing, data-operations applications may be deployed for each network slice to filter relevant data, apply security, add context, aggregate network slice metrics, etc. This addresses several of the design challenges related to the monitoring framework (for example, the approach is scalable and allows attributing the cost of monitoring to each network slice). Moreover, flexible QoE sensors may be employed, from simple aggregation and transformation tasks to inference of elaborate ML models.

Ingesting data from external sources is a crucial part of the knowledge acquisition process. ML algorithms perform and learn more efficiently with massive amounts of data (big data); this holds almost independently of the objectives or use cases being addressed. It is unlikely that network slices generate enough operations data to support their own learning processes. Thus, there is a need to combine multiple sources. However, data from other network slices or data from the underlying network slice infrastructure may be subject to confidentiality or privacy limitations. Data ingestion must enforce governance rules dictated by the data owner. In addition, external data may contain corrupt or partial data, thus, it must be parsed, validated, and cleaned before it enters the data-lake. Finally, the data ingestion must receive the data on the sender's "terms", namely, it must handle the data volume and maximal data rates.

The cognitive sub-plane supports a ML pipeline. As a starting point, and external to the pipeline, there is a data discovery and gathering phase, this is where input for ML occurs. Logically, this step represents a data source from the pipeline point of view. Internally, it is divided into six different functional areas, covering all phases from data collection to the ML models lifecycle:

1. **Ingest data**: this module enables the pipeline to read data, and its responsibility is divided into two components:
    i. **Readers**: data input can be multiple files containing observations or streaming data. Each reader abstracts the medium source of the observations and their nuances;
    ii. **Normalization modules**: data normalization is the process of combining, merging, and cleaning data, according to the knowledge gathered from the data analysis. This includes removing duplicate observations and removing invalid and/or badly formed data;

2. **Data analysis**: the initial analysis serves the purpose of gaining data insights and further problem contextualization. This module runs statistical queries (i.e., counting, averaging, grouping), to check if the dataset is balanced, incomplete or how to focus its modelling;
3. **Transform data**: data transformation depends on data analysis and problem objectives. This module transforms the data into ML-ready. This is where features are extracted and their normalization (e.g., ordinal, one-hot encoding) happens;
4. **Create model**: ML algorithms, which can cover the classification, prediction or clustering ML areas, are applied in this phase. This is where models are effectively trained, optimized (i.e., hyper parameter tuning) and where their testing strategies are put in place: cross-validation, feature importance analysis, dimensionality reduction and so on;
5. **Deploy model**: during the training/testing phase, if a model shows significant fitness metrics values it can then be deployed into production and start being used to predict, classify or cluster data in the real-time problem domain;
6. **Monitor and maintain model**: deployed models can lose their effectiveness over time, especially when the data domain is too volatile and dynamic, this means that certain models may be unfit for usage since they no longer properly represent the real world. When models show fitness metrics that are below the configured acceptable values, they are archived, and a re-training task is scheduled to update them. When such a situation occurs, the process reverts back to step 4, the "**Create model**" phase.

8.2 Cognitive module dependencies on sensing

Single-domain performance management evolves based on technology abstraction with the intention to enable domain specific technologies, infrastructures and deployments to be easily federated across multiple domains. The pool of available resources is aimed to be exploited by operations in the context of higher level, inter-domain, business patterns. As technology related capabilities are abstracted to produce the domain offerings that are made available to multi-domain entities, a similar abstraction approach can be followed to exploit these offerings in the context of the design and provision of vertical oriented service characteristics. At the level of end-to-end network slice management, slicing can be assumed as a brokering process that focuses on matching vertical requirements with efficient management and control over the available domain offerings. Cognitive processing can be a mechanism that enables efficient QoE tailored processing. For this purpose, cognitive processing should address:

- Which pieces of the sensing/actuation information that single domains expose are required for the QoE purposes.
- How these pieces should be combined.
- How the combined information should be retrieved to feed the service feature components and also how any related actuations are to be formulated to close the automation loops.

The overall requirement that can allow the evolution of multi-domain ML modules to function properly regards the harmonized and uniform exposure of each domain monitoring and actuation offerings. This in turn necessitates the definition of a common registry of actions and counters on top of which ML modules can be implemented. The common registry guarantees that technology abstractions across domains result in widely end-to-end defined counters and actions, the availability of which can play an important role in the selection of the domains to be contributing to a specific network slice setup.

The above runtime resolutions require a design phase that allows for the mapping of ML modules with the required monitoring and actuation options. ML modules in turn are exposed for vertical selection through qualitative definitions in a way similar to the technology abstraction that takes place per domain.

8.3 Strategies to derive QoE from QoS

The cognitive sub-plane provides a framework for the QoE-aware management of network slices on top of a shared 5G network infrastructure. QoE-awareness allows it to identify degradation of slice users' QoE by directly monitoring user feedback or by inferring QoE from QoS derived from measured slice infrastructure metrics. Examples of user feedback are quality metrics transmitted from UE, feedback about user satisfaction from the NSsu, etc. Examples of QoS are KPIs derived from infrastructure network traffic (latency, throughput, etc.) and resources (cpu, memory, BER, etc.).

The following sections present two strategies for leveraging machine learning to estimate the QoE of a network slice and then triggering a remedial action to re-configure the network.

8.3.1 Estimating QoE from network QoS

In this strategy, the relationship between the target application's QoE and the QoS is learned during the training phase, but at run-time only metrics from the provider's infrastructure are collected to derive the QoS. This run-time QoS is used to infer the QoE from the training models to trigger remedial actuations when required. Our premise is that although the provider can only observe partial network information, the E2E QoE is exposed in these observations. With this goal in mind, we describe the scenario for the training of the ML model. Two slices that share infrastructure resources are required. One slice supports the target application, on the other slice background network traffic is run to generate network congestion; since the slices share infrastructure resources the background traffic will interfere with the performance of the target application which will in turn affect the target application's QoE. Training data is created from simulations of various levels of traffic congestion running in parallel to the target application. While running these simulations QoS metrics are collected from the provider's infrastructure monitoring framework. At the same time QoE metrics are collected from the user application, MOS based feedback from actual end users or the user equipment (UE) device. An ML model correlating the QoS features with the target QoE metrics is generated that estimates the QoE from QoS.

At run time, this QoE estimation model combined with current QoS measurements serves as a trigger for remedial actions by the SliceNet cognitive sub-plane. In the cases of unfavourable QoE, policies defined for the network slice would specify which remedial actions need to be triggered when facing unfavourable QoE. These remedial actions would be communicated to the orchestration sub-plane to execute the desired (re-) configurations. Examples of remedial actions are adjusting the network slice bandwidth, scaling overloaded VNFs, migrating VNFs, or handing the slice to another network slice provider.

Figure 8-1 and Table 8-1 illustrate the workflow of the *estimating QoE from network QoS* strategy workflow during the ML model training phase. The main difference between the run-time and training phase is that the QoE input from the NSsu, i.e., UE quality metrics, is consumed in the training phase, but not at run-time; rather at run-time the QoE is derived from QoS. Also, no remedial actions are triggered during the training phase since they are not relevant to the generation of the QoE prediction model.

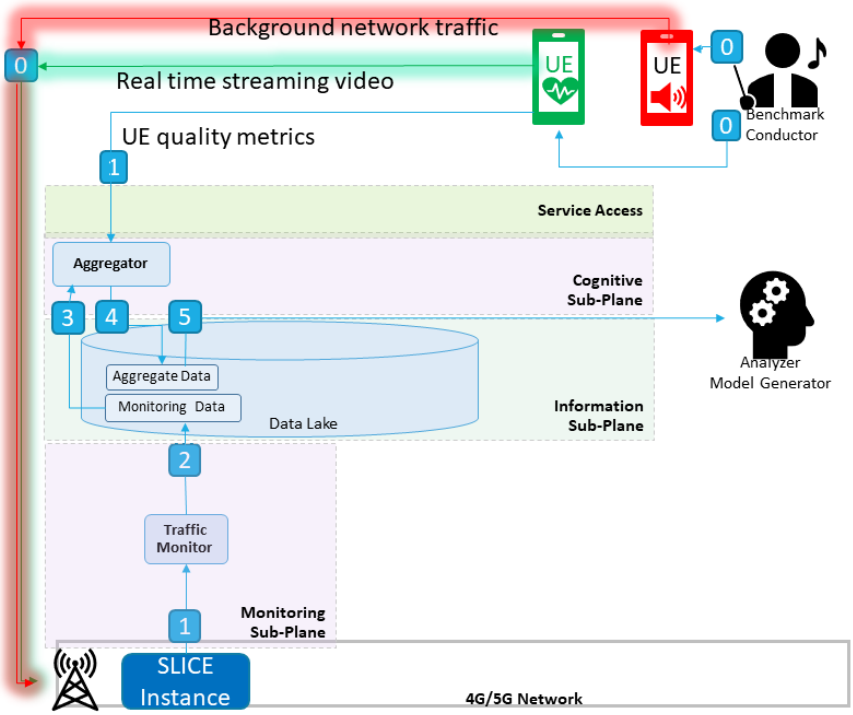


Figure 8-1 – Estimating QoE from network QoS (training)

Table 8-1 – Estimating QoE from network QoS (training)

Step	Description
0	The training phase requires the collection of many samples of the subject real time streaming slice with various levels of background network traffic (including no background traffic). A benchmark conductor is used to orchestrate the running and labelling of these benchmark samples. It runs many benchmarks. Each benchmark consists of a real time video stream and some background network traffic. Each benchmark is assigned a benchmark ID. The benchmark ID is used to map metrics from the UE's streaming video device and the network flow metrics from the slice's VNF.
1	The traffic monitor (and resource monitor) continuously collects network flow metrics related to a VNF interface servicing the stream and the rest of the network infrastructure. In parallel, the UE streams its quality metrics to the aggregator.
2	The traffic monitor stores the collected metrics into the data lake.
3	The aggregator consumes the traffic metrics.
4	The aggregator transforms the traffic metrics and UE quality metrics into QoS ML-ready features and inserts the transformed data into the shared Data Lake.
5	Once all the benchmarks are run, the analyser model generator: • Reads accumulated QoS metrics & UE quality metrics from the data lake. • Derives target QoE estimations from the UE quality metrics aggregations. • Derives QoS features aggregations highly correlated with QoE estimations. • Generates an ML model with target QoE estimations from QoS features. • Persists the ML model. The ML model is used at run-time to derive QoE estimations from the QoS metrics.

Figure 8-2 and Table 8-2 illustrate the workflow of the *estimating QoE from network infrastructure QoS* strategy during run-time.

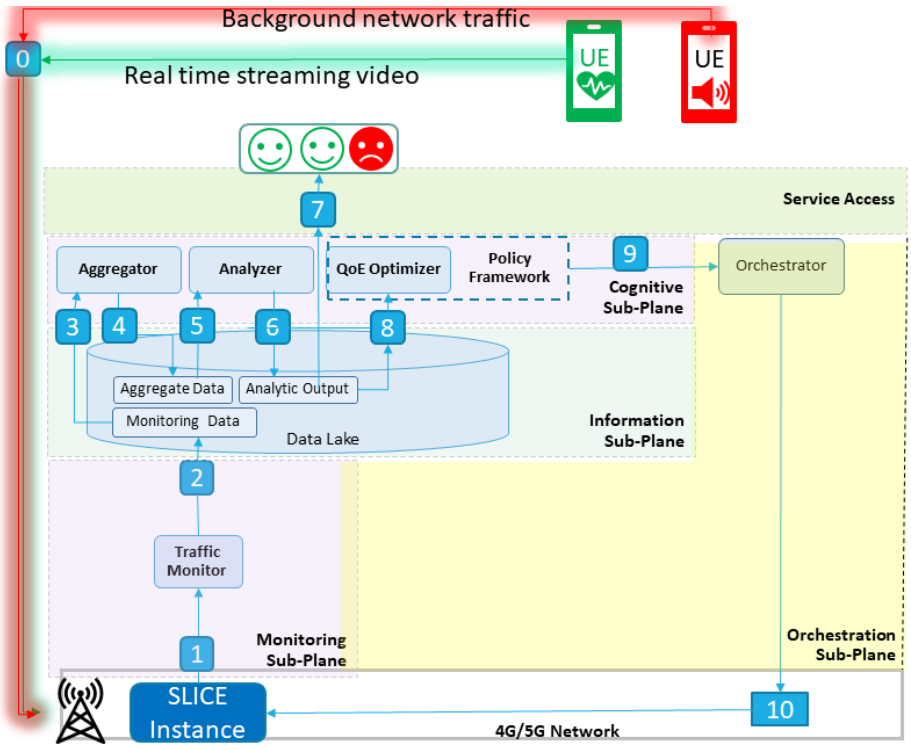


Figure 8-2 – Estimating QoE from network QoS (run-time)

Table 8-2 – Estimating QoE from network infrastructure QoS (run-time)

Step	Description
0	A UE mobile device streams data into the network slice assigned to the target application. In the background, network traffic is generated by other UE devices on another slice.
1	The **traffic monitor** (and resource monitor) continuously collects network flow metrics related to a VNF interface servicing the stream and the rest of the network infrastructure.
2	The **traffic monitor** stores the collected metrics into the data lake.
3	The **aggregator** consumes the traffic metrics.
4	The **aggregator** transforms the traffic metrics into QoS ML-ready features and inserts the transformed data into the shared Data Lake.
5	The **analyser** consumes the QoS data from the Data Lake.
6	The **analyser** (i.e., the ML model) analyses the QoS data, derives the QoE estimate of the target network slice, and produces a QoE estimate and inserts it into the data lake.
7	The QoE estimate is displayed over time.
8	The **QoE optimizer** consumes the QoE estimates.
9	When the **QoE optimizer** sees a series of unfavourable estimations it then decides, based on the policies specified for the slice when it was created, the most suitable remedial actions to overcome the undesired states. The decision is delivered to the Orchestrator.
10	**Orchestrator** orchestrates the actual network re-configuration.

8.3.2 Estimating QoE from UE quality metrics

In this strategy, the target application's QoE is inferred from UE quality metrics. The QoE is inferred from UE quality metrics learned during the training phase to generate the QoE estimation model. At run-time the same UE quality metrics are collected and streamed into to this QoE estimation model to trigger remedial actuations when required. Specifically, the UE transmits its quality metrics to the

cognitive sub-plane's aggregator which feeds the metrics into the analyzer's QoE estimation model. The QoE estimation model then triggers the QoE optimizer and policy framework to decide on the proper remedial actuations.

Figure 8-3 and Table 8-3 illustrate the workflow of the *Estimating QoE from UE quality metrics* strategy during run-time.

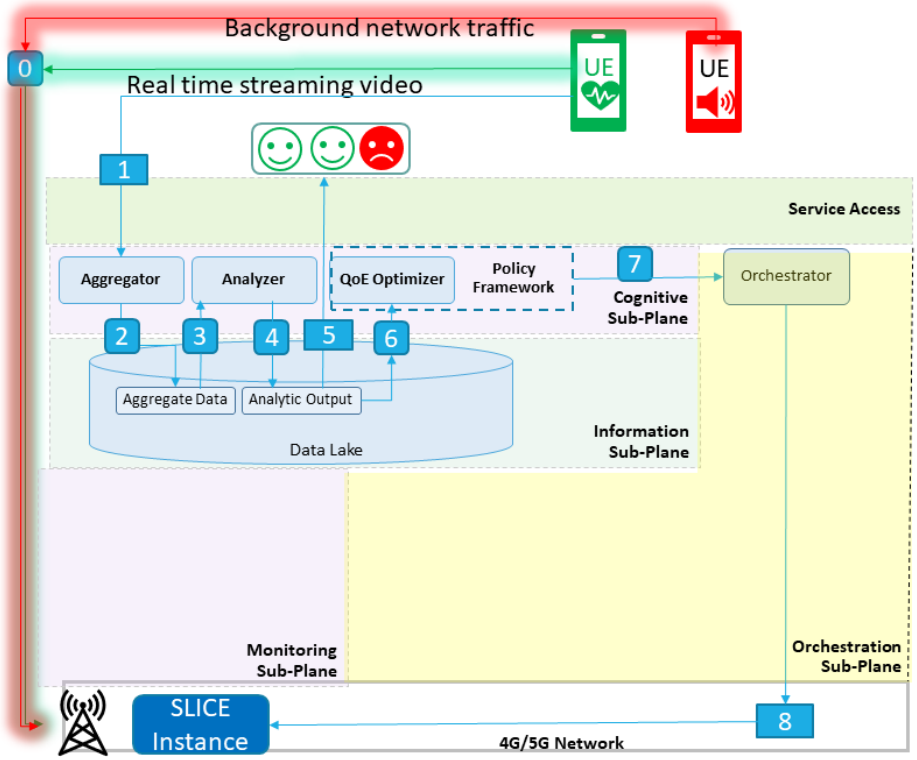


Figure 8-3 – Estimating QoE from UE quality metrics (run-time)

Table 8-3 – Estimating QoE from Measured UE quality metrics (run-time)

Step	Description
0	A UE mobile device streams data into the network slice assigned to the target application. In the background, noise is being generated by other UE devices on another slice is run background network traffic to generate network congestion.
1	The UE streams its quality metrics into the aggregator.
2	The **aggregator** aggregates UE quality metrics and transforms the data into ML ready features. Finally, it inserts the transformed data into the data lake.
3	The **analyser** consumes the Aggregator's output.
4	The **analyser** (i.e., the ML model) analyses the aggregator's output and inserts its QoE estimate into the data lake.
5	The QoE estimate is displayed over time.
6	The **QoE optimizer** consumes the QoE estimates.
7	When the **QoE optimizer** sees a series of unfavourable estimations it then decides, based on the policies specified for the slice when it was created, the most suitable remedial actions to overcome the undesired states. The decision is delivered to the Orchestrator.
8	The **Orchestrator** orchestrates the actual network re-configuration.

8.4 Security management and orchestration

The proposed multi-domain network slice orchestration approach considers security as a dedicated and specific constraint that verticals can express when they request for the provisioning of their services. In particular, these constraints are translated by the end-to-end network slice orchestration logic into specific network slice security requirements that need to be fulfilled by the security network functions to be deployed as part of the network slices providers in support of encryption as a service (vEaaS), virtual intrusion detection system (vIDS), virtual intrusion prevention system (vIPS) and virtual firewall (vFW). These security requirements are also used by the end-to-end network slice orchestration framework for the selection of the suitable network slice providers when composing the service for the verticals. As a result, such security constraints and requirements are first exposed by the network slice providers (NSp) to the network slice service providers (NSsp) in their network slice offers, and in turn by the NSsp to verticals in their end-to-end network slice offer.

8.5 Security events monitoring, detection and response

When the security functions have been deployed, either by orchestration when initiating a network slice, or by the network operators, the functions will go into operational mode where security events will be monitored, relevant data will be collected and fed to the functions for security events detection. At this point, the remediation can be triggered by the security functions themselves, e.g., running a script to add a new firewall rule to block an intruder, or the functions will trigger an alert or another event either to the orchestrator for policy-based actuation or to an admin/operator for further investigation. For a service analysis for inter-domain, the security functions/FCAPS manager will be in the NSsp domain. Information for those models will be injected from NSp to NSsp. Figure 8-4 summarize the procedure. In that case the actions are exposed as a set of configurations (e.g., firewall rules, NF config).

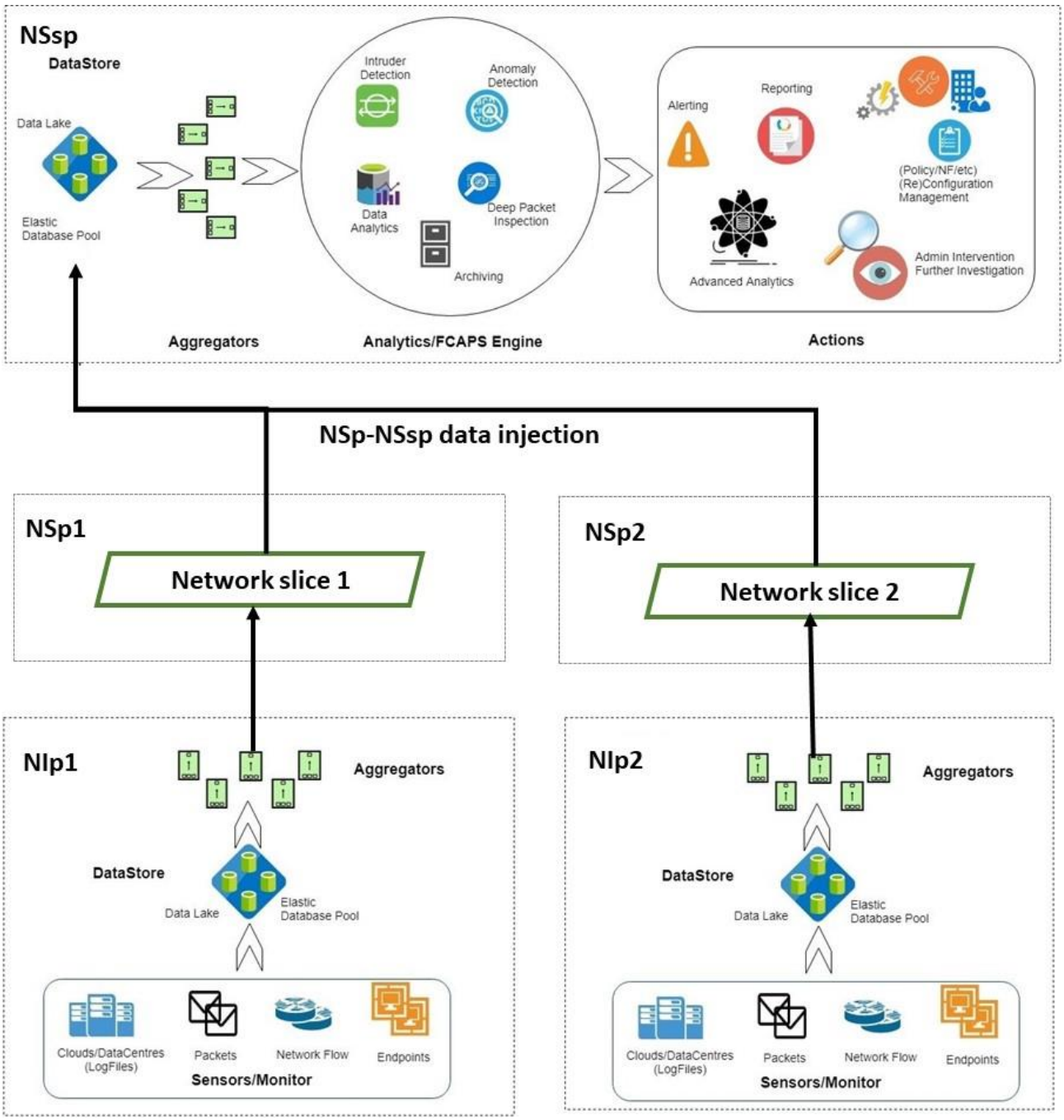


Figure 8-4 – Security events monitoring, detection and actuation

9 Security considerations in machine learning based multi-domain end-to-end network slice management and orchestration

Security aspects for consideration within the cloud computing environment, including data management, are addressed by security challenges for the CSPs, are described in [ITU-T X.1601]. [ITU-T X.1601] analyses security threats and challenges and describes security capabilities that could mitigate these threats and meet the security challenges.

9.1 Identity and access management

Identity and access management is a framework of policies and technologies for ensuring that the proper people have the appropriate access to resources and objects (hardware, applications, database, etc.). The inter-domain security management will ensure the required integrity and authorisation of the inter-domain access which includes:

- users access to the management (NSsp Admins, NSp admins, etc.) to design and onboard a network slice, or to construct policies for a network slice, or security policies for the system, components and network slices, or to configure a feature, etc.
- users access to the service offerings or applications (NSsu users, NSsp users to the NSp domain, etc.) to browse and subscribe for a service and to interact with their services at runtime, etc.

For this, the first phase will be for the system to identify the users or entities based on authentication technologies, e.g., performing identification authentication of users or entities by evaluating required login credentials (e.g., passwords, personal identification numbers (PINs), biometric scans, security tokens, etc.), or by multi-factor authentication, which requires two or more authentication factors and is often an important part of layered defence to protect access control systems.

After this identity authentication, the system will rely on an access control system which implements a process for defining security policy and regulating access to resources such that only authorized actors are granted access according to that policy. Having the access control system is fundamental to mitigating the risk of unauthorized access from malicious external users and insider threats, as well as the risk of loss or exposure of critical assets, etc. In general, access control can be rule-based/role-based/attribute-based access where the access permission for the users or entities will be depending on the conditions (rules), or the role of these users or entities, or the attributes of the actors, or it can be a combination of these types.

There are many standardisation and related documents to refer to for a strong authentication ecosystem and access management, these below are some essentials:

- [ITU-T X.1277] with universal authentication framework (UAF) describes the components, protocols and interfaces that make up the FIDO UAF strong authentication ecosystem. The goal is to provide a unified and extensible authentication mechanism that supplants passwords while avoiding the shortcomings of current alternative authentication approaches. Following the agile approach, it allows the relying parties to choose the best current authentication mechanism for the end user/interaction, while also preserving the option to leverage emerging device security capabilities in the future, without requiring additional integration effort.
- NIST attribute-based access control [b-NIST] has developed an example of an advanced access control system, attribute-based access control (ABAC), which can manage access to networked resources more securely and efficiently, and with greater granularity than traditional role-based access management. It enables the appropriate permissions and limitations for the same information system for each user based on individual attributes and allows for permissions to multiple systems to be managed by a single platform, without a heavy administrative burden.
- [ITU-T X.812] defines the basic concepts for access control, demonstrates the manner in which the basic concepts of access control can be specialized to support some commonly recognized access control services and mechanisms, defines these services and corresponding access control mechanisms, identifies functional requirements for protocols to support these access control services and mechanisms, identifies management requirements to support these access control services and mechanism and addresses the interaction of access control services and mechanisms with other security services and mechanisms.

9.1.1 Cross-domain Trust Model

This document introduces a cross-domain slicing architecture that creates end-to-end network slices operating across multiple network service provider domains. This means that a cross-domain trust model should be addressed so that the end-to-end network slices working across security domain boundaries can be trusted. The trust model should define clearly all the actors, roles and rules that are involved in the cross-domain slicing architecture to ensure the network slices can operate seamlessly

from end to end. This trust model should also support adaptation for different SLA agreements between business partners, at least to address the business use-cases it supports.

To carefully design a cross-domain trust model, it needs to address the basic questions, including:

- Who are the entities/actors involved?
- What are the resources that need to be protected?
- Who provides identity authentication service?
- Who provides authorisation services?
- What is the trust relationship among business partners?
- What is the business model?

Table 9-1 – shows a list of resources

Resources	Description
End-to-end network slice instances	End-to-end network slice instances are provided by the NSsps to the verticals, depending on the vertical needs and specified in the SLA agreement between the verticals and the NSsps.
Network slice/sub-slice instances	Network Slice (NS)/Network Sub-Slice (NSS) instances are provided by the NSps to the NSsps based on the NS/NSS offerings and the subscription from the NSsp for the NSps offerings to fulfil the end-to-end network slice requirement that the verticals require from the NSsps.
CP services	CP services that are identified by RESTful URIs, allowing clients to access to the CP services such as QoS control, NF config, etc.
Other services	Other services include monitoring services, data-plane services, management/orchestration services, etc.
Vertical data	Vertical Data that are stored or generated during the runtime of the vertical's end-to-end network slices. For example, in the eHealth use-case, the vertical data can be video streams in the ambulances, or patient information, etc.
Network data	The data related to network performance (e.g., signal strength, bandwidth, packet loss, jitter, etc.) or data related to traffic flows, etc.
Infrastructure	Network infrastructure including access network (e.g., RAN, Edge, routers, switches, etc.) and core network (core data centre, routers, etc.). Each NSp has its own network infrastructure, and it is out of scope for this contribution to secure the infrastructure.

9.1.2 Independent domain authentication

The simplest solution is that the authentication service is handled independently in each domain (Figure 9-1). An NSsu user will register with the NSsp domain to have a user account, then use this user account to authenticate with the NSsp via the NSsp one stop shop access (OSA) to access to the NSsp services (e.g., service subscription). For NSsp and NSp, the NSsp will register with the NSp to have an account in each NSp domain (e.g., "nssp_nsp_1" for NSp1, "nssp_nsp_2" for NSP2). Using the corresponding account, the NSsp can have access to the services that the NSp provides, for example, with "nssp_nsp_1", the NSsp can view the list of the network slice instances that it has subscribed to with the NSsp 1 previously and can perform NSp functions, through the OSA, on this NSp1. The NSp1 will have to maintain a repository for all user accounts it has granted and a repository for maintaining the access control for the users. Independently, the NSp2 (and other NSps) has its own identify management system.

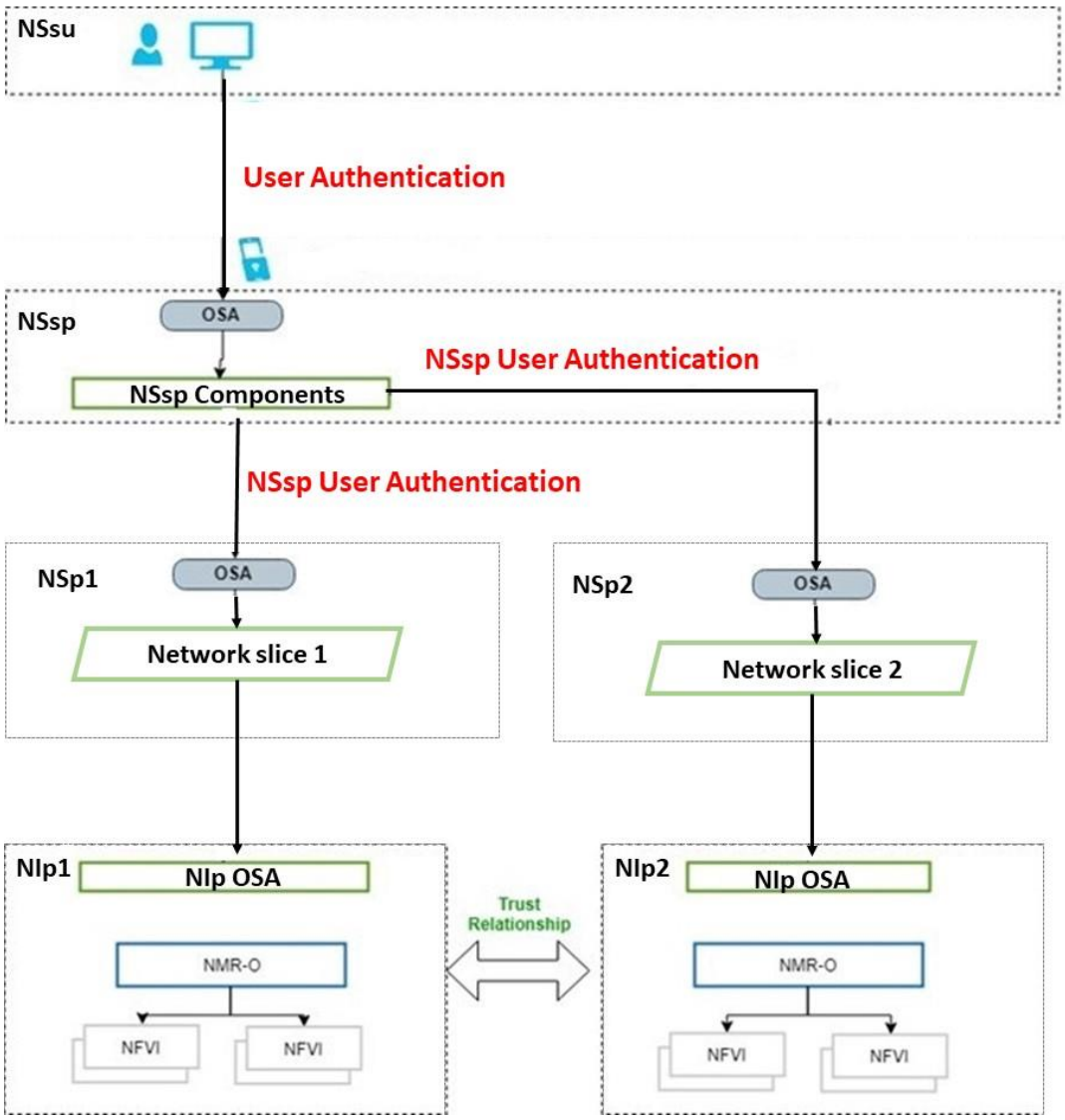


Figure 9-1 – Independent domain authentication

This model requires the NSsp to sign-on separately in each domain which is not practical.

9.1.3 Cross-domain single sign-on

This model allows movement of users between the domains with a single sign-on. With cross-domain single sign-on (CDSSO), once logged in to a domain, a user can make a request to protected resources that are located in the second domain without being forced to perform another login. In general, the CDSSO mechanism will transfer the encrypted user identity token from the first domain to the second domain, and the second domain now has the user's identity that has been already authenticated in the first domain.

9.2 Security in performance isolation through network slicing

The proposed multi-domain network orchestration and network slice management framework leverages network slicing itself as a built-in security mechanism in terms of performance isolation in the framework. In particular, two important aspects are considered. Firstly, critical signalling in the framework for control, management and orchestration purposes should be protected, and one-way to achieve such protection is through assigning critical signalling to a top priority class of network slices especially when the critical signalling has to pass by the data plane, where competing user traffic exists or when the data plane's security is compromised. By employing top priority slices, the critical

signalling traffic can be well protected against both adverse network conditions and cyber-attacks. Secondly, the user traffic in the data plane itself is protected through network slices with appropriate levels of priority as well to achieve performance isolation against potential cyber-attacks. For instance, next to the critical signalling in the system, mission-critical applications for public safety, city infrastructure such as smart grid and smart city protection, eHealth, etc. need proper performance isolation through network slicing so that the performance in terms of QoS is guaranteed against cyber-attacks such as large-scale distributed denial of services (DDoS) attacks. User traffic following a best-effort service deliver will also be confined to another less prioritised slice in order to protect the other slices of the multi-domain network. Cyber-attacks will usually be originated from such a best-effort slice and thus other high priority slices will be completely isolated.

Appendix I

SliceNet project and use case descriptions

I.1 SliceNet project

SliceNet [SliceNet Project] is a European Union Horizon 2020 (H2020) funded project as part of the 5G Infrastructure Public Private Partnership (5G PPP) 4, implementing an intelligence-based autonomic end-to-end network slicing management framework for virtualized multi-domain, multi-tenant 5G networks in order to allow a faster adoption of 5G networks by the verticals.

SliceNet satisfies the vertical requirements in terms of data processing, latency, and system performance by implementing a verticals-oriented, QoE-driven 5G network slicing framework focusing on cognitive network management and control, for end-to-end slicing operation and slice-based services, across multiple operator domains, in Software-Defined Networking (SDN)/ Network Function Virtualization (NFV)-enabled 5G networks.

For 5G verticals businesses, SliceNet offers an innovative one-stop shop solution to create customized services meeting diverging and challenging requirements.

For 5G service providers and users, SliceNet targets unprecedented guaranteed service quality, by enabling advanced users' QoE-centric service creation, delivery and lifecycle management.

For 5G network operators, SliceNet presents an integrated FCAPS (Fault, Configuration, Accounting, Performance, Security) framework for truly end-to-end management, control and orchestration of 5G slices, in a secure, coordinated, robust and verticals-oriented way, by enabling secured, interoperable, reliable and QoE-driven operations across multiple virtualized administrative domains.

SliceNet has received funding from the European Union's H2020 program, under grant agreement No. 761913. The authors thank all SliceNet partners for their support for this work.

SliceNet has two testbeds that demonstrate the proposed architecture and concepts, these are:

- Smart grid use case – The smart grid use case is implemented with an ultra-reliable low latency communication 5G network slice to demonstrate a fully decentralized high-speed self-healing solution for electric power grids. These self-healing solutions rely on distributed automation and power system protection and aim at increasing energy supply quality of service (QoS) by reducing the number of customers affected by power outages, as well as the frequency and duration of these outages. The use case highlights the potential for 5G slicing to leverage critical systems supported by 5G network infrastructures.
- e-Health use case – The eHealth use case, using a connected ambulance, aims to provide support to medical emergency first responders by developing a platform that can rapidly provision dedicated end-to-end broadband 5G network slices to advance the emergency ambulance services through the design of better-connected, integrated and coordinated healthcare. The connected ambulance will act as a connection hub for the emergency medical equipment and wearables, enabling the storing and real-time streaming of video data to the awaiting emergency department team at the destination hospital. By providing prioritized life-critical video-streaming from inside a high-speed moving ambulance, the use case achieves "reliable and dependable QoS and QoE with 'zero perceived' downtime". It will use eMBB, requiring both extremely high data rates and low-latency communication in some areas, and reliable broadband access over large coverage areas.

I.2 Smart grid vertical service UC

Table I.2 – Smart grid vertical service use case

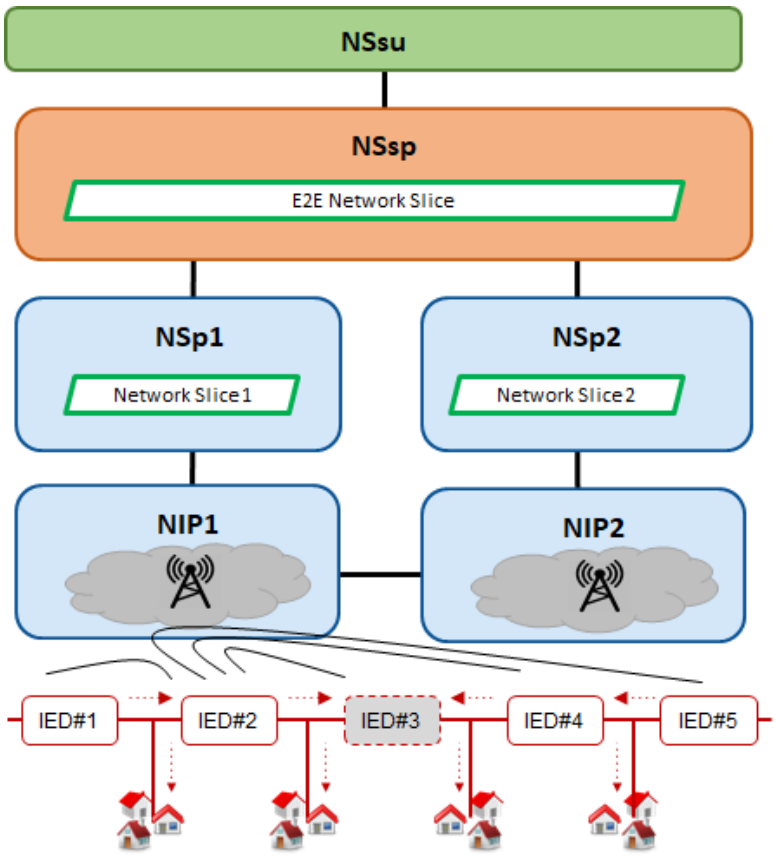
Title	Smart grid vertical service use case (See Figure I.1)
Description	• Network slice service user (NSsu) (in this case the smart grid operator) requests a URLLC service from a network slice service provider (NSsp). • NSsp is responsible for managing the service lifecycle, including its exposition to the NSsu and creation (composition of the end-to-end network slice (NS) across one or multiple administrative domains). • Network slice provider (NSp) is responsible for managing the NSs lifecycle, including their creation and exposition to the NSsp, as well as their provision, monitoring and optimization. • Network Infrastructure provider (NIP) is the owner, the provider and the manager of the network infrastructure. It is responsible for managing the network resources lifecycle (VNFs/PNFs). One NIP represents one administrative domain. The smart grid use case aims to benefit from the URLLC 5G network to implement and demonstrate an advanced self-healing solution for electric power grids. The smart grid use case comprises 3 scenarios (1) Protection coordination, (2) automatic reconfiguration and (3) Differential protection.
Roles	Player 1 – NSsu (Vertical); Player 2 – NSsp (end-to-end NS provider); Player 3 – NSp+NIP (administrative domain providing network slices to the NSsp).
Figure (optional)	

Figure I.1 – Smart grid use case from the NSsu perspective

Table I.2 – Smart grid vertical service use case

Title	Smart grid vertical service use case (See Figure I.1)
Pre-conditions (optional)	Vertical power system must be operating normally (all target sections must be energized) and all field IEDs (intelligent electronic devices) protection devices must be communicating with each other using [IEC 61850-x] R-GOOSE and with the control centre/substation.
Post-conditions (optional)	Once the use case has been executed, the entire system should maintain normal operation (communication-wise). The metrics defined for the pre-conditions should be used for the post-conditions as well. The use case scenario will be successful if the R-GOOSE events are received by the IEDs within the defined time, i.e., in time for the protection functions to coordinate.
Derived requirements	The system should be monitored for a pre-defined time period while in normal operation, before the use case scenario events take place. QoS must be evaluated for peer-to-peer communications between IEDs and for communications between the IEDs and the control centre/substation. The following metrics should be used for measuring QoS: • End-to-end latency; • Packet loss/ bit error rate (BER); • Out-of-order packets.

I.3 eHealth vertical service UC

Table I.3 – eHealth vertical service use case

Title	5G network slicing for mission-critical services (see Figures I.2 and I.3)
Description	Vertical (eHealth ambulance service) – requests an eMBB service from a NSsp NSsp – responsible for managing the service lifecycle, including its exposition to the NSsu and creation (composition of the end-to-end Network Slice across one or multiple administrative domains), aggregating multi-domain FCAPS, and hosting the optimisation over multiple administrative domains. NSp – responsible for managing the NSs and involved network resources lifecycle, including their creation and exposition to the NSsp, as well as their provision, monitoring and optimization. One NSp represents one administrative domain. The eHealth use case aims at leveraging from the public safety service that takes priority over all other network traffic (e.g., industries 4.0, smart city, ad hoc access). It is crucial to guarantee the SLA for the service, e.g., availability, delay, bandwidth, coverage, security, etc. The proposed approach meets these requirements with the approaches below: • OSA towards the vertical, with plug and play (P&P) functionalities for service monitoring, reconfiguring and auto scaling. • Cross-domain, cross-plane orchestration to provide dynamic slicing and dynamic reconfiguration based on priority level. • Cognitive, agile QoE management of network slices for service assurance of vertical business. • End-to-end network slice FCAPS management to manage fault, configuration, accounting, performance and security of all network slices across multiple planes and network operator domains.
Roles	NSsu, NSsp, NSp, Nip

Table I.3 – eHealth vertical service use case

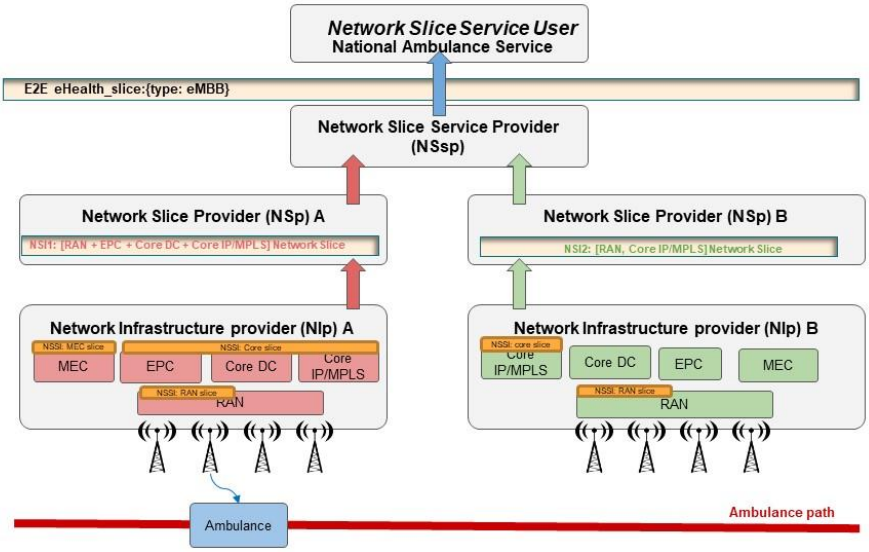
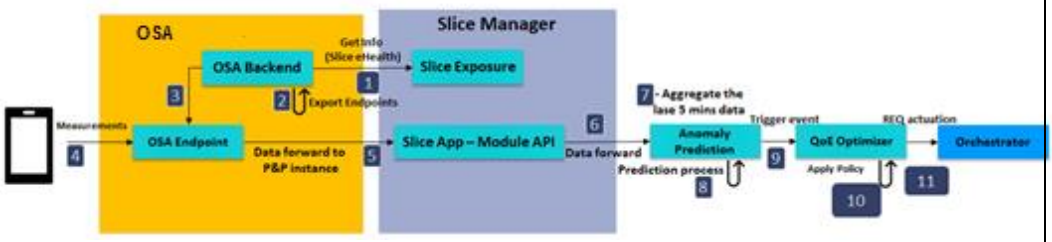
Title	5G network slicing for mission-critical services (see Figures I.2 and I.3)
Figure (optional)	 **Figure I.2 – eHealth Business Model from NSsu perspective**  **Figure I.3 – Continuance measurements** Figure I.2 shows the overall E2E deployment of the eHealth use-case. To maintain the QoE of the eHealth use-case, continuance measurements are feed from UEs to the anomaly prediction model as depicted in Figure I.3. In the NSsp domain, the OSA backend retrieves the eHealth slice information from the slice exposure in the slice manager. Then, the OSA backend creates an OSA endpoint that allows the UE to submit the collected measurements, which will be forwarded through the slice manager to the anomaly prediction model. The anomaly prediction model produces an event-based trigger to the QoE optimizer that includes a QoE value that characterizes the link quality in the next 5 minutes. Based on the QoE value and the slice policy, the QoE Optimizer may request an actuation from the orchestrator. The orchestrator uses the proper interface to request the actuation from the responsible NSp.
Pre-conditions (optional)	Subscriber identity module (SIM) cards need to be installed in ambulances and used by UEs that deliver the services.
Post-conditions (optional)	Once the use case has been executed, the entire system should maintain normal operation (communication-wise). The metrics defined for the pre-conditions should be used for the post-conditions as well.

Table I.3 – eHealth vertical service use case

Title	5G network slicing for mission-critical services (see Figures I.2 and I.3)
Derived requirements	The system should be monitored for a pre-defined time period while in normal operation, before the use case scenario events take place. QoS must be evaluated for end-to-end communications between ambulances and eHealth services (ML and streaming). The following metrics should be used for measuring QoS then later mapped to QoE metrics: • End-to-end latency; • Packet loss/ bit error rate (BER); • PHY conditions.

Bibliography

[b-ITU-T Y.1714] Recommendation ITU-T Y.1714 (2009), *MPLS management and OAM framework.*

[b-ITU-T Y.2011] Recommendation ITU-T Y.2011 (2004), *General principles and general reference model for Next Generation Networks.*

[b-ITU-T Y.3100] Recommendation ITU-T Y.3100 (2017), *Terms and definitions for IMT-2020 network.*

[b-ITU-T Y.3502] Recommendation ITU-T Y.3502 (2014), *Information technology – Cloud computing – Reference architecture.*

[b-NIST] NIST SPECIAL PUBLICATION 1800-3 (2017), *Attribute Based Access Control*.
https://www.nccoe.nist.gov/sites/default/files/legacy-files/abac-nist-sp1800-3-draft-v2.pdf