Deliverable D2.2
Techno-economic analysis of the beyond 5G environment, use case requirements and KPIs

Document Summary Information

Grant Agreement No	871780	**Acronym**	MonB5G
Full Title	Distributed Management of Network Slices in beyond 5G		
Start Date	01/11/2019	**Duration**	36 months
Project URL	https://www.monb5g.eu/		
Deliverable	D2.2 – Techno-economic analysis of the beyond 5G environment, use case requirements and KPIs		
Work Package	WP2		
Contractual due date	M12	**Actual submission date**	24/11/2020
Nature	Report	**Dissemination Level**	Public
Lead Beneficiary	OTE		
Responsible Author	(OTE)		
Contributions from	All		

Disclaimer

Copyright message

CONTENTS

List of Acronyms and Abbreviations

Acronym	Description
3GPP	Third Generation Partnership Project
AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
AR	Augmented Reality
BBU	Baseband Unit
BS	Base Station
CN	Core Network
DE	Decision Engine
ES	Experimental Scenario
MEC	Mobile Edge Computing
ML	Machine Learning
MS-SSIM	Multi-Scale Structural Similarity
NSI	Network Slice Instance
NSSI	Network sub-Slice Instance
OAI	Open Air Interface
QoE	Quality of Experience
QoS	Quality of Service
VMAF	Video Multimethod Assessment Fusion
VNF	Virtual network Function
VR	Virtual reality
URLLC	Ultra-Reliable Low-Latency Communication
eMBB	Enhanced Mobile Broadband
RAN	Radio Access Network
DDoS	Distributed Denial of Service
PoC	Proof of Concept
MNO	Mobile Network Operator
SLO	Slice Level Objectives
SDBM	Service Dominant Business Model

List of Figures

List of Tables

1 Executive summary

MonB5G aims to provide a zero-touch management and orchestration platform with the support of a large number for slices for networks 5G and beyond. It stems on principles monitoring, analysis and decision making along with AI techniques to achieve its goals. In addition, it incorporates security and energy efficiency principles to orchestrate a massive number of slices in an automated zero-touch way.

This document elaborates MonB5G selected practical use cases that have been initially considered in the most relevant research projects to MonB5G and telco initiatives and alliances such as NGMN, as well as in the directives of ITU-R and within 3GPP standardization body. While providing a thorough analysis of system requirements, corresponding KPIs, involved stakeholders and their overall impact. The document then delves into the description of the various MonB5G uses cases, emphasizing on target applications, demonstration setups and scenarios to be performed in WP6. A detailed analysis of functional and non-functional requirements and specific key performance indicators (KPIs) is also included, thus presenting specific stakeholder/actor interest and involvement in each use case, together with the corresponding societal/business impact. The use cases specified in this deliverable will drive the definition of the project's prototypes to be deployed and demonstrated in the project's testbeds to be demonstrated from CTTC and EURECOM. The specific use case scenarios to be demonstrated will be defined in WP6.

In this respect, MonB5G includes two use cases for PoCs with two Experimental Scenarios (ES) each one. In the first ES of the first PoC, zero-touch multi-domain service management with end-to-end SLAs, where storage, compute, and RAN functions of an Augmented Reality (AR) Tactile Internet slice will be hosted in different regions under the control of local NFVOs and Decision Engines, and stringent end-to-end real-time SLA must be honoured. In the second ES, elastic end-to-end slice management will be tested by considering a massive number of slices and automating their complex lifecycle management and reconfiguration via AI-driven decisions. In the first ES of the second PoC, attack identification and mitigation where different slices (eMBB, URLLC) are considered and Machine Learning mechanisms are applied to the domain of Anomaly Detection, in order to identify both in-slice and cross-slice attacks, localize them, and automatically perform corrective actions to mitigate them. In the second ES of the second PoC, MonB5G framework will be evaluated against adversarial attacks, such as poisoning attacks that target training phase and invasion attacks that target test phase, and the vulnerability of the distributed learning algorithms will be assessed in terms of their performance metrics (e.g., precision, accuracy, etc.).

The roles of the various actors and stakeholders involved in the use case scenarios are developed in this deliverable, including: infrastructure provider, network slice provider, slice operator, network slice management provider, slice template provider, VNF provider, slice tenant, service/application provider, service broker and end-user that all of them are engaged in the project's tasks to make the MonB5G ecosystem.

The requirements and KPIs of the above-mentioned PoCs will on one hand be aligned with the ones described by 5GPPP but at the same time will complement them since this deliverable provides for the first time PoCs which have been softly or not at all been dealt with in other 5G-PPP projects. Specifically, MonB5G KPIs cover the whole cross-domain network slice's lifecycle in terms of availability, admission rate, reconfiguration latency, security attack/anomaly prediction/identification accuracies and reaction time, slice isolation and SLA violation reduction, overhead, energy consumption and OPEX minimization, as well as decentralized AI performance and robustness metrics.

On the other hand, the deliverable provides a market analysis, where gaps in the existing management and orchestration platforms are identified and consequently presenting how MonB5G intends to fill them up. The document also discusses the value that AI/ML brings to both telco operators and end users, by (i) improving network efficiency, (ii) reducing network operational costs and (iv) enhance the overall quality of experience (QoE). Moreover, the document revisits business models in telco alliances and standardization bodies and

analyses their shortcomings while showcasing how MonB5G will address them via a novel business model, detailing the involved actors, the benefits and the value chain.

In addition, it provides the value proposition that the scenarios bring to the various stakeholders and gives an insight of the benefits to the above actors.

The innovation of this work is that it outlines use cases and scenarios that are discussed for the first time in the context of 5G-PPP projects and the KPIs which are provided in the deliverable have been imprecisely described in other projects.

It is noteworthy that this deliverable is submitted by M12 while the work in the architecture is still ongoing and the corresponding deliverable D2.1 will be submitted by M17. Therefore, this deliverable is based on a preliminary approach for the architecture that may be modified later in the project

2 Overview and Report Structure

The overall goal of this deliverable is multi-fold, since it (i) describes the use case scenarios that are going to be studied and tested throughout the project, (ii) provides an insight of the overall scenarios by considering requirements for each one, (iii) including the necessary evaluation metrics through specific KPIs, (iv) identifies stakeholders and focuses on business aspects together with the value proposition of the use cases currently under development in MonB5G.

In addition the deliverable identifies the interaction of various stakeholders, identified with the scope of the project in mind (i.e. infrastructure providers, resource providers and service providers), the cooperation models that they entail, and the incentives of each stakeholder.

Furthermore, this deliverable describes the relevant use cases each of them involving with two scenarios that will be deployed, it provides functional, non-functional and business requirements and provides an analysis of the KPIs that are going to be evaluated during the project. These KPIs are meticulously described, are produced through a specific, pre-defined scenario and evaluated accordingly in a pre-defined manner. The main notion behind the design of those novel KPIs is to go beyond standard QoS metrics and better capture E2E performance for different verticals/applications. This study will serve towards a “Vertical-Horizontal” integration of KPIs, e.g., by combining both network/infrastructure-level and service-level metrics, while the actual data-driven mechanisms and algorithms to achieve this are the subject of WP3.

The deliverable is divided in 5 sections. The following sections are:

Section 3 provides an overview/background of the use cases described in various SDO’s and shortly in other 5G-PPP projects to show consistency with the SDO’s and similarities and/or differences with other 5G-PPP projects.

Section 4 provides a detailed analysis of the MoNB5G use case/scenario along with the requirements to achieve the goals of these use cases and provides the KPIs that will prove and verify the functionalities and breakthroughs of the project. I also analysis the roles and interactions of the various actors involved in the use cases and discusses the benefits of them.

Section 5 provides an analysis of the value proposition brought by MonB5G and considers the benefits of the actors involved in the use cases.

The last chapter is the conclusion that provides a summary of the outcomes of this report.

3 Overview of 5G Use Cases in the Literature

This section provides an overview of the use cases which are described and analysed in 5G research projects, in initiatives and Standardization fora. Research in 5G technology has been very intensive in the past years with regards to 5GPPP [2] phase I, II and III projects. Consequently, several EU funded projects have developed their own use cases and scenarios to place requirements and KPIs in the technology and services offered to the end users. Similarly, other initiatives like NGMN, and standardization bodies, like 3GPP, ETSI and ITU-R, have identified respective requirements on the technology such that to focus the research for handling the future demands.

3.1 Overview of use cases in Industrial Initiatives such as NGMN

Short overview of the use cases discussed in NGMN. Twenty-five (25 use cases have been classified and grouped in eight families, which are:

- Broadband Access in Dense Areas
- Broadband Access Everywhere

- Higher User Mobility
- Massive Internet of Things (IoT)
- Extreme Real-Time Communications
- Lifeline Communication
- Ultra-reliable Communications
- Broadcast-like Services

3.2 Overview of the use cases in 5G-PPP projects

The use cases described in 5GPPP [2] projects are classified in 6 families which are provided below:

- Dense Urban
- Broadband (50+Mbps) everywhere
- Connected vehicles
- Future smart offices
- Low bandwidth IoT
- Tactile internet / automation

3.3 Overview of use cases in ITU-R

ITU-R defines the overall objectives of the future development of International Mobile Telecommunications (IMT) for 2020 and beyond. It has identified three directions of services, namely "Enhanced mobile broadband", "Ultra-reliable and low latency communications", and "Massive machine type communications". The above three scenarios have certain key capabilities related to data rates, latency, energy efficiency, etc.

The use cases described in ITU-R [3] are classified in three main usage scenarios which are:

- Enhanced mobile broadband
- Ultra-reliable and low latency communications
- Massive machine type communications

3.4 Overview of use cases in 3GPP

3GPP TS 28.554[4](2019-12)[4] started working on a new Study Item called "Study on New Services management and orchestration''. The use cases are related to the measurement of KPIs and are outlined as follows:

- Use case for end-to-end latency measurements of 5G network-related KPI
- Use case for number of registered subscribers of single network-slice-instance-related KPI
- Use case for upstream/downstream throughput for one-single-network-slice-related KPI
- Use case for mean PDU sessions number in network slice instance
- Use case for virtualized resource utilization of network-slice-instance related KPI
- Use case for 5GS registration success rate of one single-network-slice instance-related KPI
- Use case for RAN UE throughput related KPI
- Use case for QoS flow retainability-related KPI
- Use case for DRB accessibility related KPI
- Use case for mobility KPI
- Use case for DRB retainability related KPI
- Use case for PDU session establishment success rate of one network slice (S-NSSAI) related KPI

- Use case for integrated downlink latency in RAN
- Use case for PDU session Establishment success rate of one single-network-slice instance-related KPI
- Use case for QoS flow retainability-related KPI
- Use case for 5G Energy Efficiency (EE) KPI

3.5 Other 5G-PP projects with some similarity to the Use Cases of MoNB5G

There are too many research projects that have resulted in a large number of scenarios and UCs focusing on diverse requirements. Consequently, for the sake of simplification, we have identified the most relevant projects to MonB5G and discussed their similarities. Given the large number of UCs, it would be unrealistic to consider all the UCs proposed by the research community and the standardization bodies and fora for evaluation of the 5G RAN design

5G-PPP projects with some relevance with respect to the use cases are outlined in the following Table 1.

Table 1:5G-PPP projects with relevant Use Cases

SUPERFLUIDITY	SUPERFLUIDITY offers a converged solution to deal with the complexity emerging from three forms of heterogeneity: heterogeneous data traffic and end-points, heterogeneity in services and processing needs and heterogeneity in access technologies and their scale. Cloud-based architectures, virtualisation of radio and network processing or software acceleration are some of the topics of mutual interest for both projects. Specific assets of SUPERFLUIDITY that are relevant to MonB5G include the control framework presented in [13], which encompasses both C-RAN and MEC, the management and orchestration concepts presented in [6], which include MEC service relocation, and the monitoring tools for anomaly detection in VNFs discussed in [5]. These aspects should be considered when devising the dynamic and analytics tools and the orchestration and management infrastructure of MonB5G. The key use cases which are relevant to MonB5G are as follows: • Automated deployment of a C-RAN and MEC components at the edge and central clouds. SDN enabled advanced management and control over network resources (specifically client access control, extended network operator management, network reliability). • Policy driven service relocation from central DC to edge based MEC. • Unikernel based service deployment – DDoS service attack mitigation at the cloud edge.
SELFNET	SELFNET develops a self-organised network management in NFV and SDN, using artificial intelligence. Details about the five-layered architecture for providing SON in an SDN/NFV network are given in [8]. In addition, the distributed 5G SON orchestration may be useful for MonB5G SON functionalities. The use cases defined in the SELFNET project are: • *Self-healing*, in order to detect and predict common failures/malfunctioning in 5G network infrastructure (hw/sw failures, infrastructure/operation vulnerabilities or power supply interruptions) to apply reactive or preventive recovery, developing analyser to infer Health of Network metrics coupled with self-healing diagnosis intelligence to derive potential problems. • *Self-protection* to detect and mitigate effects of cyber-attacks and restore 5G network traffic to a steady state of security, through a deployed and chained at different locations of the network. • *Self-optimization* to automatically respond to degradation of QoE levels (either actual or predicted), coupled with end-to-end proactive energy management for optimized resource deployment across the 5G network.

4 MonB5G Use cases for PoC: Description and Detailed Analysis

4.1 MonB5G Ecosystem

The MonB5G ecosystem provides a solution for various stakeholders, whom all wish to enter to the market through the zero--touch management platform, to provide their services for current and new end-users. Task

2.1 deals with the use cases for PoC, the objectives and challenges of each experimental scenario the stakeholders and their roles which lead to the requirements and KPIs and to the market analysis and business values. This is shown in the following Figure 1. MonB5G will be demonstrated in use cases: zero touch management platform and secured platform.

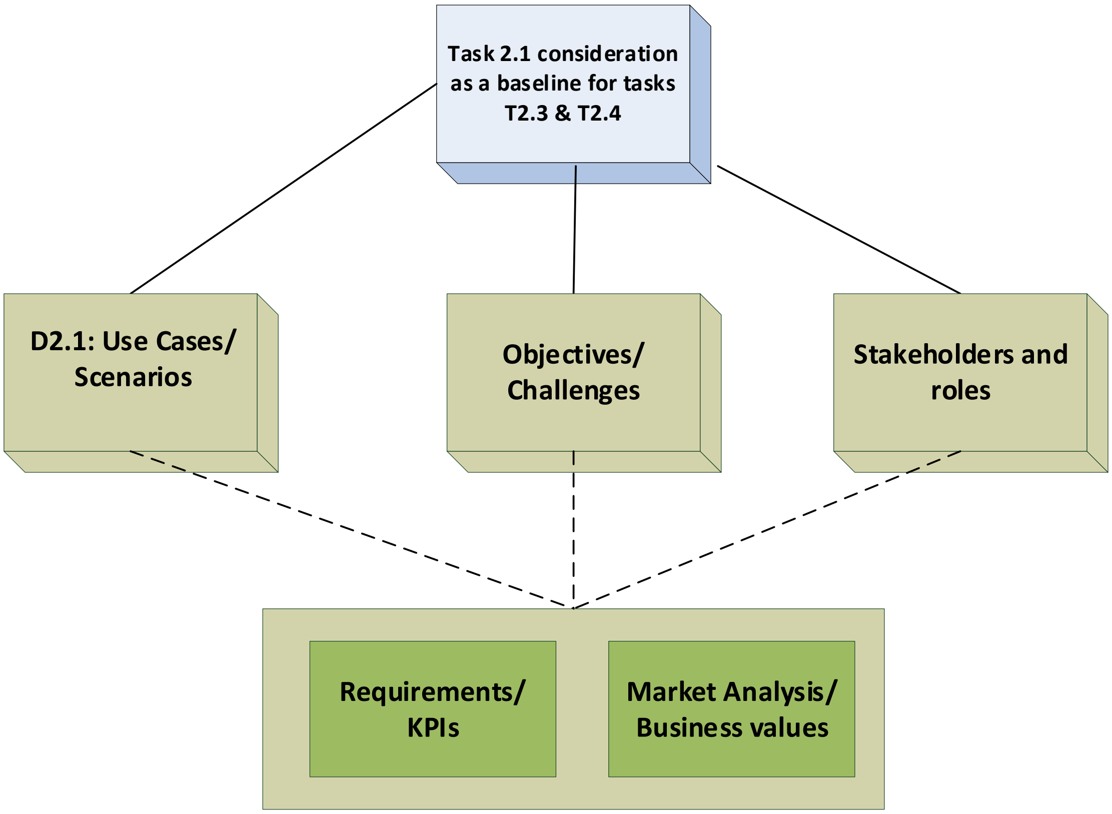

Figure 1: Overview of Task 2.1

5G Stakeholders Definitions 3GPP vs. NGMN

Before presenting the definitions of the stakeholders' roles according to our MonB5G architecture, a brief comparison of the relevant stakeholders as per the **3rd Generation Partnership Project (3GPP) [4]** the umbrella of a number of standards organizations which develop protocols for mobile telecommunications, and the **Next Generation Mobile Networks (NGMN) [1]** Alliance, a mobile telecommunications association of mobile operators, vendors, manufacturers and research institutes is provided in the following Table 2: 5G Stakeholders Definitions 3GPP vs. NGMN

.

Table 2: 5G Stakeholders Definitions 3GPP vs. NGMN

Role	3GPPP	NGMN
Network operator	Network operators are companies that provide customers with access to a telecommunications network (especially mobile phone networks) or to the internet.	**Asset Provider** One of the operator's key assets is infrastructure. Infrastructure usually is used by an operator to deliver own services to the end-customer. However, especially in the wholesale business it is common that parts of the infrastructure – so-called assets – can be

		used by a third party provider. Assets can be different parts of a network infrastructure that are operated for or on behalf of third parties resulting in a service proposition. Accordingly, one can distinguish between Infrastructure as a Service (IaaS), Network as a Service (NaaS) or Platform as a Service (PaaS). These may be summarized as Anything as a Service (XaaS). Another dimension of asset provisioning is real-time network sharing that refers to an operator's ability to integrate 3rd party networks in the MNO network and vice versa, based on a dynamic and context dependent policies (e.g., congestion/excess capacity policies).
Service provider	Service providers in telecommunication area are companies that provide its subscribers with access to the Internet or other value added services, e.g. Cloud computing, storage and e-learning. Service providers are more generally used to refer to third party or outsourced suppliers, including telecommunications service providers (TSPs), application service providers (ASPs), storage service providers (SSPs), and Internet service providers (ISPs).	Directly addresses the end customers where the operator provides integrated service offerings based on operator capabilities (connectivity, context, identity etc.) enriched by partner (3^{rd} party/OTT) content and specific applications. Integrated streaming solutions can be an example here but even services such as payments are possible.
Connectivity providers	Perform day-to-day operational activities to provide network connection via wired/wireless networks.	Another role an operator can play in the future is one of a Connectivity provider. Basic connectivity involves best effort IP connectivity for retail and wholesale customers. While this model is basically a projection of existing business models into the future, enhanced connectivity models will be added where IP connectivity with QoS and differentiated feature sets (e.g. zero rating, latency, mobility) is possible. Furthermore, (self-) configuration options for the customer or the third party will enrich this proposition.
End users	An End user is a client person who uses a product, application, solution, system, network services etc. An end user will use some services and Infrastructures designed beforehand by 5G projects.	

MonB5G Stakeholders

In the following section we will present the adapted stakeholders' roles as per our MonB5G architecture, shown in the Figure 2 and create a list of stakeholders' definition that will be used throughout the project. The following list of roles will have different interactions, as per each of the Use Case Scenarios, which will be defined accordingly within each scenario in the upcoming chapters.

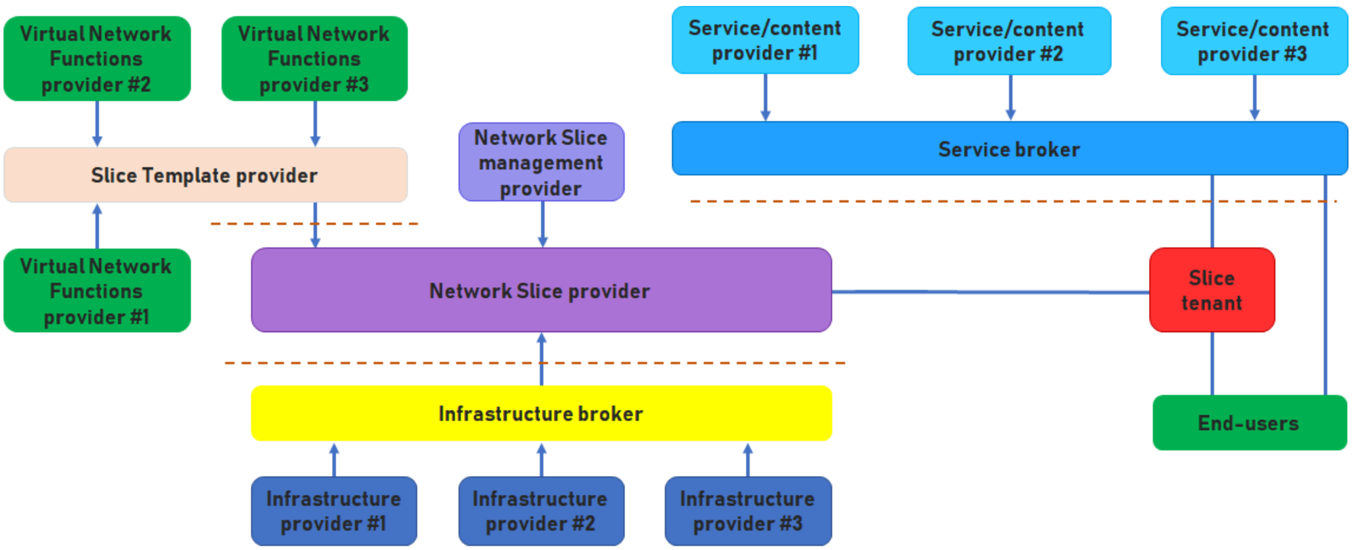

Figure 2: MonB5G's Slice lifecycle business model

List of all MonB5G Stakeholder Roles:

1. **Infrastructure provider**

Owns and manages a physical network infrastructure and its resources, including wireless radio network infrastructure, backhaul and transport infrastructure, Network Points of Presence (PoPs)? Data Centers, etc. They are responsible for the physical network deployment and maintenance. Infrastructure providers can be Mobile Network Operators (MNOs) or business third parties that interact with other stakeholders but not with end users directly.

2. **Infrastructure broker**

Integrates, enriches, and resells on the infrastructure providers' offers to the network slice providers, using also the own know-how. Exposes MANO interfaces to enable automated interworking of the entire ecosystem and smooth E2E processes flow.

3. **Network slice provider**

Creates and prepares logical, fully functional partitions of communication network architectures and provides them to slice tenants using slice templates. The slice template is provided by a slice template provider that can be different from the network slice provider. The role includes allocating/managing physical or virtual resources as needed by the kind of service dictated by the tenant, whilst guaranteeing the isolation among the allocated resources.

4. **Network slice management provider**

Provides management of the end-to-end lifecycle of network slices and services covering design, instantiation and closed-loop operations for slice providers. Slice management providers also overview the assisted learning process, policy management and enforcement, supervision of autonomic management processes, engineering rules fine-tuning or selection etc.

5. **Slice template provider**

Creates a template of a slice, which serves as a baseline for the definition of a set or service/slice types by defining the necessary interconnections between VNFs and border interfaces. Templates include integration of software of various editors with necessary E2E testing, and can also provide parametrization or engineering rules as an added value.

6. **Slice operator**

The role of the slice provider is to provide slices to tenants (customers) instantiated over its infrastructure, but also spanning different administrative domains if necessary. This entity will typically coincide with a mobile network operator, but this is not a strict requirement. Our proposed MonB5G architecture and implementation will permit recursive deployment, allowing a tenant to also act as a slice provider of its own, if needed. The slice provider composes end-to-end network slices by putting together resources and network functions of different technology domains (TDs), e.g. RAN, core, cloud, which may be managed by different entities.

7. **VNF provider**

Develops different network-oriented software components in the form of VNFs with specific functionalities to slice template providers allowing reliable scale up and/or scale out of multiple software only functions. Role includes embedding management functions in the slice templates e.g. AI algorithms, resource tuning and selection rules.

8. **Slice tenant**

This is the customer/operator of the slice. It can be a vertical service provider, which will be submitting slice deployment requests over the interface of the slice provider (part of the OSS/BSS), signing a service-level agreement on specific SLOs expressed as a set of slice or service level KPIs. Ideally, SLA conformance should be monitored in a trustworthy way, and arbitration and compensation should take place automatically, with zero or minimal human involvement.

9. **Service broker**

Rewraps, integrates and enriches the service/content providers' offer into a new product with some added value, to slice tenants, and/or end-users.

10. **Service/content provider**

Operating on top of a network infrastructure belonging to an MNO and based on a pre-defined Service Level Agreement (SLA) set of requirements, thus exploiting an MNO network infrastructure for services complementary to the telecommunication industry.

Service/content providers do not own a network infrastructure but need to interact infrastructure providers to request network resources and to negotiate SLAs. In this way operators are no longer merely suppliers of communication services, but business enablers.

11. **End-user**

An End user is an entity who will use the services/contents or slices as provided by the other stakeholders.

5G-PPP KPIs and their relevance to MonB5G is provided in the following .

Table 3: Relevance of MoNB5G to 5G-PPP KPIs

	Relevance and impact on 5G-PPP KPIs	MonB5G
	Performance KPIs	
P1	Providing 1000times higher wireless area capacity and more varied service capabilities compared to 2010.	N.A.
P2	Saving up to 90% of energy per service provided.	High
P3	Reducing the average service creation time cycle from 90 hours to 90 minutes.	High
P4	Creating a secure, reliable and dependable Internet with a "zero perceived" downtime for services provision.	High.
P5	Facilitating very dense deployments of wireless communication links to connect over 7 trillion wireless devices serving over 7 billion people.	N.A.
P6	Enabling advanced user controlled privacy.	High
	Societal KPIs:	
S1	Enabling advanced User controlled privacy;	High
S2	Reduction of energy consumption per service up to 90% (as compared to 2010);	High
S3	European availability of a competitive industrial offer for 5G systems and technologies;	High.
S4	Stimulation of new economically-viable services of high societal value like U-HDTV and M2M applications;	Medium
S5	Establishment and availability of 5G skills development curricula (in partnership with the EIT).	N.A.
	Business-related KPIs:	
B1	Leverage effect of EU research and innovation funding in terms of private investment in R&D for 5G systems in the order of 5 to 10 times;	Medium
B2	Target SME participation under this initiative commensurate with an allocation of 20% of the total public funding;	High
B3	Reach a global market share for 5G equipment & services delivered by European headquartered ICT companies at, or above, the reported 2011level of 43% global market share in communication infrastructure.	T.B.D.

.

Table 3: Relevance of MoNB5G to 5G-PPP KPIs

	Relevance and impact on 5G-PPP KPIs	**MonB5G**
	Performance KPIs	
P1	Providing 1000times higher wireless area capacity and more varied service capabilities compared to 2010.	N.A.
P2	Saving up to 90% of energy per service provided.	High
P3	Reducing the average service creation time cycle from 90 hours to 90 minutes.	High
P4	Creating a secure, reliable and dependable Internet with a "zero perceived" downtime for services provision.	High.
P5	Facilitating very dense deployments of wireless communication links to connect over 7 trillion wireless devices serving over 7 billion people.	N.A.
P6	Enabling advanced user controlled privacy.	High
	Societal KPIs:	
S1	Enabling advanced User controlled privacy;	High
S2	Reduction of energy consumption per service up to 90% (as compared to 2010);	High
S3	European availability of a competitive industrial offer for 5G systems and technologies;	High.
S4	Stimulation of new economically-viable services of high societal value like U-HDTV and M2M applications;	Medium
S5	Establishment and availability of 5G skills development curricula (in partnership with the EIT).	N.A.
	Business-related KPIs:	
B1	Leverage effect of EU research and innovation funding in terms of private investment in R&D for 5G systems in the order of 5 to 10 times;	Medium
B2	Target SME participation under this initiative commensurate with an allocation of 20% of the total public funding;	High
B3	Reach a global market share for 5G equipment & services delivered by European headquartered ICT companies at, or above, the reported 2011level of 43% global market share in communication infrastructure.	T.B.D.

KPIs which are considered in MonB5G are provided in the following Table 4: KPIs considered in MonB5G

.

Table 4: KPIs considered in MonB5G

KPI	DoW Description	Quantification	Baseline	Means of Verification (Meas. Instrument/simulator/open source meas.)	Scenario for main KPI (How we are going to prove the targeted value, explain scenario)	Comments
1.1	**3 management entities/TD**	3	Non-modular architecture	MS/AE/DE per domain (RAN, Edge, and Cloud)	MonB5G architecture	
1.2	**Reduce the reaction time to an NS malfunction**	Minimal	Absence of prediction	At the RAN level, the RAN orchestrator first detects the malfunction when it happens, then it starts to deal with it. This KPI can be measured using OAI platform and/or simulator (OAISIM).	To demonstrate this, we consider the example of managing URLLC slice resources. We will use prediction of latency at the RAN level in order to reduce the time needed for RAN resources management.	
2.1	**Improve NS performance prediction**		Absence of prediction	By predicting the future congestion of slice (high number of users and their demands) using distributed AI techniques	The slice notice congestion because of more users and their demands. Using the AI driven distributed technique (situation awareness approach) we detect the future congestion	
2.2	**Reduce SLA violations**	Minimal	A system with network slicing that uses a central algorithm to measure SLA violations.	Measured through the MS of the network slice, which has to be enabled to detect when the SLAs of the slice are not met.	Deploy a series of network slices with different SLAs into an experimental platform (either simulated or a working prototype), deploy different combinations of them, and in different orders. Deploy these network slices, enabling different types of policies for slice admission and resource allocation, and evaluate the amount of SLA violations in every instance.	Develop distributed algorithms (based on ML) for prediction of resource utilization to anticipate SLA violations, and use this prediction data to drive dynamic and distributed resource allocation policies, also based on ML.
3.1	**Reduce overhead to the central system**	Minimal	Fully centralized AI algorithm	Each data sample provided by MS or weight sent by federated learning are coded in 32 Bytes	Measure the reduction in traffic volume (MB) between the central MS/AE/AE and the distributed MS/AD/DE instances following the use of a decentralized AI algorithm (instead of raw datasets, only	

					weights are exchanged with the central entities in federated learning)	
3.2	**Maximize NS acceptance ratio**	Maximal	Offline optimization algorithms	Using a discrete event simulator to simulate network slice requests arriving on the network slice manager	Measure how many network slices requests were accepted and how many were rejected after some simulation time using MonB5G algorithms and compare it with results provided in the state of the art or by other techniques such as an offline optimization algorithm that has in advance the information of all requests that will arrive in the network slice manager	An interesting idea is to assess various scenarios with different network load conditions and compare the performance of the algorithms these different scenarios
3.3	**Minimization of NS resource consumption**	Minimal	Offline optimization algorithms	Using a discrete event simulator to simulate network slice requests arriving on the network slice manager	Measure the amount of resources consumed (e.g., CPU, RAM, bandwidth) using MonB5G algorithms and compare it with results provided in the state of the art or by other techniques such as an offline optimization algorithm that has in advance the information of all requests that will arrive in the network slice manager	An interesting idea is to assess various scenarios with different network load conditions and compare the performance of the algorithms these different scenarios
4.1	**OPEX reduction**	Minimal	Tenants/operators to manually modify or add monitoring or analytic agents to each of their services	Measure the amount of operations not performed by a human leveraging netDevOps-like operations and Container Orchestration Engines supporting MonB5G administrative components in an ETSI NFV IFA 029-fashion.	Update of MonB5G administrative element for a tenant having multiple services in his/her slice.	The proposed validation relates to an operator-initiated action to update MonB5G administrative elements for a tenant with multiple services. This action will trigger rolling updates of MS/AE/DE layer components for all services owned by a tenant.

4.2	**Reduction in the deployment time of multi-domain services**	Minimal	Absence of MonB5G architecture	Collect metrics via REST API-based telemetry agents	Implement an AR service by emulating multi-domain infrastructures Multiple instances of a Virtualized Infrastructure Manager (VIM) (i.e. OpenStack) can be deployed and in turn interconnected with an emulated transport network and instantiation of MonB5G Reference Architecture with three Technological Domains (TDs) (RAN, Edge and Cloud), as well as Central Element (i.e. Inter Domain Service Manager (IDSM) and Inter Domain Manager and Orchestrator (IDMO)),	
5.1	**Reduce time to manage RAN resources**	Minimal	Absence of prediction	RAN orchestrator measures periodically the performance of the RAN resources for example in terms of latency.	We will use prediction of latency at the RAN level to reduce the time needed for RAN resources management.	
5.2	**Improve slice performance isolation**	Maximal	Absence of per-slice queueing and scheduling mechanisms in lower part of the a 5G protocol stack.	Prototype implementations based on OAIplatforms and/or simulations. Data plane RTT measurements or latency introduced by the gNB packet processing.	Scenario related to PoC-1 (URLLC slice management)	In terms of scheduling options, several slice-ready implementations are already available. Limited contributions here?
5.3	**Reduce RAN-oriented overhead**	Minimal	Current LTE RAN deployments already collect 100s of aggregated metrics per ~15 mins. Monitoring might be relaxed or at least not augmented.	Measurement of monitoring information volumes reported to upper architectural layers.	Scenario related to PoC-2 (Slice Isolation)	Highly Related with KPI 3.1.
5.4	**Reduce slice reconfigurations**	Minimal	Absence of DRL-based slice lifecycle management	Deep reinforcement learning-based DE for lifecycle management. The inverse of the number of reconfigurations is	Compare reconfigurations between heuristic and DRL	

				included in the reward function		
6.1	**10x faster attack/ anomaly/ degradation identification**	10x - 10s	A Vanilla solution without intelligent tool to verify the positivity of a cyber event.	Measure the efficiency of the MonB5G system to immediately detect and identify a security attack/anomaly, by measuring the time taken from the start of the attack/anomaly until it is confirmed	This KPI will be measured during the validation of the scenario aLTEr attack. The realization of this scenario will leverage the components that will be implemented in T5.1/T5.2.	The anomalies raised by the detection system could be false positives, and the response system should verify the actual attack by carrying out some action plan to modify the system under protection and it requests additional detection.
6.2	**10x faster attack/ anomaly/ degradation remediation and reconfiguration 10s**	10x	Vanilla solution based on IDS and Firewall	Measure the time taken by the MonB5G system to mitigate an attack	By specific slicing-related security extensions to the ETSI MEC architecture (e.g., access control to the RNI, Location, and traffic rules services per slice), unauthorized access to sensitive information and attacks to the traffic steering procedures can be mitigated, thus achieving a level of slice (security) isolation. Second, we will showcase the efficiency of our mechanisms in identifying attacks originating from the edge and limiting their scope in order to avoid their propagation.	
6.3	**False positive rate < 1%**	<1%	No misbehaving elements that takes part in learning.	The KPI should stay below the threshold even in the presence of misbehaving elements. We can also calculate to which extent the learning algorithm should be compromised to exceed the threshold.	As in KPI 6.2	This KPIs is related to the robustness of learning algorithms in face of attacks. In PoC-2 Experimental scenario 2, leveraging trust and consensus schemes, the effect of attack on the performance of

						learning algorithm will be limited.
6.4	**E2E slice availability > 99%**	>99%	Vanilla solution based on IDS and Firewall	Measure, during an attack, the downtime of the of the end-to-end slice.	As in KPI 6.2	
6.5	**Slice isolation < 5% degradation**	<5%	A network slice instance is deployed without any control and enforcement mechanism to prevent interference generated by the activity of another NSI	Measure the efficiency of the MonB5G DE system to anticipate and avoid the inter-slice isolation violations in the case of a side channel attack carried out by another slice.	Leveraging the DE for VNFC node failure investigation and localization and taking actions to reduce the E2E downtime of the services	The degradation will be characterized by a high packet loss rate, an increase of latency or a bad QoE
6.6	**Learning robustness (values below specific thresholds)**	Maximal	Vanilla solution based on IDS	Measured by the performance of ML running by the AE and DE components, we will generate a high number of act on the high number of emulated network slices	Described in PoC2	
7.1	**Improve EE by a factor of 10**	10x	Basic scenario with no DRL	Include the energy consumption in the reward function of DRL-based DE, and perform energy-aware VNF placement	Compare the performance w and w/o DRL	
7.2	**Guarantee that the vertical application KPIs and SLA remain unaffected**	Maximal	No MonB5G Administrative components reallocating resources to guarantee SLA compliance.	When we prove the EE x10, the other KPIs should not be affected	Service-level KPIs for specific scenario should be maintained despite MonB5G Administrative Components performing ML/AI-based energy saving strategies that would normally degrade performance in legacy systems (e.g. NFVI cluster reduction, lowering MCS at RAN).	Baseline scenario should be well-known, and its respective service-level KPIs controlled. Energy-savings strategies must occur at different technological domains.

4.2 Use cases for PoC introduction

The purpose of the MonB5G use cases (UC) is to present a different approach compared to other projects, in order to demonstrate the benefits of an autonomic management platform for increasing end user experience and the external attacks mitigation into a network. In this section the four MonB5G experimental scenarios, the main KPIs, their requirements, their mapping to the architecture are thoroughly presented. MonB5G stems around two use cases that deal with zero touch and secure management platform. The Experimental Scenarios (ES) described in this project are providing the means to result to a zero-touch management and a secured infrastructure that can be bring new benefits to the involved stakeholders. This is shown in the following Figure 3.

Figure 3: MonB5G ecosystem with the two Use Cases

MonB5G project differs from other 5G-PPP projects in terms of its scope which is to provide a secured, flexible and zero touch infrastructure. The description and the key points of the two use cases are described in the following two sections.

4.3 Use Case 1: Zero-Touch Network and service management with end-to-end SLAs

In this use-case, the MonB5G distributed mechanisms (i.e., the MS, AE and DE) are evaluated and showcased in a Tactile Internet application in two different experimental scenarios (ESs) that cover i) cross-domain (i.e., RAN, edge and cloud) end-to-end SLA and ii) elastic end-to-end slice management, respectively. The considered Tactile Internet application corresponds to Augmented Reality (AR) for virtual event attendance and presents, thereby, high availability and ultra-low latency requirements as well as has a palpable impact on business and society. This use case demonstrates the AI-driven zero-touch slice management in both intra-slice or inter-slice levels and is evaluated through a plethora of practical KPIs.

After describing and analyzing each experimental scenario, the objectives as well as the functional and non-functional requirements thereof are presented. The mappings of the ES to both MonB5G architecture and the experimental platform are then provided, and the practical KPIs involved in the ES as well as their measurement means are detailed. Finally, the societal and business impact of the ES is reported.

4.3.1 UC1/ES1: ZERO-TOUCH MULTI-DOMAIN SERVICE MANAGEMENT WITH END-TO-END SLAS

4.3.1.1 DESCRIPTION AND DETAILED ANALYSIS

In this experimental scenario, multiple NFVIs, hosted in both project testbeds as well as on AWS infrastructure will be combined to demonstrate Zero-Touch service management in complex multi-domain services. The Storage, Compute, and RAN functions of the Tactile Internet application will be hosted in different regions under the control of local NFVOs and Decision Engines, but end-to-end SLA must be honoured. Continuous monitoring and closed-loop autonomic control mechanisms, which will be common across regions and testbeds, will ensure self-healing, self-configuring and self-scaling of services, to address faults and performance issues in any of the service technological domains.

The main objective of this ES is to satisfy the stringent requirements of a real-time Tactile Internet application leveraging MonB5G Administrative Elements (i.e. MS, AE, DE), which autonomously make local decisions (e.g. service adjustments, resource orchestration requests, etc) in order to significantly reduce the number of end-to-end SLA violations. Such service-level violations are identified and prevented by MonB5G Central Element, which can compose end-to-end service KPIs from MonB5G Measurement Systems at each Technological Domain (i.e. NSSI).

This use case will be implemented and evaluated at CTTC's testbed. The target application to be considered is Augmented Reality (AR) for virtual event attendance. Indeed, AR is a representation scheme wherein objects in the real world are either masked or supplemented with artificial digital objects in a constructive way. The Pokémon Go game released in 2016 [32] is an example of a constructive AR experience wherein computer-generated characters are overlaid into the real-world view. The latency requirement for the overall end-to-end chain in AR is stricter as visual changes are not only triggered by the motion of the user but also by any change (e.g. lighting or natural object movement) in the surrounding environment. Some sources identify the limit for a perfect AR experience at below 10ms [9] which is in line with the standardized 5G Quality Class Identifier (5QI) 80 introduced in 3GPP TS 23.203 Rel. 15 [10], and summarized in Table 5.

Table 5: 3GPP 5QI for AR

5QI	Resource Type	Priority	Packet Delay Budget	Packet Error Loss Rate	Example Services
80	non-GBR	6.8	10 ms	10^{-6}	Low latency eMBB applications (TCP/UDP-based); Augmented Reality

Therefore, this use case will leverage the highly distributed MonB5G mechanisms (i.e., the MS, DE and AE) to provide automated, zero-touch service management across domains, and ensure end-to-end cross-domain SLA by continuously monitoring the QoS (in this case the 99%-percentile of latency is more accurate than average) across all technological domains by the local monitoring systems, E2E objective QoE metrics for AR such as MS-SSIM and VMAF [11] and energy consumption. Based on all these measurements that serve as an environment input, Decision Engines (DEs) decentralized reinforcement learning algorithms will manage the life-cycle of the AR slice created according to an URLLC template and automate the control of the intra and inter-slice interactions and adaptation to optimize the multi-objective cost of both E2E SLA violation and energy-efficiency. Note that this URLLC slice will rely on edge cloud to address the latency/jitter requirements by bringing the computing resources in close proximity to the user equipment.

In terms of equipment, the AR use case will be based on a hardware AR kit (e.g., Microsoft HoloLens [33]) that, thanks to its integrated camera, will transfer the current field of view to the remote attendee. Through a desktop application, the remote attendee can also provide his remarks by writing/drawing 3-Dimensions (3D) objects into the field of view. These holographic objects will appear to be real for the person wearing the kit and are used for guidance. Note that there are many open source software project implementations that can work with AR kits. Otherwise, Microsoft Dynamics 365 [34] can be used.

4.3.1.2 OBJECTIVES/CHALLENGES FOR UC_1/ES_1

The main objective of this experimental scenario is to assess the data-driven management systems in a multi-domain scenario with respect to their ability to guarantee the stringent end-to-end SLA of the Tactile Internet application. Automated zero-touch service management and multiple redundancy mechanisms must ensure practically zero downtime due to the critical, high-availability operation of the platform.

Realising such a collection of mechanisms implies the involvement of a considerable amount of service and resources managers, not only for each technological domain (i.e. Cloud, Edge, RAN) but also for each tenant. Current specifications recommend a somehow centralised approach, where AI/ML techniques are leveraged to learn from previous events and prevent SLA violations. Such intelligence can be produced leveraging extensive telemetry and big data techniques, creating control loops between components of the Assisted System (AS) and the central manager (which may directly trigger operations or only provide recommendations to AS).

The caveat with such approach (detailed at ETSI ENI system architecture v1 [35]) is its centralised nature. Managing a great number of geographical locations (i.e. NFVI), tenants and multi-PoP Slices (e.g. URLLC slices) not only would require a considerable increase in telemetry traffic traversing the network segments connecting each managed element of the AS and the central manager, but would potentially raise scalability and security considerations of the central manager entity itself (e.g. how bit can it get, single point of failure, etc).

For solving this issue, MonB5G architecture proposes a completely distributed and decoupled architecture. The former feature enables different control loops (e.g. ranging from small at VNF level, to big at end-to-end Slice level), while the latter feature guarantees compliance with current specifications, e.g. realising a common MonB5G administrative layer for one or more tenant services and resources as VNF Common/Dedicated Service [36]. Under these specifications, MonB5G administrative layer could be orchestrated alongside a tenant's Network Slice Service Instance (NSSI), ensuring close control loops (i.e. MS, AE and DE) which minimize out-of-slice control traffic. Such properties help enable challenging low-latency scenarios such as this experimental scenario.

4.3.1.3 MAPPING OF THE EXPERIMENTAL SCENARIO TO MONB5G ARCHITECTURE

The initial architecture concept of MonB5G was presented and described in Annex-1 of the DoA. Despite provision of the analysis of the architecture is out of the scope of this deliverable, it is depicted in Figure 4 for clarity purposes. It also serves for mapping the functionality and operation of the experimental scenario with respect to the architecture and presents how the architecture will provide solution to the objectives/challenges of this ES. The main actors and stakeholders that can be involved in the commercial deployment have been also presented in this figure. From this it can be found the pros and cons of the different players for the MonB5G implementation. More details regarding the architecture of MonB5G will be presented and analyzed in the upcoming deliverable D2.1.

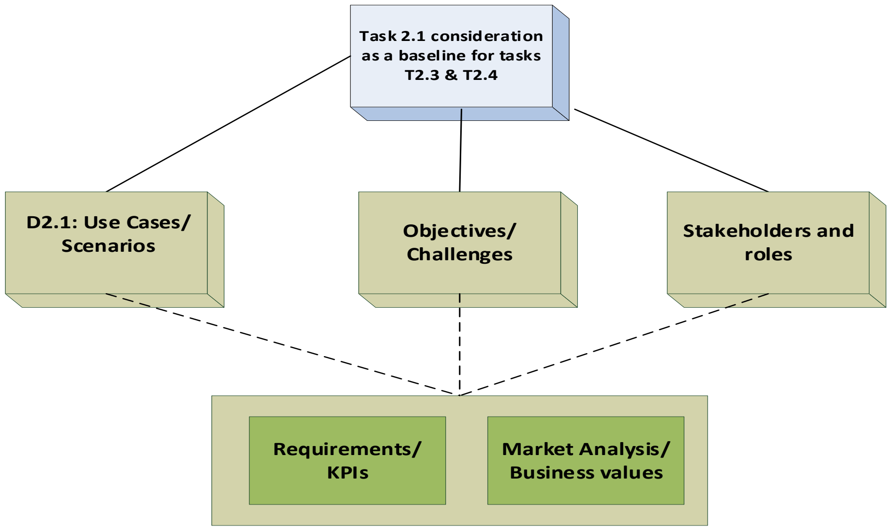

Figure 4: Mapping of scenario and Interaction of stakeholders for UC_1/ESES1 in MonB5G

4.3.1.4 REQUIREMENTS ANALYSIS AND KPIS FOR UC1/ES1

This section describes the functional and non-functional requirements for the management of the MonB5G platform and discussion of KPIs/KQIs as outlined in [37]. These are detailed as follows:

FUNCTIONAL REQUIREMENTS

This paragraph includes functional requirements that specify the behaviour or function of the solution that will be studied and any other functionality that defines what the experimental scenario is supposed to accomplish. These requirements consider functional blocks on the user terminals, on the wireless connectivity, on the BB and virtualization, on MEC, slicing, and management and orchestration. They are as follows:

- **[UC1/ES1 F.req-1] UE category:** A high capability UE is required to support both DL/UL high throughput and ensure the low radio interface latency needed by AR. Some possible choices:
  - DL Category 13: 4 layers, 256QAM
  - DL Category 14: 8 layers, 256QAM
  - UL Category 16/ UL Category 17: Both 64-QAM and 256-QAM

  Note that UE category is defined by the couple (ue-CategoryDL, ue-CategoryUL).
- **[UC1/ES1 F.req-2] Wireless connectivity:** To ensure low latency at the radio interface, it is required to combine both MIMO and large bandwidth at sub-6 GHz, eventually using carrier aggregation. The corresponding baseband unit (CU) will run on a high-performance computing hardware to process the traffic in fractions of milliseconds.
- **[UC1/ES1 F.req-3] Baseband unit virtualization:** To allow for a flexible allocation of computing resources for baseband processing, it is required that BBUs be virtualized and instantiated on-demand, at any location where resources are available.
- **[UC1/ES1 F.req-4] MEC:** To achieve low latency for AR service, it is also required that MonB5G leverages MEC to bring the computing capabilities to the edge and exploit them to dramatically reduce latency.
- **[UC1/ES1 F.req-5] Slicing:** Slicing is required to achieve the necessary isolation of the AR slice and guarantee its quality of service (especially the low latency as defined in the UC1 description).

- **[UC1/ES1 F.req-6] Management and orchestration:** A management and orchestration (MANO) platform is also required to monitor and manage both the physical and virtual network functions, perform the orchestration of VNFs for the AR tenant/slice/service.

NON-FUNCTIONAL REQUIREMENTS

The non-functional requirements (scalability, interoperability, modularity and security) have been identified. As for the functional requirements, the following high-level requirements have been identified:

- **[UC1/ES1 NF. req-1] Scalability:** The implementation of the use case should be scalable and allow for the automatic admission of multiple concurrent AR users/sessions/slices without modifying the architecture.
- **[UC1/ES1 NF. req-2] Interoperability:** The setup of the use case should be interoperable with any type of user device/AR kit.
- **[UC1/ES1 NF. req-3] Modularity:** The implementation should be modular to a great extent in such a way the upgrade of one component can be independently done without affecting the other components.
- **[UC1/ES1 NF. req-3] Security:** This experimental scenario leverages the zero touch Management which is based on data collection and also on AI/ML models. Protecting these models against evasion or poisoning attack is a major requirement. The robustness of federated learning algorithm is an essential meet.

KPIs

- **[UC1/ES1 KPI-1]: Reduction of SLA Violations**

An SLA is defined as a formal agreement between two or more entities that is reached after a negotiation with the scope to assess service characteristics, responsibilities and priorities of every part of the infrastructure. This agreement establishes different aspects of the service, such as performance, availability, among others. In general, the SLAs include the metrics that quantitatively characterize the communication service. According to the 5G architecture specification provided by the 5G PPP Initiative [2] the SLA requirements could be:

- End-to-end latency and bandwidth requirements, necessary for the service to function
- Metrics related to the dimensioning of the service (number of users supported, area of coverage)
- Availability and Reliability
- Optimization targets for service deployment, that could include deployment time and energy efficiency

Each network slice is specified with its own SLA constraints [[15], [17]-[20]], that determines the resources it needs. Depending on the SLA specifications, it is necessary to map a resource configuration to the slice requirements that fulfils the SLA constraints or to a series of QoS specifications. For example, according to the 5G PPP Initiative, bandwidth requirements at the SLA level can be mapped to bandwidth requirements on the links of the slice [2], which in turn translates to physical bandwidth resources being allocated to those virtual links.

By making this resource mapping as accurate and precise as possible, the resource utilization of the infrastructure can be increased by deploying the largest number of slices that the infrastructure can support [21]. Very generally, an SLA violation occurs whenever the infrastructure is not able to provide the resources that slices need in order to fulfil their performance constraints (latency, bit rates, etc). When multiple slices are running simultaneously, there are multiple factors that can cause the SLA of a slice to be violated, among which we can mention:

- Overbooking of infrastructure resources [21], [22].
- Sudden spike in resource demand, which could manifest itself as traffic overload, or an increase in the requirements of the VNFs of the slice.
- Physical failure of substrate links [24]

Sharing the resources to support multiple slices might increase the chance of a SLA violations, because it is possible that resource overbooking might occur [13]. If resource overbooking is not carefully managed, then the performance penalties suffered by the slice owners could be very high. However, if some degree of resource sharing is not considered, then the efficiency of the resource utilization of the infrastructure will drop, reducing the profitability and the revenue of the infrastructure operator. For this reason, a solution that over allocates resources to all slices regardless of their SLAs is not desirable. Therefore, a careful balance needs to be struck between the sharing the resources and ensuring that the SLAs are always met.

There are many solutions and techniques that can be implemented in order to ensure SLA compliance and efficient resource allocation [5, 10–12, 14, 15]. In [11], the authors propose a slice recovery mechanism for deterministic traffic demands that models the failure recovery problem as a mixed integer program (MIP). The recovery mechanism works by rearranging the virtual-to-physical resource mapping whenever the physical resource fails and remapping the virtual resource to a different physical counterpart.

In [5], the authors propose a 5G Network slice orchestrator that seeks to increase the amount of slice requests that get deployed in a physical 5G infrastructure, while ensuring that their respective SLAs are met. Their orchestrator models the slice admission control problem as a D-dimensional Multiple Choice Knapsack problem [38], constrained by the limits corresponding to each slice.

In [10], the authors make use of Deep Neural Network consisting of Convolutional Neural Networks to predict the slices' traffic load through the RAN cells of the infrastructure, in order to ensure that the RAN resources would get allocated to the slices in an efficient way. With this DNN, the authors designed customized loss functions that model the costs associated with SLA violations and resource over allocations, in order to optimize the resource allocation.

In the context of MonB5G, we will predict the occurrence of SLAs at different technological domains, making contributions to the state-of-the-art in this respect. Subsequently, we would use our best-performing predictor, make a distributed version to provide a MonB5G compliant implementation based on its architectural components (MS/AE/DE), and then use those predictions to mitigate the occurrence of SLA violations, or reduce them entirely.

The measurement of an SLA violation takes place at each domain, where the status of every component of the slice (VNFs/PNFs) is monitored, in order to make sure that the performance metrics that they present are compliant with the SLA constraints. A violation will be observed whenever a component is not able to maintain the performance metrics within a certain threshold.

Other KPIs which are going to be considered and assess the performance of the MonB5G infrastructure are:

- **[UC1/ES1 KPI-2]:** Reduce the reaction time to an NS malfunction
- **[UC1/ES1 KPI-3]:** Improve NS performance prediction
- **[UC1/ES1 KPI-4]:** Reduce Overhead to the central system
- **[UC1/ES1 KPI-5]:** Support more NS instances
- **[UC1/ES1 KPI-6]:** OPEX reduction.
- **[UC1/ES1 KPI-7]:** Reduce time to manage RAN resources.

- **[UC1/ES1 KPI-8]:** Improve of slice performance isolation.
- **[UC1/ES1 KPI-9]:** Reduce RAN-oriented overhead

4.3.1.5 MAPPING OF UC1/ES1 TO THE EXPERIMENTAL PLATFORM

The following Figure 5 describes the components and functions of the platform that are going to be used to assess and verify the objectives/challenges, requirements and KPIs of the experimental scenario. This POC will be implemented in IQU. The platform proposed in Figure 5 can be configured to emulate multi-domain infrastructures. Excluding the RAN domain, multiple instances of a Virtualized Infrastructure Manager (VIM) (i.e. OpenStack) can be deployed and in turn interconnected with an emulated transport network. This is also true for the Edge Point-of-Presence (PoP). Sticking to a single NFVI, Figure 6 provides and instantiation of MonB5G Reference Architecture (as proposed in Figure 4) with three Technological Domains (TD), a Central Element (i.e. Inter Domain Service Manager (IDSM) and Inter Domain Manager and Orchestrator (IDMO)), and Administrative Elements tailored to specific goals at each TD as an example.

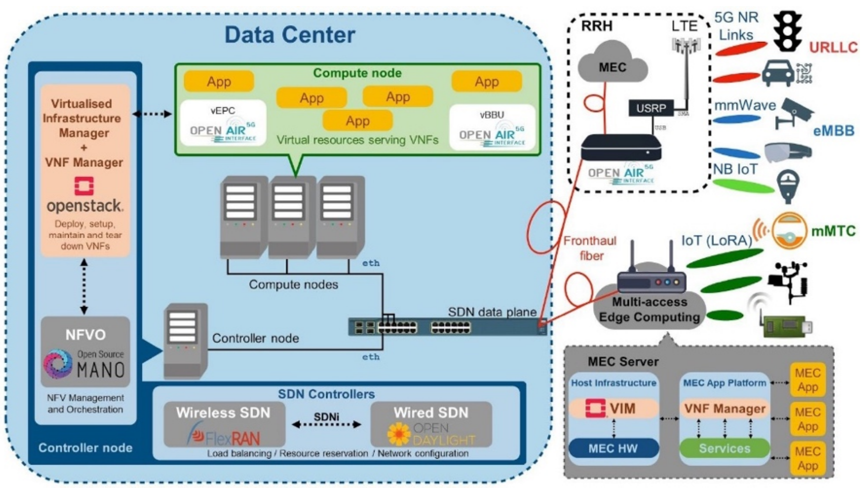

Figure 5: CTTC´s PoC platform in Barcelona

Technological Domains in the figure are isolated in three boxes for descriptive purposes. These are connected together via emulated transport networks segments as overlays running on top of a 10Gb Ethernet network, i.e. with enough capacity and flexibility to recreate a wide variety of network conditions. Said emulated transport network is managed by an SDN Controller, which is aware for the virtual network services thanks to APIs exposed by VIM (i.e. OpenStack Neutron ML2 plugin[1]) and is able to coordinate the horizontal stitching of NSSI at each NFVI PoPs with the help of NFVO Wide-Area-Network Infrastructure Manager (WIM) plugins[2].

The Tactile Internet application is supported by MonB5G Administrative Elements at:

- **RAN:** its configurability stems from the RAN Controller service, which interacts with RAN PNFs (gNB in Figure 6:) and exposes reference points to admit management actions, policies, and interaction with local TD MS/AE/DE elements.
- **Edge:** enabling local KPI monitoring and advanced LCM operations (e.g. MEP LCM, MEC Applications

[1] https://wiki.openstack.org/wiki/Neutron/ML2

[2] https://osm.etsi.org/wikipub/index.php/WIM

LCM). Moreover, reference points among MonB5G Administrative Elements and local MEO enables autonomous infrastructure management (e.g. virtual resources scaling).

- **Cloud:** hosting application elements requiring greater computation or storage resources. This TD also hosts the tenant's MonB5G Central Element (IDSM and IDMO), which collects from all TDs' MS/AE in order to compute end-to-end service KPIs. Furthermore, IDSM/IDMO leverage AI/ML algorithms to perform preventive operations and avoid SLA violations (e.g. creation of replicas, slice-wide reconfigurations).

In the proposed Instantiation for this experimental scenario (Figure 6), each TD per tenant is composed of Functional and MonB5G Administrative layers. The former holding the tenant's VNFs, while the latter holds VNFs for supporting a Platform as a Server (PaaS) in the form of a Container Infrastructure Service Instance (CISI) and Container Infrastructure Service management (CISM), an option proposed in ETSI NFV IFA 029[3].

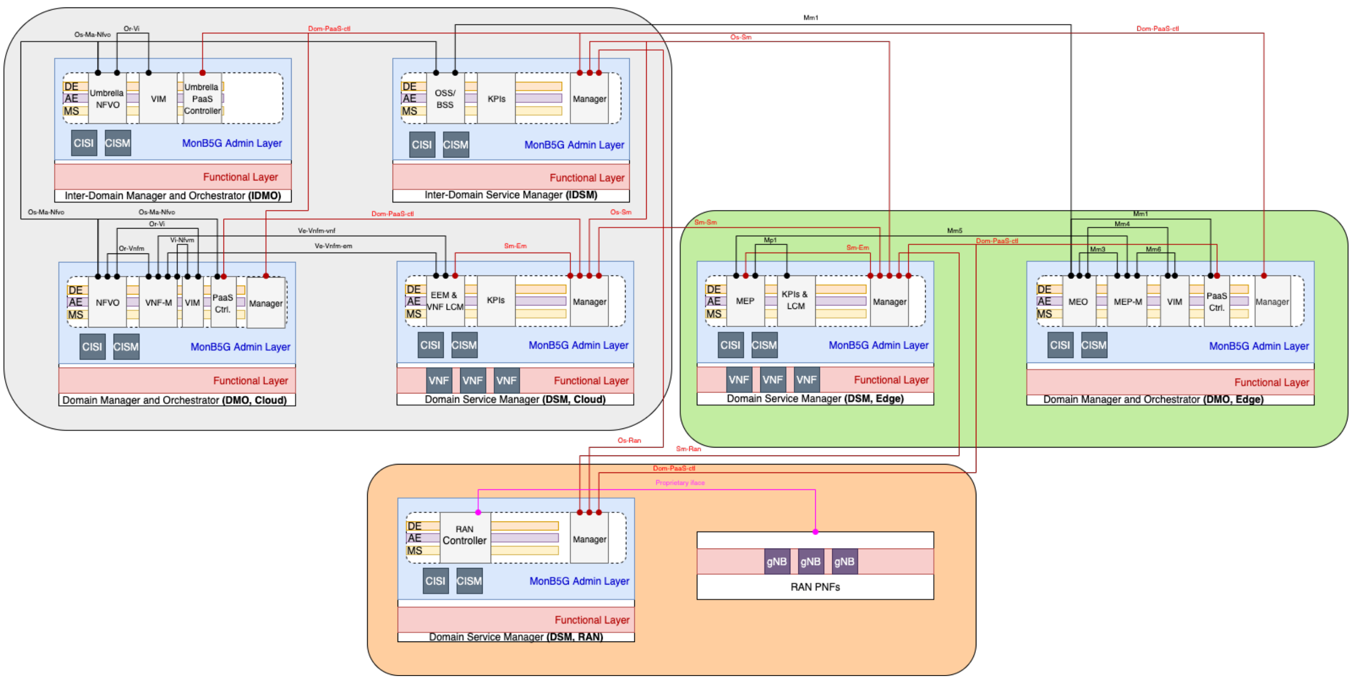

Figure 6: MonB5G Architecture Instantiation for a Single Tenant and three Technological Domain

MonB5G Management Elements avoid SLA violations by performing service-level operations by DSM (e.g. KPI monitoring, data plane security, VNF LCM), and infrastructure adjustments leveraging a Domain Manager and Orchestrator (DMO). DSM and DMO, as well as the Inter-Domain equivalents (i.e. IDSM, IDMO) possess the same structure as the DSM in Figure 7. The overall deployment can be observed in Figure 7. That is, MonB5G Admin Layer is orchestrated alongside the Functional Layer of each TD (i.e. NSSI). Administrative Elements (i.e. MS/AE/DE) are realized as namespaces in the CISI, holding agents performing specific functions (e.g. measure % CPU utilization for a MS agent). Microservices leverage interaction among agents (i.e. MonB5G Administrative Elements) to e.g. avoid SLA violations at each TD, provide security, among other objectives for the TD. In turn, IDSM/IDMO can feed from TDs Administrative Elements to compute end-to-end KPIs or perform slice-wide reconfigurations.

[3] https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/029/03.03.01_60/gr_NFV-IFA029v030301p.pdf

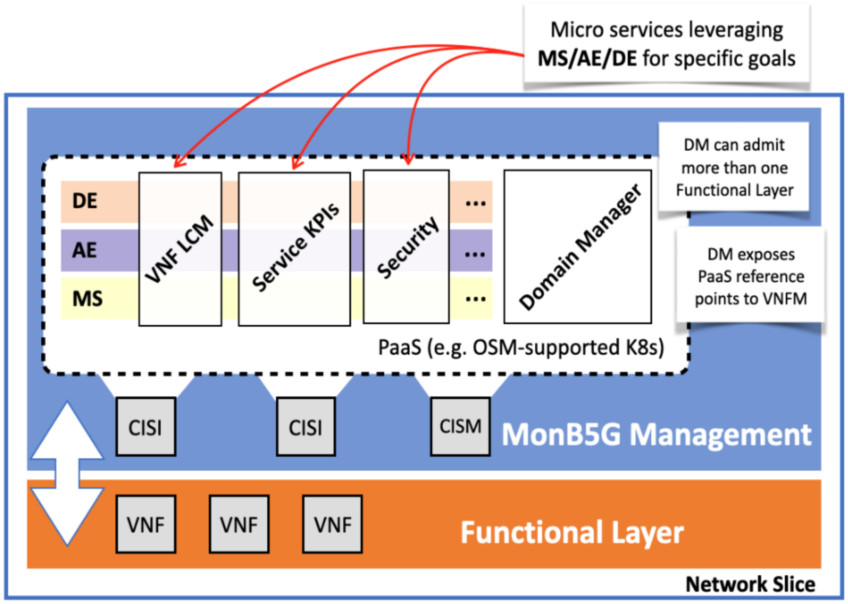

Figure 7: MonB5G Domain Service Manager (DSM) instantiation example

4.3.1.6 PLANNING MEASUREMENT MEANS FOR KPIs/KQIs

This section discusses the means of how it is planned to measure/assess each targeted KPI. It is crucial to have a knowledge from the initial steps before the integration of the various elements of how to perform measurements for test

- **[UC1/ES1 Meas. KPI-1]:** Reduce SLA violations: It may be measured through the MS of the network slice, which has to be enabled to detect when the SLAs of the slice are not met. In addition, it can be performed through the analysis of tradeoff between resources allocation and violation of KPIs.
- **[UC1/ES1 Meas. KPI-2]:** Reduce the reaction time to an NS malfunction: at the RAN level, the RAN orchestrator first detects the malfunction when it happens, then it starts to deal with it. This KPI can be measured using OAI platform and/or simulator (OAISIM).
- **[UC1/ES1 Meas. KPI-3]:** Improve NS performance prediction: by predicting the future congestion of slice (high number of users and their demands) using distributed AI techniques.
- **[UC1/ES1 Meas. KPI-4]:** Reduce Overhead to the central system: It measures the reduction in traffic volume (MB) between the central MS/AE/AE and the distributed MS/AD/DE instances following the use of a decentralized AI algorithm (instead of raw datasets, only weights are exchanged with the central entity in federated learning).
- **[UC1/ES1 Meas. KPI-5]:** Support more NS instances: Using a discrete event simulator to simulate network slice requests arriving on the network slice manager.
- **[UC1/ES1 Meas. KPI-6]:** OPEX reduction: It measures the amount of operations not performed by a human leveraging NetDevOps-like operations and Container Orchestration Engines supporting MonB5G administrative components in an ETSI NFV IFA 029-fashion.
- **[UC1/ES1 Meas. KPI-7]:** Reduce time to manage RAN resources: The RAN orchestrator measures periodically the performance of the RAN resources for example in terms of latency. This KPI may be measured through simulation and/or prototype using OAI platform or simulator (OAISIM).
- **[UC1/ES1 Meas. KPI-8]:** Improve of slice performance isolation: it will be measured via a prototype

implementation based on OAI/srsLTE platforms and/or simulations.

- **[UC1/ES1 Meas. KPI-9]:** Reduce RAN-oriented overhead: measurement of monitoring information volumes reported to upper architectural layers.

4.3.1.7 SOCIETAL AND BUSINESS IMPACT OF UC1/ES1

This section discusses the business impact/benefit of this experimental scenario that brings to the actors, stakeholders' involvement and interaction, OPEX/CAPEX reduction, impact to the society in general.

For this ES, the defining assessment of impact/benefit is through **Tactile Internet Application**, which from the societal/end-users' perspective will be an **Augmented Reality (AR)** application for a virtual event attendance.

AR combines digital information with the physical environment, in real-time. The use of **AR** for the casual user has grown in recent years through mobile applications and games, such as the famous Pokémon Go game played by 1 Billion users as of February 2019. The growth is on-going and exponential, where the augmented (**AR**) and virtual reality (**VR**) **market size** worldwide is **18.8bn USD**, on June 2020, essentially tripled since 2016, and doubled since 2019 according to a recent Statista report[4].

Additionally, the recent global **COVID-19 pandemic** has impacted the world unexpectedly, and some of the most affected areas are public events. Major global conferences and events have been cancelled or delayed, while others moved them virtually. This situation brings a whole new level of urgency and value to applications such as AR for event attendance such as in our **MonB5G UC1/ES1**, moving this solution rapidly from luxury or novelty into an essential need. Where stakeholders of various sectors do not have the possibility of joining an event physically, some of the events that simply switched "virtual" by using a traditional teleconference approach have lost substantial impact and reach, thus AR/VR would fill this gap by placing the user within the event, as if physically there, whilst still remote. Such Tactile technology applications' quality are governed by high **QoS**, and **QoE** metrics, thus achieving the targeted KPIs in our experimental scenario; **QoS** (99%-percentile of latency), and **QoE** (e.g. MS-SSIM and VMAF), would in turn prove an elevated and enhanced experience to societal usage.

4.3.2 UC1/ES2: ELASTIC END-TO-END SLICE MANAGEMENT

This section describes the second experimental scenario of use case 1 along with the requirements, KPIs and challenges involved in this case.

4.3.2.1 DESCRIPTION AND DETAILED ANALYSIS

In this experimental scenario, multiple NFVIs, hosted in both project testbeds as well as on AWS infrastructure will be combined to demonstrate Zero-Touch service management in complex multi-domain services. In this ES the Storage, Compute, and RAN functions of the Tactile Internet application will be hosted in different regions under the control of local NFVOs and Decision Engines, but end-to-end SLA must be honored.

[Req. Continuous monitoring and closed-loop autonomic control mechanisms, which will be common across regions and testbeds, will ensure self-healing, self-configuring and self-scaling of services, to address faults and performance issues in any of the service technological domains.

[4] Augmented reality (AR) - statistics & facts (June 2020), Statista.com

The goal of this use case is the evaluation of the MonB5G architecture and algorithms in a Tactile Internet application, with very stringent availability and ultra-low latency requirements. In particular, the specific application that will be considered is Augmented Reality (AR) for virtual event attendance. In this experimental scenario, a massive number of slices will be emulated. As the number of network slice instances (NSIs) grows, the complexity of the lifecycle management increases, making it mandatory to automate the whole slice management and reconfiguration process.

Each NSI will consist of multiple Network Slice Subnet Instances (NSSIs), generally one per technological domain, while each technological domain in MonB5G has its own AI-enhanced, data-driven components (i.e., MS, AE, DE). The massive number of slices in addition to the Tactile Internet application NSI will be considered, with the aim of demonstrate the following points:

- Continuous monitoring by the local Monitoring Engine in each Network Slice Subnet Instance to detect performance problems at appropriate time scales.
- DE at each domain able to solve local faults and provide model updates to the central DE.
- Proactive actions will be taken at the DE, based on traffic pattern forecasts and on sub-slice real-time data, in order to prevent the end-to-end SLA violations. Furthermore, proactive actions and decisions will be implemented in the respective domain controllers.

4.3.2.2 OBJECTIVES/CHALLENGES

The main objective of this experimental scenario is to assess the data-driven management systems in a multi-domain scenario w.r.t. their ability to guarantee the stringent end-to-end SLA of the Tactile Internet application. Automated zero-touch service management and multiple redundancy mechanisms must ensure practically zero downtime due to the critical, high-availability

The ability of the MonB5G techniques to react to local performance issues in multiple technological domains will be assessed, as well as at changes in the traffic patterns in different timescales. The goal is to guarantee almost zero latency in Tactile Internet applications by proactively acting against peaks of user demand, instead considering classical reactive strategies. Furthermore, special emphasis will be put on new AI-enhanced, data-driven Radio Resource Management (RRM) techniques to optimize the RAN sub-slice.

4.3.2.3 MAPPING OF THE EXPERIMENTAL SCENARIO TO MONB5G ARCHITECTURE

Despite provision of the analysis of the architecture is out of the scope of this deliverable, it has been depicted in Figure 8 for clarity purposes. It also serves for mapping the functionality and operation of the experimental scenario with respect to the architecture and presents how the architecture will provide solution to the objectives/challenges of this ES. The main actors and stakeholders that can be involved in the commercial deployment have been also presented.

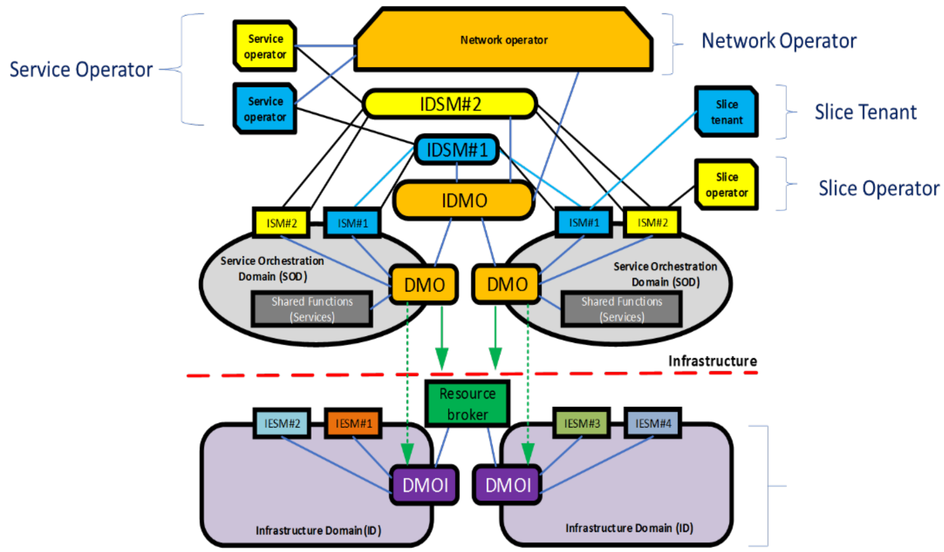

Figure 8: Mapping of experimental scenario and Interaction of stakeholders for UC1/ES1 in MonB5G

4.3.2.4 REQUIREMENTS ANALYSIS AND KPIS

The MonB5G architecture aims to provide an end-to-end solution for the deployment and management of network slices. In this experimental scenario, zero-touch management solutions are key features to tailor the network resource allocation with the actual traffic demand in a dynamic and automatic way, finally enabling higher levels of energy efficiency. At the same time, more accurate resource allocation would allow for a reduction of operational expenditures (OPEX) as well as for higher Quality-of-Service provisioning towards the end-users.

FUNCTIONAL REQUIREMENTS

- **[UC1/ES2 F.Req.-1]** Monitoring information: it is required to have the limited set of physical (and virtual) resources deployed along the network to determine the state of the network before enforcing any slice allocation decision. However, considering the multi-domain and multi-tenant framework represented by the 5G mobile infrastructure, and some of its unique features like mobility and wireless communication.
- **[UC1/ES2 F.Req-2]** NSI and NSSI continuous monitoring: It is required to have a multi domain uninterruptible monitoring for NSI and NSSI to offer a high level SLA to the end users.
- **[UC1/ES2 F.Req-3]** The functional elements should prove by means of ML-based solutions e.g., Variational Auto-Encoders (VAE), to compress the multi-dimensional information, and generate synthetic KPIs related to both network slice management and network slice orchestration. Such synthetic KPIs are particularly suitable for the zero-touch network management envisioned in the MonB5G project, as many ML-based solutions (including VAEs) can be pre-trained offline before deployment, i.e. ensuring an adequate level of performances since their first instantiation.
- **[UC1/ES2 F.Req-4]** This ES should develop closed feedback loops (backpropagation), the overall encoding process could be refined at runtime to optimally adapt it to the real-time slice traffic demands.
- **[UC1/ES2 F.Req-5]** An autoencoder architecture should be developed to provide the possibility to bring the encoded synthetic KPIs (which may be inherently difficult to be interpreted by humans) and

recover them to their original state. This decoding operation would eventually allow both tenants and service providers autonomously and promptly assess specific service performances in case of need.

- **[UC1/ES2 F.Req-6]** New monitoring elements should be developed to be able to analyse a massive set of running slices, focusing on inter-slice isolation aspects. In this context, multi-dimensional information is used to obtain a multi-slice and multi-domain representation of the system. On the one hand, this allows to overcome the limitations of legacy approaches that are typically linked only with a partial view of the overall system. On the other hand, this opens at the definition of a new set of KPIs specifically designed for the massive slice support commitment, e.g., KPIs to measure the numbers of slices that the management system can support while attaining specific latency/responsiveness requirements, KPIs to evaluate the efficiency of the infrastructure sharing, and KPIs to evaluate technical and economic benefits such as energy efficiency, OPEX savings, and reduction of control plane overhead thanks to zero-touch slice MANO.

KPIs

- **[UC1/ES2 KPI-1]:** Reduce the number of SLA performance violations by 20%
- **[UC1/ES2 KPI-2]:** Improve network energy efficiency by a factor of 10
- **[UC1/ES2 KPI-3]:** Reducing Static Slicing overhead will result in 30% higher utilization (will be achieved with dynamic reconfiguration techniques)
- **[UC1/ES2 KPI-4]:** Compared to Static Slicing, demonstrate the same or better SLA tolerances (or risk of missing SLAs) when dynamic slicing techniques are used
- **[UC1/ES2 KPI-5]:** 10x reduction in signalling / monitoring overhead with the use of federation techniques

4.3.2.5 MAPPING OF UC1/ES2 TO THE EXPERIMENTAL PLATFORM

This section presents the components and functions of the platform, see Figure 9, that are going to be used to solve the objectives/challenges of the experimental scenario. The platform is based on a USRP element for the access network and on the Open Air Interface (OAI) software that is used for the implementation of the core functions.

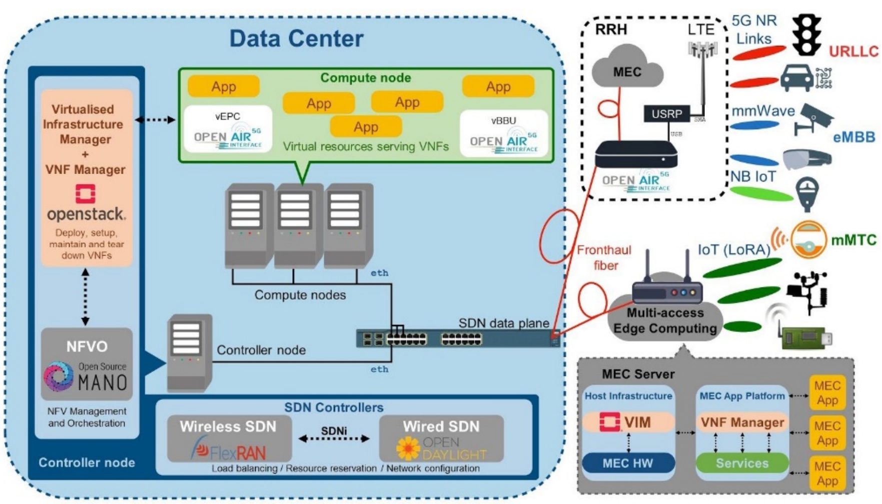

Figure 9: CTTC´s PoC Platform in Barcelona

3.1.1.8 SOCIETAL AND BUSINESS IMPACT OF UC1/ES2

This section discusses the business impact/benefit of this ES that brings to the actors, stakeholders involvement and interaction, OPEX/CAPEX reduction, impact to the society in general. These impacts are shortly outlined below:

- **[UC1/ES2 Imp.-1]:** The reliability of connectivity of this use case will soar, and the dropped call will become a thing of the past.
- **[UC1/ES2 Imp.-2]:** The edge computing proposed in the project will bring new capabilities while generating massive amounts of data.

The radar diagram that presents the baseline KPIs and the improved values on these KPIs proposed by MonB5G for UC1 is provided in the following Figure 10. The squared dots present the baseline KPI values while the circled dots present the proposed KPI values that can be achieved via MonB5G.

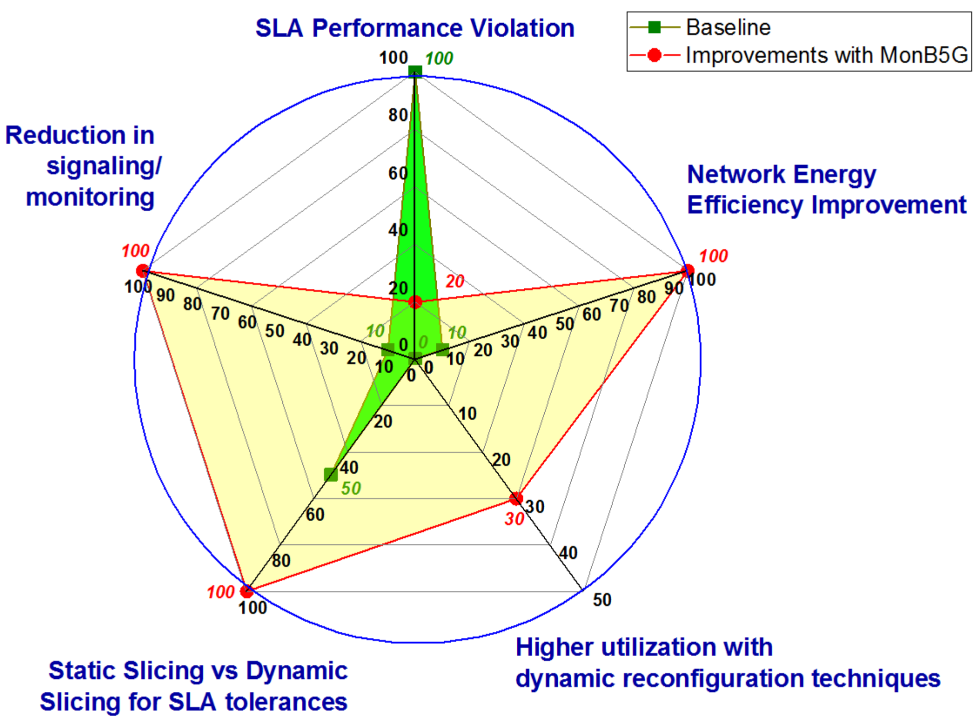

Figure 10: Radar chart for UC1 presenting the KPI bounds

4.4 Use Case 2: AI-assisted policy-driven security monitoring & enforcement

This use case describes the isolation offered by Network Slicing not only offers performance guarantees to the executed applications but is also vital for security. Slicing should ensure isolation, such that attacks (e.g., leakage, breach, DDos) remain contained and don't propagate to the network. However, many vulnerabilities are inherent in the slice selection and isolation mechanisms. Hence the use of AI is necessary to support the integration of millions of devices while still guaranteeing infrastructure security and optimal resource usage. This use case aims to demonstrate how MonB5G with its highly decentralized MS, AE and DE addresses security aspects in two main directions: First, it will show how its decentralized, data-driven management mechanisms help towards accurate and timely identification of security problems and their automatic mitigation, ensuring slice security and isolation. At the same time, it will focus on the security issues of the aforementioned AI-based mechanisms themselves, which stem from their distributed nature and the potential lack of trust and the incentives to cooperate among the parties that are involved in the monitoring and learning processes. Both cases feature the technologies and mechanisms MonB5G will develop for trust management and secure and robust decentralized monitoring. This use case will be demonstrated in Sophia Antipolis' testbed.

4.4.1 UC2/ES1: ATTACK IDENTIFICATION AND MITIGATION

This section describes the first ES of use case 2.

4.4.1.1 DESCRIPTION AND DETAILED ANALYSIS OF UC2/ES1

This experimental scenario applies Machine Learning mechanisms in the domain of Anomaly Detection, in order to identify attacks to slices, localize them, and automatically perform corrective actions to mitigate them. Specifically, MonB5G monitoring engines are used to monitor user behavior in all technological domains and enforce a set of per-slice policy rules and SLAs. These policy rules can protect the network from common attacks, such as DDOS, privilege escalation, etc. In this use case multiple coexisting slices will be deployed over the testbeds, located at CTTC and EUR premises, including all technological domains (RAN, core, etc.), each corresponding to potentially different service types (uRLLC, and eMBB) and each with different performance and security SLAs. In this ES, the response of MonB5G will be evaluated against the following attacks:

- In-slice attack: A subset of the UEs that have been attached to a specific NSI generate malicious traffic towards the infrastructure services, trying to exploit management interfaces. A typical example is compromised MTC devices generating a massive number of network attachment requests. This generates events both at the RAN and the core network level, where separate components of the monitoring, decision, and analytics engines reside. These attacks need to be quickly identified, and the system should ensure that it has not incorrectly classified normal network traffic as malicious (i.e., false positive). Moreover, the decision engine should apply the appropriate policy to mitigate them.
- Cross-slice attack: Sub-slices in individual technological domains (e.g., VNFs instantiated at the MEC) can be shared by multiple Network Slice Instances (NSI). These shared VNF instances may engage in malicious activities against other coexisting slices, acting as "Trojan Horses" that are used to break the slice isolation. The management system needs to ensure the appropriate level of isolation in order to protect other slices from attacks by untrusted shared sub-slices. First, we will demonstrate that by specific slicing-related security extensions to the ETSI MEC architecture (e.g., access control to the RNI, Location, and traffic rules services per slice), unauthorized access to sensitive information and attacks to the traffic steering procedures can be mitigated, thus achieving a level of slice (security) isolation. Second, we will showcase the efficiency of our mechanisms in identifying attacks originating from the edge and limiting their scope in order to avoid their propagation.
- This use-case experimental scenario will demonstrate the robustness of MonB5G for identifying, detecting and then mitigating the in-slice and cross-slice attacks. Different malicious events would be detected thanks to various MonB5G's components including monitoring, analytics, and decision engines. This experimental scenario will assist the efficiency of MonB5G, in terms of response time, for detecting both in-slice and cross-slice attacks while preventing both false negative and false

positive detections

4.4.1.2 OBJECTIVES/CHALLENGES

This use-case experimental scenario will demonstrate the robustness of MonB5G for identifying, detecting and then mitigating the in-slice and cross-slice attacks. Different malicious events would be detected thanks to various MonB5G's components including monitoring, analytics, and decision engines. This experimental scenario will assist the efficiency of MonB5G, in terms of response time, for detecting both in-slice and cross-slice attacks while preventing both false negative and false positive detections.

One of the important challenges to address in this ES is to rapidly identify and mitigate security attack/anomaly, thanks to the MonB5G components AE and DE. Building a system that can quickly react and mitigate attack and anomaly will allow sustaining the end-to-end slice availability and hence support the consumer-slice provider SLA. Technically speaking, this experimental scenario, will demonstrate the MonB5G efficiency to quickly detect and mitigate attacks that come from (1) inside the slice and involving different technological domains (RAN and cloud), and different entities such as (the RAN orchestrator and the NFVO), and (2) cross slice attacks using principally MEC as a technological domain. Accordingly, this will demonstrate that MonB5G components can be used for different technological domains, and runs in complex scenarios interacting with different entities.

4.4.1.3 MAPPING OF THE UC2/ES1 TO MONB5G ARCHITECTURE

The initial architecture concept of MonB5G was presented and described in DoA. Although it is not the scope of this deliverable to provide analysis of the architecture, it is depicted in Figure 11 for clarity purposes and for mapping the entities, functionalities and operations of the experimental scenario which are used in the architecture and discuss how the architecture will provide solution to the objectives/challenges of this scenario. Present the stakeholders/actors in the architecture and show the architectural elements which are involved to realize the experimental scenario under consideration

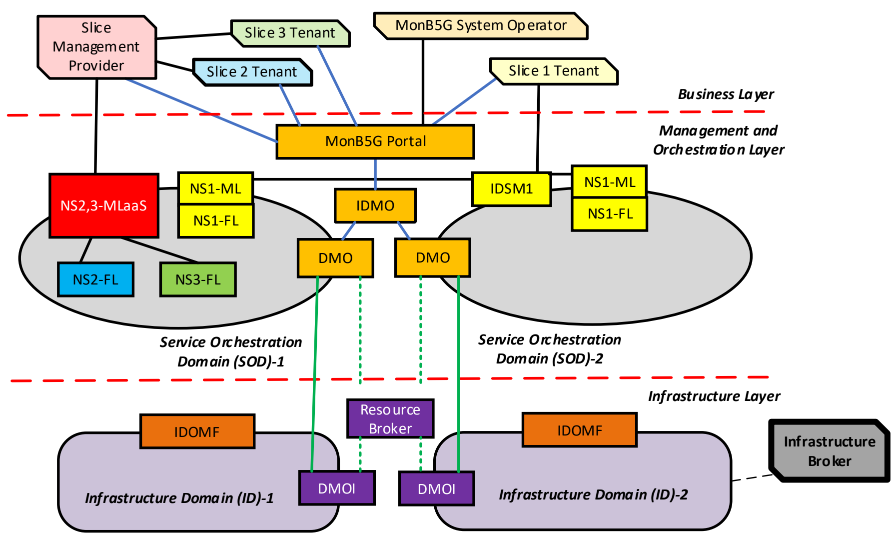

Figure 11: Mapping of experimental scenario and Interaction of stakeholders for UC2/ES1 in MonB5G architecture

4.4.1.4 REQUIREMENTS ANALYSIS AND KPIS

FUNCTIONAL REQUIREMENTS

- **[UC2/ES1 F.req-1]** MonB5G should provide slice isolation.
  As it is stated in [28], network slicing is the technique of isolating the end-to-end performance of a network portion compared to another. A consumer may request a network slice as a service with criteria including a level of isolation that guarantees more or less interfered functioning with the activities of other instances. Isolation refers to the degree of managed object sharing [29], for instance, there can be a network slice subnet instance common to multiple network slice instances, or two network service instances using the same VNF instance, or two VNFC running as pods on a same K8s worker node. In order to fulfil the requested isolation level, it must be clearly defined in order to be measured and enforced We can perceive isolation of network slice through the prism of security as a decomposition in key security criteria such as confidentiality, integrity, availability. Confidentiality means no information about the resources such as data, applications, assets, and services of a network slice instance should be disclosed or made available to another unauthorized instance. Integrity means ensuring the NSI resources are not altered or deleted by the activities of another network slice instance. Finally, availability control should ensure the functioning of services and applications of a NSI at the performance level agreed by the consumer regardless the activities of the other instances. As 3GPP places network slicing concept at the OSS/BSS layer and it is enabled by NFV, isolation is a complex goal to achieve because all the aforementioned security criteria must be considered across all the layers that comprise the NFV architecture. For instance, at the top layer where the vertical or communication services are offered via network slice instances, the service provider who aims to mutualize the underlying network slice subnet instances to maximize its benefits should implement mechanism to ensure their partitioning. Similar consideration at the network service layer where instantiated VNFs can be shared to offer a common service to multiple network service instances, a load balancing should be enforced to handle the promised user traffic.

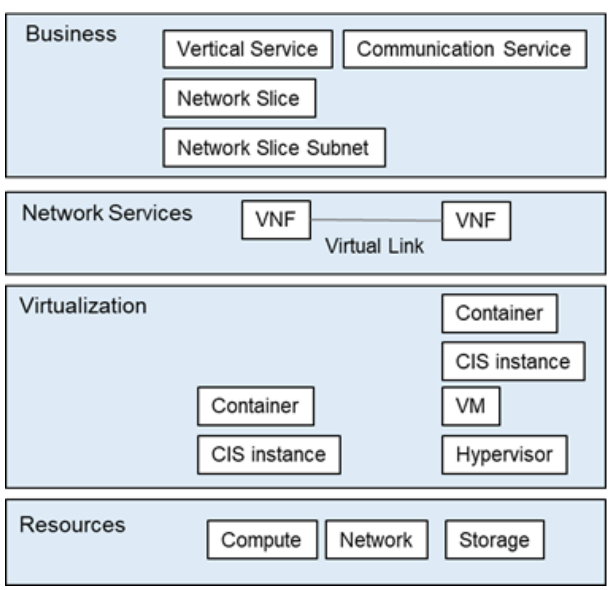

Figure 12: Isolation should be assessed at each layer of the NFV architecture [48]

- **[UC2/ES1 F.req-2]** The MonB5G should provide network security that can allow level provider to provide to its consumer (Communication Service Provider or Vertical) to choose a security level for the protection and defense of their networks and services. The level of security allows to rate a system from the most to the least secure. The network slice provider defines its security levels based on the consumer needs and desired objectives. It can leverage existing and well-known guidelines and frameworks provided by security agencies to define its objectives. For instance, the NIST Cyber Security Framework [30] provides a detailed strategic vision to manage cybersecurity risks though the five core functions: identify, protect, detect, respond and recover. The decomposition of these functions in categories and sub-categories gives full details on the recommendations of the framework. In addition, the framework also offers a guidance to develop security profiles as a prioritization of category and subcategory functions to put in place based on customer requirements and the acceptable risks. ENISA also, [31] provides guidance to assess risks and take appropriate measures to reach a target security objective as shown in Figure 13.

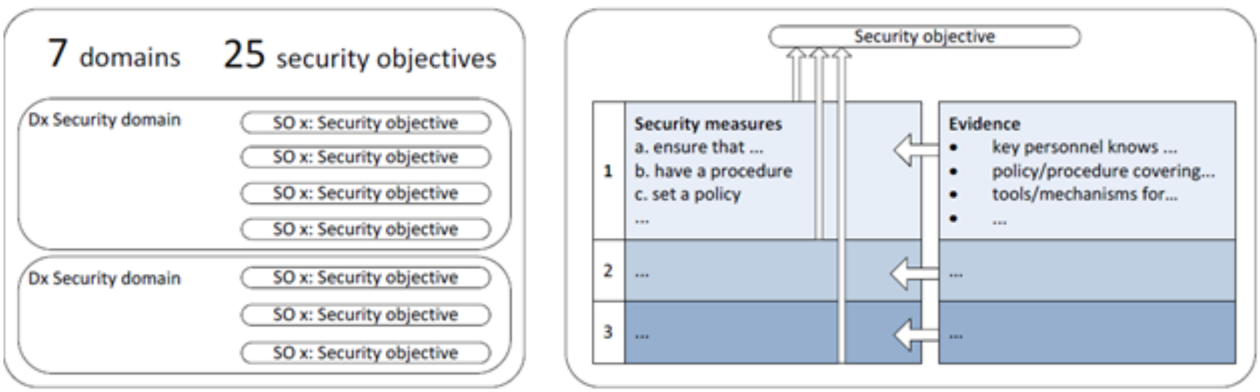

Figure 13: Example of structure of the security objectives and security measures [31]

As a result, a level of security will be translated into a plan of measures to reduce the risks to the level that the consumer can accept. The plan comprises a) the safeguard measures put in place to defend the network slice instance from cyber-attack, b) the detection of cyber security events and the security monitoring and c) the incident handling functions aiming to minimize the adverse impact of the incident and ensure the normal operation of the network slice instance.

KPIs

This section discusses the KPIs as a result of the requirements discussed above. These KPIs are:

- **[UC2/ES1 KPI-1]:** 10x faster attack/anomaly identification
- **[UC2/ES1 KPI-2]:** 10x faster attack remediation and reconfiguration in the order of 10s.
- **[UC2/ES1 KPI-3]:** E2e slice availability > 99%
- **[UC2/ES1 KPI-4]:** Per slice component availability (probability that the service is available) > 99%
- **[UC2/ES1 KPI-5]:** False positive rate in attack classification (false classification of events as attacks) below 1%.

4.4.1.5 MAPPING OF UC2/ES1 TO THE EXPERIMENTAL PLATFORM

This section describes the components and functions of the platform, see Figure 14, that are going to be used to solve the objectives/challenges of the experimental scenario. Referring to MonB5G terminology, EURECOM platform is composed by three technological domains: RAN domain managed by a RAN controller, a Cloud Domain and an Edge Domain compliant to ETSI MEC specification. Both Cloud and Edge domain are managed by the same NFVO. Regarding the infrastructure, the RAN is based on OpenAirInterface (OAI) elements, deploying a 5G gNB and 4G base stations. The Cloud domain is based on OpenShift/Kubernetes PAAS, which allow deploying VNFs and MEC applications using Docker containers. For the Edge, the MEC Edge Platform is run as a VNF, and provides RNIS as well as traffic redirection MEC services. The Core Network is based also on OAI and virtualized using Docker containers. To deploy an end-to-end network slice, EURECOM platform provides a Web Portal that allows a vertical to describe its trial, upload the Network Service Descriptor (NSD) and its set of Virtual Network Functions (VNF) or Application Descriptor (AppD) to deploy on top of MEC or Cloud. Note that, the equivalent of the portal functions is available via a Northbound API (NBI), connecting directly to the Slice Orchestrator (SO) of the platform. The latter is in charge of deploying an end-to-end Network Slice to run the trial. Now the mapping with MonB5G architecture is that the SO is equivalent to IDMO, the NFVO to a DMO, and the RAN controller as a DMO.

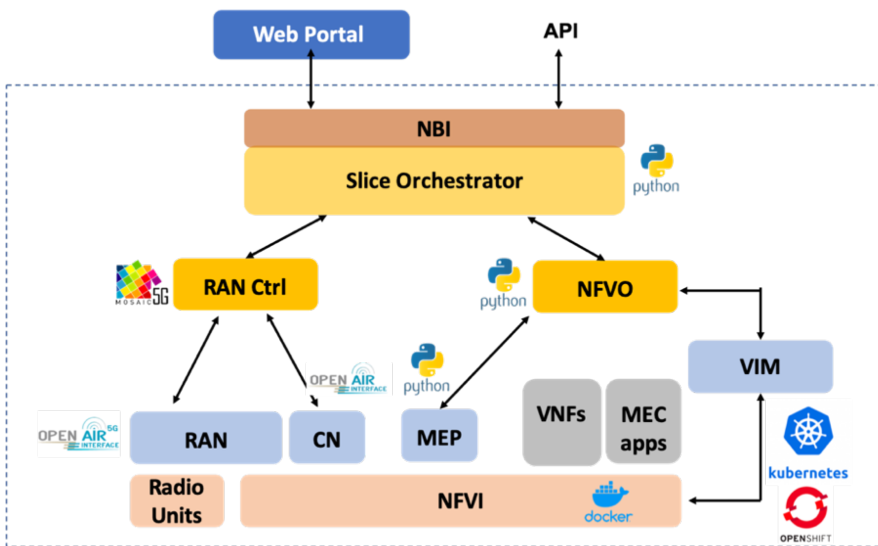

Figure 14: EUR´s PoC platform in Sophia Antipolis

For ES1 of UC2, we distinguish two cases: in-slice attack and cross-slice attack. The elements that are involved in both sub-cases are different. In the first, the attack should be detected at both the RAN and CN elements. We assume that the CN is slice specific and deployed inside the slice. The current version of the platform does not allow to do so, but it will be extended during the project lifetime. In Figure 15, we illustrate a possible mapping between MonB5G elements and the platform elements to run PoC2/ES1/sub-scenario1.

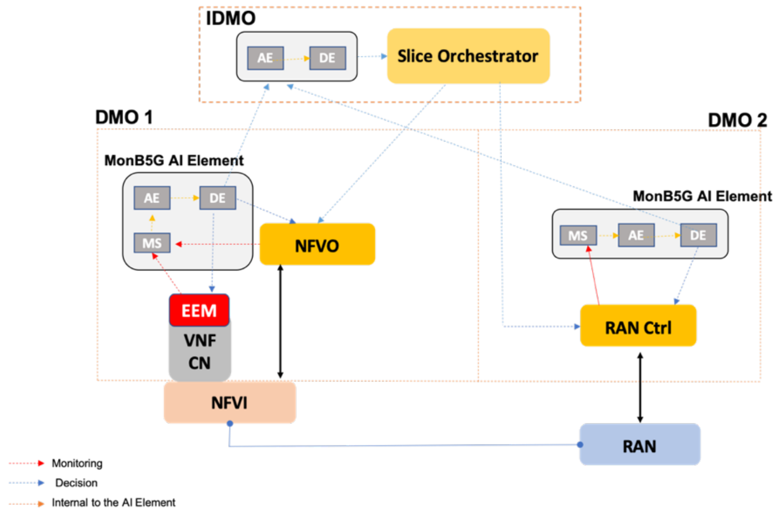

Figure 15: UC2/ES1 Inslice attack mapping

We envision that MonB5G AI Element to be developed during the project is composed of MS, AE, and DE. As the experimental scenario involve two DMO and one IDMO, three technological domains AI Element needs to be devised and developed. Mainly, in thisES, they will be in charge of detecting in-slice attacks by monitoring: (i) the VNF CN, through the Embedded Element Manager (EEM) as well as the NFVO regarding the system level performances; (ii) the RAN through specific monitoring interfaces to be provided by the RAN Ctrl; and hence consider decisions to mitigate the attacks. It should be noted that all the interfaces, represented as non-continue line arrows, will be specified and developed during the project lifetime.

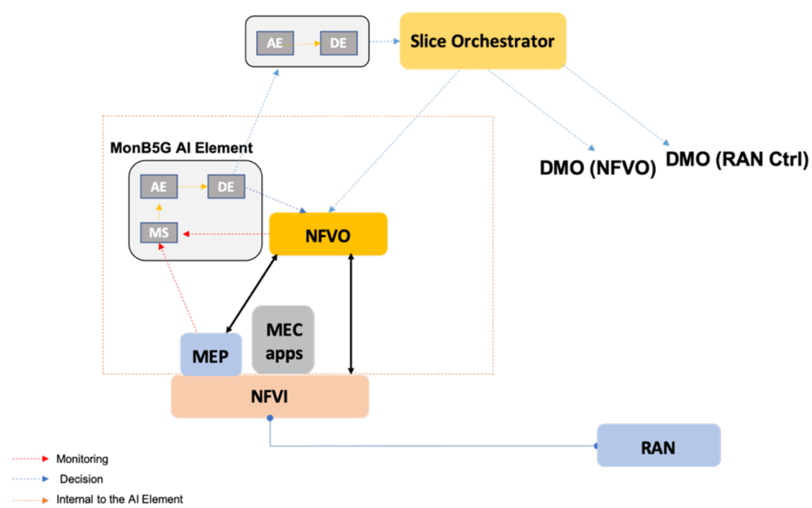

Figure 16: UC2/ES1- cross slice attack mapping

Figure 16, represents the case of the cross-slice attack, which is concentrating on the MEC domain. The attack concerns the MEP, and particularly MEC services. The attack needs to be mitigated in order to ensure isolation among running slices (and MEC applications). Here, only one Technological Domain is concerned, as shown in the figure, the IDMO may involve other domains if the decision affects the end-to-end slice, e.g. stop the slice. Like in the precedent experimental scenario, all the interfaces that need to be implemented are represented as dashed arrows. The MonB5G AI Element will interact with the MEP to monitor it, and possible action will be enforced through the NFVO or the IDMO.

4.4.1.6 PLANNING MEASUREMENT MEANS FOR KPIs/KQIs

The KPI that need to be measured in this experimental scenario are all related to security of running slice. We are planning to measure the following KPI:

- **[UC2/ES1 Meas. KPI-1]:** 10x faster identification of security attack/anomaly
  This KPI will measure the capacity of the MonB5G system to detect and identify a security attack/anomaly, by measuring the time taken from the start of the attack/anomaly until it detections. This KPI will be compared to a vanilla solution that might be relying on existing IDS, such Snort.
- **[UC2/ES1 Meas. KPI-2]:** 10x faster attack remediation and reconfiguration in the order of 10s
  As the precedent KPI, this KPI will measure the time taken by the MonB5G system in order to mitigate an attack. It will be also compared to a vanilla solution based on IDS and Firewall.
- **[UC2/ES1 Meas. KPI-3]:** End to end slice availability (the probability that the service is available at the level specified in the consumer-slice provider SLA) > 99%
  This KPI will measure the ability of MonB5G to rapidly detect an attack and remediate it aiming at keeping a high availability of the end-to-end slice. It is related to the precedent KPI, and will measure, during an attack the downtime of the of the end-to-end slice.
- **[UC2/ES1 Meas. KPI-4]:** Per slice component availability (the probability that the service is >99.9999)
  This KPI is similar to the precedent one, but mainly focus on the availability of the service provided by the network slice.
- **[UC2/ES1 Meas. KPI-5]:** Slice isolation: <5% performance degradation in application-level terms (throughput, end-to-end delay) during attack episode (from start of attack to the moment attack has been handled) on coexisting slice
  Full protection against cross-slice confidentiality and traffic steering attacks at the mobile edge: this KPI measure the capacity of the MonB5G system to ensure isolation and avoid attack coming from other slices, or from inside the slice. The KPI measure during an attack the capacity of isolation of the system, and measure the impact of the attack on the performance degradation. The KPI will be compared to the cases: (1) no attack (i.e. no degradation); (ii) attack without using the MonB5G system.

4.4.1.7 3.1.1.8 SOCIETAL AND BUSINESS IMPACT OF UC2/ES1

The telecoms network have long being the target of hackers that try to breach the networks and prove that they are vulnerable, in other cases the attackers want to gain faulty income. Today, businesses have access to more valuable information than ever before—and protecting that information is vital to business success. MoNB5G also studies the use of large amount of data to be used for analytics which shows the importance of keeping data in safe depositories. Solarwinds MSP[5] reports that the global cost of dealing with cybercrime is estimated to reach $6 trillion by 2021.

[UC2/ES1 Imp.-1]: Network security has always been looked upon by operators in order to boost client and consumer confidence
[UC2/ES1 Imp.-2]: Network security protects the telecoms business from the reputational and legal fallout

[5] https;//www.solarwinds.com

of a security breach and the data shared across the network.

4.4.2 UC2/ES2: ROBUSTNESS OF LEARNING ALGORITHMS IN THE FACE OF ATTACKS

This subsection presents the second experimental scenario of use-case 2.

4.4.2.1 DESCRIPTION AND DETAILED ANALYSIS

This experimental scenario will first assess the vulnerability of the distributed learning algorithms to adversarial attacks and how their performance metrics (e.g., precision, accuracy, etc.) could be affected by those attacks. Then, the model robustness will be improved accordingly, and the new resulting weights and configurations will be propagated to the decentralized elements (DE) of the proposed hierarchical architecture of MonB5G. We assume that each decentralized element will be operating a local learning module as part of a federated learning algorithm while the centralized element, through its component AE, will serve as a sink for data submission (model updates). In this ES, MonB5G framework will be evaluated against adversarial attacks, such as poisoning attacks that target training phase and invasion attacks that target test phase. In this use case, poisoning adversarial attack is considered where a malicious agent can poison some fraction of the data in order to ensure that the learned model satisfies some adversarial goal. In addition, MonB5G will adopt another variant of poisoning attack where an agent aims to poison the parameter updates (of the federated learning algorithm) to be sent to the centralized element. A typical example would be to consider migration of VNFs scenario where the ML model, used by the MonB5G platform, generates migration decisions (select the appropriate target nodes) based on required system and network resources to optimize slices' resources wastage. If metric data are not manipulated, the ML model will decide to perform migration operations to both reduce the bandwidth usage and load balance system resources (e.g. CPU, RAM, disc). However, if an attacker/malicious- agent is able to manipulate the metric data, it could fool the ML model into tacking randomized migration decision, which will result in adding additional network and system overhead. Moreover, if the attacker decides to target a given TD, this will result in ravaging all the slices using resources form that particular TD. To do so, the following yardsticks need to be considered where:

- A large number of parallel slices will be emulated.
- Synthetic samples are generated from the emulated experimental scenarios.
- The automatic selection of a branch of distributed learning techniques that will engage in misreporting attacks, submitting fraudulent model updates and other monitoring information in order to fiddle with the outcome of the learning algorithm.

By means of the trust-based report evaluation mechanisms (e.g. blockchain technologies) and consensus schemes that will be in place, the effect of these attacks on the performance of the learning algorithms (e.g. precision will stay at an accurate level) will be limited.

4.4.2.2 OBJECTIVES/CHALLENGES

The purpose of this use-case experimental scenario is to demonstrate that even under a significant number/ratios of misbehaving entities, distributed learning can be carried out in a robust way. This will be validated using standard metrics used in machine learning, such as true/false positive rates, precision/recall, and the area under curve (AUC).

Two main challenges need to be addressed to reach this objective. First, this experimental scenario needs the involvement of many components of the MonB5G system, i.e. MS, AE and DE, which will be studied and developed in WP3, WP4 and WP5. A specific focus will be on the ML algorithms used by AE and DE, which will be evaluated in terms of ML-specific KPI using a large number of emulated Network Slices. Second, another challenge will consist of how to realistically emulate real network slice traffic, using platforms that were not built for this specific purpose.

The initial architecture concept of MonB5G was presented and described in DoA. Although it is not the scope of this deliverable to provide analysis of the architecture, it is depicted in Figure 11 for clarity purposes and for mapping the entities, functionalities and operations of the scenario which are used in the architecture and discuss how the architecture will provide solution to the objectives/challenges of this scenario. The figure also presents the interaction of stakeholders/actors who are used in the architecture and show the architectural elements which are involved to realize the scenario under consideration

4.4.2.3 MAPPING OF THE UC2/ES2 TO MONB5G ARCHITECTURE

This use-case experimental scenario does not concern any specific TD. The main objective is to study the impact of attacks on the MonB5G elements, AE/DE. For AE, it studies how attackers can disturb the learning process and may alter the model's accuracy. While, for DE, this scenario will shed light on the non-optimal decision that may be taken by the DE due to bad AE accuracy. Unlike the precedent use-cases, this UC has no direct mapping with the MonB5G architecture, as it is agnostic to the TD, and it focuses mainly on the AE/DE performances.

4.4.2.4 REQUIREMENTS ANALYSIS AND KPIS

The KPIs for this experimental scenario are mainly related to the performances of the AI/ML algorithms used by the AE and DE. The KPIs are as:

KPIs

- **[UC2/ES2 KPI-1]:** False positive rate in attack classification (percentage of false classification of events as attacks) below 1%.
- **[UC2/ES2 KPI-2]:** Learning robustness: Precision, recall (true positive rate), fall-out (false positive rate), Area Under Curve values above/below specific thresholds vs. specific ratios of misreporting slice components.
- **[UC2/ES2 KPI-3]:** The accuracy loss should remain neglected. Adversarial training protects the model against adversarial attacks however this additional training may affect negatively the original model accuracy due to this a balance should be established between the adversarial training and the model accuracy.

4.4.2.5 PLANNING MEASUREMENT MEANS FOR KPIS/KQIS

The KPI that need to be measured in this scenario are all related to security of running slice. We are planning to measure the following KPI:

- **[UC2/ES2 Meas. KPI-1]:** This KPI-1 will be measured by generating several attacks on the high number of emulated network slices, where the objective is to see the behavior and the capacity of the AE to have a low false positive rate, particularly acting in an environment with a high number of emulated network slice. The MonB5G system will be compared to a vanilla solution based on IDS, such as Snort.
- **[UC2/ES2 Meas. KPI-2]:** This KPI-2 is planned to be measured by the performance of ML running by the AE and DE components, we will generate a high number of act on the high number of emulated network slices. Again, the MonB5G system will be compared to a vanilla solution based on IDS.

4.4.2.6 SOCIETAL AND BUSINESS IMPACT OF UC2/ES2

This section discusses the business impact/benefit of this scenario that brings to the actors, stakeholders' involvement and interaction, OPEX/CAPEX reduction, impact to the society in general. The impacts are:

- **[UC2/ES2 Imp.-1]:** The reduction on the number of attacks will increase the QoE to the end users
- **[UC1/ES2 Imp.-2]:** The ML will be crucial to the offered customer service. The end user can enjoy a service with higher quality.

The radar diagram for this use case is provided in the following Figure 17. Again, the squares show the baseline KPI bounds while the dots show the MoNB5G proposed to be achieved KPI bounds

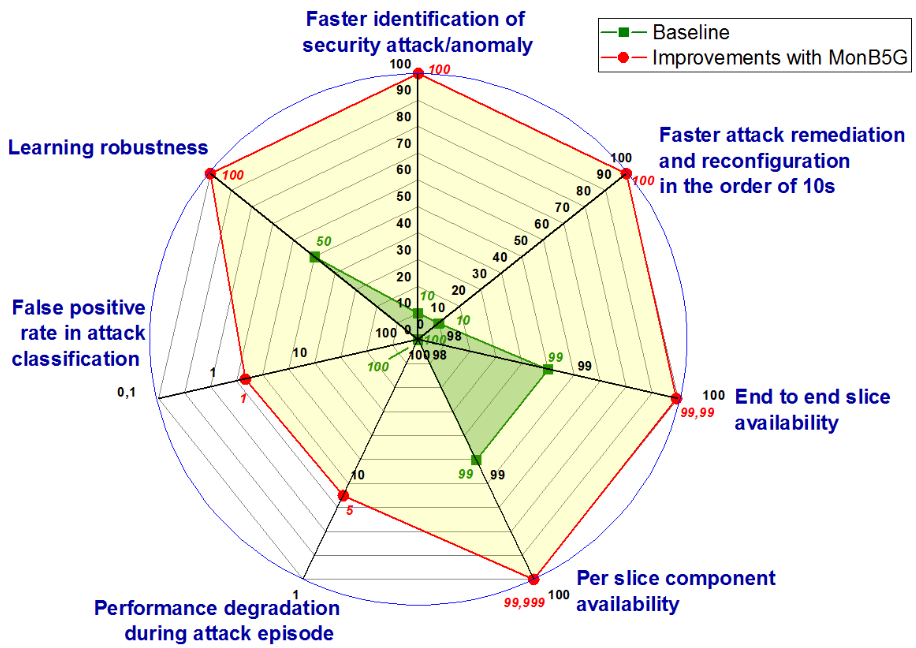

Figure 17: Radar chart for use case 2

5 Market Analysis

5.1 Introduction

In the previous sections we have identified the different stakeholders involved in MonB5G, their roles and interrelation to each other. The connection of the stakeholders to the architecture is identified and presented as well. The purpose of the following market analysis is to show the benefits and values created to these stakeholders by the MonB5G initiatives.

5.2 Open Source and Market analysis overview

This section provides an overview of the management and orchestration platforms currently used by telecom operators. Then, we will discuss the existing gaps, how MonB5G aims to cover these gaps, and the value that MonB5G brings to the stakeholders.

Telecom operators are starting to adopt MANO solutions for their NFV deployments. In the MANO space, one can find several Open Source proposals together with additional relevant solutions developed by vendors. Some of these implementations are identified and described below[6].

OSM (OPEN SOURCE MANO) [40] is one of the most important Open Source initiatives to highlight. It is an Open Source MANO stack aligned with ETSI NFV Information Models. OSM supports and automates service orchestration to simplify the NFV life cycle. Some of the key features offered by OSM are related to EPA (Enhanced Platform Awareness)-based resource allocation capabilities, which enable high-performance VNF deployments with lower Total Cost of Ownership (TCO) for the operator. OSM also provides multi-site support to automate service delivery across multiple sites and underlay network management functionalities. The SDN controllers and VIMs supported in the second release of OSM are ODL (OpenDaylight), ONOS (Open Network Operating System) and Floodlight from the SDN perspective, as well as OpenStack, Vmware vCloud Director and AWS (Amazon Web Services) as VIMs.

ONAP (OPEN NETWORK AUTOMATION PLATFORM) [41] is an Open Source MANO solution that comes from the merger of AT&T's ECOMP and the OPEN-O projects. ONAP provides capabilities for the design, creation, orchestration, monitoring and lifecycle management of VNFs, network services and SDN solutions. ONAP provides all these services in a dynamic and real-time virtualized infrastructure environment in addition to different graphic design studio and monitoring tools.

CLOUDIFY MANO [42] is an Open Source solution with a TOSCA-based cloud orchestration framework that can be leveraged as both the NFVO and G-VNFM in the context of the ETSI MANO architecture. It is able to interact with multiple VIMs, containers, and even external and non-virtualized infrastructure and devices, OSS and BSS, since it enables communication with any northbound or southbound APIs. Currently, Cloudify supports OpenStack, vCloud VMWare, Azure Cloud Stack, SoftLayer and AWS as VIMs and ODL, OpenContrail and ONOS as SDN controllers.

THE ERICSSON MANO [43] solution is called Ericsson Orchestrator. It is aligned to the MANO, ETSI, 3GPP, TMF, IETF standards and it is based on a model driven design based on TOSCA. The Ericsson Orchestrator plays the role of NFVO, G-VNFM, EVNFM. It also allows orchestration and management of network services and cloud resources like storage, networking and virtual machines. It provides management and orchestration of the PNF & VNF life cycle management. This MANO solution offers on-boarding and instantiation of virtual applications using the Open Virtualization Format (OVF) package standard, capabilities of integrating with SDN for L3 VPN connectivity, service chaining of functions and infrastructure and VNF monitoring (fault and performance management).

NOKIA'S MANO [44] proprietary implementation is called CloudBand and is an ETSI NFV MANO system with commercially proven reliability, automation, repeatability and security. It is flexibly deployed for any combination of NFVI / VIM, generic VNF Manager (G-VNFM), and NFV Orchestrator, and serves VNFs from Nokia and other suppliers. It is also the cloud management system for the Nokia Government Cloud Enablement Platform. Commercial implementations may not be as ambitious as the open models in terms of features or functionalities covered, but they are ready to be used in a real environment. As a MANO is key in

[6] ALTRan, 2017, *MANO and Monitoring (White Paper)*

a NFV implementation, it is not rare to observe some carriers that are taking both approaches: first, using a commercial MANO stack to go to market in the short term, but at the same time evaluating and even participating in open source MANO initiatives to adapt them as soon as they are sufficiently mature.

5.3 Creation of end user value from AI/ML

This section discusses the value that AI/ML brings to the operators and the end users. We will point up the envisioned AI/ML's benefits at different levels, from network efficiency to services and finally to the end customer.

ML/AI techniques have evolved over the years and have gradually been adopted into various types of systems, including telecommunications networks management systems. Figure 18, depicts this process of evolving adoption of more complex ML techniques into the telecom network, along with the development of new generations of management systems. In addition, the picture shows the role of what happens in the network towards understanding the root causes and finally looking forward towards automatic proactive adjustments in the network to support autonomous operations.

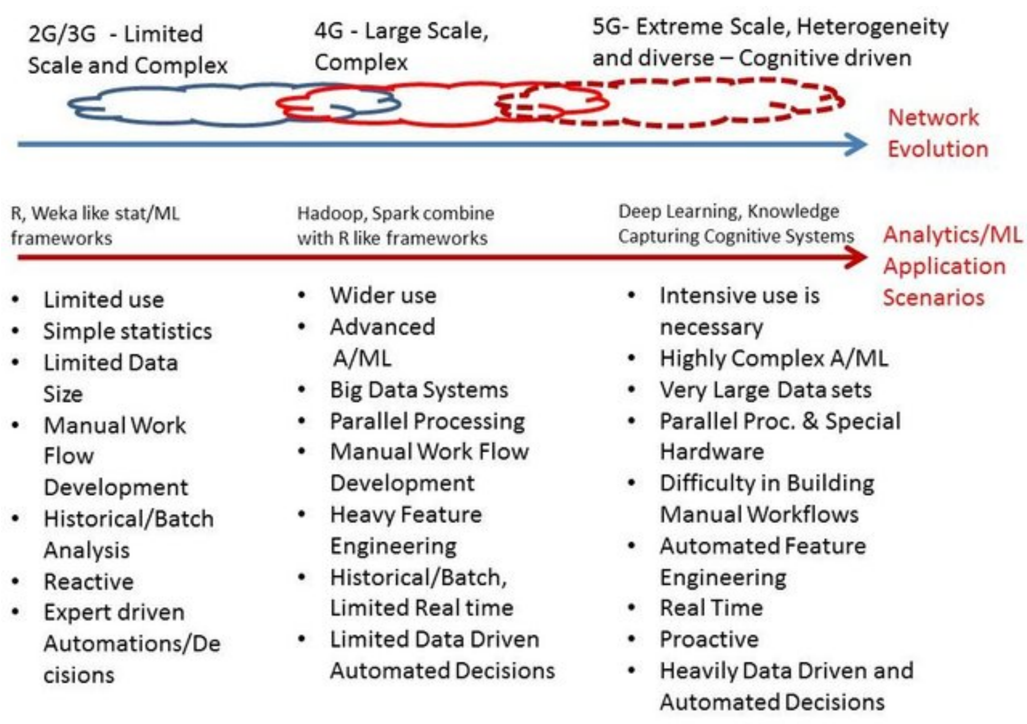

Figure 18: Parallel changes in network management systems and ML application scenarios and techniques [49]

The experimental scenarios addressed in MonB5G relate to the creation and operation of a large number of end-to-end slices over several domains. As explained in the previous sections, we look at this relation both from a scalability and automation point of view (in UC1, ES1 and ES2), as well as from a security and trust point of view (in UC2, ES1 and ES2). In these scenarios, we focus on scalability in terms of number of slices, dynamic behaviour in terms of slice admission and slice operation, ability to react in near-real time for re-configurations and optimisations, timely detection and mitigation of security threats, and resiliency of distributed learning techniques to adversarial attacks.

The AI techniques, tailored by MonB5G for management and orchestration of massive numbers of parallel network slices in an evolved 5G environment, create considerable value to operators and end users with ground-breaking advances in terms of resilience, scalability, cost, coverage and robustness.

As MNOs are considered as key players in the 5G technology market, they will benefit from the adoption of AI techniques which are distributed among the MS/AE/DE elements of the MonB5G orchestration platform and enable various functionalities and characteristics within these engines, towards the final goal of efficient, scalable and proactive orchestration.

While a lot of work has been placed until now on gathering large amounts of historical data and uncover patterns in them in a centralised manner, localised ML was not introduced into proposed architectures until recently {\ref ZTA}. In MonB5G, a lot of effort is focused towards the adoption of distributed ML models, as well as minimising the amount of data to be transferred between the different engines at different layers, which brings advantages related to speed of re-configuration, efficient operation and scalability.

At the same time, while evaluating where simple ML models can give quick insights, we are also exploring more complex techniques, such as deep learning, federated learning under constraints and reinforcement learning.

Our architectural design, coupled with the right choice of ML techniques, achieve: (1) improved operational efficiency by autonomous slice management and configuration of a very high number of network slices corresponding to diverse services and vertical applications, (2) reduced resource consumption, such as energy, communication and computing units, by sharing resources among complementary slices and avoiding the traditional bottlenecks and congestion, (3) decreased network monitoring load, due to distribution and network-aware mechanisms, which is critical towards achieving carrier grade performance, (4) enhanced trust between network tenants and infrastructure operators by having full visibility in the network, especially when tenants compete for resources in a virtualized environment, (5) upgraded robustness and security by automatic identification and mitigation of novel threats and vulnerabilities. As such, the AI techniques enable telecom network operators and service providers to bring value to their end users with resilient, uninterrupted, and ultra-low latency services at low cost to customers.

5.4 Actors in MonB5G

The 3GPP vision [46] of communication services distinguishes between B2C (customer), B2B (business), B2H (household) and B2B2X (anything) and in case of the last category the unlimited recursive chain is allowed. The communication services are built on the network, which is split into the levels of network slice, network slice subnet and network function. Moreover, a network slice/slice subnet or even network function may be wrapped and exposed as a service (a network slice subnet is defined as one or more interconnected network functions, while network slice is defined as a complete set of network functions composing of a complete network architecture able to support an end-user’s communication service). This vision is followed by the management architecture framework [47] in which each of the four abovementioned relational levels has its own management functions: Network Function Management Function, Network Slice Subnet Management

Function and Communication Service Management Function. Of course, this functional split has also the hidden layer of resources, which are subject to various virtualization technologies.

Despite the fact that 3GPP still uses the term “Mobile Network Operator”, this cannot be any longer considered as a singular role. It should be either mapped to certain roles of the new business model, which results in dealing with all the consequences of technology evolution over the years, or treated as a conventional designation for a certain category of business unit that performs activities related to selected roles of the telco business model. When the technology (i.e., hardware) implied a service, the business model consisted of customer and the network operator, which sometimes used leased resources (e.g., cables or long-distance transmission). During the IP revolution in telecommunications, network (access) and service layers were separated, and OTT operators appeared starting to compete into offering incumbent services. The additional split of network view came with regulatory enforced Local Loop Unbundling and Bit-Stream Access, which led to leasing of last mile resources on a massive scale. Finally, virtualization and softwarization imposed a network split into several layers and silos (e.g., MANO). After that evolution, traditional Network Operators needed to redefine themselves in the new multi-layer ecosystem in order to decide which roles they are going to play. Currently many Network Operators decide to sell out their physical resources (to other businesses), because they prefer to operate communication networks at the functional level. As network slices are logical communication networks (or at least slice subnets are their functional building blocks), the role of slice provider may be considered as the closest role to the traditional Network Operator position. However, the Network Operator may play multiple roles of the business model – infrastructure provider, infrastructure broker (if using others’ infrastructure), slice provider, slice broker (if using slice subnets from others’ domains), service provider (because of interactions with customers and e.g. operating IMS or other service platforms) or service broker (if selling others’ content or services – e.g. triple play with IPTV/VoD).

The initial outline of the business model shown in Figure 1, can now be extended to cover the important role of a Slice broker. As the MonB5G architecture distinguishes between domain and inter-domain levels of management and orchestration processes, the role of integrator of sub-network slices into a stitched (sub-)network slice is necessary. This integration (at the level of IDSM and IDMO in the MonB5G architecture) will be played by the Slice broker, whose role is similar to the roles of Infrastructure brokers or service brokers. In a simple case, the entity acting as both e.g. RAN Slice provider and Core Slice provider may integrate these slices within its domain of ownership into one inter-technological slice, further exposed to higher layers of the business model. It should be noted, however, that a broker role may be in a general case recursive, because each broker integrates lower level “products”, reshapes them and includes its know-how into a new, value-added “product”.

Additionally, the management processes at the slice broker level can be also delegated to a separate entity. Hence, similarly to Slice provider level, the separate role of Network Slice Management provider will be present here.

From the point of view of MonB5G project’s goals, the work is focused on the automated and AI-driven management of large amounts of parallel network slices. According to the MonB5G architecture, the management functions are implemented as VNFs and interact with communication services related VNFs with regard of specificity of the latter. The outcomes of the MonB5G project will therefore impact the following roles of the business model:

- VNF providers – as they deliver management VNFs;

- Slice Template providers – as they integrate VNFs into further tested communication networks and include management functionality and engineering for slice template users;
- Network Slice providers and Network Slice brokers – as they onboard the slice templates and instantiate them, using their (I)DMOs;
- Network Slice Management Providers – as they are the users of management functions, further customizing them according to their internal policies, based on the experience and know-how, and on requirements of Slice tenants driven by communication service requirements.
- Slice tenants – as they can request more robust SLAs supported by more efficient management mechanisms.
- Service providers and Users – as they may define end expected elevated QoS targets.

5.5 MonB5G Actors' benefit

The proposal of the MonB5G project provides the following benefits to the actors of the model described above:

- AI-based automation of network operation processes, supporting large numbers of concurrent networks, which cannot be efficiently managed with humans in the control loop or with currently available means of automation using fixed algorithms (Network Slice providers, brokers and management providers);

- Higher robustness, stability, security and performance of communication networks, implying increase of services quality and customers experience (Slice tenants, as responsible for SLAs, Service brokers/providers and End-users);
- Automation of identification, analysis, interpretation, classification and response elaboration for new phenomena (both behavioral and security-related) appearing in networks (Network slice management providers and other actors consider the business model, which can lead to the benefit of improvement of network performance);
- Higher energy costs savings achieved through higher agility of AI-driven power management (Infrastructure providers and the rest of the actors upwards the business model, due to the lower factor of energy cost in prices of products/services delivered at successive levels of business relations of the business model);
- Higher efficiency and flexibility in resources utilization implying lower demand for installation of new resources, thereby lower CAPEX and less power consumption (Infrastructure providers and the rest of the actors could consider the business model, similarly to the abovementioned impact of costs);
- Coherent and scalable network management architecture (VNF providers, Slice Template providers, Network Slice providers, brokers and management providers, and – indirectly – higher-level consumers of Network Slice-based products/services paying the related cost components);
- New class of know-how, to be commercially exploited (providers of management-related VNFs, Slice Template providers, and other actors providing operational slice management).
- Openness to solutions developed by the respective software providers (VNF providers, Slice Template providers, Network Slice providers), which drives the competition in the market resulting in higher quality and wider spectrum of innovation-driven products and services offered to clients (Slice Tenants, End-users).
- Flexibility of implementation of slice management components enabling less complex, more efficient and faster launch of services utilizing slices from multiple technological domains as well as facilitating

co-operation between different providers in terms of development and delivery of the final product (VNF providers, Slice Template providers, Network Slice providers, Network Slice management providers) Value proposition.

5.6 Value chain of MonB5G

A value chain can be defined by the full range of activities that companies and workers perform, to bring a product from its conception to commercialization and finally to a profitable for the company product. This includes activities such as design, production, marketing, distribution and support to the final consumer. The activities that comprise a value chain can be contained within a single company / provider or divided among different providers. Value chain activities can produce goods or services and can be contained within a single geographical location or spread over wider areas. In **Error! Reference source not found.**, the value brought by MoNB5G to the stakeholders is presented.

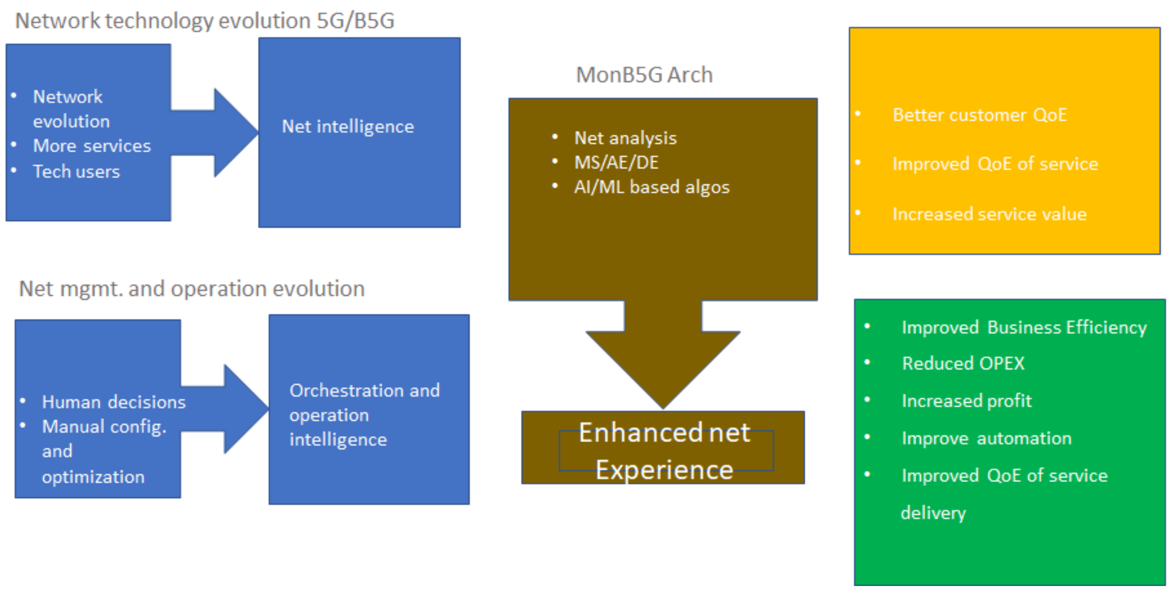

Figure 19: Value chain that MoNB5G brings to the stakeholders

The benefits brought by the different drivers and value-added elements of the MoNB5G architecture are described below.

5.6.1 BUSINESS BENEFITS OF AUTOMATIC MONITORING AND OPTIMIZING SLICING IN MONB5G

Automatic slicing and slice management can be the change required in the industry to sustain growth using 5G-related network technologies. Automatic network slicing offers both customization and optimization of the network to fast meet market demand for specific end-to-end requirements with reduced operation costs. In fact, slices can be customized to meet the needs of operators to offer vertical services and can be also optimized by a huge variety of characteristics, including latency or bandwidth requirements. To give a simple example, applications like remote operation of machinery, tele-surgery and smart-metering all require connectivity, but with vastly different characteristics.

The potential benefits of slicing are discussed along the following dimensions: cost reduction, product management improvement and greater agility.

Build, develop, and maintain services: Slicing can help Service Providers (SPs) to develop services that are more relevant to their customers, opening up new revenue opportunities. In this way, network slicing will benefit many industries by offering a smart way to segment the network and support particular services. It will enable SPs to better control the QoS delivered to their clients and better penetrate the enterprise segment. The upcoming offered highly differentiated services will be able to attract this enterprise segment and bring more profit to a variety of industries.

Agility and flexibility: Slicing provides for the logical separation of virtual networks for different services, which would be independent of each other. This means that teams within the organization could similarly become more independent in how they work; changes to a particular slice could be made without having to consider the implications across the entirety of the network and all stakeholders or customers involved. Development efforts on new and existing services could be decoupled from each other and from the underlying infrastructure. They could each progress under their own constraints and timelines, but still benefit from common infrastructure. For example, if the operator wanted to introduce something new to the current core services, there is no need to redesign the whole network every time.

Cost and efficiency: New networks aim to support a variety of potential applications and use cases and must be built to serve needs not well met by existing networks. Therefore, the network of the future needs to be designed in a way that addresses this feature by removing unnecessary functionalities and adding new technologies. Using virtualized slices could enable the cost-effective use of resources by having a common underlying infrastructure, where resources can be partitioned, as well as shared, in an optimized way, so that total cost of ownership (TCO) may be reduced. By decoupling the physical and virtual infrastructure, an operator would be able to have a single, shared physical infrastructure and run slices on top for each of their operations, maximizing the use of their resources.

Management: Having separate slices for different types of customers could allow the operator to better manage services, rather than have to operate them all within one network. This could have further benefits when it comes to designing revenue models and the associated systems for particular sets of customers. In the short term, there is an opportunity for operators to apply this concept to better manage transitions from legacy services. Also, supporting hybrid networks (2G/3G/4G/5G) and hybrid virtualized/non-virtualized legacy networks will also present notable transition challenges.

5.6.2 BUSINESS BENEFITS USING A DISTRIBUTED MANAGEMENT PLATFORM IN MONB5G

The distributed MoNB5G platform can lead to the following benefits for the MoNB5G stakeholders:

Optimum utilization of resources:

The distributed management infrastructure proposed by MoNB5G aims to optimize the utilization of resources in opposition to the centralized one that needs a vast amount of resources to cope with the huge amount of data needed. However, there is always the tradeoff between processing power and transmission load & consistency which is an issue that is considered in this project.

In addition to this, an efficient distributed management system enables quick and smooth operations that ultimately saves time. MonB5G also looks at optimizing resource utilization in the network and researches means with the aid of AI of improving it.

Scalability

The operator/SP that launches a service is making a business plan that includes an assessment on how scalable it will be from the beginning. That is because, if the number of data in the network grows, so will the load on the central systems. In the case of a centralized network, an increasing load will slow down the system. Scaling that network up would necessarily mean a hardware upgrade or more efficient sharing of resources which leads to more complex management strategies and this is a main concern of all stakeholders involved. With distributed networks, scaling is simpler, as new administrators can be added throughout the network to add more processing power.

5.7 Business Model Method

The business model method used in MonB5G considers the work that is proposed in [39] and particularly the service dominant business model (SDBM) design described in it. This method is based on the interaction of the actors of MonB5G, the costs/benefits involved from this interaction and the value proposition from each actor. Figure 20 presents, in a diagram with concentric circles, the financial actions and benefits brought to the actors of MonB5G, the value proposition that is resulted from the interaction of the actors and the value chain resulted in MonB5G.

Error! Reference source not found., depicts the actors, their involved costs/benefits, their value propositions and the created value chain from MonB5G to the ecosystem. The first concentric layer is the center point of the analysis since it provides the value from each actor, which is then connected to a final value chain. The second concentric frame defines the value proposition produced by each actor. The third concentric frame depicts the financial costs and the benefits to the actors and MonB5G. Lastly, the last frame presents the actors involved in MonB5G. The main actor is the one that sets-up the business model and is actively involved in the final solution. There are other partners that provide additional values to the main actor. Each of the actors collaborate in a way where each of them brings a value to the model but at the same time each actor is benefited financially. Table 6: SDBM business model values for MonB5G

, presents some of the actors involved in MonB5G, their financial costs and benefits from their interaction with the ecosystem, the value proposition produced by each actor and the final value chain brought to MonB5G.

Table 6: SDBM business model values for MonB5G

<table>
<tr><th>Actor</th><th>Cost/benefit</th><th>Value proposition</th><th>Value Chain in MonB5G</th></tr>
<tr><td>VNF providers</td><td rowspan="5">All the actors need to put some investment in staff for SW production / they receive a share of income</td><td rowspan="5">• VNFs are a crucial part of MonB5G
• Network establishment
• Templates are providing the structure of VNFs and thus is another significant part of MonB5G</td><td rowspan="5">• End User QoE increase
• Flexible platform
• Stable platform
• Secure platform
• Scalable platform</td></tr>
<tr><td>Slice Template providers</td></tr>
<tr><td>Network Slice providers and Network Slice brokers</td></tr>
<tr><td>Network Slice Management Providers</td></tr>
<tr><td>Slice tenants</td></tr>
</table>

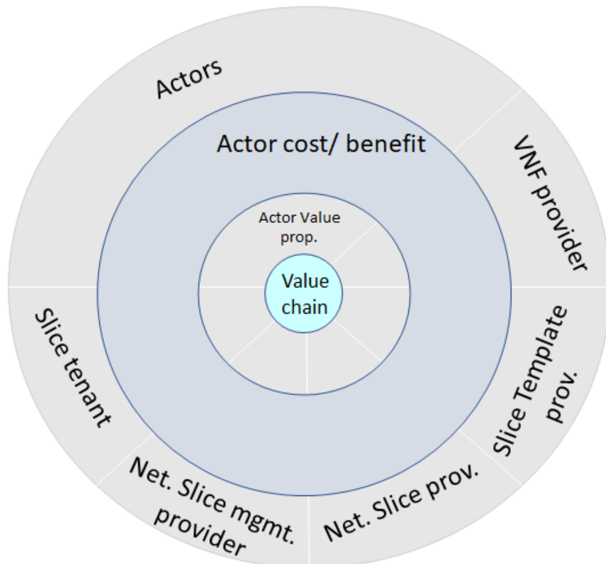

Figure 20: Diagram showing the business modeling method

6 Conclusions

MonB5 intends to enable significant breakthroughs to achieve an infrastructure that provides network operators with a zero-touch management system that orchestrates resources and manages a large number of slices in a flexible and adaptive way. It provides a highly flexible and scalable platform, able to support new business cases and revenue streams by proposing new business actors, by providing new opportunities and revenues.

MonB5G addresses the means for a platform that can flexibly handle a range of services and resources. It enables paradigms of ML algorithms to achieve a zero-touch flexible and secure infrastructure that can offer better guarantees to the various stakeholders and play a key role in the 5G ecosystem. MonB5G supports a scalable platform able to support new business models and revenue streams by creating a neutral host market and reducing operational costs, by providing new opportunities with more stakeholders involved in the telecoms arena.

MonB5G along with the other projects of this 5G-PPP program proposes networks that become sufficiently flexible and agile to handle a range of applications and services pertaining to different domains/verticals. At the same time, it provides a transformation towards a significant reduction in cost and at the same time results to the optimal allocation of available resources to satisfy the bounds of the initial proposed KPIs for driving an autonomous infrastructure and coping with the numerous security attacks on the infrastructure and management domains.

In this deliverable, we presented the state of the art on 5G ecosystems and technology trends. Furthermore, we gave an overview of new use cases including requirements, KPIs, definition of new stakeholders and the impact to the society and the proposed proof of concepts to achieve them.

The first use case describes how to achieve a zero-touch management platform for achieving a high level of SLA. In particular, MonB5G will implement this use case on a platform in Barcelona in order to assess the proposed KPIs.

The second use case analysis a 5G secure platform with a target to prove that such secured systems are necessary to be achieved in future networks, since security is going to be one of the most crucial capabilities of networks.

For each of these use cases, two experimental scenarios are described with the involved stakeholders, the deployment topology and the KPI assessment process.
The detailed MON-B5G architecture, briefly shown in this deliverable, will be the subject of deliverable D2.1.

References

[1] NGMN Alliance, "5G white paper", 2015. Available online: https://www.ngmn.org/uploads/media/NGMN_5G_White_Paper_V1_0.pdf.

[2] 5G-PPP, "Living document on 5G PPP use cases and performance evaluation models". Available online: https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-use-cases-and-performance-evaluation-modeling_v1.0.pdf.

[3] Recommendation ITU-R M.[IMT.Vision], document 5/199, "Framework and overall objectives of the future development of IMT for 2020 and beyond" [adopted in July 2015].

[4] 3GPP TS 28.554 v16.3.0 (2019-12), ''Technical Specification Group Services and System Aspects; Management and orchestration; 5G end to end Key Performance Indicators (KPI), (release 16)

[5] H. Daniel, L. Peer, "Initial Design for Control Framework", SUPERFLUIDITY Deliverable 6.1, 2016.

[6] Segall, E. Biton, "Initial Design SLA based deployment", SUPERFLUIDITY Deliverable 6.2, 2016.

[7] C. Raiciu, "Modelling and Design for Symbolic Execution and Monitoring", SUPERFLUIDITY Deliverable 6.3, 2016.

[8] R. Calé, P. Neves, "Definition of APIs and Interfaces for the SELFNET framework", SELFNET Deliverable 2.2, March 2016.

[9] GSMA, ``Cloud AR/VR Whitepaper'', 2019.

[10] 3GPP, TS23.203, `` Policy and charging control architecture '', Rel15, 2018.

[11] M. Orduna, C. Díaz, L. Muñoz, P. Pérez, I. Benito and N. García, "Video Multimethod Assessment Fusion (VMAF) on 360VR Contents," in IEEE Transactions on Consumer Electronics, vol. 66, no. 1, pp. 22-31, Feb. 2020.

[12] M. Habibi, B. Han, M. Nasimi, and H. Schotten, "The structure of service level agreement of slice-based 5G network," in IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2018.

[13] H. Daniel, L. Peer, "Initial Design for Control Framework", SUPERFLUIDITY Deliverable 6.1, 2016.

[14] 5GPPP Architecture Working Group, "View on 5G architecture." https://5g-ppp.eu/wp-content/uploads/2019/07/5G-PPP-5G-Architecture-White-Paper_v3.0_PublicConsultation.pdf. June, 2019.

[15] V. Sciancalepore, X. Costa-Perez, and A. Banchs, "Rl-nsb: Reinforcement learning-based 5G network slice broker," IEEE/ACM Transactions on Networking, vol. 27, pp. 1543–1557, 2019.

[16] E. Kapassa, M. Touloupou, A. Mavrogiorgou, and D. Kyriazis, "5g SLAs: Automated proposition and management of agreements towards QoS enforcement," in 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN) 2018.

[17] N. Salhab, R. Rahim, R. Langar, and R. Boutaba, "Machine learning based resource orchestration for 5G network slices," in IEEE Global Communications Conference (GLOBECOM) 2019.

[18] X. Shen, J. Gao, W. Wu, K. Lyu, M. Li, W. Zhuang, X. Li, and J. Rao, "Ai-assisted network-slicing based next-generation wireless networks," IEEE Open Journal of Vehicular Technology, vol. 1, pp. 45–66, 2020.

[19] C. Marquez, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Pérez, "Resource sharing efficiency in network slicing," IEEE Transactions on Network and Service Management, vol. 16, no. 3, pp. 909–923, 2019.

[20] E. Kapassa, M. Touloupou, and D. Kyriazis, "SLAs in 5G: A complete framework facilitating VNF- and NS- tailored SLA management," in 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2018.
[21] T. V. K. Buyakar, H. Agarwal, B. R. Tamma, and A. A. Franklin, "Resource allocation with admission control for GBR and delay QoS in 5G network slices," in International Conference on COMmunication Systems NETworkS (COMSNETS), 2020.
[22] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Perez, "Deepcog: Cognitive network management in sliced 5G networks with deep learning," in IEEE Conference on Computer Communications (INFOCOM) 2019.
[23] R. Wen, G. Feng, J. Tang, T. Q. S. Quek, G. Wang, W. Tan, and S. Qin, "On robustness of network slicing for next-generation mobile networks," IEEE Transactions onCommunications, vol. 67, pp. 430–444, Jan 2019.
[24] J. X. Salvat, L. Zanzi, A. Garcia-Saavedra, V. Sciancalepore, and X. P. Costa, "Over-booking network slices through yield-driven end-to-end orchestration," in CoNEXT'18, 2018.
[25] B. Han, V. Sciancalepore, D. Feng, X. Costa-Perez, and H. D. Schotten, "A utility-driven multi-queue admission control solution for network slicing," in IEEE Conference on Computer Communications, (INFOCOM) 2019.
[26] H. Chergui and C. Verikoukis, "Offline SLA-constrained deep learning for 5g networks reliable and dynamic end-to-end slicing," IEEE Journal on Selected Areas in Communications, vol. 38, no. 2, pp. 350–360, 2020.
[27] B. Khodapanah, A. Awada, I. Viering, D. Oehmann, M. Simsek, and G. P. Fettweis, "Fulfillment of service level agreements via slice-aware radio resource management in 5G networks," in IEEE 87th Vehicular Technology Conference (VTC Spring), pp. 1–6, 2018.
[28] Americas, 5. (2020). Security Considerations For The 5G Era -. White Paper.
[29] Chatterjee, A. S. (Feb 2017). Cloud Security Issues and Challenges. Journal of Network and Computer Applications 79 (Feb. 2017), pp.88 115, issn: 1084-8045, doi: 10.1016/j.jnca.2016.11.027.
[30] European Union Agency for Network and Information Security. (October 2014). Technical guidance on the security measures in Article 13a - Version 2.0.
[31] National Institute of Standards and Technology. (April 2018). Framework for Improving Critical Infrastructure Cybersecurity - Version 1.1.
[32] https://www.techradar.com/news/how-nintendo-fared-in-2016
[33] https://www.microsoft.com/en-us/hololens
[34] https://dynamics.microsoft.com/en-us/
[35] https://www.etsi.org/deliver/etsi_gs/ENI/001_099/005/01.01.01_60/gs_ENI005v010101p.pdf
[36] https://www.etsi.org/deliver/etsi_gr/NFV-IFA/001_099/029/03.03.01_60/gr_NFV-IFA029v030301p.pdf
[37] https://www.monb5g.eu/
[38] M. Akbar, et. al., ''solving the multidimensional multiple-choice knapsack problem by constructing convex hulls., Computer Oper. Res. 33(5), 1259-1273 (2006)
[39] Turetken, O., Grefen, P., Gilsing, R. *et al.* ''Service-Dominant Business Model Design for Digital Innovation in Smart Mobility,'' Bus Inf Syst Eng **61,** 9–29 (2019).
[40] https://osm.etsi.org/
[41] https://www.onap.org/
[42] https://cloudifynetwork.com/tag/mano/

[43] https://www.ericsson.com/en/portfolio/digital-services/automated-network-operations/orchestration/ericsson-orchestrator
[44] https://www.nokia.com/networks/solutions/cloudband/
[45] Ghulam Bhatti, ‘’ Machine Learning Based Localization in Large-Scale Wireless Sensor Networks,’’ MDPI Sensors 2018.
[46] https://www.etsi.org/deliver/etsi_ts/128500_128599/128530/15.00.00_60/ts_128530v150000p.pdf
[47] https://www.etsi.org/deliver/etsi_ts/128500_128599/128533/15.00.00_60/ts_128533v150000p.pdf
[48] https://www.etsi.org/deliver/etsi_gr/NFV-EVE/001_099/012/03.01.01_60/gr_nfv-eve012v030101p.pdf
[49] Bosneag, Anne-Marie, S. Handurukande and M. Wang. “Fighting Fire with Fire : Survey of Strategies for Counteracting The Complexity of Future Networks Management.” (2016).