**5G smarT mObility, media and e-health for toURists and citizenS**

# Deliverable D3.2
# Technologies, architecture and deployment initial progress

***Project Details***

| | |
|---|---|
| ***Call*** | H2020-ICT-19-2019 |
| ***Type of Action*** | RIA |
| ***Project start date*** | 01/06/2019 |
| ***Duration*** | 36 months |
| ***GA No*** | 856950 |

***Deliverable Details***

| | |
|---|---|
| ***Deliverable WP:*** | WP3 |
| ***Deliverable Task:*** | Task T3.1 – T3.5 |
| ***Deliverable Identifier:*** | 5G-TOURS_D3.2 |
| ***Deliverable Title:*** | Technologies, architecture and deployment initial progress |
| ***Editor(s):*** | Carlos Barjau, Manuel Fuentes, Álvaro Ibáñez (UPV) |
| ***Author(s):*** | Marco Gramaglia, Albert Banchs, Ginés Garcia-Aviles (UC3M), Galini Tsoukaneri (SRUK), Bruno Foresi, Pietro Scalzo, Giorgio Calochira (TIM), Felipe Vicens, Ignacio Labrador (ATOS), Christophe Burdinat (EXP), Irene Alepuz (UPV), Linas Maknavicius, Bessem Sayadi (NOK-FR), Eleni Rigani, George Mitropoulos (NOK-GR), Ewelina Szczekocka (ORA-PL), Mara Piccinino, Giancarlo Sacco (ERI-IT), Ioannis Belikaidis (WINGS), Ioannis Stenos (WINGS), Vangelis Kosmatos(WINGS), Stefania Stavropoulou (WINGS), Michalis Mitrou (WINGS), Iason Bitchavas (WINGS), Ioannis Chondroulis (WINGS), Thomas Ferrandiz (BCOM), Sofiane Imadali (ORA-FR) |
| ***Reviewer(s):*** | Galini Tsoukaneri (SRUK), Christophe Burdinat (EXP), Panayiotis Verrios (ACTA), Pascal Allain (PRE), Silvia Provvedi (ERI-IT) Belkacem Mouhouche (SRUK), Eleni (Nelly) Giannopoulou (WINGS) |
| ***Contractual Date of Delivery:*** | 31/07/2020 |
| ***Submission Date:*** | 31/07/2020 |
| ***Dissemination Level:*** | PU |
| ***Status:*** | Final |
| ***Version:*** | 1.0 |
| ***File Name:*** | 5G-TOURS_D3.2_Technologies_architecture_deployment_initial_progress_v1.0 |

| ***Disclaimer*** |
|---|
| *The information and views set out in this deliverable are those of the author(s) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.* |

***Deliverable History***

| Version | Date | Modification |
|---|---|---|
| ***V1.0*** | *31/07/2020* | *Initial version, submitted to EC through SyGMa.* |

# Table of Content

# List of Acronyms and Abbreviations

| | |
|---|---|
| ***3GPP*** | 3rd Generation Partnership Project |
| ***5GPPP*** | 5G Public-Private Partnership |
| ***4G*** | 4th Generation mobile network |
| ***5G*** | 5th Generation mobile network |
| ***5GC*** | 5G Core |
| ***5G EVE*** | 5G European Validation platform for Extensive trials |
| ***5G-Xcast*** | Broadcast and Multicast Communication Enablers for the Fifth Generation of Wireless Systems |
| ***5G-MoNArch*** | 5G Mobile Network Architecture for diverse services, use cases, and applications in 5G and beyond |
| ***5GT*** | 5G-Transformer |
| ***5GT-VS*** | 5GT Vertical Slicer |
| ***ABR*** | Adaptive Bitrate |
| ***AI*** | Artificial Intelligence |
| ***AMF*** | Access and Mobility Function |
| ***API*** | Application Programming Interface |
| ***AR*** | Augmented Reality |
| ***BBU*** | Baseband Unit |
| ***BM-SC*** | Broadcast Multicast Service Centre |
| ***BSCC*** | Broadcast Service and Control Centre |
| ***BSS*** | Business Support Systems |
| ***CN*** | Core Network |
| ***CNN*** | Convolutional Neural Network |
| ***COTS*** | Commercial Over The Shelf |
| ***DVB*** | Digital Video Broadcasting |
| ***E2E*** | End to end |
| ***EE*** | Execution Environment |
| ***eMBMS*** | Enhanced Multimedia Broadcast Multicast Services |
| ***ENI*** | Experimental Network Intelligence |
| ***EPC*** | Evolved Packet Core |
| ***FEC*** | Forward Error Correction |
| ***GPRS*** | Generic Packet Radio Service |
| ***GSM*** | Global System for Mobile communications |
| ***GTP*** | GPRS Tunnelling Protocol |
| ***GUI*** | Graphical User Interface |
| ***HPHT*** | High Power High Tower |
| ***IaaS*** | Infrastructure as a Service |
| ***IoT*** | Internet of Things |
| ***IRP*** | Integration Reference Point |
| ***IWL*** | Interworking Layer |
| ***KPI*** | Key Performance Indicator |
| ***KQI*** | Key Quality Indicator |
| ***LTE*** | Long Term Evolution |
| ***MANO*** | Management and Orchestration |
| ***MBB*** | Mobile Broadband |
| ***MDAF*** | Management Data Analytics Function |
| ***MEC*** | Mobile Edge Computing |
| ***mMTC*** | Massive MTC |
| ***MooD*** | Multicast on Demand |
| ***MAE*** | Mean Absolute Error |
| ***MTC*** | Machine Type Communication |
| ***NEF*** | Network Exposure Function |
| ***NSD*** | Network Service Descriptor |
| ***NFV*** | Network Function Virtualisation |
| ***NFVI*** | NFV infrastructure |
| ***NR*** | New Radio |
| ***NRF*** | Network Repository Function |
| ***NSA*** | Non Standalone |
| ***NSO*** | Network Service Orchestrator |
| ***NSSF*** | Network Slice Selection Function |
| ***NWDAF*** | Network Data Analytics Function |
| ***PM*** | Performance Monitoring |
| ***PSO*** | Particle Swarm Optimisation |
| ***O-DU*** | O-RAN Distributed Unit |
| ***O-RAN*** | Operator Defined Next Generation RAN Architecture and Interfaces |
| ***O-RU*** | O-RAN Radio Unit |
| ***ONAP*** | Open Network Automation Platform |
| ***OSM*** | Open Source MANO |
| ***OSS*** | Operation Support System |
| ***OVS*** | OpenVSwitch |

| | |
|---|---|
| ***QoE*** | Quality of Experience |
| ***QoS*** | Quality of Service |
| ***RNN*** | Recurrent Neural Network |
| ***SA*** | Standalone |
| ***SA*** | System Architecture |
| ***SDN*** | Software Defined Networking |
| ***SDO*** | Standards Developing Organisation |
| ***SDR*** | Software Defined Radio |
| ***SLA*** | Service Level Agreement |
| ***SMF*** | Session Management Function |
| ***UC*** | Use Case |
| ***UDM*** | Unified Data Management |
| ***UMTS*** | Universal Mobile Telecommunications System |
| ***UPF*** | User Plane Function |
| ***URLLC*** | Ultra Reliable Low Latency Communications |
| ***VDU*** | Virtual Deployment Unit |
| ***vEPC*** | Virtual EPC |
| ***VIM*** | Virtualized Infrastructure Manager |
| ***VM*** | Virtual Machine |
| ***VNF*** | Virtual Network Function |
| ***VR*** | Virtual Reality |
| ***WEF*** | Wireless Edge Factory |

# List of Figures

# List of Tables

# Executive Summary

This report marks the second update out of four planned for the project duration, conglomerating the overall work performed in WP3 (network architecture and deployment). The scope of this deliverable is to report the progress made in the 5G-TOURS architecture for the 5G network deployment. The three trial sites (Turin, Rennes, and Athens) have experienced improvements and changes, and initiated their integration with 5G EVE infrastructure and application portal.

This report focuses on three main aspects: i) the description of the overall project network architecture (which stems from the state of the art proposals and the 5G EVE one), ii) its enrichment with specific components that are propaedeutic to the correct operation of the use cases developed in the different sites, and iii) the overall integration of these technologies (through the architecture) into the 5G-EVE infrastructure. In detail:

Section 2 covers the progress on the high level 5G-TOURS architecture, the physical infrastructure update per node, and the interaction with the portal and the interworking layer of 5G EVE. This section documents the progress made in Task 3.1 “Overall architecture and Security” in Section 2.1. In order to design the logical architecture, several requirements have been considered, especially the ones coming from the verticals that envision new functionality from the network such as suggestions for the re-configuration of the network. A collaboration between WP2 (use cases design) and WP3 concluded that, depending on these requirements, a bypass of the 5G EVE Interworking Layer was essential due to the necessity of fine grain orchestration control of the physical infrastructure and the related Network Functions, specific to each use case. In close synchronisation with WP4 (Touristic city), WP5 (Safe city) and WP6 (Mobility-efficient city), the physical infrastructure description for the three sites is also covered in Section 2.2, and has been organised by phases, depending on their current capabilities (phase 1) and future 5G capabilities and 5G-TOURS novelties incorporation (phase 2).

Next, Section 3 is composed by the 5G-TOURS technological innovations update, with five major enhancements that encompass all the domains of the network, going beyond the current 5G standards and provide the required KPIs to support the innovative use cases of the project. They include: i) Enhanced Management and Orchestration (MANO), focused on the latest Open Source MANO (OSM) release 8 (see Section 3.1), ii) Artificial Intelligence (AI) orchestration, providing intelligent solutions and algorithms developed inside the project for the network slice management (see Section 3.2), iii) Broadcast support, detailing the inclusion of broadcast in the 5G-TOURS infrastructure and covering the progress made in 3GPP for broadcast communications (see Section 3.3), iv) Service Layer description with the requirements, expectations and implementation methods, also featuring a cross collaboration between WP7 (system integration and evaluation) and WP3 for Key Performance Indicator (KPI) monitoring (see Section 3.4), and v) 5G Radio Access Network (RAN) enhancements involving the motivation for an end-to-end orchestration including the radio resources, existing open-source solutions and standardisation status (see Section 3.5). Additionally, 5G-TOURS is expected to create footprint with these innovations beyond the pure support to the use cases, by promoting the technology to relevant working groups such as ETSI ENI, 3GPP or OSM.

Finally, Section 4 describes a summary of 5G EVE cross-site architecture, development status and implementation strategies for 5G-TOURS technology in order to provide full support to the verticals. The 5G EVE [1] application portal was made public in summer 2020 exposing the Interworking Layer for the vertical on-boarding to 5G experiments. The different approaches for the inclusion of 5G-TOURS novelties into 5G EVE infrastructure are analysed within.

# 1 Introduction

The exceptional situation worldwide has highlighted the importance, role and the shortcomings of the telecommunications network ecosystem. It is undeniable that 5G is here and has an important role to play in the telco world, as a service itself and as a complementary wireless network for private business. The 5G commercial equipment is already reaching the stores worldwide, with industry studies reporting that more than 20 chipsets are commercially available providing Release 15 connectivity featuring peak speeds of gigabytes orders of magnitude [2] [3]. In Europe, more than 10 countries are offering their 5G connectivity to their customers [4].

However, the early 5G connectivity often relies in the so-called Non-Standalone 5G, where the Long Term Evolution (LTE) Core Network (CN) is providing all the transport signalling but the radio technology used is New Radio (NR). This deployment does not really accomplish the 5G potential: a customizable, one-size-fits-all network, ready to serve many different verticals simultaneously with latency in the range of milliseconds and gigabyte levels of throughput. In this sense, the 5G-TOURS project brings the verticals and the customers closer to the realities of 5G by trialling a diverse set of use cases leveraging the work perform in ICT-17 project 5G EVE in both the application portal front-haul and the underlying cutting-edge 5G network deployment of the three nodes involved in the consortium. To create an architecture and framework that can support all the use cases, it is necessary to consider and include the latest technology trends, including customisation and monitoring (Service Layer), smart management of network resources (AI-based Orchestration and Enhanced MANO) and point-to-multipoint support for an effective way to deliver ultra-high quality media content.

In parallel to 5G-TOURS, 5G EVE also evolved during these months. While continuously upgrading their physical infrastructure, the most relevant change is the launching of the application portal to the public, which interfaces the 5G equipment distributed in the European nodes to the interested verticals, allowing them to design, deploy and monitor their own 5G experiments [5]. The developments of 5G EVE are being closely monitored by 5G-TOURS and their implementation requirements are being factored in the design philosophies for the 5G-TOURS innovations.

The first deliverable [6] describes the initial design ideas for the architecture, technologies and implementation. This deliverable marks the first progress update on the aforementioned topics, complemented with a new section on 5G EVE on-boarding and technological integration strategies. In detail, Section 2 includes the high-level architecture of 5G-TOURS and a description of all the underlying layers that form it. Additionally, the status and improvements of the physical nodes of Turin, Rennes and Athens are illustrated. Section 3 is dedicated to the technological progress of 5G-TOURS and the mapping to the physical nodes, organised in five sections: Enhanced MANO, covering the 8$^{th}$ release of OSM; AI Orchestration, showcasing the novel algorithms for smart network resource and slice management; Broadcast Support, describing the status in standardisation fora for point-to-multipoint cellular communications and their inclusion in 5G; Service Layer, encapsulating the 5G EVE portal and providing KPI monitoring and decision making; and 5G RAN enhancements, providing the need for E2E orchestrators that also manage the RAN and the possible solutions in standardisation forums and open-source groups. Section 4 focuses on 5G EVE, with a description of the application portal, interworking layer and blueprint syntax, and a summary of the insertion strategies between 5G-TOURS technology and 5G EVE infrastructure. Section 5 closes the document with the main conclusions derived from this deliverable.

# 2 Progress on 5G-TOURS Architecture

## 2.1 5G-TOURS Architecture

In this section, we describe the first version of the 5G-TOURS overall architecture, which represents the up to date status of the Task 3.1 work, capturing the input from several other work packages, and more specifically:

- The requirements for the specific use cases, gathered from WP2;
- The advances of the different physical deployments happening in the test sites, which have an impact on the physical architecture of the system;
- The inputs coming from WP7 in terms of KPI assessment.

In the following sections we thoroughly discuss all these points, starting from the requirements (Section 2.1.1), how we addressed them in the overall architecture (Section 2.1.2) and their implication in the physical setup (Section 2.1.3).

### 2.1.1 Requirements

A modern network architecture, that embraces the novel concepts of softwarization, programmability and inclusion of artificial intelligence, necessarily relies on the recent advances that relevant Standard Developing Organizations such as 3GPP and ETSI, and industrial associations are currently proposing. Among them we can list the Open Network Automation Platform (ONAP) [7] (a piece of software that is currently integrated in 5G-EVE) or the Operator Defined Next Generation RAN Architecture and Interfaces (O-RAN) [8] (which will be used for prototyping advanced radio access solutions, not included in the final deployment). By analysing these solutions [6] together with the inputs coming from the WP2 work on the requirement definitions of the use cases [9], we defined the initial version of the project's overall architecture, discussed in Section 2.1.2. In the following, we discuss the requirements that drove the design of the architecture, analysing their impact on the architectural definition.

#### 2.1.1.1 Broadcast support

For some of the use cases in the touristic city testbed in Turin, broadcast support is a must, as the increasing number of targeted users in the network requires an efficient way of distributing ultra-high quality content to them. Broadcast support is orthogonal to our definition of the overall architecture, which is impacting only one specific site: Turin. However, broadcast has to be supported by the overall architecture and especially provide interactions with other modules. For instance, the enhanced MANO shall correctly handle the broadcast specific Virtual Network Functions and orchestrate them in the network for the use cases where they are required and also with the Service Layer, to allow verticals leveraging the broadcast capabilities. This also effects the 5G EVE infrastructure that shall provide the required assets and configurations. More details of this functionality are provided in Section 3.3.

#### 2.1.1.2 Enhanced MANO and Service Layer

Being based on the 5G EVE, 5G-TOURS inherits the Management and Orchestration infrastructure provided by the project [10]. However, the requirements coming from the vertical service providers in the project allowed us to define a more complex management system, that includes the Service Layer developed in the project (Section 3.5), MANO elements with enhanced functionality (especially tailored to support the Service Layer, (Section 3.1) and the 5G EVE portal [11] (especially for the service blueprint definition). Thus, our management and orchestration architecture flexibly takes into account situations in which the architecture shall support enhanced functionality that is currently out of scope for 5G EVE, and cannot be directly supported without a deeper refactorization of the available assets, while collaborating with the 5G EVE project to enhance the current infrastructure with the 5G-TOURS inputs.

Namely, the features that shall be supported by the Enhanced MANO and Service Layer are related to the capability of controlling the network service lifecycle that goes beyond the possibility of on-boarding the network slice serving it, and steering parameters such as boundary condition (e.g. introduced latency or background traffic). Instead, 5G-TOURS verticals require the following:

- Live feed for the KPIs: Verticals need to always have under control the KPIs they require for the service they are provisioning (e.g., user throughput, latency, number of users) to eventually take control decisions with the metric levels experienced by the network services. These metrics can be either continuous flow of data or alarms, to which they subscribe in an event-based fashion;

- Monetary information: traditionally, service providers and tenants defined and enforced QoS levels and the monetary costs through interactions that involved human intervention (mainly through Operation Support Systems, OSS and Business Support Systems, BSS). Hence, the involved timings, may be too slow for rapidly changing environments such as the ones envisioned by 5G (and 5G-TOURS), 5G networks allow for a much more flexible management of this interaction, eventually allowing dynamic Service Level Agreement (SLA) between service providers and network operators by means of software interfaces among them. Hence, one of the fundamental features that need to be enabled is the capability to dynamically provide information about the current expenditure and possible alternatives that may be selected to service providers (thus, taking into account business metrics that are only known to the service provider);
- Suggestions to improve quality: together with the live feed for the KPIs, as discussed before, service providers also require suggestions (coming from AI modules deployed in the Network Operator infrastructure) about re-orchestration triggers, to effectively take decision about the service provisioning that go beyond metrics, directly exposed to the service providers (i.e., the load of other services currently hosted in the infrastructure).

For all these reasons, the 5G-TOURS management and orchestration extends the one of 5G EVE, adding support for these requirements. All these requirements drive the interface definition and the specific algorithms that address them as, for instance, the AI based one discussed next.

#### 2.1.1.3 Big Data and AI

Network control, management, and orchestration entail the dynamic placement, configuration, and resource provisioning of Virtual Network Functions (VNFs) within the Network Function Virtualisation (NFV) infrastructure.

The complexity of these operations exceeds substantially that of equivalent tasks in legacy 4G LTE networks. There, the relatively limited amount of variables in the one-size-fits-all core and Radio Access Network (RAN) domains accommodates management models that mainly rely on expert monitoring and intervention. Instead, the traditional human-based approach is hardly viable in virtualised 5G networks where the coexistence of heterogeneous mobile services, diversified network requirements, and tenant-defined management policies create a need for specialised and time-varying infrastructure deployments. This, in turn, calls for automated solutions in the control, management, and orchestration of the network. Hence, Artificial Intelligence (AI) and Big Data are a natural choice to support the emerging need in autonomous network operation and management. 3GPP and other Standards Developing Organisations (SDOs) have started delineating the road for the integration of AI into the mobile network architecture.

More specifically, network orchestration shall be driven by two goals: i) the fulfilment of the service requirements and ii) the proper usage of the resources. While this can be achieved with an excessively overprovisioned network, it is of interest of the network operator to provide such resource assignment in a sustainable way, to minimise the footprint on the utilisation, as depicted in Figure 1.

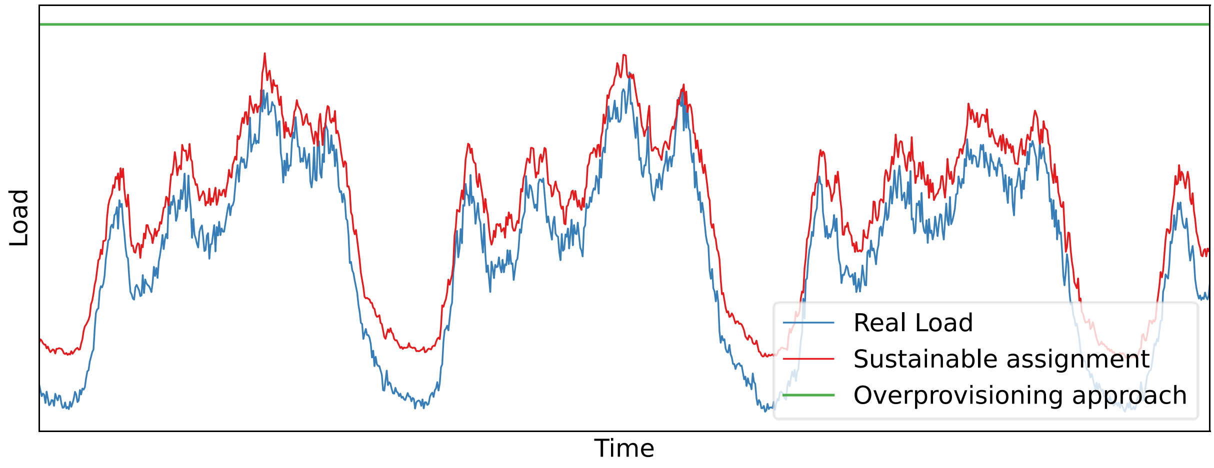


**Figure 1. AI based resource assignment**

In Figure 21, the AI-based resource assignment depicted in red can fulfil the service level agreements between the service provider and the network operator incurring a much lower overhead than a blind overprovisioning approach (which is traditionally employed in the network). In order to do so, AI algorithms that depend on information coming from network elements, such as the Network Data Analytics Function (NWDAF) [12] from the 5G Core and the Management Data Analytics Function (MDAF) [13] from the 3GPP Management functions, will automatically steer the network decisions on behalf of the operator and the vertical. 5G-TOURS considers this by integrating elements coming from relevant initiatives such as ETSI Experimental Network Intelligence (ENI) into the architecture [14].

### 2.1.2 Overall Architecture

The 5G-TOURS overall architecture has been designed with the goal of fulfilling the requirement posed by the different use cases, analysed in WP2 (Section 2.1.1), while gathering input from all the ongoing standardisation activities happening at each domain of the network. We followed a layered approach that builds on the findings of the 5G EVE project architecture, enriching it with specific elements that support the 5G-TOURS view.

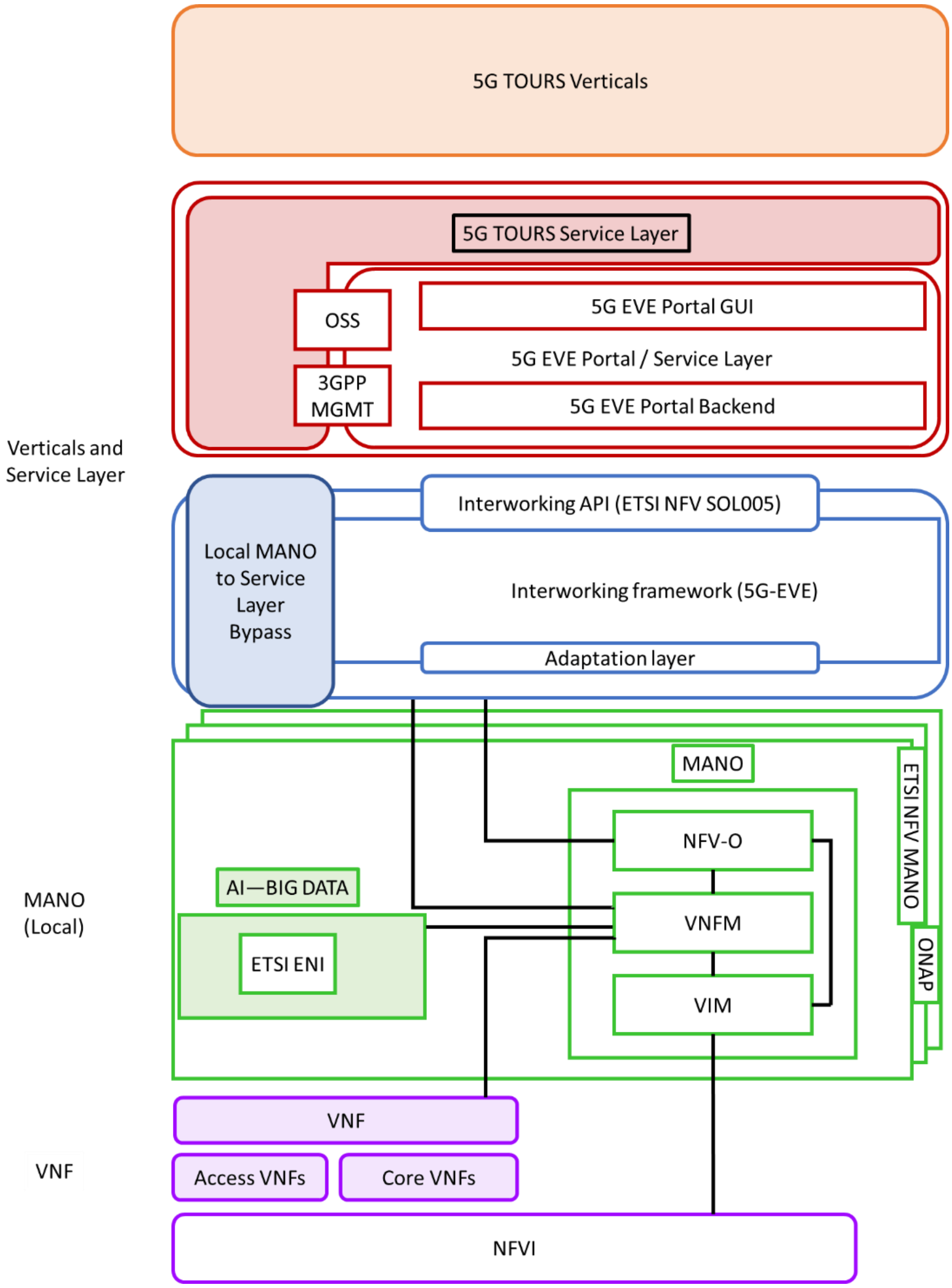


**Figure 2. High level 5G-TOURS architecture. Solid filled blocks are native to 5G-TOURS**

#### 2.1.2.1 Architecture Description

The overall 5G-TOURS architecture is depicted in Figure 2. It is composed by the following five main layers:

- The Verticals: The aim of 5G-TOURS is to provide a flexible infrastructure for the verticals, thus verticals provide the utmost feedback regarding the usability of the architecture and the network, as it must fulfil the requirements set for their use cases through the service layer in a user friendly way. Specific details about this part are reported within WP4, WP5 and WP6 deliverables;
- The Service Layer: The service layer is the “door” that is used by the vertical to access the 5G-TOURS network. It must provide advanced functionality for the service on-boarding, integrating AI and Big Data concepts, to support the automated operation of the network. The Service Layer has to consider its integration with the 5G EVE portal, which can provide a subset of the needed functionality;
- The Interworking Layer (Central MANO) and its bypass: The 5G EVE interworking layer glues together the different sites that provide infrastructure around Europe. It also allows for some functionality for network on-boarding and KPI monitoring, as well as the definition of the templates that are used by the underlying network infrastructure. However, as some of this functionality is not enough for the 5G-TOURS system, we envisioned a bypass that directly interfaces with the local Management and Orchestration (MANO) of each site, to provide enhanced functionality that would have been too complex to provide when using the Interworking Layers included in 5G EVE. Still, 5G-TOURS will use the portal and other software assets provided by 5G EVE as much as possible;
- The local MANO: As each site leverages on different infrastructure (both physical and for the available network function), so the MANO needs to be tailored to the different sites. Each of the MANO interfaces to the interworking layer (as part of 5G EVE) also offers specific APIs for the bypass to be used by the 5G-TOURS service layer. The MANO offers all the functionality exemplified by the ETSI NFV MANO reference architecture [15] composed by the NFV Orchestrator (NFV-O) VNF Manager (VNFM) a, enhanced by some 5G-TOURS specific functionality such as the AI-based orchestration algorithms. This also includes specific architectural design such as the ones designed by the ETSI Experimental Network Intelligence (ENI) group [16], to which 5G-TOURS is contributing with a proof of concept;
- The VNF (infrastructure layer): which represents all the assets available in the sites.

#### 2.1.2.2 Service Layer

The 5G-TOURS Service layer has to provide several features to the vertical providers, which have been reviewed in [6]. Among them, we list the continuous KPI Monitoring, the possibility of on-boarding services through network slicing with an intent-based interface and possible AI-based suggestions about re-orchestration and network plan selection. Thus, the Service Layer requires a wide access to many domains in the network infrastructure, ranging from the RAN to the Orchestration modules.

Still, 5G-TOURS does not build the full architecture from scratch, as it leverages on the 5G EVE one, which includes the “Interworking Layer” [17], a feature-rich layer that includes some of the functionality envisioned by the 5G-TOURS Service Layer. Hence, we performed a thorough gap analysis to understand the features that shall be implemented (either on top, or together) the “interworking layer and the Portal. Mainly, the 5G-TOURS Service Layer will improve the “Interworking Layer” by adding the following new functionality:

1. **Enhanced lifecycle management:** The 5G-TOURS Service Layer envisions a holistic lifecycle management for the service which includes, for instance, scaling, re-orchestration, and relocation of VNFs on the fly, by possibly including such decisions in the vertical business intelligence. This is currently not possible with the 5G EVE portal, which automatically handles this. Currently, the 5G EVE infrastructure provides Experiment lifecycle management. That is, the 5G EVE portal offers access to the 5G EVE infrastructure, which is shared across several projects (5G EVE and several other ICT-19 project). Thus, for the sake of security, the experiment lifecycles only offer limited primitives, such as scheduling the creation of an experiment or removing it (eventually modifying it) before its execution, which take place asynchronously. The aim of 5G-TOURS is to provide full control of the lifecycle management to verticals, hence there is a need for a richer API.
2. **Automatic exposed run-time management** (i.e., scaling, re-orchestration): while 5G EVE handles re-orchestrations and scaling function internally, it does not expose such APIs to the verticals. For the reasons explained above, this functionality is planned in the 5G-TOURS portal.
3. **KPIs monitoring:** currently 5G EVE supports KPI monitoring only through the web portal, but not through an API (which is however, planned in a later version). Still, 5G-TOURS plans to employ AI

and Big Data driven solutions, which may require early access to such metrics (and many more, which include infrastructure load). Therefore, this functionality has to be extended.

4. **AI Suggestions to Verticals**: In 5G EVE, there is currently no support for enhanced AI-based suggestions to the verticals (more on this in Section 3.2). This is fundamental for the closed control loop on network slices that involve the vertical service provider as well. That is, quality of experience metrics is only gathered by the applications run by the vertical, as well as other business metrics, such as current revenues, expenditures or subscriber churn rate. Thus, all this information shall flow between the infrastructure provider (5G EVE in this case) and the vertical (the 5G-TOURS vertical).

For these reasons we decided to have a hybrid service layer that can be very thin for some use cases, when the required functionality is already provided by the 5G EVE modules, up to a very thick one that also embodies its own implementation of the 5G EVE functionality when very enhanced features are needed. This selection will happen on a per use case basis, as discussed in Section 3.4.

#### 2.1.2.3 Interworking Layer and its bypass

The 5G EVE Interworking Layer provides the required abstraction to allow the 5G EVE portal to be agnostic of the different underlying technologies, and to permit the independent deployment of the different sites without compromising the development of the single frontend for verticals. Briefly, the Interworking Layer is composed by the following elements:

- The Multi-Site Network Service Orchestrator (NSO), Multi-site Catalogue and Multi-site Inventory, which is responsible for the management of the lifecycle of the deployed components, jointly allowing for multisite slices supporting the 5G-TOURS verticals' experiments.
- The Data Collection Manager, which collects the required performance metrics to ensure the correct operation of the infrastructure and to validate the targeted KPIs.
- The Runtime Configurator, which applies the required runtime configurations to the provisioned services.

However, this infrastructure may not expose all the capabilities needed by 5G-TOURS, which can require direct access to the underlying MANO infrastructure to perform direct operation, such as scaling, migration or resource assignment. For these reasons, we also envisioned a bypass of the Interworking Layer to be used directly by the Service Layer to interface with the local MANO deployed at each site. The details of such bypass will be UC-dependent and detailed in deliverable D3.3.

#### 2.1.2.4 MANO / Enhanced MANO Layer

The management and orchestration layer deployed at the different sites is specifically tailored to the infrastructure that is handled. Therefore, 5G-TOURS (by leveraging on the 5G EVE infrastructure) has different deployments of Orchestrators available at each site, which summarised as follows:

- Turin: two orchestrators have been deployed, one OSM-based and another one based on the Ericsson EVER Orchestration framework.
- Rennes: the orchestration in Rennes is based on the ONAP software.
- Athens: the OSM orchestrator is used for the Greek site.

In the overall architecture, the mapping of the requirements to the fulfilling functionality is tackled at the UC level, as different requirements coming from the verticals are very specific to the management plane of each site, making it very difficult to generalise in a centralised way. Thus, 5G-TOURS will develop enhanced MANO functionality that are applicable to each orchestration reference framework, depending on the use case. The enhancement of the MANO will thus take place with specific additions to e.g., OSM (as discussed in detail in Section 3.1) and will be directly related to one of the following items:

- Extensions to support the service layer functionality, such as the automatic translation of vertical services templates into network slice descriptors that can be orchestration platforms;
- Enhancement to support the AI-based autonomous operation of the network, including enhanced gathering of data, monitoring techniques and alarm-based interfaces.

All of these enhancements, being very software oriented, will be pushed to relevant initiatives (such as the OSM deployment or the ETSI ENI PoC framework) as much as possible, to show the impact potential of the project technologies.

#### 2.1.2.5 Analytics and AI

One of the major contributions of the WP3 work is the activity on the Data Analytics from the Network and its AI-based management. AI is a natural choice to support the emerging need in autonomous network operation and management. 3GPP and other SDOs have started delineating the road for the integration of AI into the mobile network architecture.

Such a process starts with an efficient collection of data in the network infrastructure and knowledge inference from that data, which are paramount to effective AI-assisted decision-making. In this sense, SDOs are pushing efforts towards defining AI-based Data Analytics frameworks that are suitable for autonomous and efficient control, management and orchestration of mobile networks. For instance, 3GPP has incorporated into its standardized architecture the modules (i) Network Data Analytics Function (NWDAF) [12], and (ii) Management Data Analytics Function (MDAF [13]). Other organizations, such as the O-RAN alliance, envision similar entities in their architectures [8]. ETSI has also defined comparable assisting elements within the Industry Specification Groups (ISGs) on Experiential Networked Intelligence (ENI) and Zero touch network & Service Management (ZSM [14]).

All these ongoing efforts are, however, at an early stage. The frameworks they propose, and the solution designs they foster are preliminary and mainly aim at introducing several key building blocks at a very high level of abstraction. They are still far from detailed, full-blown network data analytics that are ready for deployment.

Thus, 5G-TOURS accommodates this data-driven management of the network by incorporating in the overall architecture, and specifically in the orchestration architecture, an explicit block that is empowered with specific algorithms such as the ones defined in Section 3.3. Finally, through the orchestration framework, the outcomes of AI-driven management of the network will be fed to the service layer, in order to provide enhanced intelligent functionality to the verticals.

#### 2.1.2.6 Network Function Layer

The Network Function Layer includes all the (virtualised or not) network functions that are used by 5G-TOURS to provide support for the different use cases implemented in the project. The network function technology differs from site to site (as it depends on the specific deployment of the 5G EVE infrastructure), but as baseline it includes 3GPP Release 15 network functions. In particular, 5G-TOURS will focus on two kinds of network functions that are characteristic of the use cases and the deployments planned in the project:

- **Broadcast Network Functions:** as discussed deeply in Section 3.3, 5G-TOURS is pioneering the European research in the field of broadcasting, both by showcasing bleeding edge technologies (such as the High-Power High Tower trial in Turin) and close-by monitoring of the standardisation efforts for the Release 17 broadcast support in 5G;
- **The Data Analytics Network functions**, both in the core network (NWDAF) and in the RAN (where such an analytics function is yet to be defined by the standards). As the data analytics from the network is one of the main innovation of 5G-TOURS, major attention will be devoted on how to leverage such data and provide optimal operation of the network (especially in terms of orchestration).

The specific details of the available network functions are provided in Section 2.1.3.

#### 2.1.2.7 Infrastructure Layer

The infrastructure deployment available for 5G-TOURS is fully described in Section 2.2, and contains information about the different edge sites and NFV Infrastructure (NFVI) deployments provided in each site. Still, the infrastructure topology follows common patterns composed by an aggregation site, where the core functions and the orchestration framework usually run, and a set of edge sites where the use cases are located. The high level instantiation pattern is depicted in Figure 3.

The different edge site deployments host the specific RAN deployments, that could be either small cell based or macro cell, depending on the specificities of the use cases. For instance, the thick walls of Palazzo Madama require an indoor RAN deployment, while the optimal Ambulance Routing in Rennes, relies on a city-wide coverage. Core Network Functions are instead deployed in a centralised location, usually managed by the reference Network Operator in the site (TIM for Turin, OTE for Athens) or by a Network Equipment Provider such as B-COM in Rennes. Still, in the French site, the Orchestration is hosted in the Network Operator (ORA) premises. This centralisation allows for a better resource utilisation and a better provisioning of features such as authentication of users, which do not require stringent latency requirements.

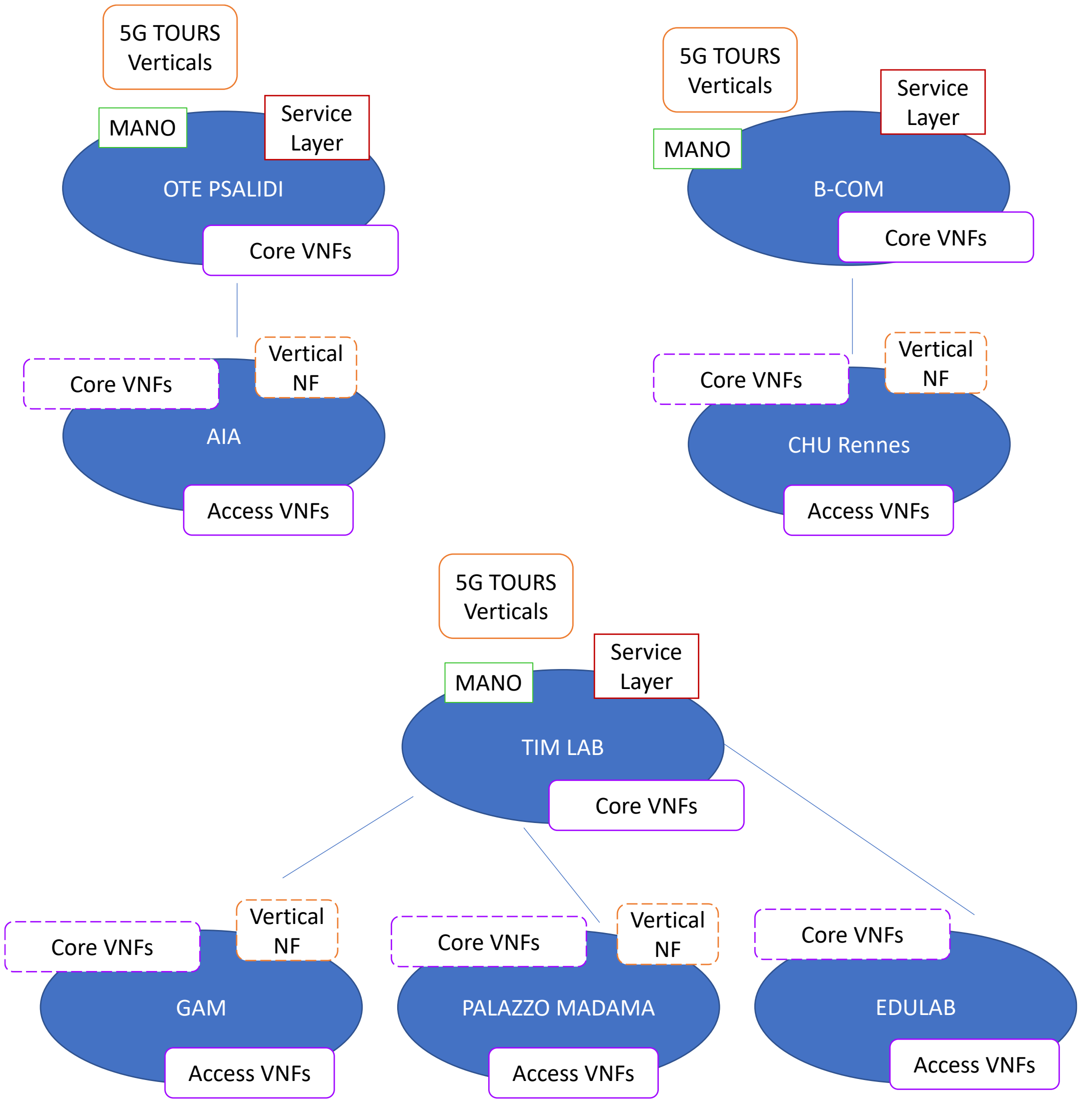


**Figure 3. Infrastructure deployment used by 5G-TOURS use cases**

Analogously, the network management and orchestration take place in these locations, as there is a global view of the available resources in all the network locations. Here, the AI modules and the Big Data analytics gather information about the current status of the network and provide intelligent policy to the orchestration framework for scaling and relocation decisions.

For these reasons, some functions of the Core Network may be relocated at the Edge (together with the Vertical Service Functions) to provide support for extreme requirements on e.g., latency. These relocations will be piloted by the orchestration framework from each centralised location.

Finally, the service layer can be deployed at each site if the bypass is used, or on top of the 5G EVE portal, if it supports all the features needed by the use case.

### 2.1.3 5G security-by-design for verticals

Wireless communications are inherently vulnerable and needs specific protection against interception and tampering. Consequently, ever since Global System for Mobile Communications (GSM)—the second generation of mobile networks—, encryption has been used on the radio interface to secure user communications. In the next two generations of mobile networks, Universal Mobile Telecommunications System (UMTS) and LTE, respectively, the security architecture was significantly enhanced. Besides encryption of user traffic, these networks have also provided mutual authentication between mobile terminals and the network, as well as integrity protection and encryption for all control and management traffic. Overall, UMTS and LTE security features ensure not only high level of security and privacy for subscribers, but, very importantly, assurance of resilience required to combat various forms of attacks against the integrity and availability of the services these networks provide. This raises the question: Are new security concepts required for the next mobile network generation? The answer is yes. On the one hand, the support of a variety of new use cases and, on the other, the adoption of new networking paradigms have made it necessary to reconsider some current elements in the approach to security. Figure 4 visualises the main drivers for 5G security.

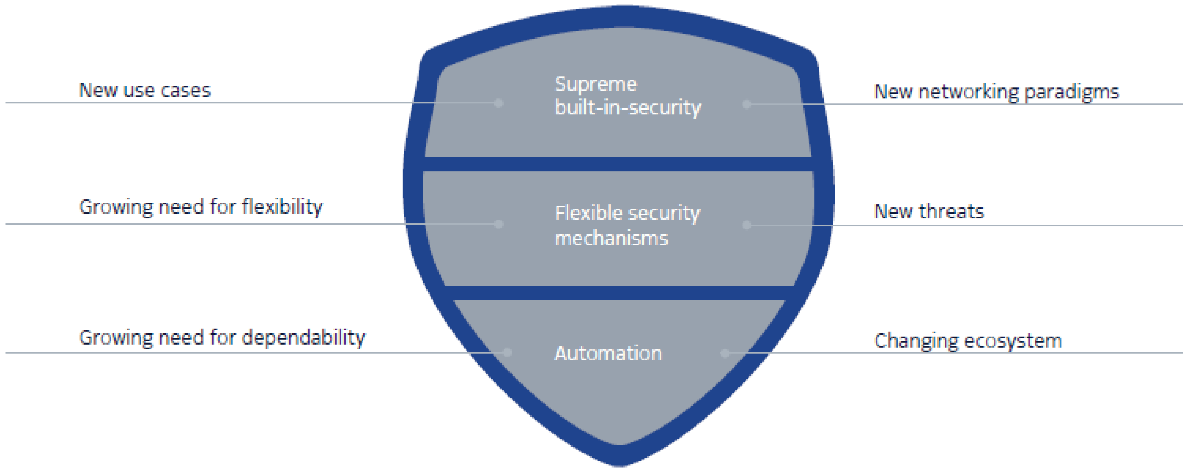


**Figure 4. 5G security drivers**

While LTE was designed primarily to support the mobile broadband use case (i.e., broadband access to the Internet), 5G targets a number of additional use cases with a variety of specific requirements. These cases include support of enormous numbers of mobile devices, as well as the need for ultra-low latency in the user communication. Use cases, such as vehicular traffic control or industry control, place the highest demands on the dependability of the network. Indeed, human safety and even human lives depend on the availability and integrity of the network services (especially for the safe city and in particular for the ambulance use case).

To support each use case in an optimal way, security concepts will also need to be more flexible. For example, security mechanisms used for ultra-low latency, mission-critical applications may not be suitable in massive Internet of Things (IoT) deployments where mobile devices are inexpensive sensors that have a very limited energy budget and transmit data only occasionally.

To efficiently support the new levels of performance and flexibility required for 5G networks, it is understood that new networking paradigms must be adopted, such as Network Function Virtualization (NFV) and Software Defined Networking (SDN). At the same time though, these new techniques also bring new threats. For example, when applying NFV, the integrity of VNFs and the confidentiality of their data may depend, to a large degree, on the isolation properties of a hypervisor. More generally, they will also depend on the whole cloud software stack. Vulnerabilities in such software components have surfaced in the past quite often. In fact, it remains a major challenge to provide a fully dependable, secure NFV environment. SDN, for its part, bears the threat that control applications may wreak havoc on a large scale by erroneously or maliciously interacting with a central network controller.

Another driver for 5G security is the changing ecosystem. LTE networks are dominated by large monolithic deployments, each controlled by a single network operator that owns the network infrastructure while also providing all network services. By contrast, 5G networks may see a number of specialised stakeholders providing end-user 5G network services, as illustrated in Figure 5.

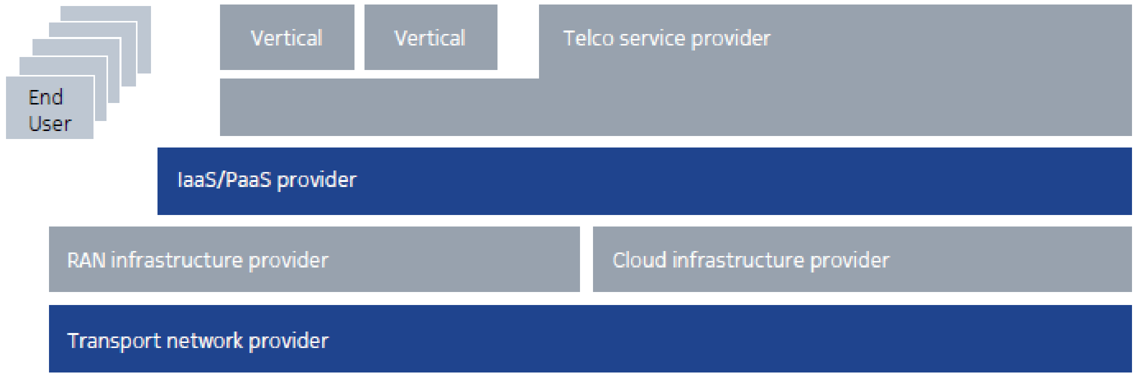


**Figure 5. Example of multiple stakeholders involved in providing end-user 5G**

In particular, there may be dedicated infrastructure providers decoupled from telco service providers that host several service providers as tenants on a shared infrastructure. In another case, telco service providers may offer not only end-user communication services, but also provide complete virtual networks or “network slices” specialised for specific applications, (such as IoT applications). These may be operated by verticals. For example, a manufacturing company could run a virtual mobile network specialised for industry control applications for its own plants. The relevant security issue here is the building and maintenance of new trust relationships among all stakeholders. The aim would be to ensure a trusted and trouble-free interaction resulting in secure end-user services.

Obviously, 5G networks must support a very high level of security and privacy for their users (not restricted to humans) and their traffic. At the same time, networks must be highly resistant to all kinds of cyber-attacks. To address this two-fold challenge, security cannot be regarded as an add-on only; instead, security must be considered as part of the overall architecture and **built into the architecture right from the start *(“security-by-design”)***. Based on a secure architecture, secure network function implementations are also essential in order to ensure a highly secured network. Security assurance methods are therefore essential so that operators can ensure the required security level for a variety of network functions.

5G security must be flexible. Instead of a one-size-fits-all approach, the security setup must optimally support each application. This entails the use of individual virtual networks or network slices for individual applications, as well as the adjustment of the security configuration per network slice. Security features subject to this flexibility may comprise the mechanisms for identifying and authenticating mobile devices and/or their subscriptions, or for determining the way that user traffic is protected. For example, some applications may rely on security mechanisms offered by the network. These applications may require not only encryption, as in LTE, but also user plane integrity protection. However, other applications may use end-to-end security on the application layer. They may opt out of network-terminated, user-plane security because it does not provide additional security.

## 2.2 Physical Implementation

### 2.2.1 Touristic City Deployment Update

The network deployed in the touristic city will be composed of a combination of (i) commercial deployments based on infrastructure owned by TIM and provided by Ericsson, and (ii) experimental/pre-commercial facilities deployed to test Rel-16/17 equipment and 5G-TOURS innovative functionalities, relying mostly on indoor environments. The integration with the 5G EVE infrastructure and platform in terms of areas to be covered, deployment requirements, timeline and deployment roadmap will also be considered. Moreover, the nodes deployed for 5G-TOURS specific use cases will be complementary components to the 5G EVE infrastructure in order to take advantage of its platform functionalities and for the innovations studied in 5G-TOURS.

All the touristic city use cases requirements have been collected to start the network design activity, and can be summarised as follows:

- **5G coverage requirements:** the use cases will be implemented at Palazzo Madama (UC1.a: In the very heart of Turin, UC2.a: Palazzo Madama exclusive exhibitions for all, UC2.c: Surveillance of the museum, UC3: Robot-assisted museum guide and UC5: Remote and distributed video production) and GAM (UC1.b: GAM and Edulab – Gamification, let’s play artist, UC2.b: Play and visit modern art from museum to school and UC3) museums. In more detail, Palazzo Madama will require a mixed outdoor / indoor coverage in order to support the UC5 of the itinerant orchestra, while the other use cases will be implemented only inside the museum, as for the use cases at the GAM. The outdoor coverage will rely on the commercial network, while for the indoor one there are two different deployment approaches that will be reflected in the phases described below. The solution that is currently being evaluated for indoor coverage is based either on Ericsson DOT or 4422 radio systems, inside Palazzo Madama and GAM museum, while the outdoor coverage extension will be implemented by installing the AIR 6488, an Advanced Antenna System (AAS). The frequency that will be used is the TIM licensed portion of band in the 3.7 GHz band (i.e. 80 MHz from 3720 MHz to 3800 MHz);
- **5G service requirements:** the network design will be based on the most stringent requirements coming from the different use cases. Such requirements are the downlink bitrate, the uplink bitrate and the latency, which have been described in WP2 and WP4 deliverables;
- **Broadband connection requirements:** the network infrastructure requires the support of a broadband wired connection over optical fibre to the museums. The optical fibres provide the connection between

the equipment that will be installed both in the museums for the indoor coverage and in the site identified for the outdoor coverage and the appropriate Ericsson baseband units;

- **Planning and operational requirements:** an important aspect that must be taken into account during the design phase of the network is related to the network planning, which consists in determining the coverage provided by the different cells. An optimal network planning will guarantee a good service provision in terms of connection stability and performance.

According to the requirements that have been defined, the large number of rooms that will be covered at Palazzo Madama and GAM, the time plan of the different use case, and, last but not least, the need for integration with the 5G EVE infrastructure (at the moment still under deployment), it would not be possible to deploy a network that would meet all these aspects at its first implementation. Therefore, it has been decided to follow a two-phase implementation approach. The main aspects of the two implementations phases are summarised hereafter:

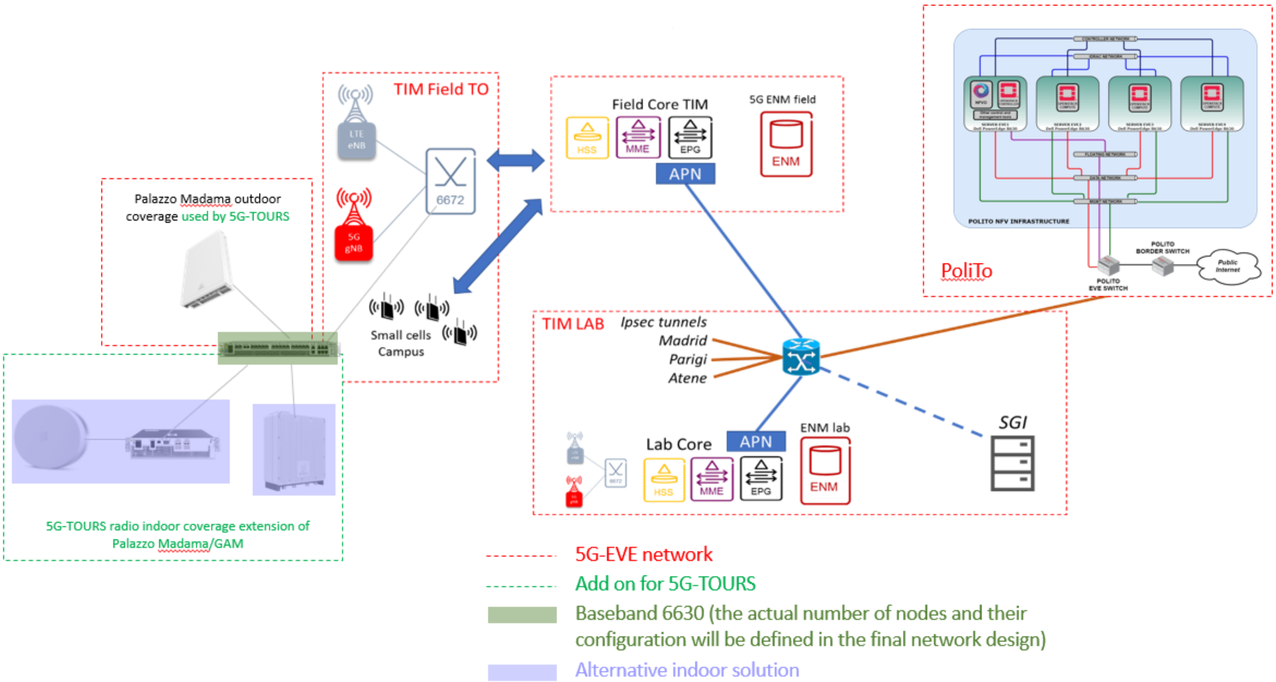


**Figure 6. Preliminary Network architecture for Phase 1**

### Phase 1

The phase 1 of the network implementation foresees that the 5G indoor coverage of Palazzo Madama and GAM will be connected to the TIM commercial network whose CN node is located in Milan (i.e. Field Core TIM). From this perspective, the indoor coverage will be a full-fledged extension of the outdoor coverage for those use case that will require both of them in an early phase of the project. The network implementation of the phase 1 will address only the objective to validate the need of 5G. From the 5G EVE infrastructure / platform integration point of view, the possibility to include the 5G-TOURS coverage infrastructure in the 5G EVE ecosystem will be evaluated. Additionally, even if the use of the 5G EVE platform (i.e. portal) will not be required for this phase, at the same time, it will not be precluded (depending on the previous point). Figure 6 depicts the current infrastructure functional blocks Details of the implementation will be available at later stages, after the completion of the undergoing network design process.

### Phase 2

The phase 2 of the network implementation foresees that the 5G indoor coverage of Palazzo Madama and GAM will be connected to the TIM laboratory network that will be provided by the 5G EVE infrastructure (i.e. 5G EVE lab CN). In this case, the outdoor and the indoor coverage will practically be two independent networks coexisting in the same area and operating on the same frequency range. The network implementation of phase 2 will address both objectives to validate the need of 5G and demonstrate the benefits of the 5G-TOURS innovations, taking advantage of the 5G EVE laboratory network functionalities that will be made available to 5G-TOURS. Phase 2 will therefore provide an almost full integration with the 5G EVE infrastructure (i.e laboratory network). Therefore, the 5G EVE platform (i.e. portal) could be used depending on the needs. Phase 2 roadmap will strictly depend on the status of the 5G EVE infrastructure deployment, and its feasibility needs to be carefully investigated in all the concerned aspects. In Figure 7 the Phase 2 infrastructure functional blocks currently considered are depicted., The option to connect the indoor radio installations to TIM Lab Core is under analysis.

Deployment roadmap foresees to implement Phase 1 running network by the end of 2020 / beginning of 2021 while Phase 2 will strictly depend on the status of the 5G EVE infrastructure.

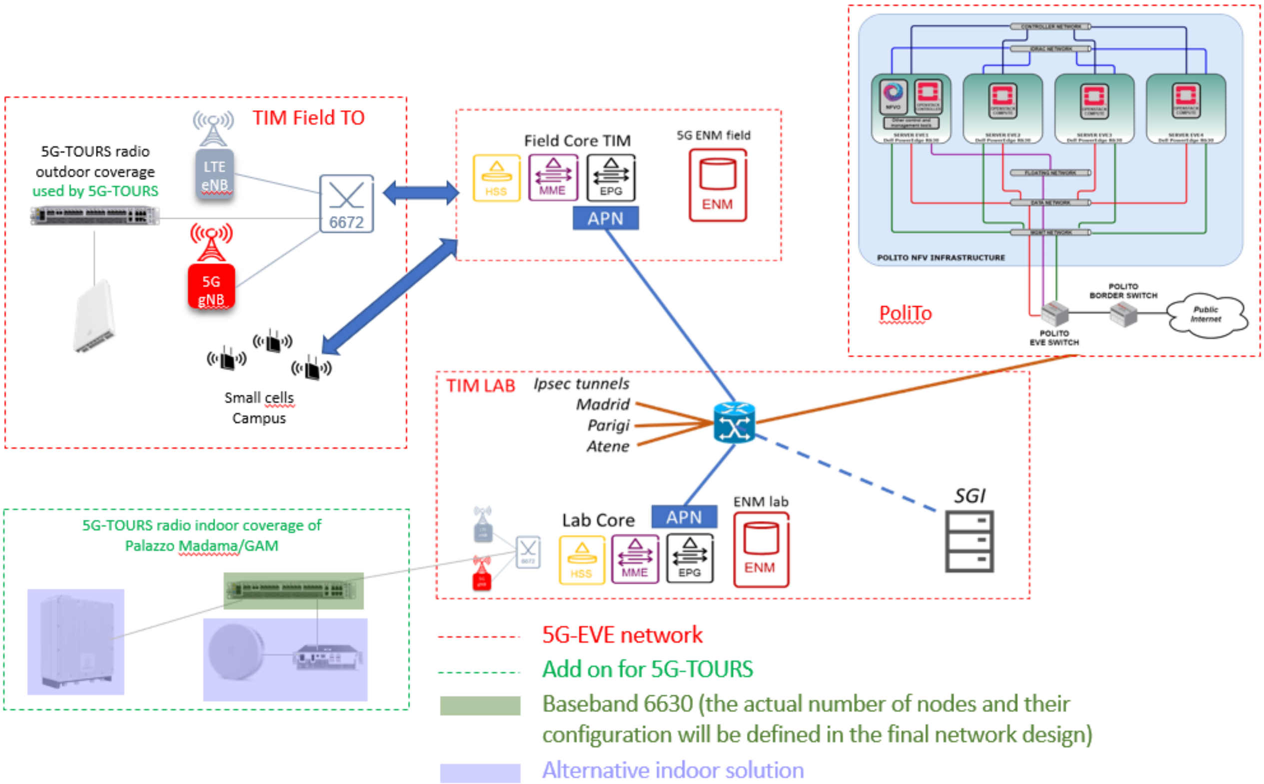


**Figure 7. Preliminary Network architecture for Phase 2**

### Network Equipment

The two-phase network enhancements implementation approach consists of the reuse of pre-existing network infrastructure (5G EVE and TIM commercial network), adding or extending 5G radio coverage at sites where the UCs will be implemented. Radio equipment will be connected to existing Ericsson Baseband 6630 node/nodes via fibre link to ensure outdoor and indoor coverage. Baseband 6630 (Figure 8) flexibly handles compute functionalities for all generations of radio access technologies, LTE and NR.

**Figure 8. Baseband 6630**

Indoor coverage solution study for Palazzo Madama and GAM museums is at the advanced phase and it is taking into account two different alternative options based on:

- Ericsson Radio DOT System 4479 and Indoor Radio Unit IRU 8846 (Figure 9): The Radio DOT solution provides indoor connectivity through remotely powered active antenna elements using standard enterprise LAN cabling, while sharing centralised baseband and radio resources;

**Figure 9. Internal Radio Unit 8846, Radio DOT4479**

- Ericsson Radio 4422 (Figure 10): The macro Radio 4422 is a 4 Transmitters and 4 Receivers radio with a high radio performance and power efficiency and it can be used also for indoor coverage. The Radio 4422 will be located near an antenna (e.g. Antenna Kathrein 80010922, Antenna Kathrein 80010564), adopting a temporary installation solution.

**Figure 10. Radio 4422 and Antenna Kathrein 80010922 and Antenna Kathrein 80010564**

The final position of the indoor equipment inside Palazzo Madama and GAM museums will be subject to final technical solution and Italian laws (e.g. Cultural Heritage Superintendence, etc) for these reasons some preliminary hypothesis has been done, but they will be finalized at the end of the low level design activity. Figure 11 represents two hypotheses of radio and antenna positioning on pole in case of Radio 4422 solution:

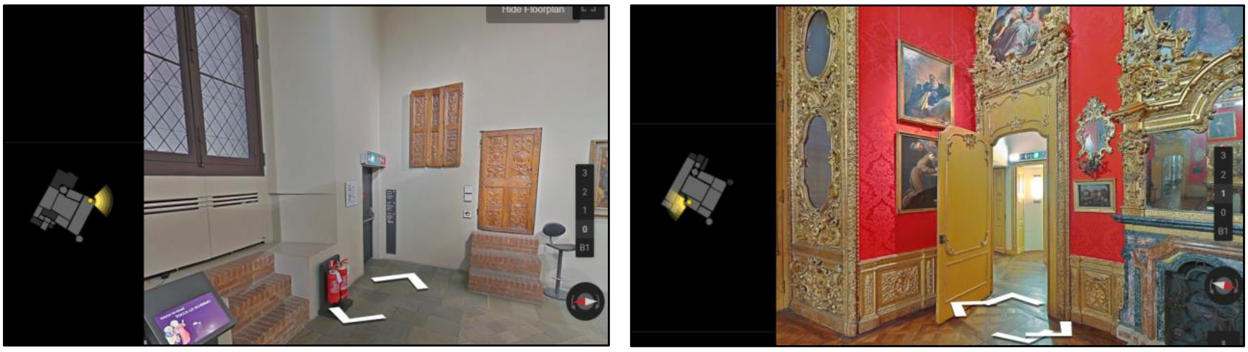

**Figure 11. Radio and Antenna positioning**

**Figure 12. Radio and Antenna installation on pole**

Outdoor coverage will be implemented by the AIR 6488 (Figure 13) mounted on the pole (Figure 12) an Advanced Antenna System (AAS), with 64 transmitters and 64 receivers.

**Figure 13. AIR6488**

#### 2.2.1.1 Broadcast Architecture

The base of this use case is the network deployed for San Giovanni in June 2019. Unlike some previous trials, the network is based on a Wi-Fi multicast distribution, were users are able to appreciate the new technology directly on their own mobile phone equipped with a special app. The available radio-frequency channel for the trials is the VHF 11 channel (216-223 MHz, centre frequency 219,5 MHz) in vertical polarisation. There will be one High-Power High-Tower (HPHT) transmitter located in *Torino-Eremo* (the main transmitting site of RAI in Turin).

The Broadcast Service and Control Center (BSCC) is provided by Rohde & Schwarz, and enables the delivery of multicast flows over 5G networks in broadcast mode. It could be feed by a multicast server from ENENSYS-Expway, providing multicast streams for OTT live or on-demand contents. The 5G broadcast transmitter is a product of Rohde & Schwarz that receives the IP signal from the BSCC and gives a fixed RF output at 730 MHz (ch. 53, UHF). This RF signal is then converted at 219.5 MHz (ch.11, VHF), amplified and sent to the antenna system. The signal is transmitted from Torino-Eremo towards the city and will be received both in indoor (at Palazzo Madama) and mobile scenario (in the city centre) as described below.

***Terminal equipment***

Since 5G broadcast terminals are not available in the market, the signal reception at the end user will be achieved by using a Software Defined Radio (SDR) equipment. In the ***indoor scenario***, the 5G broadcast signal will be received at Palazzo Madama by means of a VHF antenna for indoor reception. The receiver station will be composed by:

- 1 SDR receiver, provided by Rohde & Schwarz and iFN;
- 1 server and access point by Global Invacom to perform multicast Wi-Fi redistribution;
- Smartphones.

To reach the end users, the last mile towards the smartphones will be covered via Wi-Fi. The IP stream from the SDR receiver will feed a server which is able to receive, transcode (if requested) and apply Forward Error Correction (FEC) at application level. The server is also responsible for configuring and monitoring an ad hoc WiFi access point. The ad hoc access point, whose firmware has been modified by Global Invacom, will receive the stream from the server and send it via multicast Wi-Fi to a high number of simultaneous terminals.

On the other hand, in the mobile ***in-car scenario***, a professional van will be used, equipped with:

- 1 VHF car roof antenna;
- 1 SDR receiver by Rohde & Schwarz and iFN;
- 1 access point to cope the “last mile” towards the end user;
- 1 smartphone for in-car handling.

***Application components***

In the ***indoor scenario***, the communication between the SDR equipment and the end users will be ensured by using a special mobile app (Bx-WiFi-GI for Android/iOS), developed by Global Invacom. In the mobile ***in-car scenario,*** a single terminal will be used, connected to the SDR via Wi-Fi. In this case the final user can use a free downloadable client app (such as VLC, Good Player, etc) to watch the video.

### 2.2.2 Safe City Deployment Update

All use cases in the “Safe City” will be trialled in Rennes, using the mobile network infrastructure of Orange (for use cases 6 and 9) and the experimental network of BCOM based on Amarisoft and Nokia hardware (for use cases 7 and 8).

**Phase 1**

During the first phase, use cases 6 (remote health monitoring and emergency situation notification) and 9 (optimal ambulance routing) will use the existing LTE-M mobile network of Orange since they rely on IoT sensors and devices. They will also make use of the 4G LTE network for mobile broadband (MBB) communication of Orange. For use cases 7 (teleguidance for diagnostics and intervention support) and 8 (wireless operating room), an experimental 5G network will be deployed to be used in Ultra Reliable Low Latency Communication (URLLC) use cases. There will be two deployments of 5G NR wireless coverage:

- Outside: at the BCOM premise, for the connected ambulance (Figure 14). A suitable 5G NR antenna will be installed on the roof of the BCOM building, using primarily the 26 GHz frequency band;


**Figure 14. 5G-TOURS 5G NR NSA/SA wireless coverage at BCOM**

- Inside: at the wireless operating room at CHU Rennes to provide high-speed, low-latency wireless access for medical imaging equipment, using 26 GHz for data transmission and 2.6 GHz as the anchor frequency band (Figure 15).

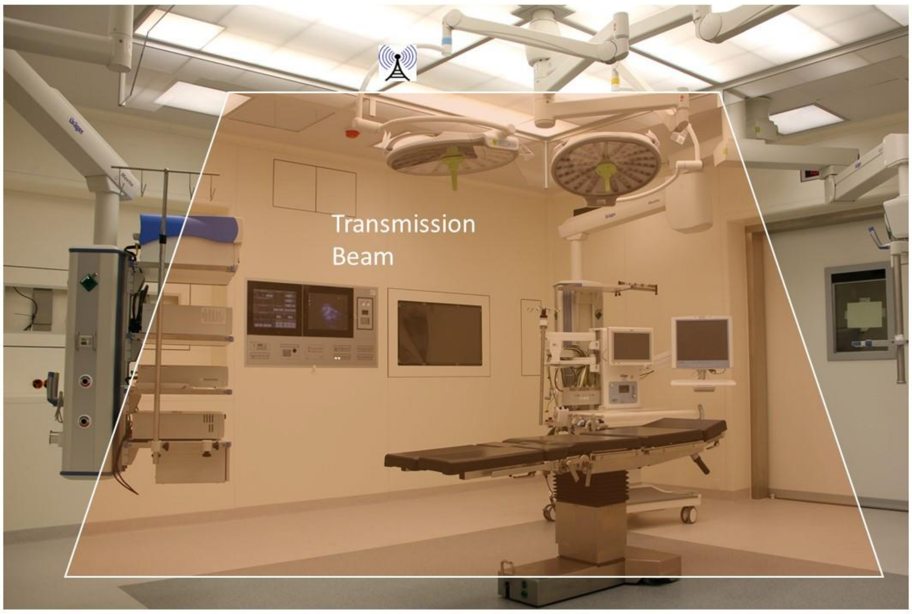


**Figure 15. 5G-TOURS 5G NR wireless coverage in the Wireless Operating Room at CHU**

At the BCOM premises there will be a 5G base station with a local virtual User Plane Function (UPF), part of the so-called “Wireless Edge Factory” (WEF) [18]. Similarly, there will be a WEF UPF at the hospital that connects to the WEF core network hosted in the BCOM datacentre through a dedicated Virtual Private Network (VPN) backbone. This will enable the setting of end-to-end (E2E) network performance KPIs and the prioritisation of data traffic between the ambulance and the hospital to guarantee the required quality of service (QoS). Furthermore, the WEF Core Network deployed in BCOM data centre will manage the WEF UPF at the hospital to connect the 5G terminals of the Wireless Operating Room.

The network equipment consists of the following three layers: control plane, user plane, and RAN network equipment. The purpose of these layers in the 5G-TOURS project is described below.

The control plane is a virtual 4G Core Network compatible with the 5G Non-Stand Alone (NSA) standard. It is deployed as a set of Docker containers managed by a Kubernetes cluster. This cluster is hosted on the Flexible Netlab platform [19] in the BCOM data centre.

The user plane equipment provides connectivity between the RAN equipment and the data network (Internet). The main component is the UPF component of the WEF provided by BCOM. Two instances of the UPF will be deployed as part for 5G-TOURS.

The first one will be a purely virtual UPF deployed in BCOM data centre as a virtual machine hosted on an OpenStack cluster provided by Flexible Netlab. This virtual machine hosts an OpenVSwitch (OVS) virtual switch that acts as a tunnel endpoint for the GTP tunnels coming from the RAN equipment deployed at BCOM for UC7. It is thus used to connect this RAN equipment to the Rennes CHU through the VPN between BCOM and the CHU.

The second one is an appliance built from a commercial off-the-shelf (COTS) network switch and a COTS mini-ITX PC. The PC is a KVM hypervisor that hosts an OVS-based virtual machine similar to the one deployed in Flexible Netlab. It will be deployed in the technical room of the Rennes CHU and will interconnect the RAN equipment deployed there with the various components required by UC8. The same WEF Control Plane will manage this switch through the VPN established between BCOM and the CHU.

For 5G-TOURS, we will use Nokia Small Cell technology as our RAN equipment. Two Small Cells will be deployed, one at the Rennes CHU to provide coverage for the Wireless Operating Room, and one at BCOM premises to cover the outside area for UC7. Both will use the 26GHz/2.6GHz bands in 5G NSA mode. In addition, we will conduct the first integration tests for UC8 in BCOM showroom. This will be done with an Amarisoft Classic Callbox RAN equipment. This equipment uses the 3.5GHz band and is also compatible with 5G NSA mode.

All the medical equipment that require 5G wireless connectivity will connect to this RAN equipment through 5G Customer Premise Equipment (CPE) component. A high bandwidth mmWave equipment is required, and BCOM is analysing the options between Samsung, Huawei and Sony commercial equipment working in this frequency range. The exact model is yet to be determined.

**Phase 2**

During the second phase, use cases 6 and 9 will rely on a 5G massive MTC (mMTC) network (release-15) once it becomes available. For use cases 7 and 8, the experimental 5G network will be integrated with the 5G EVE infrastructure. The WEF core network and user plane will be deployed by the ONAP orchestrator hosted by Orange in their Châtillon datacentre as part of their 5G EVE infrastructure. The data centers of BCOM and Orange are already connected in the scope of the 5G EVE project (Figure 16). If possible, the Nokia RAN could also be orchestrated through 5G EVE.

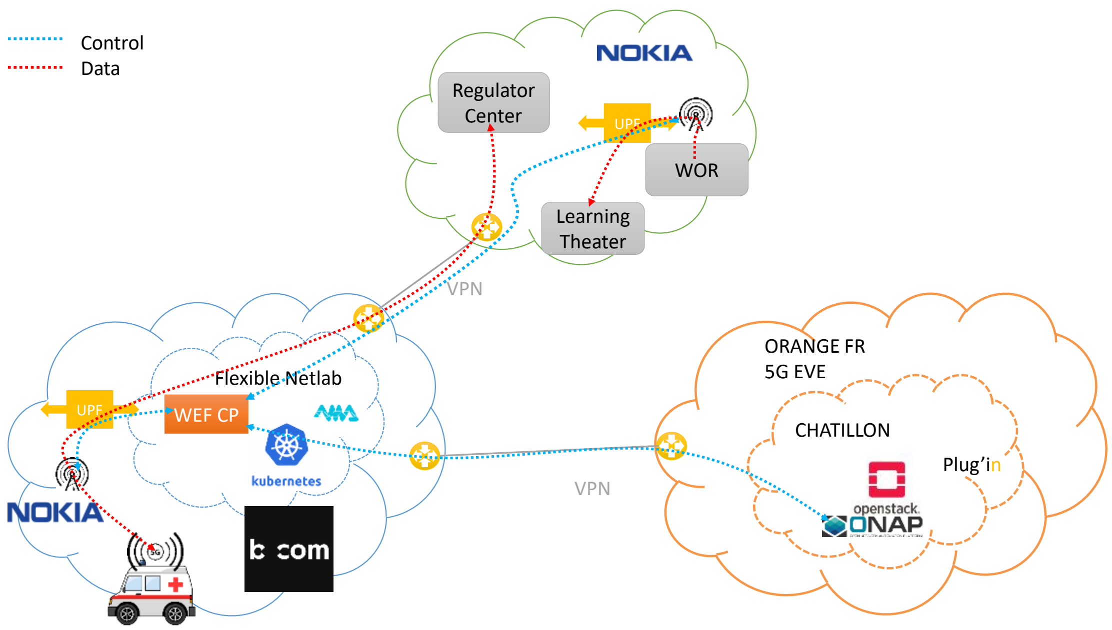


**Figure 16. Overall network architecture and physical deployment of network equipment and functions**

### 2.2.3 Mobility-Efficient City Deployment Update

The Mobility Efficient City network deployment will largely rely on the available Core Network infrastructure from 5G EVE Athens for the vertical trials. It relies on a 4G Core in NSA mode with a 5G migration in phase 2, using equipment provided by Nokia. The infrastructure consists of RAN, offering 5G NSA (Option 3X) support for phase 1, and 5G SA support for phase 2 of the project, virtual Evolved Packet Core (vEPC) (phase 1) and 5GC (phase 2) elements and tools for the management and orchestration of the NFV infrastructure.

**Mobility-Efficient City RAN deployment**

Three locations will be used for RAN equipment deployment in order to support the use cases described, as given below.

Regarding the UC10, smart parking management use case and the UC13b: VR Bus Excursion, an air scale Baseband Unit (BBU) will be placed in the control room of building B11 in Athens International Airport. One pair of 4G/5G outdoor antennas (non-standalone solution) will be placed at the rooftop of the same building. A backhaul connectivity will run from B11 to the main control room of the airport. From there, a 10Gbit Ethernet connection is already established to Packet Core network (Figure 17).

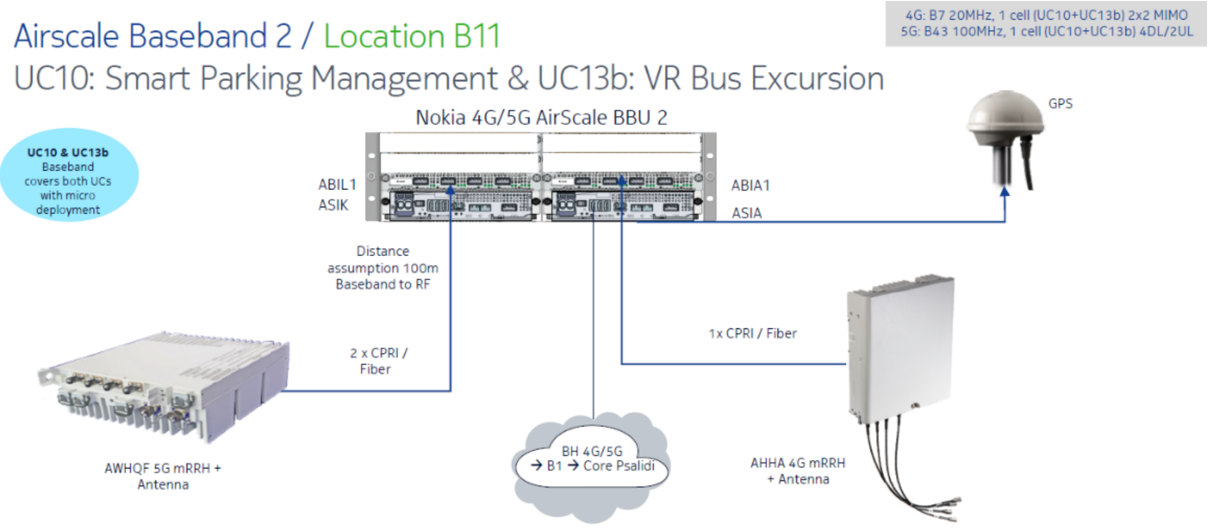


**Figure 17. Architecture deployment for UC10 and UC13**

Figure 18 illustrates the control room of building B11 and the antenna mast on the roof top of the same building in Athens International Airport.

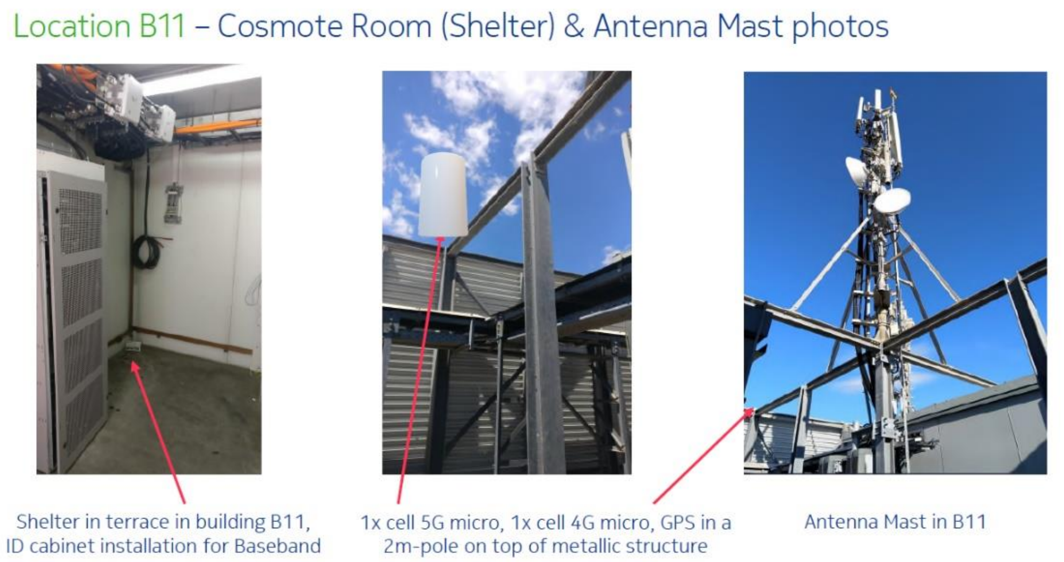


**Figure 18. 5G RAN equipment in the Athens International Airport**

Regarding the UC11 airport evacuation use case, an air scale BBU will be placed in the control room of building B2 in the Athens International Airport. Three pairs of 4G/5G indoor antennas (non-standalone solution) will be placed on the second floor of the same building (red square in Figure 19). Regarding the UC13a, the AR Myrtis Exhibition use case, one pair of 4G/5G indoor antennas (non-standalone solution) will be placed at the predefined Exhibition area Figure 20). A backhaul connectivity will run from B2 to main control room of the airport. From there a 10 Gbit Ethernet connection is already established to the packet core network.

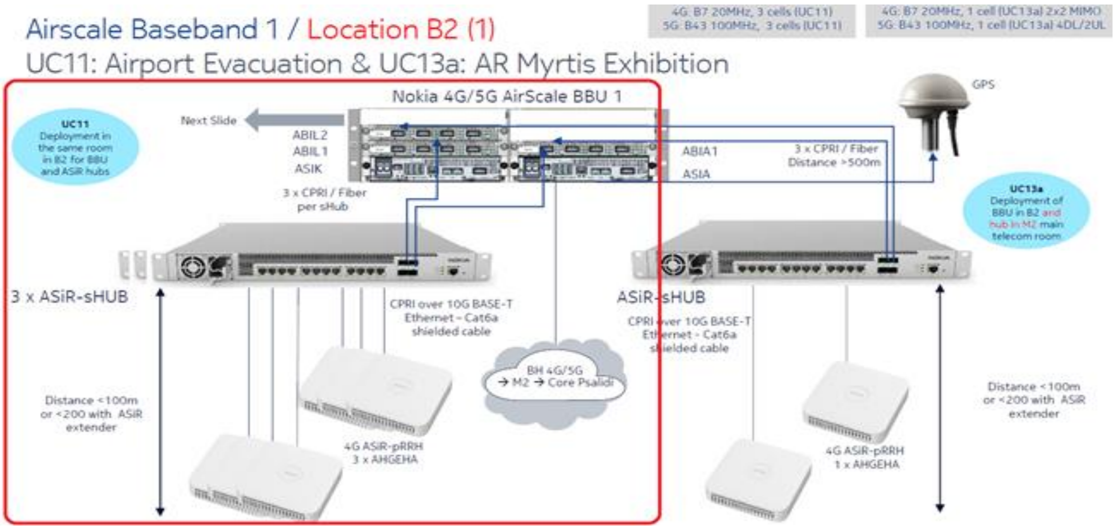


**Figure 19. Backhaul 5G equipment in the Athens International Airport**

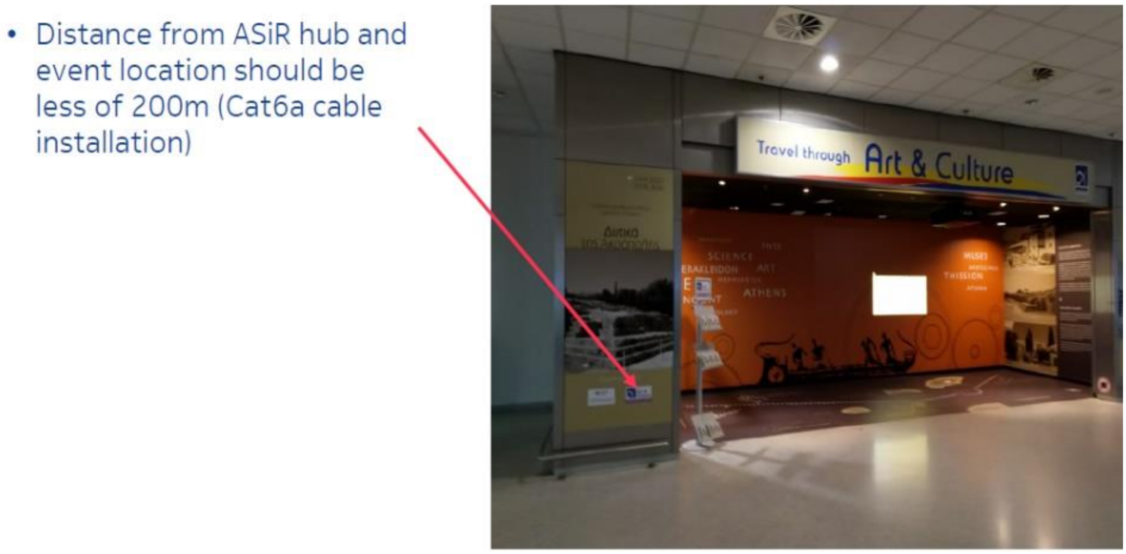


**Figure 20. 5G equipment deployment for AR Myrtis art exhibition**

Regarding the UC12, the Video enhanced runway vehicles use case, an air scale BBU will be placed in the control room of building B2 in the Athens International Airport (Figure 21). One pair of 4G/5G outdoor antennas (non-standalone) will be placed on the rooftop of the same building. A back-hall connectivity will run from B2 to main control room of the airport.

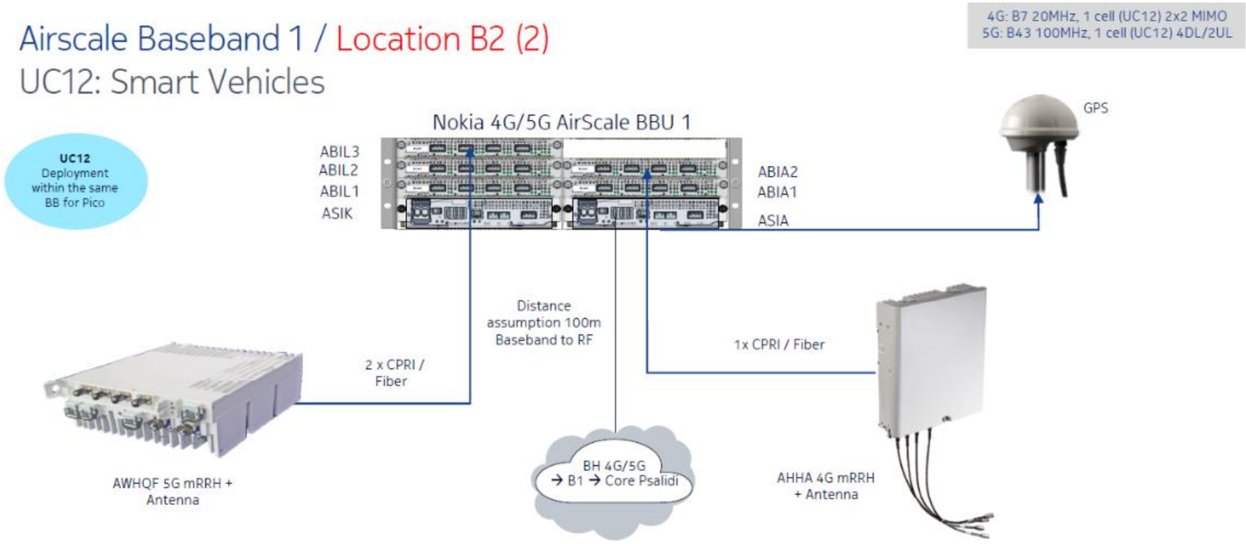


**Figure 21. 5G Architecture for Smart Parking use case**

Figure 22 illustrates the control room of building B2 and the antenna mast in the rooftop of building B2.

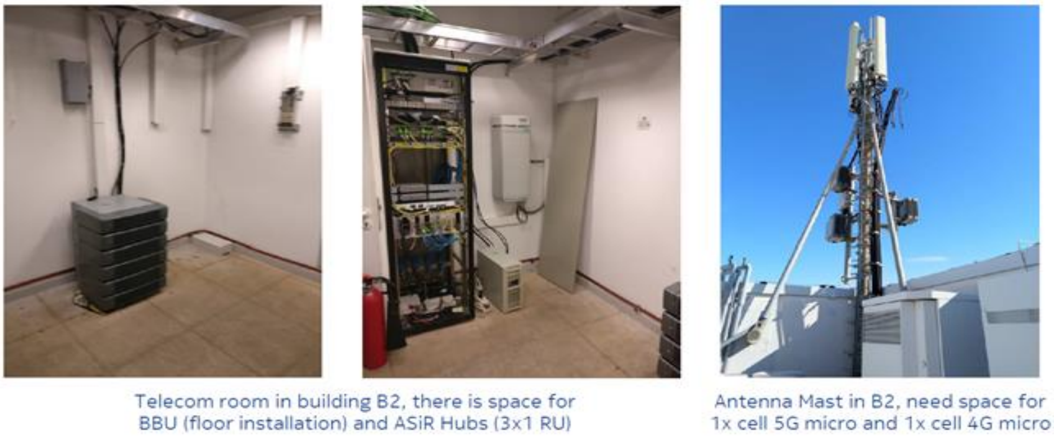


**Figure 22. 5G Equipment for Smart Parking use case**

Note that the same RAN equipment will be used in both phase 1 and phase 2 without needing to change installation location, where in phase 2 the equipment will be used as NR Standalone node.

**5G-TOURS Core Network deployment**

The 5G-TOURS Core Network will reuse 5G EVE Core Network deployment for both phase 1 and phase 2 of the project.

**Phase 1**

Regarding phase 1 deployment, a vEPC Core Network will be used and connected with 5G NSA RAN as described in other deliverables and shown in Figure 23. Currently the 5G EVE team is working on setting up the platform at OTE premises in Psalidi, Attiki. In phase 1 of the project, 5G-TOURS will reuse 5G EVE platform management and configuration layer for network configuration and Use Cases setup.

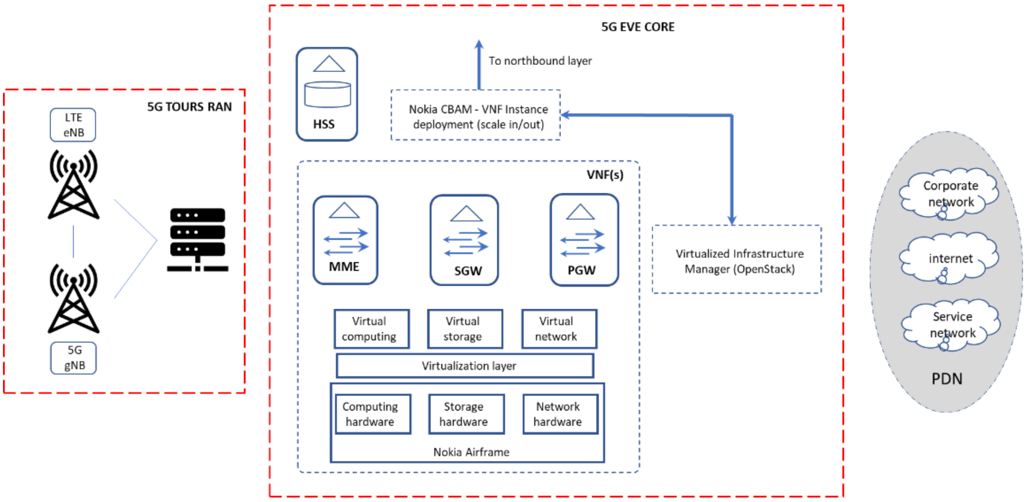


**Figure 23. Current Core Network deployment for the Mobility-Efficient City**

### Phase 2

Regarding the phase 2 deployment, the vEPC will be replaced by 5GC consisting of Access and Mobility Function (AMF), Session Management Function (SMF), Unified Data Management (UDM) function, and possibly Network Slice Selection Function (NSSF) and Network Repository Function (NRF), as shown in Figure 24. The 5GC will be connected with 5G RAN acting in Standalone (SA) mode. In phase 2 of the project, 5G-TOURS will use enhanced MANO capabilities offered by the project. Currently the team is investigating how much of these MANO extensions will be used during demonstration of the Mobility Efficient City Use Cases.

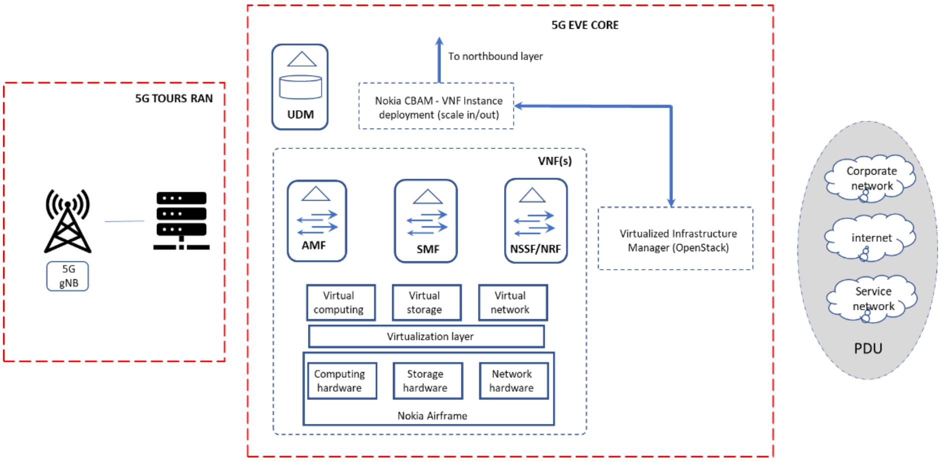


**Figure 24. Phase 2 Core Network deployment for the Mobility-Efficient City**

The mobility-efficient city network deployment will largely rely on the available Core Network infrastructure from 5G EVE Athens for the vertical trials. It relies on 4G (and eventually 5G) equipment provided by Nokia. The infrastructure consists of RAN, offering 5G NSA (Option 3X) support for phase 1, and 5G SA support for phase 2 of the project, virtual Evolvev Packet Core (vEPC) Core Network (phase 1) and 5GC (phase 2) Core Network elements and tools for the management and orchestration of the NFV infrastructure.

All use cases for the mobility-efficient city will be deployed at Athens International Airport site, as follows:

- Use Case 10: Smart Parking Management system for real time assistance to visitor on available parking spaces with information about location of empty parking spots;
- Use Case 11: Airport Evacuation offering automated guidance to emergency routes via mobile phone application;
- Use Case 12: Video enhanced runway vehicles providing high definition (4K) live stream to airport control centre;
- Use Case 13a: Augmented Reality (AR) Myrtis Exhibition offering high quality digital (images, 3D models, video) content to airport visitors;
  Use Case 13b: Virtual Reality (VR) bus Excursion offering to passengers of school/excursion bus high quality and personalized rich content using VR technology.

# 3 5G-TOURS Technology Evolution

Within the 5G-TOURS Architecture there are novel technology evolutions that will i) enhance the state of the art solution for 5G Networks and ii) are necessary enablers for many of the use cases that will be presented by the project [20] [21] [22]. Specifically, 5G-TOURS will provide enhancements to the MANO framework (see Section 3.1), bleeding edge solutions for the Broadcast Support for media use cases (see Section 3.3), the Service Layer (Section 3.4,) and the AI Orchestration (Section 3.2, with its extensions to the RAN support, Section 3.5).

As already partially discussed in the Architecture description (Section 2.1), although the 5G-TOURS framework will leverage on the 5G EVE infrastructure, enhancing it to support specific functionalities, some technologies will be native to certain sites, as their main goal is to support use cases that will be showcased in one site only. The mapping is represented in Table 1.

**Table 1. Technology mapping to physical sites**

| Technology | Site | Notes |
|---|---|---|
| **Enhanced MANO** | Turin | Integration of the OSM enhancements discussed in Section 3.1 |
| | Rennes | Integration of the ONAP extensions to the ORA-FR deployment. |
| **Broadcast Support** | Turin | Development of the Broadcast Support solution in TIM LAB under study (see Section 3.3). HPHT trial in Eremo (see Section 3.3) |
| **Service Layer** | Athens | Implementation of the KPI exposure functionality to tenants (see Section 3.4) |
| | Turin | Dynamic SLA requirements from the tenant (see Section 3.4) |
| **AI Orchestration** | Turin | Proactive scaling of VNFs, as part of the ETSI ENI PoC (see Section 3.2). Extension to the RAN to support AI. |

Some of the solutions (e.g., some of the Service Layer as specified in Section 3.4.2.3), will bypass 5G EVE and directly act on the local MANO, however we will work with the 5G EVE colleagues to integrate the solutions in the portal as much as possible.

## 3.1 Enhanced MANO

In order to support the novel functionality envisioned by other innovative modules in the 5G-TOURS architecture, such as the Service Layer or the AI management of the network, the MANO shall be enhanced to take into account these peculiarities, providing interfaces and procedures that can be leveraged for that purpose.

Nowadays, there exists a collection of 5G orchestrators exist in the market. Some of them come from research projects (e.g. SONATA [23], 5G-Transformer [24], 5G-MoNArch [25]), while others are proprietary (blueplanet [26]) or from telco vendors (Nokia CloudBand [27], Ericsson cloud manager [28]). On the other hand, first-class NFV orchestrators exist, which have a significant impact on the telco landscape for the following reasons: a) high configurability, b) transparent roadmap, c) no license required, d) multi-vendor, among others. It is important to highlight ONAP [7] and OSM [29] as two prominent solutions in the scope of open source of NFV orchestrators. Both are being used in 5G EVE to orchestrate the deployments for 5G-TOURS.

5G-TOURS has been investigating how to cover the aspects related to smart mobility for tourism in Europe. In this process, a set of developed use cases have been analysed. The deployment of the network services requires two main characteristics: a) Novel network service management mechanisms b) Predictive algorithms for network service optimisation, both to support a massive demand of tourism services and to improve the flexibility, scalability, reliability of the use cases.

The identified features to experiment as an extension of the MANO orchestrators are:

- Deployment of Artificial Intelligence agents for network service optimisation;
- Application of the network slice primitives for vertical slices in day 1-2 operations (configurations performed after the initial setup and launch);
- Extended monitoring to extract information from mobile devices towards service optimisation.

In [6], 5G-Transformer and 5G-MoNArch orchestrators were analysed, where the benefits of orchestrating network services were highlighted. The use of OSM as the NFV Orchestrator in 5G EVE gives the opportunity to enhance OSM towards providing features for the 5G-TOURS use cases. At the same time, it is possible to create an impact in the OpenSource community by contributing to ETSI in the standardisation and evolution of the OpenSource MANO (OSM).

The possibility of including these enhancements in both the OSM codebase and deployed in the sites using OSM will be considered according to the different phases of the sites (as discussed in Section 2.2).

### 3.1.1 Enhancements to support AI management of the Network

The first planned extension to the MANO is related to the native support of AI and Big Data algorithms. The release-EIGHT of OSM (released in mid-July 2020) implements the feature #8157 [30] (VNF Indicator collection using exporters from Prometheus, an open-source monitoring solution [31]) that includes a definition of an Execution Environment (EE). This EE is the base implementation for the feature regarding the deployment and management of artificial intelligence agents. In particular, it is the deployment of the AI agents that will be executed in this EE.

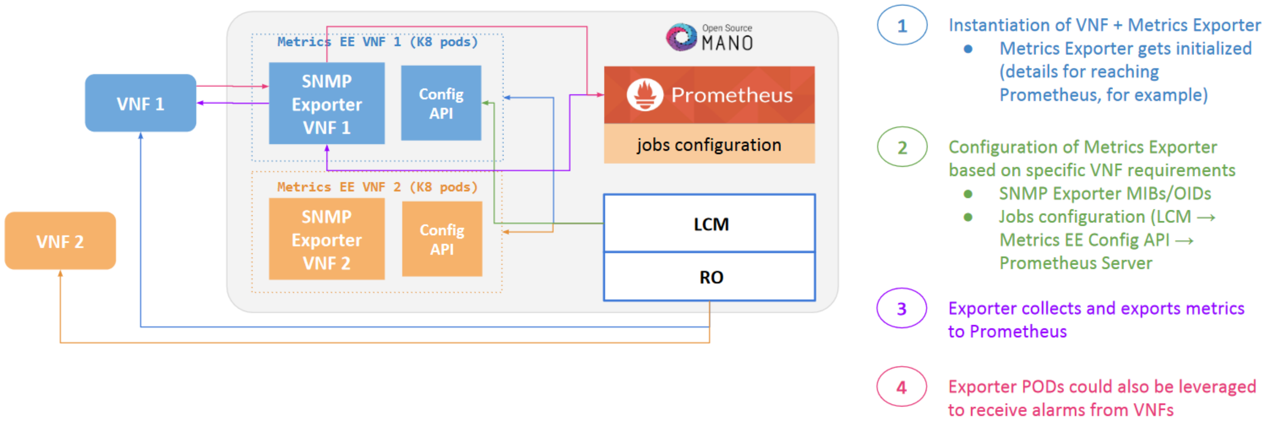


**Figure 25. OSM Feature 8157 – General Flow**

The MANO enhancement planned in 5G-TOURS is aligned with the OSM feature #8157 presented in Figure 25. The application on the user device is sending data streams with anonymised user information to the VNF. Later, the information is collected by specific exporters and processed by artificial intelligence agents.

In OSM#9 the feature #9063 [32] was presented which includes artificial intelligence agents for network service optimisation. Currently, the Service Assurance module in OSM reactively manages the monitoring close loop. This Service Assurance module is made up of the so-called "MON" and "POL" modules. The first one covers the basic monitoring of VNFs with a solid architecture able to expand the monitoring parameters using exporters; the second one is used to process auto-scaling operations and alarms. For triggering alarms thresholds can be defined per KPI, and the system reacts when a threshold is surpassed. However, this current behaviour represents some limitations, since it imposes a reactive model where the decision is taken by a fixed limit and is not adjusted by the evolution of the KPI itself.

While the reactive model being used might be a reasonable approach for many cases, it might be suboptimal for complex network services, and mainly, in cases where it is required to sense multiple and not-clearly related metrics. To overcome this, OSM feature #9063 [32] proposes to include Artificial Intelligence agents for dynamically detecting anomalies in the network service using machine learning instead of a static based technique, thus optimising the network service accordingly. It also predicts any defects that may happen in the future by applying anomaly detection techniques in the forecasted metrics value. The feature includes: a) libraries to collect the metrics from Prometheus, b) libraries to set the context of the model (NS VNF IDs), c) libraries to produce the alarms to POL, d) Northbound Interfaces (NBI interfaces) and OSM client commands for Agent lifecycle, e) Life-Cycle Management agent handling.

### 3.1.2 Enhancements to support the service layer

Besides supporting the execution of AI algorithms such as the ones described in Section 3.2, an enhanced MANO shall also provide interfaces for their correct operation and the configuration of policies such as, for instance, the ones envisioned by the Service Layer (see Section 3.4).

So, when enabling Artificial Intelligence agents in enhanced MANO frameworks, it is imperative to have an execution environment to run the agents. As is shown in Figure 26, there are three key elements when designing AI agents:

- **Interface to observe:** The observation will take place between the AI agent and the metrics produced by the environment. In Figure 26, the metrics will be fetched from the Prometheus server.
- **Interface to act** based on the algorithms developed in the agent: The actuation will be between the AI agent and the Policies module of OSM. The reason for this way of actuation is because it is planned to have a scoring decision table that does not take the same action for the same alarm.
- **Context:** This is the information that is injected to the agent to provide the environment of operations. The AI Agent will receive the VNFs/Network Services/Slices identifiers to lookup the metrics in Prometheus and execute the model towards the interface towards the POL module.

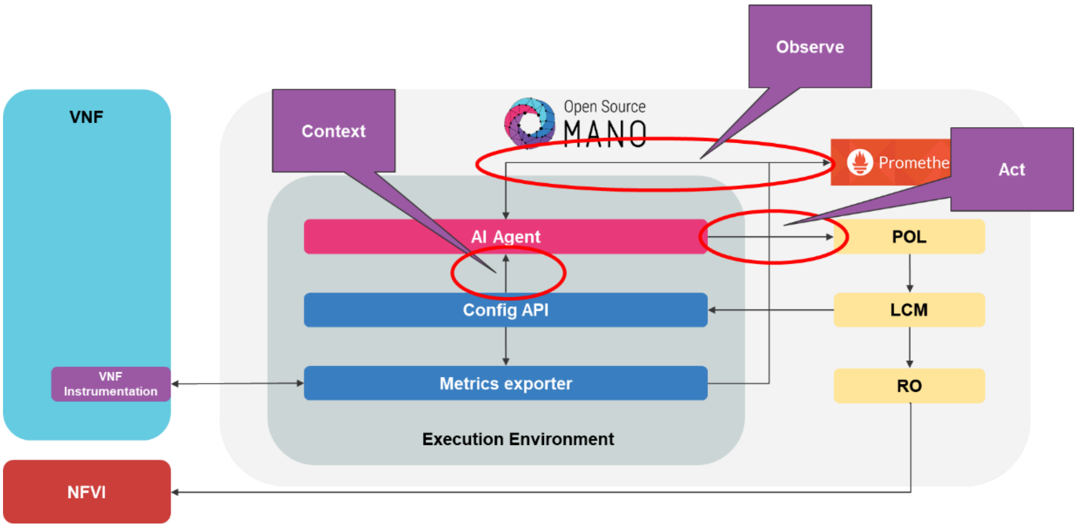


**Figure 26. Enabling Artificial Intelligence Agents in EE - OSM**

In 5G-TOURS, another architectural option is being explored (Figure 27), where the VNF is not instrumented to send metrics to a metrics exporter. Instead, there is an AI agent deployed in the VNF as ancillary Virtual Deployment Unit (VDU). The AI agent deployed as VDU of the VNF will collect the information and metrics directly from the VDU to be optimised, and it will send the actions to the AI agent proxy. Finally, the AI agent proxy will translate the actions from the VDU AI agent to POL.

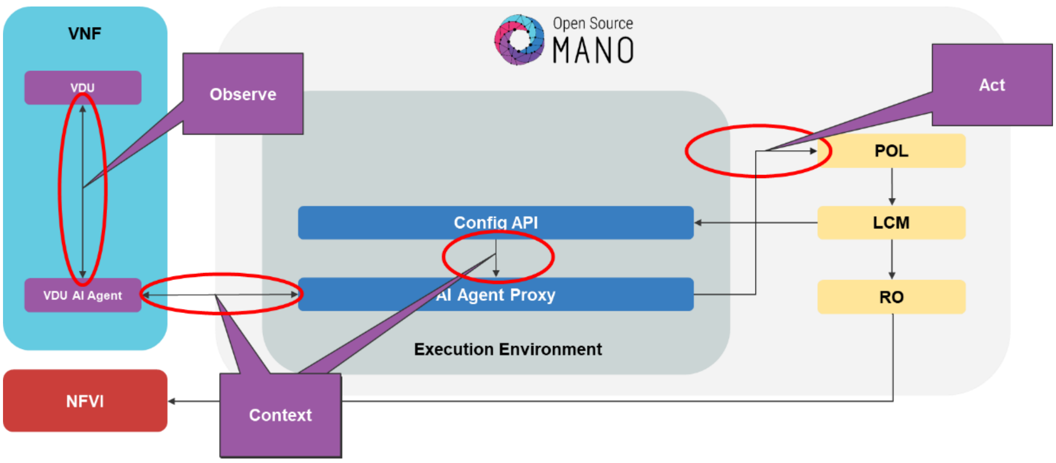


**Figure 27. Enabling Artificial Intelligence Agents in with AI Agent proxy - OSM**

As is shown in Figure 28, the process of creating the models is not part of the enhanced MANO and takes place outside the service platform. Existing OpenSource communities focused on producing the Software Development Kit (SDK) and marketplaces to distribute AI models such as Acumos [33]. 5G-TOURS could take advantage of pre-created models and use them in the AI agents to optimise their network services.

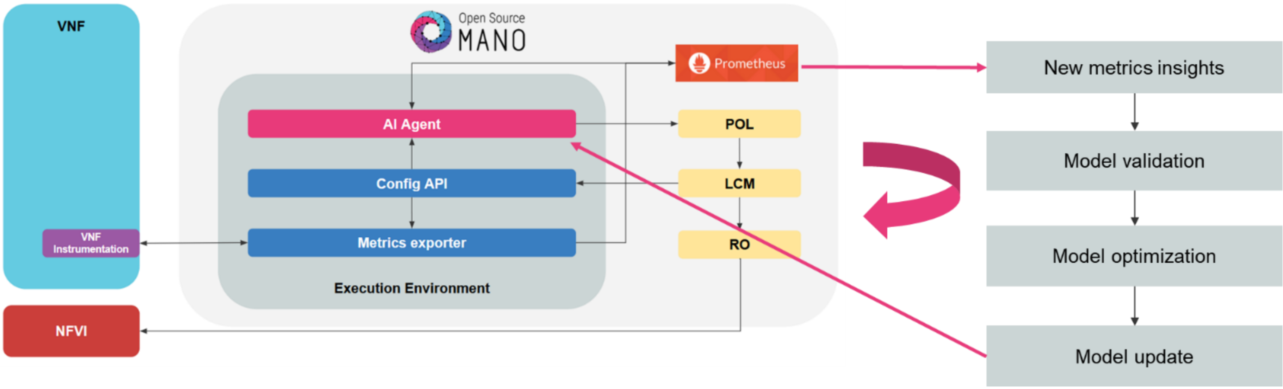


**Figure 28. Open loop for AI model optimisation**

The last optimisation required in 5G-TOURS for the use cases from the MANO point of view is about the usage of network slice primitives, which makes possible to manage Network Slices as single entities (instead as a set of separated VNFs). OSM feature 7187 [34] related to network slice primitives is now pending for design and development. This feature, in combination with the AI agents and extended monitoring, will allow 5G-TOURS to produce specific actions supporting the outcomes of the AI agents towards service optimisation of the vertical slices. Note that the primitives that will be developed and handled are at vertical slice level and will help the management of day 1-2 of our vertical slices for the use cases.

## 3.2 AI Orchestration

5G-TOURS envisions a novel set of use cases for current and future cities that are expected to greatly improve the way people live and experience them and enhance people's quality of life within those cities. To achieve this goal, the use of 5G networks has been envisioned in order to take advantage of the lower latency and increased data rate (among others) that the 5G technology promises. The envisioned use cases are based on verticals, such as multimedia, e-health, and transportation, that present different traffic patterns and have diverse traffic requirements, such as latency and throughput. Although 5G makes a shift towards a vertical service model, where the provided services are tailored to the needs of the industry, mainly through the use of NFVs, network slicing and MANO, the current state-of-the-art approaches only provide a generic functionality that is not able to take into account the diverse requirements of the different verticals and use cases. The envisioned use cases call for accurate resource provisioning, high resource efficiency, as well as minimised network latency, while respecting the agreed SLAs between the network operator and the vertical industries.

Furthermore, the deployment of the different use cases envisioned for 5G-TOURS will take place over the existing infrastructure provided by 5G EVE. The aspiration of 5G EVE is to create the foundations for a pervasive roll-out of end-to-end 5G networks in Europe. 5G EVE supports this fundamental transition by offering facilities to validate network KPIs and services to vertical industries and the 5G-TOURS project. The 5G-TOURS architecture (Section 2.1) is built on top of the 5G EVE platform which, in turn, is based on commercial and pre-commercial equipment provided by the vendors of the consortium. As a result of all the above, in order to be able to efficiently serve the envisioned use cases, as well as the different vertical industries, it is important to design and implement a new and efficient AI-assisted MANO entity that will be able to operate in a multi-service, multi-slice and multi-tenant environment (Figure 29).

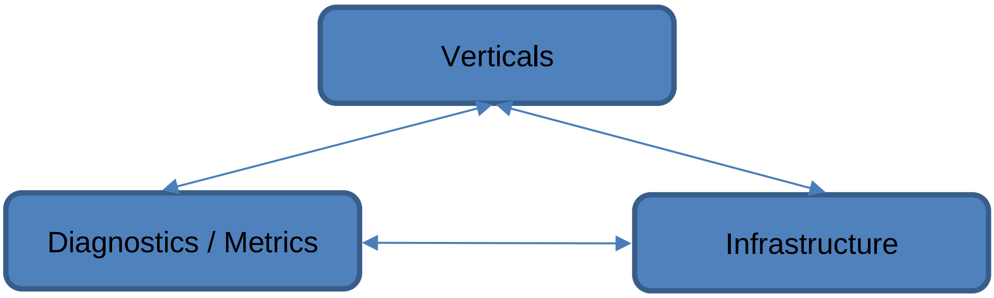


**Figure 29. AI orchestration**

In order to design such an AI-assisted MANO entity different aspects and directions will be followed. To begin with, a VNF profiling mechanism will be implemented (Section 3.2.1). VNF profiling can be of significant importance in order to model the relations between the VNFs and the different KPIs that need to be met (e.g.

resource allocation). Such models can then be used to generate input knowledge to different AI-based predictive mechanisms to aid the decision making for resource management. VNF profiling can also be used to fine-tune the SLA agreements between the network operator and the vertical industries served by the network.

Further to the VNF profiling, we will work on AI-assisted capacity forecasting (Section 3.2.1), as well as efficient network slicing based on data analytics (Section 3.2.2). Capacity forecasting will allow the system to have a clear view of the current and expected resource requirements in order to adequately allocate and reschedule the available resources. By predicting future demands, it will be possible to decide on an accurate amount of resources to serve the different verticals, thus increasing the system efficiency and reducing the operating costs. As network slicing is an important aspect of 5G networks, we will also explore the use of data analytics to develop mechanism that provide a holistic solution for both resource and service orchestration.

Given the diverse requirements of the several verticals that 5G networks are expected to support, it is also important that the most efficient AI mechanism is used for different verticals and use cases that is better tailored to that vertical's requirements. Therefore, as part of the AI Innovation and Orchestration, we will also look into mechanisms that are able to compare and contrast between the different AI-based algorithms that aim to serve different use case requirements, and provide recommendations on the most efficient algorithms to use and approaches to follow (Section 3.2.4).

All the above lines of work aim at developing novel algorithms that address the different requirements of the use cases and verticals that will be served within 5G-TOURS. In order to be able to communicate the requirements among the use cases and the network, we will finally look into building an enhanced MANO platform that includes different interfaces to allow the communication between the vertical industries and the network. The following sections provide more details into the different lines of work that will be followed.

### 3.2.1 5G-TOURS vertical's VNFs profile building

In 5G-TOURS project, we have different use-cases related to three main vertical industries, which are developing VNFs without knowing the hardware (HW) capabilities in 5G EVE platform that has to support and deploy many VNFs coming from different verticals and projects. These vertical's VNFs have different hardware requirements. For that reason, we propose to profile the vertical's VNF in order to fine tune the SLA level that should be injected in the 5G EVE platform via the experiment blueprint. The VNF profiling will be built based on the feedback from 5G EVE platform via the monitoring framework. It is important to highlight, that profiling and benchmarking are a new focus point of ongoing NFV related research [35]. In the next section, we propose to clarify both techniques.

#### 3.2.1.1 Profiling vs Benchmarking

A VNF is typically a packet processing function, deployed as part of a chain in a bigger network service (e.g. router, firewall, transcoding). This is in contrast to a common cloud application which has a limited architecture consisting of a frontend with backend application and/or database server. Benchmarking and profiling characterise the performance of a VNF and generates input knowledge for resource management decisions taken by management and orchestration systems. When characterising the performance of a software component, the terms profiling and benchmarking are often used interchangeably, although we can state some differences:

- **Benchmarking:** A benchmark is typically the result of a well-defined test procedure, under strict controlled boundary conditions. By defining edge cases and maximising the workload, the limits of the software processing are explored. One important aspect here is that the benchmark is a single result value, testing the maximum value (e.g. time or operations/sec) in a defined environment (e.g. hardware/software configuration, different versions). The goal of the benchmark is to compare and select the best option out of multiple benchmarked possible solutions;
- **Profiling:** A profile is used to understand the behaviour of a software component. Profiling takes the whole spectrum into account, not necessarily to compare with others. In software development it is typically used to understand in which function the program spends most of the time. In general, profiling helps to indicate where the bottlenecks are, not to select the best component, but to optimise a single component.

Both approaches make sense to characterise a certain VNF, but in the operational context, we prefer the term "VNF profile" to indicate the relation between several performance indicators of the VNF. A complete description of a VNF's behaviour spans many different metrics and tests. An assessment of the VNF test coverage can be done as described in [36] and [37]. We summarise this method in Table 2 for our envisioned profiling approach.

The activation and deactivation are out of scope since what happen in 5G EVE platform is seen as a black box from the vertical perspective. For that we focus on the operation stage. 5G EVE platform allows a vertical to define a context blueprint in addition to the vertical service blueprint. Speed, accuracy, and scalability of the VNF are tested by imposing a ‘normal’ workload, varying between expected operational use-cases. Edge cases such as attacks, stress or duration testing are out of scope for the reliability study.

**Table 2. Classification template for a certain test method to evaluate the test coverage**

| | Speed | Accuracy | Scalability | Reliability |
|---|---|---|---|---|
| Activation | - | - | - | - |
| Operation | X | X | X | - |
| De-activation | - | - | - | - |

#### 3.2.1.2 VNF profiling methodology

The purpose here is to create a 5G EVE profile of the 5G-TOURS vertical VNFs. To implement this, the 5G EVE portal design is used.

The Experiment Blueprint Builder is a component of the 5G EVE Portal web Graphical User Interface (GUI), which provides a simple wizard to guide the experiment developer in the definition of an Experiment Blueprint. This wizard is organised in steps where the user selects the components of the Experiment Blueprint, defines its parameters (e.g. metrics and KPIs) and on-boards the associated Network Service Descriptor and translation rules.

The steps of the wizard for building a new experiment blueprint are described in [11]. In particular, the experiment developer initially provides general information about the experiment blueprint, e.g. its name, version and a textual description. The second step requires the selection of the site to execute the experiment and the corresponding vertical service. This is followed by the selection of the context blueprint(s) to be used in the experiment. Both the vertical service and context can be selected from the list of available blueprints in a specific site catalogue, proposed in the web GUI.

The number of context blueprints will depend on the number of contexts the 5G-TOURS vertical wants to consider in the experiment to collect the KPIs and the monitored data of the behaviour of the 5G-TOURS VNFs in 5G EVE domain. Figure 30 depicts this interaction between 5G-TOURS layer and 5G EVE platform to build such kind of profile.

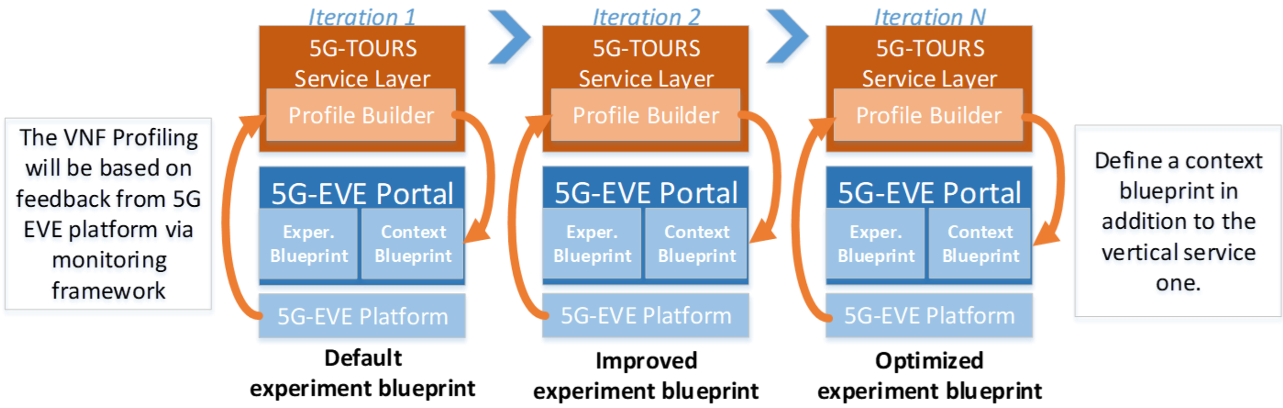


**Figure 30. 5G-TOURS vertical’s VNFs profile building based on 5G EVE Platform**

The VNF profile can be used as a baseline to compare monitored performance during operation or deployment phase. Any deviation from the pre-measured profile can indicate an anomaly, requiring a healing action from the 5G EVE management and orchestration layer or notification of SLA violation. Another application, which will be developed in the course of 5G-TOURS project, is the calculation of the correct resource allocation based on the VNF profile in order to meet the performance requirements.

As explained earlier, our goal is to derive a model which explains the relation between several performance indicators of a VNF. We classify three different metric categories which characterise the VNF:

- **Workload metrics** are used to quantify the amount of 'work', which is presented at the VNF's input. These reflect the generated load (e.g. packet size and request(s), the variety in payload content);

- **Resource metrics** quantify the payable/physical/scalable hardware resources obtained from the IaaS provider (e.g. vCPU, memory, storage or network);
- **Performance metrics** monitor the KPIs. These are (often SLA-defined) measurements of the processed workload, thus taking the output of the VNF into account (e.g. delay, loss, and throughput).

5G EVE defined two context blueprints, as default at this stage. The first one is the background traffic context blueprint, which be used to derive the workload metric. The second one is the delay context blueprint and it could be used to derive the performance metrics. Many others context blueprints will be generated.

### 3.2.2 AI-based capacity forecasting

Once network services are associated with a slice from the service layer and the orchestration, slices must be allocated sufficient resources. Due to the prevailing softwarization of mobile networks, such resources are mainly of computational nature. This holds both at the RAN where they map to, e.g. CPU time for containers running baseband units (BBU) in Cloud Radio Access Network (C-RAN) datacentres, and in the Core Network (CN) where, e.g., virtual machines run softwarized 5G Core (5GC) entities in datacentres. In this case, ensuring strong KPI guarantees often requires that computational resources are exclusively allocated to specific slices, and cannot be shared with others [38].

The dynamic allocation of network resources to the different admitted slices is, in fact, a chief management task in network slicing. In this context, the network operator needs to decide the amount of resources that should be dedicated to the different slices in advance, so as to ensure that the available capacity is used in the most efficient way possible and thus minimise operating expenses (OPEX). The key trade-off is between:

- Under-provisioning: if the operator allocates less capacity than that required to accommodate the demand, it incurs into violations of the SLA established with the tenant;
- Over-dimensioning: excess resources assigned to a slice imply a cost in terms of unnecessarily allocated resources that go unused.

Finding the correct operational point requires (*i*) predicting the future demand in each slice [39], and (*ii*) deciding what amount of resources is needed to serve such demand. These two problems are complex per-se. Forecasting future demands at service level requires designing dedicated, accurate predictors. On the other hand, allocating resources in a way that under-provisioning and over-dimensioning are poised to minimise the OPEX of the operator requires estimating the expected (negative and positive) error of the prediction. Moreover, addressing (i) and (ii) above as separate problems, risks to lead to largely suboptimal solutions, since legacy predictors do not provide reliable information about the expected error they will incur.

While the complexity of the complete solution may be daunting with traditional techniques, AI can be leveraged to address both aspects at once, by solving a *capacity forecast* problem. This can be realised by training a typical Convolutional Neural Network (CNN) architecture for time series prediction with a dedicated loss function that, instead of simply minimising the error, accounts for the respective costs of SLA violations and capacity over-provisioning, as done by the DeepCog algorithm [40]. DeepCog's architecture is depicted in Figure 31 below.

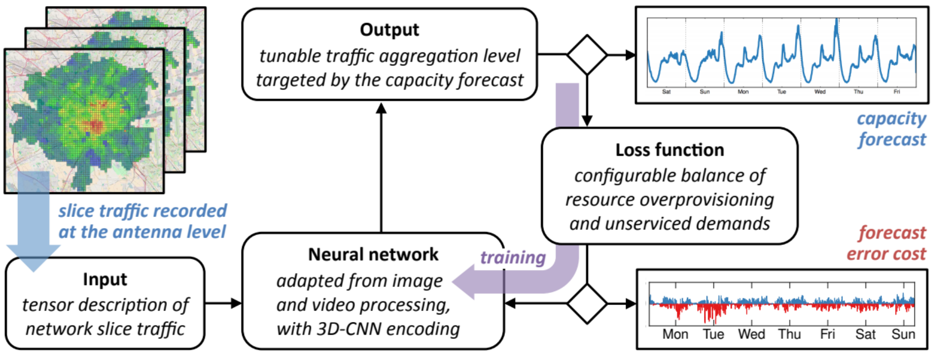


**Figure 31. Outline and interaction of the DeepCog components**

Starting from real data coming from the network, such as the one described in Section 2.1.1.3, we leverage DeepCog to perform capacity forecasting. In Figure 32, we show the deep learning architecture used by DeepCog.

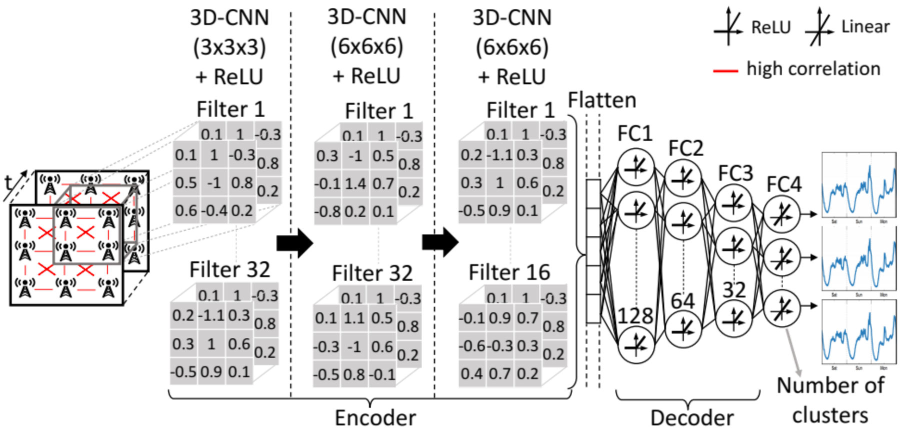


**Figure 32. DeepCog neural network encoder-decoder structure**

Note that we employ a CNN-based architecture instead of Recurrent Neural Network (RNN) models traditionally used for regression problems. Unlike RNN, CNN allows exploiting inherent spatial correlations in the traffic generated at different geographical locations. Moreover, we adopt a 3D-CNN architecture that can accommodate a tensor input. By having time as the third dimension, the model can properly account for relevant temporal autocorrelations.

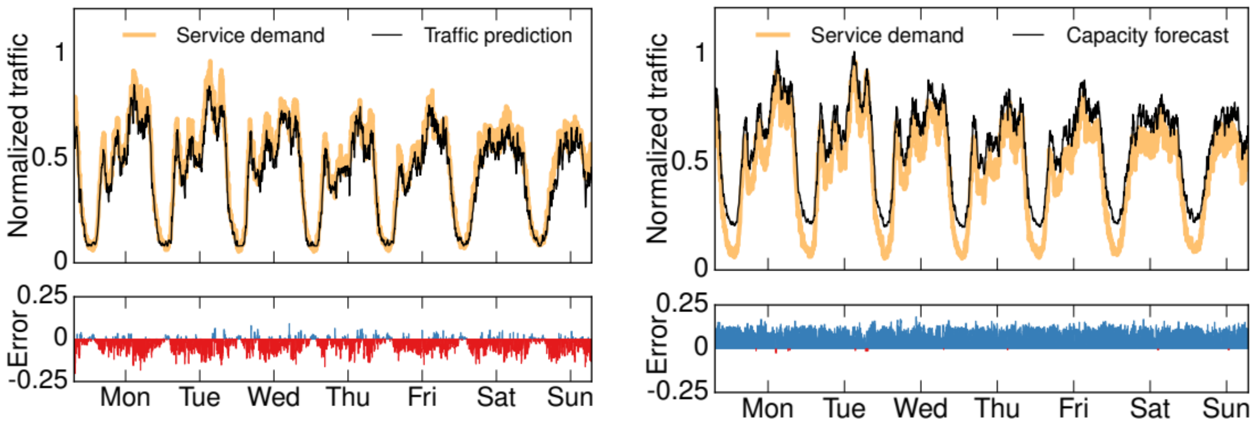


**Figure 33. Traditional traffic prediction (left) vs DeepCog capacity forecast (right)**

The operation of the proposed approach is exemplified in Figure 33. On the left, we have a typical Mean Absolute Error (MAE) function, while on the right we have a function especially tailored for network resource allocation that accounts for the actual costs of under-provisioning and over-dimensioning. The top plot rows illustrate the loss functions that drive the back-propagation process during the neural network training in each case. The key of DeepCog is the loss function (depicted in the Figure 34) which takes into account the cost components coming from the two sources.

Specifically, we model those cost with two parameters: a fixed penalty for under-provisioning β and a linearly increasing penalty for over-dimensioning γ. By allowing the parametrization of these two variables we can take into account different kind of resource deployments. For instance, very high β values indicate that the service has very low tolerance for unserved demands (i.e., a very inelastic network slice such as the URLLC one). On the other hand, very high γ mean that the kind of resources that shall be orchestrated are very expensive (i.e., such as the spectrum or an edge deployment). The vertical is allowed to modify such parameters according to its needs, for instance, by using the service layer (see Section 3.4).

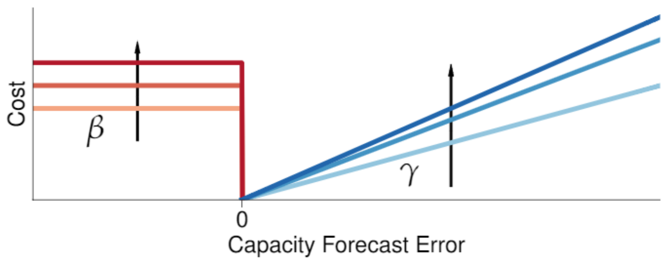


**Figure 34. DeepCog Capacity Forecast Error**

However, Stochastic Gradient Descent methods used to train neural network do not work with constant or step functions, which force us to introduce slopes for x<0 and x=0, and make the function differentiable over all the domain, as depicted in Figure 35 below.

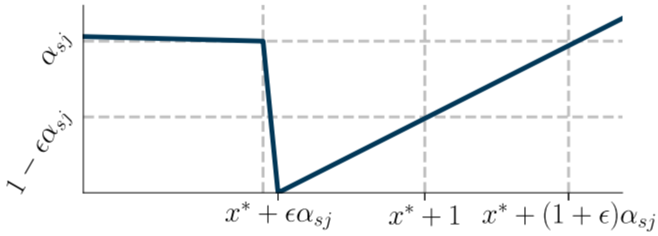


**Figure 35. DeepCog cost function implementation**

The network trained with MAE tries to anticipate the exact demand and incurs in expensive under-provisioning during a substantial portion of the time, as shown in the bottom-left error plot of Figure 34. Instead, the network trained with DeepCog can learn how to dimension capacity in the next time slot to avoid SLA violations, while keeping overprovisioning at a minimum, as illustrated in the bottom-right plot. Such a properly tuned AI-based solution allows determining the resources that should be proactively allocated to each slice to accommodate their future demand. Ultimately, this approach solves (i) and (ii) at once via AI, and yields major monetary savings for the operator.

To assess the performance of AI for the orchestration of sliced network resources, we consider three representative case studies: (i) a slice dedicated to Facebook traffic in a core network datacentre that controls all 470 4G eNodeBs deployed in a large metropolitan area; (ii) a slice allocated to Snapchat traffic in a Mobile Edge Computing (MEC) datacentre that handles the traffic of around 70 eNodeBs; and (iii) a slice accommodating traffic generated by the YouTube app at a C-RAN datacentre located in proximity of the radio access, which performs baseband processing and scheduling for 11 eNodeBs.

As a benchmark we use a legacy mobile traffic predictor, which minimises the MAE of the forecast. The prediction returned by this model is over-provisioned by 5% (a reasonable figure in presence of decently accurate prediction) to try avoiding SLA violations.

The results are summarised in Table 3 below. The AI-based approach to network resource orchestration outlined here achieves a substantial reduction of the monetary cost associated with the allocation of resources, with savings above 50% in all cases.

**Table 3. Cost Savings using DeepCog**

| Service | Deployment | Cost Savings (wrt to MAE) |
|---|---|---|
| Facebook | Core | 81.6% |
| Snapchat | MEC | 59.2% |
| YouTube | C-RAN | 64.3% |

### 3.2.3 Empowering network slicing with Data Analytics

There are few existing assets from other European projects concerned with works on 5G approach, that are perceived as a starting point to further works on AI approach. Namely, there are platform enablers from 5G EVE, as well as solutions and technologies from 5G-MoNArch.

Our particular goal concerning the “intelligent network” is to contribute to novel AI-based algorithms, which introduce an intelligence to network management and orchestration. This work will follow two directions, aiming at introducing a data-based intelligence to the network. First, some research investigation will be provided in order to go beyond the state-of-the art approaches. This is a twofold direction that includes a contribution towards global project research on AI-approach with a goal of designing complex solution, and enriching “AI-based data analytics”. Second, our work will be based on available experimental data, aiming at tuning state-of-the art algorithms and proposing possible extensions (called “AI-based data analytics”).

Trial use cases for 5G-TOURS will be the element binding those two approaches, as a trade-off between an ideal solution with a possible implementation on a real 5G-TOURS platform, where the proposed solution can be provided as a common project effort, taking into account all platform limitations and potential workarounds.

In reference to the latter, proposing a detailed evaluation methodology related to use case outcomes is foreseen, where end-points metric(s) will be considered as primary metrics, and network/slice QoS as the secondary ones. In general, it can be assumed that AI-data based analytics will support smooth orchestration (relevance to the changing needs, support of optimal decisions) for both resource orchestration, and service orchestration. This is illustrated in D3.1 (see [6], Figure 13).

As part of the practical summary of the above paragraph, this work will cover such activities as:

- in the scope of enriching AI-data based analytics:
    - analysis of business goals and requirements on network slicing and orchestration functionalities from the usage perspective (use case(s)) – in cooperation with use case(s) (as a part of AI lifecycle);
    - identifying data that will be crucial for Data Analytics in the identified business scope;
    - data acquisition from the use case(s): from the network (MANO layer) and from the services respectively as well as data analysis and preparation (cleaning, normalization etc. – all usual data transformations necessary before performing analytics algorithms);
    - selecting candidate algorithm/set of algorithms (in the first step state-of-the-art one(s)) for performing experiments and in parallel preparation of the experimental environment (for providing experiments);
    - analysis and summary of the results of experiments with conclusions and possibly proposals of algorithms' tuning and enrichment;
    - description of the results with a set of recommendations.
- in the scope of a contribution towards global project research on AI-approach with a goal of designing complex solution:
    - participation in research (perceived as a collaborative task for partners) aimed at designing a global solution (architecture, recommendations/proposals of novel algorithms, proposal of supportive technologies and tools to be involved, etc.);
    - providing description of research results and recommendations.

In the second part of these works (research) proposing a showcase can be considered (as a collaborative work of project partners) aimed at illustration of the global vision, results of research and proposal of a solution.

In the ideal world it would be useful, and even more - it would be demanded, that AI-based architecture would support automate orchestration (enabling understanding and communication between layers such as: MANO layer and service layer) resulting in automate decisions on re-allocation of resources and set up of network parameters (taking into account service needs, priorities etc., several factors). Currently adopted exemplary solutions (e.g. for MANO), imply to what extent AI solutions can be used in practice and what are their limitations (at least at this stage as probably these solutions can be extended in the future in cooperation with their authors).

Probably one of the project outcomes could be a set of recommendations for further development of the orchestration layer (for their authors). It can be a matter of further consideration.

Regarding orchestration there are two types considered:

- resource orchestration: is the assignment of the needed resources to each tenant and slice, their proper configuration (i.e., spectrum) independently of the "semantic" of the deployed VNFs;
- service orchestration: at this level the service needed by the tenant has to be understood (in terms of what network elements are necessary) and translated into VNFs. Additionally, the chaining and relation among VNFs needs to be provided.

AI-based solution should provide optimal and automated decisions on both levels in a holistic way.

AI-based data analytics framework of 5G-TOURS can adopt an intent-based approach. It should allow:

- to understand policies, coming from the vertical customer (though the service layer) and continuously align to them;
- to orchestrate end-to-end service chains deployed across heterogeneous SDN/NFV through intent-based interface;
- to use context and analytics to constantly learn and adapt to changing needs and conditions.

Our work will rely on requirements coming from other work packages (coming from use cases and their needs) which allow to plan a detailed functionality of AI-based data analytic framework.

### 3.2.4 Auto-ML system for efficient AI network selection

The envisioned used-cases of 5G-TOURS present a largely diverse set of parameters and requirements, and even minor differences among the requirements and parameters of these use cases may have an important impact on the service that is delivered. AI and ML mechanisms are envisioned to significantly enhance the MANO architecture of future 5G networks, and they will be of particular importance to 5G-TOURS MANO architecture. However, the significantly diverse requirements and parameters that the envisioned use cases may present, make it difficult to apply the single AI/ML mechanism that is optimal for all use cases. In order to improve the performance of the AI/ML algorithms it is important to tailor them to the problem or use case at hand.

Since mechanisms that have been providing efficient results for one use case might not necessarily be as efficient for another, different aspects of the AI/ML parameters space (e.g. the model selection, dataset selection) needs to be individually configured on a use case/vertical basis. This may impose additional burden on the humans that configure the networks and can increase the time required to find the optimal solution. Apart from the advances on the different AI/ML mechanisms, recent research is also focusing on AI-based solutions on the ever-growing challenge of applying AI/ML algorithms in complex problems. Auto-ML is a new research area that tries to automatically select and parameterise AI/ML algorithms in order to achieve optimal performance for the task in hand. By reducing the role humans play in the configuration and parameterisation of AI/ML algorithms, it may be possible to enable faster deployment without compromising performance.

As part of the AI Orchestration work, we plan to implement an auto-ML system that will be able to propose the most appropriate and efficient AI/ML algorithms for the different verticals that can be integrated in the considered architecture (Section 2.1). Our system will take into account a highly diverse set of parameters and requirements for the different verticals, as well as different scenarios for the state of the underlying 5G network (e.g. current resource allocation). The proposed algorithms can either be from a pool of existing AI/ML algorithms, or from the ones proposed and developed within the 5G-TOURS project.

In our auto-ML system, we will initially consider no relations between the different VNFs set up to provide the service. This will allow us to ensure that our auto-ML system is able to provide the most efficient solutions in the most common scenarios. However, in real systems, it is expected that relations between different VNFs will exist. Such relations can significantly influence the decisions taken by the AI/ML algorithms, since the configuration of one VNF may impact another VNF, resulting in sub-optimal results for the vertical. Therefore, as a second step, we intend to enhance our auto-ML system to consider different relations between the VNFs. This approach will significantly increase the complexity of the auto-ML system, but we believe that relations between VNFs will be prevalent in 5G networks, and thus it is important to be taken into account.

### 3.2.5 AI-enhanced MANO

There are a number of new technologies and applications that are being developed and will be deployed from different industries, many of which are brought together under 5G networks. A step to change the levels of process automation is needed to enable the new services. The range of processes where automation can be exploited is very broad and covers the full spectrum, from new service creation, correct and efficient placement of the various services to in-service and network monitoring. However, the more radical promise of 5G is built on developments which are less visible. At the heart of these is an Orchestrator which can collect requirements from multiple sources, analyse the data and create the best plan to deploy and run every service.

In order to create a system that manages and orchestrates this procedure for the verticals in a heterogeneous 5G environment, a software stack aligned with ETSI NFV will be utilised. Open source MANO is a critical point in the proper functioning of NFVI and VNFs, and provides the requisite requirements for provisioning and configuration of VNFs. As an extra step an OSM with AI embedded capabilities will enable advanced automation and intelligence to better manage and take actions when necessary based on the specific use case requirements. An interface from OSM to Verticals will be created in order to interact with each other and transfer information. OSM receives requirements per application and use case from the verticals to create the optimal deployment of VNFs as illustrated in Figure 36. As an example, VNFs could be deployed either on MECs or Cloud depending on the service requirements provided from the verticals. Initially OSM will connect to the available infrastructure through specific interface and acquire various information (e.g. hardware specifications, network specifications, current status and conditions, resource availability, location, etc.).

The algorithms take as input the provided information for the different nodes of the network to reach optimal decisions. These algorithms are part of the embedded AI of OSM that enable the intelligence and automation of the system, through optimal placement and resource utilisation. After the decision for deployment based on the current status of the infrastructure is reached, the vertical's VNFs are deployed and initialised. When deployed, network and application services will continue to be orchestrated and managed by OSM, and be available to the customers without changing any of the existing business services.

Metrics from deployed VNFs will be collected and reported back to the OSM in order to re-evaluate the allocation and deployment of the VNFs to always be compliant with the requirements provided by verticals and infrastructure. Those metrics will be exposed through exporters and transferred reliably using data shipper software such as the Elastic Beats [41]. The different solutions for exporters and data shippers will be tested to ensure the best solution for integration and usage. A message broker will gather the data from data shippers and push them to a data processing pipeline that could ingest data from a multitude of sources, then transform them if needed, and send them to databases. The data will be saved in either raw format or in time series databases depending their origin and usage. Finally, OSM's embedded AI will continue to analyse the data collected, historical information that could be provided from verticals and other data from infrastructure conditions to find out if changes to the deployment of VNFs need to be done. If a metric does not meet the vertical requirement for a specific VNF, OSM will make the necessary changes to fix it, either by relocation of the VNF, scaling or other appropriate actions.

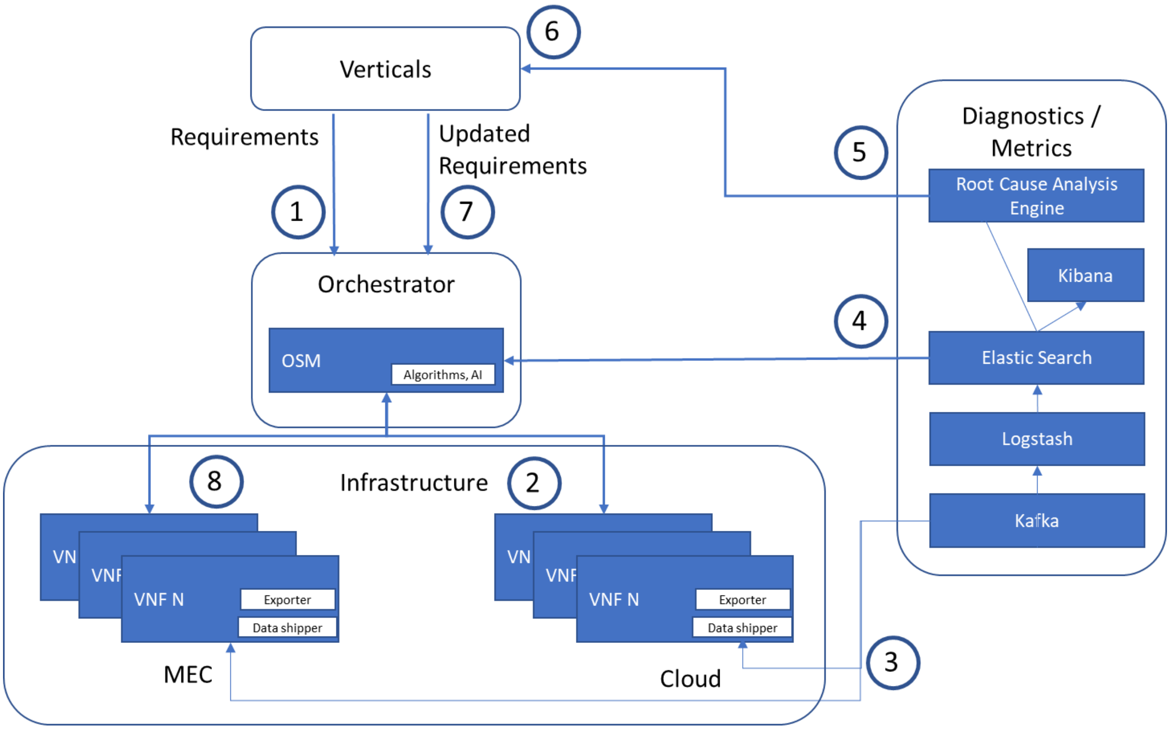


**Figure 36. Role of OSM**

Figure 36 illustrates the steps described already that will be taken so that the management and intelligence of the system to work flawlessly and in an autonomous way. The overall flow of the system respectively for each step will be:

1. OSM acquires vertical requirements, plus extra information (i.e. historical data);
2. OSM finds mapping to the infrastructure and vertical application (VNF) is deployed;
3. Exporters and data shippers send the data of specified application metrics to the Kafka broker;
4. After storage and initial processing, the data will be pushed to OSM to re-evaluate the deployment of VNFs;
5. Diagnostics will be also notified and provide an analysis of any problems or slowdowns that may have occurred back to the verticals;
6. Verticals will get the analysis and update the requirements if necessary;
7. Updates will be given back to OSM when needed;
8. With the new requirements and the metrics collected from VNFs, OSM will make the proper changes. For example, to deploy a VNF to a MEC (that was initially at the cloud) for better E2E latency by moving the specific VNF closer to the users.

For the realisation of this system there is a much-needed work to be done so that OSM can communicate with all entities and trigger changes to VNF deployments, specifically interfaces between:

- OSM and Verticals
- OSM and Infrastructure
- OSM and Diagnostics / Metrics

These interfaces will be specified and developed, talking also into account as the algorithms that will be deployed to help OSM enable the intelligence to the system. Algorithms from metaheuristics to swarm intelligence approaches will be studied, to provide a good solution to the optimisation problem. A comparison of these types of algorithms, such as Simulated annealing (SA) and Particle Swarm Optimisation (PSO) that are a probabilistic technique (SA) for approximating the global optimum of a given function and a population-based stochastic optimisation algorithm inspired by social behaviour of bird flocking or fish schooling respectively, will be tested for reaching the fastest optimal solution. With these interfaces and functions enabled at the orchestration level, OSM will bring the intelligence and highly automated procedures that are needed at the heterogeneous and multi-service 5G environment.

## 3.3 Broadcast Support

5G realised the importance of supporting multicast/broadcast delivery, both at radio and transport levels. The use of point to multipoint enables a cost-effective solution to a plethora of use-cases, constrained by device density, audience reach or demanding bitrates. However, the existing commercial solutions for broadcast services in cellular systems are based on LTE and not on 5GS. Inside 5G-TOURS, the use of broadcast is focused on the multimedia distribution, specifically Use Case 4, with two different variants: using state-of-the-art Rel'16 LTE commercial equipment and beyond state-of-the-art 5G Multicast, based on previous 5G PPP project 5G-Xcast [42]. This section details both of them, summarising the work perform in LTE 5G Broadcast trials, and the Rel-17 multicast/broadcast pre-standardisation work alongside the implementation philosophies for the multicast network functions.

### 3.3.1 LTE 5G Broadcast

This use case targets the distribution of enhanced high-quality video and immersive services. Users will be able to use their smartphones, tablets or VR devices to receive educational and informative content. This video content is transmitted via the broadcasting network of RAI. The use case will use LTE-based 5G Terrestrial Broadcast technology to transmit the video (downlink-only mode) to all users at once. The performance does not change regardless of the number of devices receiving the signal. The use case utilizes an HPHT topology to transmit the content, currently using 3GPP Release (Rel)-14. The network architecture will be updated from the current Rel-14 to Rel-16 for a second phase demonstration. Specifically:

- ***Phase 1:*** Rel-14, indoor scenario (Palazzo Madama or Rai CPTO premises) and Wi-Fi multicast distribution to a large panel of users of live or pre-recorded video signal;
- ***Phase 2:*** Rel-16 mobile in-car scenario in the city centre and unicast Wi-Fi distribution in the car of live or pre-recorded video signal.

### 3.3.2 5G Multicast

**<u>3GPP standardisation</u>**

3GPP is in the middle of standardisation effort for Rel-17 across RAN and System Architecture (SA) groups, with the new features defined and fixed in December 2019. One of these features involves the support for multicast and broadcast communications, with a Work Item affecting the radio part, and two of them located at the 5G Core network. The scope of the Study Items at Core level has been approached by two different angles, one of them considering Multicast/Broadcast as a Resource Optimisation technique over Unicast delivery, and the other one considering Multicast/Broadcast delivery as a service by itself.

The feasibility of the first approach is being studied inside "*Architectural Enhancements for 5G Multicast-Broadcast*", targeting the Public Safety, V2X, Transparent IP Multicast, IPTV, Software Delivery and IoT use cases [43]. The enhancements are grouped into nine topics, covering session management, authorisation methods for multicast groups, QoS support for multicast transmissions, fast switching between unicast to multicast and fast switching between 4G unicast to 5G multicast. Their on-going proposed solutions for each topic are reflected in TR 23.757 [44]. 3GPP will analyse the conclusions derived from this document and decide if they proceed with actual normative work to incorporate multicast in the Rel-17 Core.

On the other hand, the second approach which sees Broadcast as a service is being evaluated in "*Multicast Architecture Enhancements for 5G Media Streaming*". This Study Item was approved in March 2020 and targets the verticals for Linear TV, Live/Smart TV, Over-the-Top content delivery and Digital Radio Services. These services are characterised by audience reach and free-to-air services. It includes the integration and convergence with the multicast adaptive bitrate (ABR) capabilities lead by Digital Video Broadcasting (DVB). The impact of this multicast addition to the 5G unicast architecture will be described in TR 26.802 [45].

Finally, RP-193248 "*New Work Item on NR Multicast and Broadcast Services*" [46] focuses on providing the required radio support and procedures demanded by the two 5G multicast Study Items. The work performed will not introduce new numerologies or changes to the NR waveform in order to ease the compatibility between 5G RF chipsets, incentivising the market adoption of this feature. Any required change will happen at higher layers, i.e. Radio Resource Management and gNB related protocol and interfaces.

**Implementation Strategies**

The inclusion of broadcast support in 5G-TOURS will leverage the work done in 5G-Xcast [47] [48] as detailed in the previous deliverable [6]. The main conclusions derived were that two new network functions residing inside the 5G Core are necessary, one at the Control Plane and other at the User Plane. In this sense, the 5G-TOURS architecture enhancements are aligned with the multicast as resource optimisation. Note that, the broadcast as a service is also considered in the project and will be performed over commercial enhanced Multicast Broadcast Multimedia Service (eMBMS) Rel-16 RAI network.

The multicast network functions will be included as part of a commercial 5G Core Network, namely Open5GCore. Open5GCore is a fully compliant, software-based, licensed and modular Rel-15 Core. It is deployed in UPV campus premises, connected to several vendors gNB, both commercial and SDR-based. This laboratory environment offers the best setting for beyond state-of-the-art NF development. The multicast architecture can be seen in Figure 37.

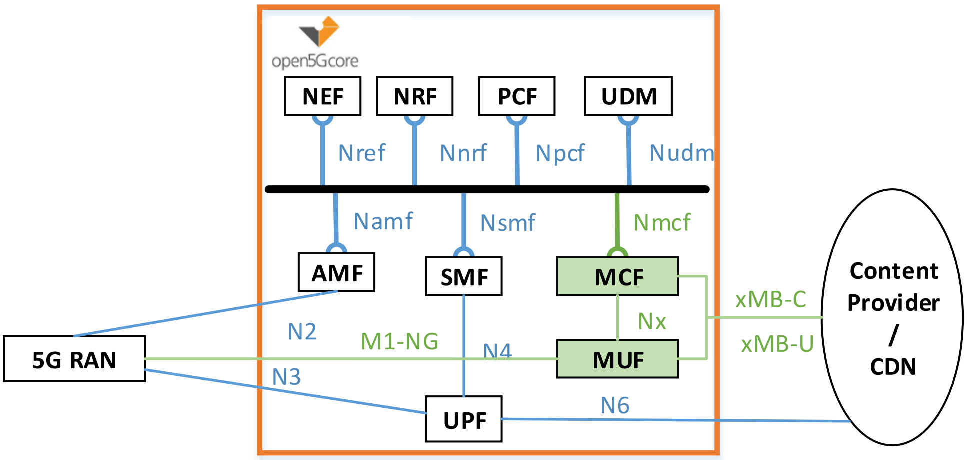


**Figure 37. Broadcast support in 5G proposed architecture**

The software development process has already begun and it is expected to last the entirety of 2020 and 2021. In order to efficiently implement the desired functionalities, a centralised approach with a single host under a virtual machine (VM) has been taken. In the latest stages of the project, the VM image will be redeployed in 5G EVE partner premises with commercial equipment.

The Multicast Control Network Function (MCF) represents the peer endpoint to the content provider for the xMB-C reference point, i.e. the control plane part of xMB interface. The MCF functionalities related to xMB-C reference point include the following:

- Authentication and authorisation
- Status notification and query.

The MCF interacts with other 5G Core network functions through service-based interfaces and uses the services offered by them to manage network resources for the xMB session. The following (non-exhaustive list) functionalities are supported in the MCF:

- Network resource management for xMB session using SMF services including:
    - Allocation of Multicast User Network Function (MUF) resources and maintenance core network tunnels between MUF and RAN node(s);

  - Allocation of (R)AN resources by RAN upon SMF request in the geographical area.
- Allocation of reference point for multicast data transport to the UE (i.e. a multicast IP address);
- Reception of consumption and reception report about a service;
- Control multicast (or broadcast) transport availability based on the consumption reporting (i.e. functionality similar to 3GPP Multicast on Demand, MooD, in LTE);
- Estimation of QoS parameters for data transfer via each available link.

Most of these control functions are similar to the functions provided by the Broadcast Multicast Service Centre (BM-SC) in LTE. It's noted that the current specification of xMB reference in 3GPP Release 14 may not fulfil the requirements for 5G multicast/broadcast capabilities and may need to be enhanced. On the other hand, the MUF represents the peer endpoint for the content provider through the xMB-U reference point, i.e. the user plane part of xMB interface. The MUF functionalities related to xMB-U reference point include delivery of content and retrieval of content. MUF sends the multicast IP packets to the RAN over M1-NG reference point

The multicast network functions will not handle anything related to the service layer, following the architecture in [44]. Unlike eMBMS, where the BM-SC performs the service layer functionality, it has been chosen for 5G-TOURS multicast implementation to delegate these functionalities outside the VNF and Infrastructure Layer; interconnected using a 5G revised version of the xMB interface. More info can be found in section 3.4.

## 3.4 Service Layer

### 3.4.1 Assessment of needs

To identify the expectations of verticals for the service layer and the requirements not fulfilled by 5G EVE, a questionnaire was circulated to verticals in each one of the three sites.

Feedback for the pilot use cases of the three sites (UC5, UC7, UC10) is described in the following subsections. The requirements expressed in responses to the questions can be broken out into three categories: Network Slice Management, Slice Monitoring, and Network Exposure.

**Network Slice Management**

- *Requirement NSM.1*: SLA management. Vertical users can negotiate the QoS considering the cost of certain network KPIs, QoS settings/configurations (UC5, UC7). Finally, the service interface could give suggestions to the application (and optionally do auto configuration itself) of the remaining network KPIs for a given target: e.g. to maximise performance, minimise costs, or maximise performance for a certain allowable cost.
- *Requirement NSM.2*: Usage of AI Orchestration to find the best possible system orchestration (i.e., function allocation, network slices/devices/resource configuration, etc.) based on the vertical service requirements and the anticipated network status (UC10).
- *Requirement NSM.3*: The vertical requirements can be set by an underlying API (UC7).
- *Requirement NSM.4*: Service layer and AI orchestration can make an opportunistic use of network resources (UC7).
- *Requirement NSM.5*: Service layer and AI orchestration can pilot the migration of VNFs to the edge or core, possibly dynamically (UC7).
- *Requirement NSM.6*: Vertical users can dynamically change the QoS requirements and the configuration.

**Network Slice Monitoring**

- *Requirement NSMo.1*: KPI validation/QoS validation: the service layer shall ensure that the service requirements (including the QoS) are continuously met. (all UCs).
- *Requirement NSMo.2*: Performance diagnosis: in case of poor performance results, a diagnostic tool can be triggered responsible to collect more measurements and/or make a deeper analysis of the performance issues.
- *Requirement NSMo.3*: The service layer shall fire events to the application in case network KPIs are getting below a threshold.
- *Requirement NSMo.4*: Output of the diagnostic tool is presented to the vertical over a GUI, proposing alternatives for improved performance.

**Network Exposure**

- *Requirement NE.1*: The service layer shall expose the device location to the vertical users (UC10).
- *Requirement NE.2*: The service layer shall monitor and expose the device battery level and charging state (UC10).
- *Requirement NE.3*: Broadcast capability exposure. The service layer shall expose a northbound API to the vertical allowing the allocation of network resource for broadcast, and a corresponding API for ingestion (UC4).

### 3.4.2 Implementation strategies

The required functionalities are not fully provided by 5G EVE modules and some of them cannot be implemented on top of the 5G EVE portal. Two complementary strategies will be used:

- Implementation on top of 5G EVE portal: the 5G EVE project is willing to take our requirements as input. A step-by-step collaboration with 5G EVE can help the adaptation of 5G EVE modules to 5G-TOURS vertical needs. In addition, 5G-TOURS will work on extending the 5G EVE portal functionalities;
- Broadcast support (*NE.3*) implies independent developments of new network functions (MUF, MCF) in the 5G core (see 3.3.2), and their associated northbound interfaces. Support of a dedicated multicast protocol stack over IP, and a set of additional procedures for announcement, repair and reporting, designed at 3GPP as 'the full service mode' or 'the broadcast service layer' will also imply the implementation of dedicated network functions;
- Specific implementation of 5G EVE functionality: Specific implementations of the service layer will focus on run-time management (scaling, re-orchestration) linked to *NSM.5* and may act as an extended Communication Service Management Function (CSMF) or a Network Slice Management Function (NSMF) (see D3.1 [1]). New dedicated blueprint format may be specified to address use case specificities (e.g. *NSM.6*).

#### 3.4.2.1 Implementation on top of 5G EVE portal

Network slice monitoring requirements can be implemented on top of 5G EVE portal. The list of KPIs monitorable by 5G EVE must be completed to offer the list of required KPI, as listed per Use-Case by WP2 and WP7.

5G EVE currently supports *Latency*, *User Data Rate*. Other KPIs are expected before the end of 5G EVE project including *Peak Data Rate*, *Capacity*, *Device Density*, *Reliability*, *Availability*. In addition, 5G EVE can support a set of KPIs that can be derived by the aforementioned KPIs based on a simple formula. This formula should be defined in the blueprints. In addition, to collect further KPIs e.g. KPIs on the application/service layer appropriate new software (data shippers) must be implemented in order to collect the metrics and publish them. This will be the case to expose device battery level and location (*NE.1* and *NE.2*).

The ability to retrieve and expose the metrics/KPI in real time is an essential prerequisite, to achieve Network Slice Monitoring requirements. A new API from 5G EVE, which has recently been implemented by the 5G EVE project, will be used. Performance diagnosis (*NSMo.2* and *NSMo.4)*: the performance diagnosis module which is developed in 5G EVE is executed after the finalisation of an experiment and provides insights regarding any performance issues while it tries to identify the cause of the problem using route cause analysis techniques. In 5G-TOURS WINGS can extend/redeveloped the performance diagnosis module in order to realise the route cause analysis in a real time manner.

Outcomes of the performance diagnosis cannot be directly fed to the 5G EVE network orchestrator, as it would require to break the interworking layer of 5G EVE. One option is to extend the blueprints to define the parameters which could be changed as response to the alerts received from the performance diagnosis.

#### 3.4.2.2 Broadcast support

Broadcast support (*NE.3*) is fulfilled by the implementation of the xMB reference point. An enhancement would be to expose the broadcast capability through the Network Exposure Function (NEF) as illustrated in Figure 38. The NEF is a function form the 5G core, aiming to expose securely the network capabilities to external applications through restful APIs. In such a case, the content provider will not directly implement the xMB API, but the API proposed by the NEF.

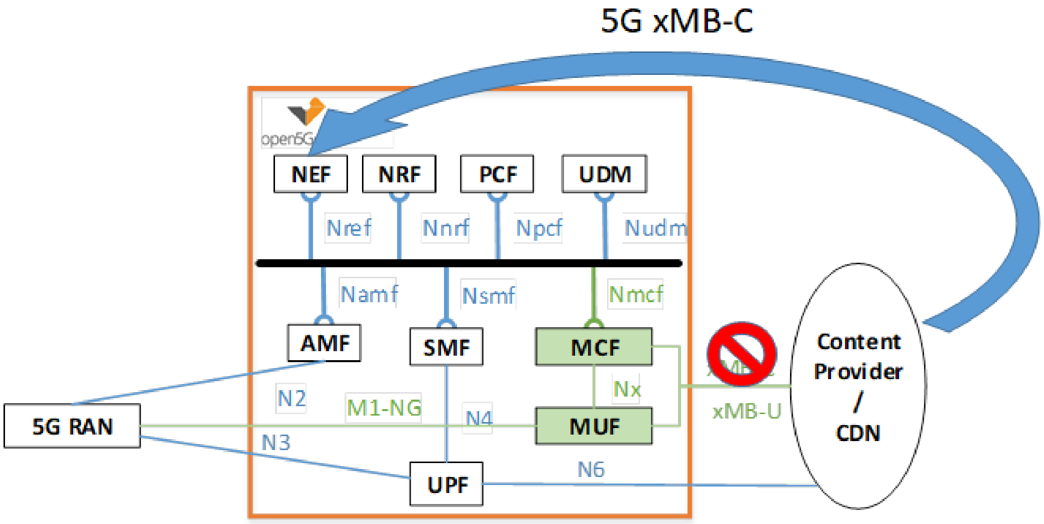


**Figure 38. xMB interface exposed by the NEF**

Exposing broadcast capability through the NEF is expected at 3GPP as depicted in the 3GPP SA2 study "*Architectural Enhancements for 5G Multicast-Broadcast*" [44]. An extension of Open5GCore (see 3.2.5) will be implemented following advancement of the 3GPP standard.

Whether the broadcast service layer (delivery protocol over UDP, signalling/service announcement, forward error correction, etc.) will be included within the 5G Core has not been decided yet at 3GPP. In particular, the 3GPP SA2 study considers two baseline architectures, where only one includes the broadcast service layer in the core (A.2 in [44]), while the other (A.1) only transparently forwards packets sent by application servers. The new 3GPP study "*Multicast Architecture Enhancements for 5G Media Streaming*" [45], which includes integration and convergence with the multicast ABR capabilities lead by DVB, will specify new network functions or, out of the 5G core, handle the broadcast service layer. 3GPP activity on these studies need to be closely monitored in collaboration with WP8. One option could be to extend the Media Application Server (AS) and Media Application Function (AF) of the current 5G Media Streaming Architecture to include the broadcast service layer.

The new/adapted network functions for multicast 5G Media Streaming will be implemented based on the broadcast service layer of the current virtualised BM-SC.

#### 3.4.2.3 Specific implementation of the service layer

In the following we discuss some specifics of the implementation we envision in the project which are relying on the specific bypass envisioned by the project to directly interact with the MANO of each site.

**AI-based scaling of VNFs**: As discussed in Section 3.2.1, resources assigned to different VNFs have to be assigned according to the real load in a proactive way, without incurring into too high overprovisioning costs. To this end, we will implement an AI-based management system which is leveraging ETSI ENI principles, as described in Section 2.1.2.5, leveraging the data coming from the network orchestrator, which is directly accessed through the bypass, discussed in Section 2.1.2.

Figure 39 below depicts the overall architecture of the modules and its interaction with the Service Layer, as discussed above.

The AI-Based scaling of VNF implements an algorithm similar to the one described in Section 3.2.1, but operates at coarser granularity as the minimum decision step is one VNF instance. This algorithm takes as input the overall load of the VNF (i.e., the CPU consumption, obtained from a monitoring software such as Prometheus).

The target VNF is yet to be decided, but it will be used by UC1 in the Touristic city site. The monitoring information comes from the bypass and the scaling decision, which proactively decide the number of instances of the given VNF are simultaneously running follow the same path.

The ad-hoc service layer that will be implemented for this purpose allows the vertical configuration that steers the algorithm operation, fixing an operational point that either consumes more resources but it is more resilient with the possible outages or the other way around. This corresponds to the *NSM.1* and *NSM.2*.

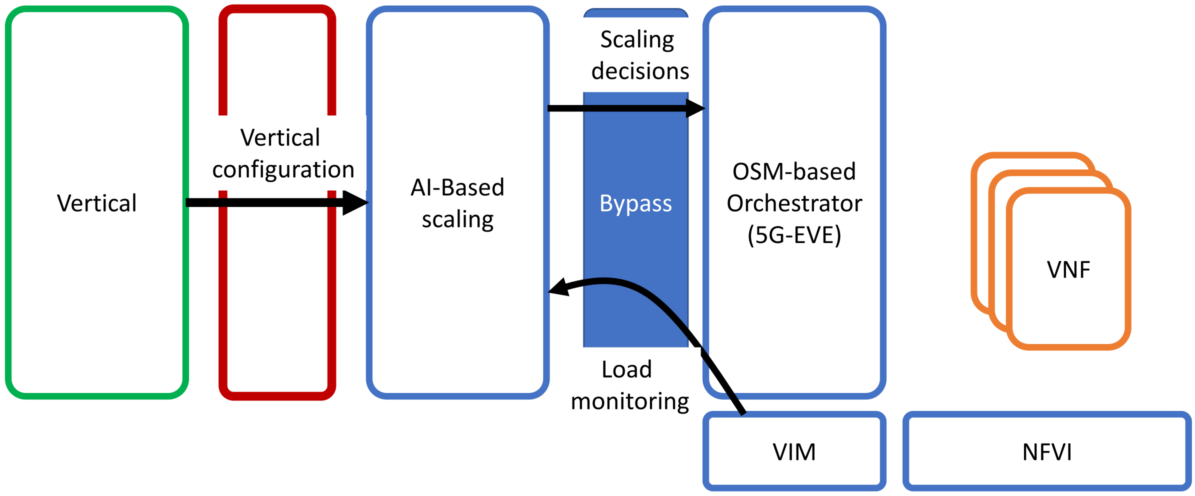


**Figure 39. Architecture of the AI-based scaling of VNF**

## 3.5 5G RAN enhancement considerations

As stated in previous sections, one of the main objectives of 5G-TOURS is the enhancement of management and orchestration capability in order to guarantee service level assurance, in particular for innovative services, highly demanding in terms of e.g. latency, throughput and reliability, while efficiently allocating network resources and taking into account the level of performance of other services.

Performance management of 5G RAN is a key component of the overall contribution of the network performance into the SLA assurance, due to intrinsic limitation of radio resources. Scope of the work is to study solutions in the area of Network Management and orchestration of Radio Access Network and to provide domain-specific enhancements to 5G-TOURS architectural and functional design. This section will explain the principles and requirements; and existing solutions in standardization groups and open source communities.

### 3.5.1 High level principles and requirements

The design of an enhanced architecture for automation exploiting management and orchestration capability in the 5G RAN domain is based on the following high-level principles and requirements:

**Performance Monitoring (PM) allowing the characterization of both Quality of Service and of User Experience**

One of the main areas of improvement on performance monitoring, is the possibility to provide more flexible definition of SLAs (e.g. session level / real time monitoring and correlation in addition to the statistical cell-level PM counters) so as to differentiate the service offering, and at the same time ensure efficient use of network resources. In particular, for what concerns the RAN domain, mapping E2E service level performance metrics into RAN Domain (Subnetwork Slice level resource KPIs) and inter-domain coordination and consistency has to be ensured. This will require Enhanced MANO functionalities for PM to identify and manage the 5G RAN KPI, taking into account the work within 3GPP and O-RAN in the end-to-end SLA assurance, according to the service characteristics provided by Service Layer in 5G-TOURS.

On the Network Function Layer, more accurate performance measurements capability is required, extending from the traditional cell-level long-term statistical counters, in order to provide UE or slice specific counters with greater time granularity also in order to support short term analysis and prediction in MANO layer with improved AI capability.

**Radio Access Network resource modeling and exposure**

Full coordination and integration of RAN domain in an E2E orchestration framework requires an accurate network resource modelling, exposing control over sub-network slice resources and radio resource management capability. Modelling should also take into account the enhanced RAN capability, providing more flexible control layer mechanism, also enhanced with support of AI/ML, either with direct control from centralised functions (for near real time adaptation) or by provisioning on node-level policy-based control for short term adaptation.

**Advanced Management and Orchestration functions**

In order to fully exploit enhanced network function, exposure and monitoring, interworking integration of RAN domain management with E2E service management and orchestration is needed to ensure accurate mapping of SLAs into Network policy and performance objectives.

Improvement of 5G RAN resource utilization should also be achieved in an automatic manner with support of both predictive analytics to plan orchestration actions in advance (e.g. predicting overload in a given area/cell) and reinforcement learning based correlation between network performance and automated corrective actions.

Moreover, tighter integration of network creation and design functions in the automation workflow shall be achieved, to ensure a fully automated and coordinated process covering all the different network management phases (design, provisioning, assurance, etc). Finally, new capabilities shall be introduced at the network orchestration layer to provide AI/ML lifecycle management, allowing (re-)training and deployment of models at different layers for network status prediction, closed loop control, etc.

A high level description of an orchestration process highlighting the involvement of 5G RAN domain, is provided below:

- A Communication Service request is translated to a 5G slice request from the Service Layer to the Management and Orchestration layer, with service and resource description (QoS profile, resource type, users/group of users, area (tracking areas, cells)) and SLA for Assurance.
- For the 5G RAN domain, the Management and Orchestration layer checks the availability of the necessary network capability and resources.
- The Management and Orchestration layer provides 5G RAN network functions with the baseline network resource model configuration and the related PM configuration to monitor KPIs for Service Assurance.
- The new 5G slice is activated.
- Based on the observation of SLA related KPIs, the Management and Orchestration layer may:
  - predict or detect a potential overload in a given area that would affect either the new service with guaranteed QoS or the existing services;
  - determine the need to change the resource allocation via explicit resource allocation/deallocation to the slice, and/or changing the resource allocation policy to be enforced at control/network function level;
  - set new KPI monitoring criteria both for the new service and the existing ones.
- The network controllers/functions enforce the new resource allocation / policy by executing the necessary actions (e.g. scheduling / traffic shaping, carrier aggregation/dual connectivity, load balancing, coverage/capacity optimisation).
- When the new service/slice is stopped or removed, the network management and orchestration layer adopts a suitable strategy to revert to the new condition.

### 3.5.2 State of the art in Standard Development Organizations (SDO) and Open Source Communities (OSC)

Evolution of Network Management and Orchestration framework involving 5G RAN is being assessed and developed by several SDO and OSC. In the following the main bodies and the ongoing activities relevant for 5G-TOURS are described next:

O-RAN defined a new RAN framework based on two main objectives:

- **Openness:** open / fully interoperable framework targeting flexible design and management of network functions based on cloud platforms to cope with different deployment scenarios;
- **Intelligence:** introduce AI/ML principles applied at various network levels, enabling dynamic local radio resource allocation and network-wide efficiency.

The O-RAN architecture relies on two new principles [49]:

- an Open Fronthaul, decoupling L1 functions of eNB/gNB into the O-RAN Distributed Unit (O-DU) and O-RAN Radio Unit (O-RU) components;

- the introduction of a Near Real Time RAN Intelligent Controller (Near RT RIC), offering cloud native applications to introduce AI/ML enabled algorithms to closely control the RAN functions via E2 interface and exposing an open, policy-based control towards a Service Management and Orchestration layer, for e.g. Traffic Steering and QoE/QoS optimisation.

The principles for AI/ML application to RAN domain include ML methodologies applicable to RAN algorithms, mapping of ML into RAN control loops, and lifecycle of ML models (including data acquisition, model training/testing and deployment into network controllers, nodes and functions) [50] [51] [52].

Evolution of both Open RAN architecture, Service Management and Orchestration and support of AI/ML as proposed by O-RAN may be considered in the design of 5G-TOURS architecture.

**ONAP**

5G network automation is an important part of the content of Open Network Automation Platform (ONAP) Release 6 (Frankfurt), with particular focus on 5G RAN use cases including key components for orchestration, i.e. Network Resource Model, Performance Management, centralized SON functions. Future developments are foreseen for Release 7 with special focus on End-to-end slicing, including RAN domain. Implementation principles defined by ONAP and their evolution may be considered by 5G-TOURS in the context of Service Layer and MANO enhancements.

**3GPP SA5**

SA WG5 has the target to specify the requirements, architecture and solutions for provisioning and management of the network (RAN, CN, IMS) and its services. In particular, in release 16 (and in release 17 for the enhancements) of the standard the main work items related to our interests are the following:

- Management of QoE measurement collection. The objective of the work item is to specify the function Management of QoE measurement collection. Specifically:
  - Concepts, use cases and requirements;
  - Collection control and configuration management;
  - Information definition and transport;
  - Requirements for QoE measurement collection Integration Reference Point (IRP);
  - Information Service (IS) for QoE measurement collection IRP;
  - Solution Set (SS) for QoE measurement collection IRP.
- Enhancement of performance assurance for 5G networks including network slicing. The objective is to enhance the following aspects of performance assurance for 5G networks including network slicing:
  - Management services for collecting and reporting the E2E performance measurements and NF performance measurements, and management service for threshold management;
  - Performance measurements for the Rel-16 network feature for 5G NFs; - E2E performance measurements for the Rel-16 network feature for 5G networks including network slicing;
  - Performance measurements of the ng-eNB in terms of connectivity with 5GC, and performance measurements of the EPC Network Functions in terms of connectivity with NR - KPIs (including E2E KPIs and other types of KPIs) for Rel-16 network feature for 5G networks including network slicing.
- Integration of ONAP and 3GPP 5G management framework;
- Management aspects of 5G Service-Level Agreement. The objective of this work item is to specify a closed loop assurance solution that helps an operator to continuously deliver the expected level of communication service quality. The closed loop assurance solution allows an operator to create a closed loop management service that automatically adjusts and optimises the services provided by NG-RAN and 5GC based on the various performance management and Quality of Experience (QoE) input data, and the state of the 5G network, using data analytics.

These novel management capabilities will be integrated as much as possible in the 5G-TOURS implementation of the local MANO platform, to maximize the alignment with 3GPP management architecture and its evolution.

**TM Forum**

TM Forum [53] provides a description of SLA definition, characteristics and requirements. For SLA definition and SLA assurance activities consistent with 5G system architecture, it is important to identify the service phases and, for each one, the performance metrics related to the customer's perception of the QoS. It is also relevant to find a correlation map among these metrics and the performance measurements related to the network and its processes at various levels (at Network Slice, Network Domain and Network Function level). A top-down approach with regard to the service offered may be combined with a bottom-up approach to assess transactions necessary to deliver the service. This leads to a framework that facilitates the identification of the mapping between the network (functions, processes, and their performance measurements) and service performance aspects. This approach will define the Key Quality Index (KQI) and the network performance measurements to be included in the SLA document [53]. TM forum concepts may be taken into account e.g. in the 5G RAN related Service Layer and MANO enhancements.

# 4 Technology Integration into 5G EVE

## 4.1 5G EVE status

### 4.1.1 Architecture

The main 5G EVE objective is to create a 5G E2E facility that will be useful for a number of vertical industries to implement, test and validate their 5G-ready use-cases in real on-air transmission. From the beginning of the project, the 5G EVE facility has evolved with new features issued from 3GPP Rel-15 releases, complying with 3GPP-Rel-16 specifications at the end of the project.

The 5G EVE E2E site facility is composed of four European operational and interconnected site facilities (France, Greece, Italy and Spain). Each site facility has its own 5G full-chain from physical infrastructure to orchestration. Then, all sites are interconnected and, on top of them, a homogeneous testing and validation framework for verticals is developed, integrated and offered to verticals through a common portal allowing them to define, plan, execute and validate test cases related to their 5G-based applications.

Reference [10] gives the general 5G EVE architecture overview as also seen in Figure 40, starting at the top from the web portal application. This portal is a web interface where the verticals will be able to define their use-case (design, run and validate experience by collecting the performance evaluation) for experimentation operation. While the four site facilities are implementing different orchestrator types, an interworking layer (IWL) has to be implemented that manages multi-site services orchestrator, sites interconnection, multi-site catalogue, etc.

The IWL is defined so that each site facility is connected via its northbound interface to the IWL southbound. Each of the four sites is based on different functional components, technologies and physical infrastructure where the verticals are locally deployed.

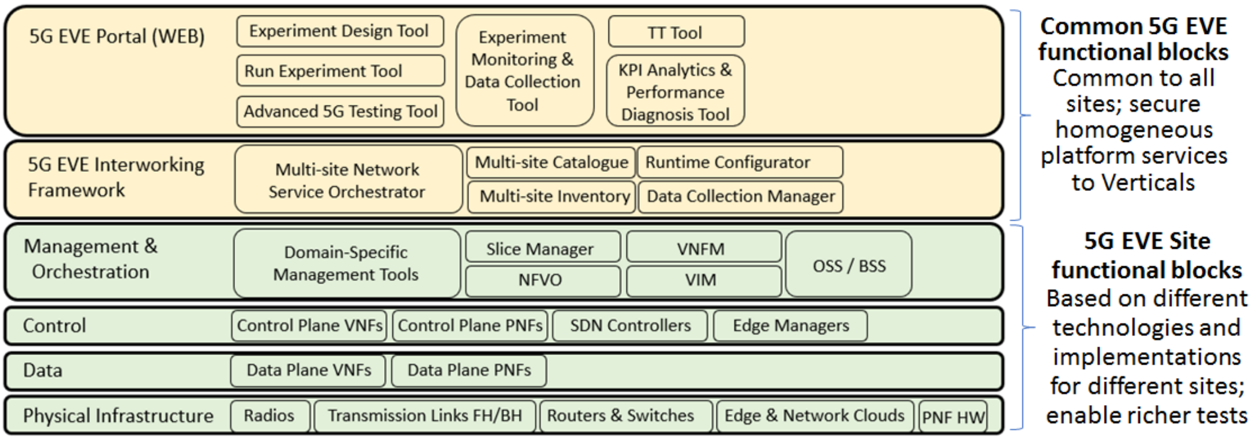


**Figure 40. 5G EVE general architecture overview**

The IWL is hosted in Turin in the TIM premises. The multi-orchestrator part for the verticals will be implemented to have direct access to any distant sites facilities for starting the integration and use-case execution.

### 4.1.2 Services overview

The objective of the different site facilities is to expose to the verticals, services to deploy the different use-cases. During the first project year, an initial general planning in terms of 5G EVE E2E facility specifications has been concluded. The different phases, with some technical specifications, are reported in Table 4.

**Table 4. 5G EVE capabilities roadmap**

| Capabilities | Features | 2019/MAY | 2020/JAN | 2020/JUL | 2021/JAN |
|---|---|---|---|---|---|
| 5G Services | Enhanced MBB (eMBB) | O | O | O | O |
| | URLLC (URLLC) | (Pre-sched) | Y(Rel-15) | O(Rel15) | O(Rel-16) |
| | Massive IoT (mMTC) | O (LTE-M+NB-IoT) | O (LTE-M+NB-IoT) | O (LTE-M+NB-IoT) | O(Rel-16) |
| 5G Architecture Options | Option-1 (Legacy) | O | O | O | O |
| | Rel15-GNR + EPC in NSA mode | | O | O | O |

| | | | | | |
|---|---|---|---|---|---|
| | Rel15-5GNR + Rel15-5GC in SA mode | | | O | O |
| | Rel16-5GNR + Rel16-5GCore (in NSA & SA modes) | | | | O |
| 5G Access Features | Flexible Numerology | | O | O | O |
| | Massive MIMO | O | O | O | O |
| | Multi-User MIMO | | O | O | O |
| | RAN Virtualization | | | O | O |
| | Latency Reduction | O (pre-scheduling) | O(Rel-15) | O(Rel15) | O(Rel-16) |
| Core Network | vEPC supporting 5G | O | O | O | O |
| | 5GC | | | O | O |
| | CUPS | O | O | O | O |
| | SBA | | | O | O |
| | Interworking with LTE | | | O | O |
| Slicing | Network Slicing (std 5G Services: eMBB, URLLC, mMTC) | | O | O | O |
| | Service Slicing (cloud orchestration level) | | | O | O |
| | Multi-site Slicing | | | O | O |
| Virtualization | NFVi support | O | O | O | O |
| | SDN control | | TBD | O | O |
| | Vertical Virtualized Application deployment support | O | O | O | O |
| Edge Computing | 3GPP Edge Computing | | O | O | O |
| | ETSI MEC | | (optional) | (optional) | (optional) |

The mapping of the capabilities, highlighted in Table 4 has been decided in terms of services availability in the different sites.

### 4.1.3 5G EVE Interworking model implementation and deployment

The IWL framework provides a set of key features to the 5G EVE experimentation platform as it abstracts the specific logic and tools available in each site facility, aiming at exposing the entire multi-site infrastructure through a unified and site-agnostic information model towards the 5G EVE portal. The heterogeneous capabilities adopted in each site facility are properly abstracted by the IWL framework to enable the realisation of a common and unified end-to-end experimentation and validation facility, thus ensuring interoperability of validation services offered by the various sites while achieving their end-to-end integration. In terms of main features, on the one hand the IWL framework provides a common interworking model for the abstraction of the heterogeneous capabilities and orchestration solutions implemented in each 5G EVE site facility, exposing seamless services and APIs to the 5G EVE portal to facilitate the experimentation of vertical services. On the other hand, the IWL framework enables the site facility interconnection at different levels (from orchestration to control and data plane layers) to allow the execution of experiments (and thus instantiation of vertical services) in different sites (and when and where possible and required, in multiple sites) in a transparent way.

Figure 41 illustrates the 5G EVE IWL framework architecture. The five main components building the IWL framework are the Multi-site Network Service Orchestrator, the Multi-site Catalogue, the Multi-site Inventory, the Runtime Configurator and the Data Collection Manager. All of them are specified in terms of main features in [54] of the 5G EVE project (section 4 of the document), including northbound and southbound APIs. As a general principle, the 5G EVE IWL framework architecture re-uses (and enhances where required) existing relevant standards and solutions provided by other European funded projects to not design and develop from scratch such a complex architecture.

The interworking APIs are built by the collection of the APIs exposed by each of the core IWL framework components. These APIs are reported using OpenAPI representations considered as the REST implementation of the APIs. Most of these interworking APIs highly leverage on existing ETSI NFV SOL005 APIs for the lifecycle management of NFV Network Services implementing the vertical experiments. These SOL005 APIs (and the related data models) are enhanced where needed to support the IWL reference model.

At the southbound, the IWL framework adaptation layer plays a key role as it implements the required abstraction features on top of the site facilities, harmonising the heterogeneous capabilities offered by each site in terms of orchestration, catalogues, monitoring and runtime configuration under a common information model and set of APIs. Where applicable, the common APIs exposed by the adaptation layer to the other IWL framework services are based on existing standard specifications, like those offered to the Multi-site Network Service Orchestrator and the Multi-site Catalogue, which are based on ETSI NFV IFA specifications.

Two different options are proposed for the interaction and integration of each IWL framework component with the adaptation layer for their access to the different sites' services (from orchestration to catalogues and monitoring). In the first option, each of the IWL framework components independently interacts with the per-site services and tools (e.g. orchestrators, SDN controller, catalogues, monitoring platforms, etc.) through the adaptation layer. In the second option, the Multi-site Network Service Orchestrator acts as a proxy for all the IWL framework interactions with the local per-site services, thus relaxing the complexity of the abstraction layer. The decision is left open for the software implementation stage to have each component to choose the best option.

This means the Multi-site Catalogue provides on its own the whole set of IWL features for what concerns the management of Network Services and VNF Descriptors, i.e. unified northbound APIs based on ETSI NFV SOL005 [55] and adaptation (with data model and API logics translation) towards the heterogeneous per-site orchestrators and catalogues.

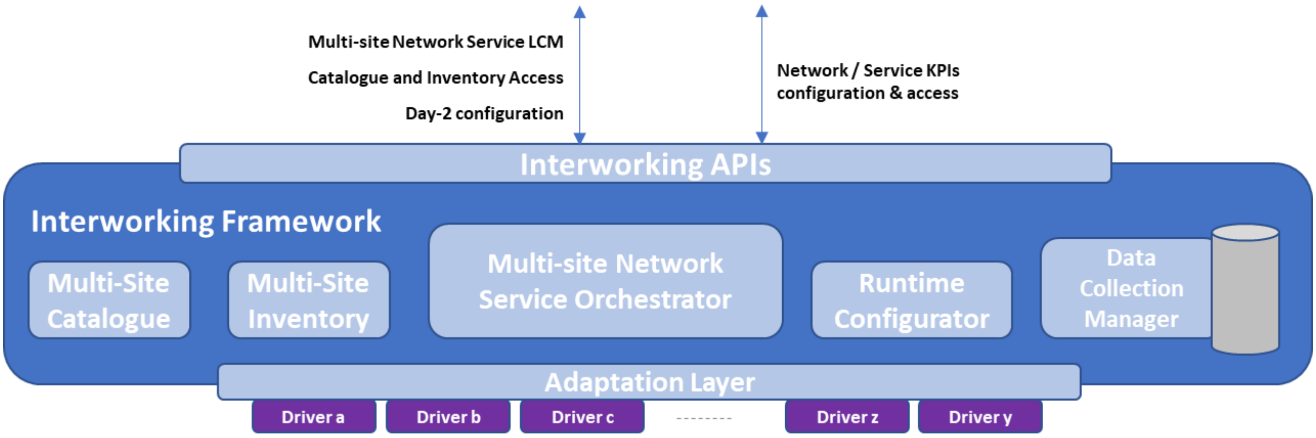


**Figure 41. IWL Framework Architecture**

## 4.2 Integration Strategies

5G EVE and 5G-TOURS are working closely to guarantee the success of on-boarding the different 5G-TOURS vertical use-cases. In particular, the network deployments provided by 5G EVE is monitored and assessed and if necessary, expanded to be able to support trialling the 5G-TOURS use-cases, which includes the possible coverage expansion for some use cases. The integration strategies identification is an important step to identify the required updates or extension of 5G EVE platform in order to succeed to deploy 5G-TOURS use-cases.

### 4.2.1 Overall methodology

The methodology is explained in Figure 42:

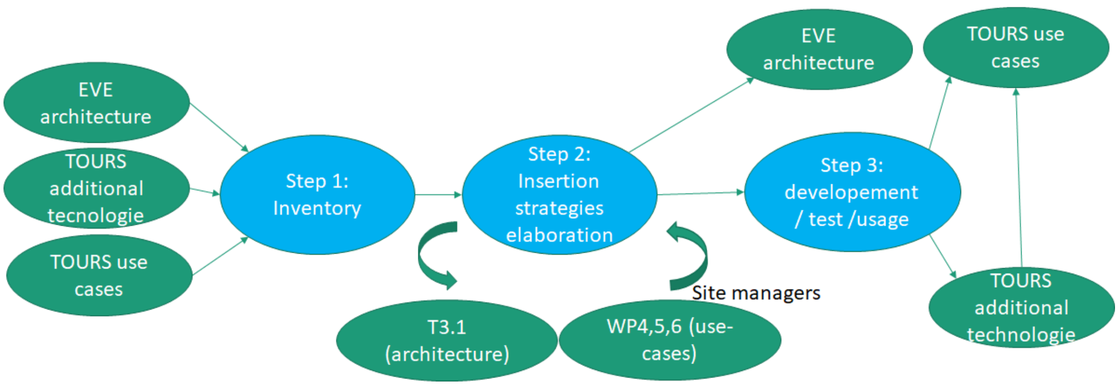


**Figure 42. Insertion Methodology**

The methodology is composed by three steps: inventory, insertion strategies and development.

For step 1, we have three types of inputs:

- 5G EVE architecture;
- 5G-TOURS additional technologies;
- 5G-TOURS use-cases.

In Step 2, we elaborate and list all possible "insertion strategies" from list of use-cases. This list can feed the architecture exercise. Finally, in Step 3 we develop any needed additional technologies.

## 4.2.2 Inventory

### 4.2.2.1 5G EVE architecture

The architecture of 5G EVE is depicted in Figure 43. Three main components are clearly identified. An inter-working layer, in blue colour, sits between the Experimentation Portal, frontend for experimenters, and the different deployed sites, where experiments have to be executed.

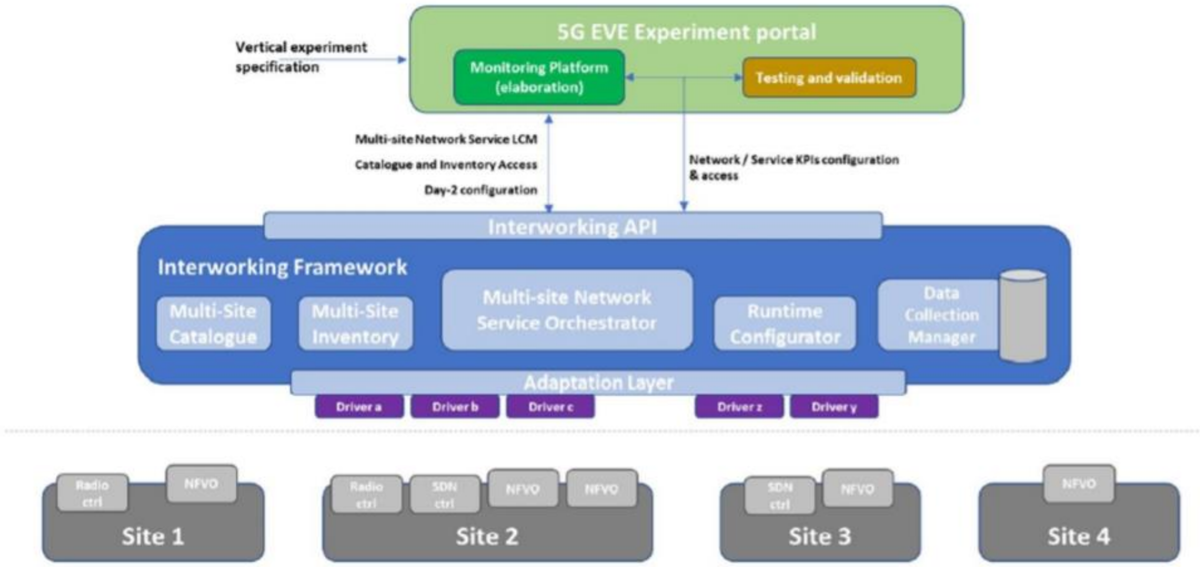


**Figure 43. Interworking Framework Architecture**

The experimental portal implements a workflow to design, define, prepare and execute experiments. Each step is implemented using a template. The experiment developer will collect the vertical requirements and translate them into a collection of VNFs, named as blueprint. Of course, the VNFs should be available in the catalogue, otherwise, we have to ask for the missing ones from the VNF providers. Once the experiment blueprint is ready it is pushed on the portal.

As a second stage, the experimenter implements the experiment lifecycle (configuration, execution on the site, monitoring and collection results). At this stage, we notice that the portal is the first level of interaction between 5G-TOURS and the 5G EVE platform since it is designed for that. If all the components of 5G-TOURS use-cases are available in the VNF catalogue, it is sufficient to inject the blueprint to the portal and our use-case is executed. Otherwise, we have to push the missing VNFs to the 5G EVE VNF's catalogue.

The portal is composed of a front-end and a backend. The backend is communicating with the inter-working layer, which connects the French, Greek, Italian and Spanish sites. This inter-working layer is not accessible by 5G-TOURS. Therefore, 5G-TOURS, as a project, cannot inject any intelligence there. It is clear that most of the use-cases of 5G-TOURS are mono-site and someone could suggest bypassing the inter-working layer. This is not possible, since all the functional blocks between the portal frontend and the site is under the control of 5G EVE.

In the sequel, a description of what is behind the site is provided. The site refers to a geographical location. In our case, we are interested in the French, Italian and Greek sites. Each site implements CORE and RAN domains and orchestrates them using different platforms, namely OSM or ONAP. 5G EVE invested to cover some specific areas of interest for their use-cases. Concerning 5G-TOURS, the radio sites should be extended to include the area of interest. This is in the context of 5G-TOURS plan. The radio extension should be connected to the

CORE used in 5G EVE, which means that the CORE services should be updated to make this feasible (add new IP@, new SIM, etc).

To summarise, the possible extension or entry points where 5G-TOURS could act are depicted in Figure 44 and noticed by the red pictogram. In 5G-TOURS it will be possible to:

- inject new experiment blueprint;
- upload new VNF to the catalogue to build the experiment blueprint;
- enhance some site using a different frequency;
- extend the radio coverage by adding new sites.

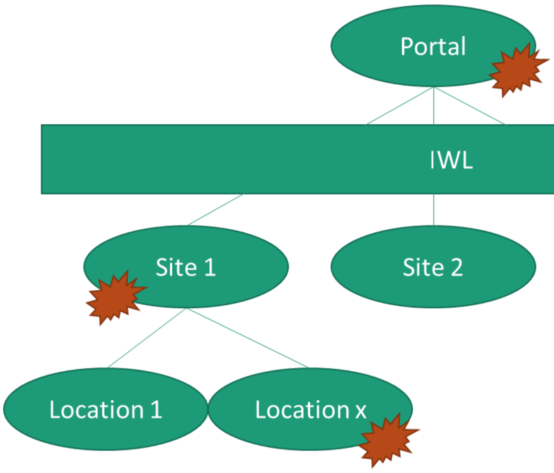


**Figure 44. Possible Injection points of 5G-TOURS and 5G EVE**

#### 4.2.2.2 **5G-TOURS additional technologies**

In the following the main technologies developed by 5G-TOURS that would need to be “inserted” are listed. Specifically:

- Broadcasting solution: Some of the use cases addressed in 5G-TOURS require sending simultaneously video content, as well as other types of content, such as software updates for IoT devices, to an extremely large number of users. Since the use of one-to-one unicast transmissions would be highly inefficient, the support for broadcasting and multicasting technologies providing point-to-multipoint (PTM) transmissions is essential. As the broadcast technology is not supported by 5G EVE, 5G-TOURS will address the design and implementation of this technology within its deployment. 5G-TOURS will follow a stepwise approach in the deployment of the broadcast architecture, where each step provides additional capabilities but also involves more evolved equipment, as analysed below:
    - 5G Non-Standalone (LTE) Core with LTE eMBMS RAN, where both Core and RAN networks are supported by existing 4G infrastructure and the Core will be used from 5G EVE;
    - 5G Standalone Core with LTE eMBMS RAN, with a fully 5G Core infrastructure involving 5G broadcast network;
    - 5G Standalone Core and NR Broadcast RAN, where the 5G Core is combined with an NR Broadcast RAN prototype. This step is subject to the evolution of the related standards and their adoption by the vendors providing the equipment:
        - The feasibility of the first two steps could be realised by combining the 5G EVE platform with the additional inputs provided by the 5G-TOURS: eMBMS middleware (i.e., MBMS client), virtualised BM-SC and MBMS-GW;
        - The mobile core network architecture enables multicast and broadcast capabilities based on the 5G architecture defined in 3GPP Rel. 15. For that, two new NFs will be added into the 5G Core: Multicast User Plane Network Function (MUF) and 5G-Xcast Control Plane Network Function (MCF).
- Orchestration & AI innovation: This requires an elastic management and resource orchestration solution which could be offered by AI. Some examples of performance boosting capabilities that could be provided by AI techniques are the following: (i) learning and profiling the computational utilisation patterns of VNFs, thus relating performance and resource availability, (ii) traffic prediction models for proactive resource allocation and relocation, (iii) optimised VNF migration mechanisms for orchestration using

multiple resource utilisation data (CPU, RAM, storage, bandwidth), and (iv) optimised elastic resource provisioning to network slices based on data analytics.

- Service layer: The 5G-TOURS service layer focuses on tenants that require a customised network slice from the infrastructure that meets their specific requirements for the service that the tenant is providing on top of the network slice. The vertical industries are prominent examples of tenants that may be purchasing a network slice. For instance, a hospital may request a URLLC network slice for its remote operation or first-aid care, an airport may request an mMTC slice for its operation, or a museum may request an eMBB slice to improve the experience of its visitors. In order to request, establish and maintain a network slice, a tenant requires an interface from the infrastructure that includes the following functionalities:
    - **Instantiate a network slice:** When a tenant needs a network slice, it has to issue a request to the infrastructure indicating the following information: (i) the area that needs to be covered by the network slice, (ii) the load that needs to be supported by the slice, (iii) the requirements that need to be satisfied in terms of rates, latencies and reliability, and (vi) the mobile terminals and devices that belong to the slice, among other data;
    - **Orchestrate application-layer virtualised functions:** In case the tenant needs to place some of its application layer functions within the network infrastructure it has to provide the virtualised function to the infrastructure indicating the constraints associated to the placing of such function. This could be the case, for instance, when providing an AR/VR server with very low latency requirements in the museum;
    - **Monitor the network slice:** After its network slice has been instantiated, the tenant needs to be able to monitor the service provided by the network slice in order to confirm that its requirements are being satisfied, assess the service quality that is being provided to its clients, and take corrective actions when required – e.g., if the load of the slice is higher than initially expected, a larger slice may be requested;
    - **Operate the network slice:** The tenant needs to be able to perform some operations on a running network slice, such as adding new users to the slice, increase or decrease its coverage, change the requirements or the load, re-orchestrate application-layer virtualised functions, etc. Note that some of these operations may be taken as a reaction to the monitoring information collected for the slice.

These technologies will not be used on all sites but need to have a proper "insertion strategy".

#### 4.2.2.3 5G-TOURS use-cases

The use cases addressed by 5G-TOURS are grouped around three main themes that represent different aspects of the city as it is summarised in the Figure 45:


**Figure 45. Vision of 5G-TOURS improving the experience of citizens and tourists**

- **The touristic city**: The visitors of museums and outdoor attractions are provided with 5G-based applications to enhance their experience while visiting the city. This includes VR/AR applications to complement the physical visit with additional content, involving interactive tactile communications. The experience of the visitors is also enhanced with robot-assisted services, telepresence to allow for remote visits, as well as live events enabled by mobile communications such as multi-party concerts.
- **The safe city**: 5G technology greatly improves the safety in the city by providing means to better assist health-related care in all the phases of an incident, ranging from the health monitoring for prevention and early detection, to diagnosis and intervention at the ambulance, and surgery at the wireless operation

room in the hospital. Also, 5G enables monitoring and remote assistance to tourists from their home country.

- **The mobility-efficient city**: Mobility to reach and move inside the city is made more efficient and comfortable. This involves smarter cities, gathering information about the city and using it to improve navigation systems as well as parking. Traveling is also made more enjoyable, providing AR/VR services to passengers, and airports become logistically more efficient by relying on 5G for their operation.

### 4.2.3 Cumulative list of insertion strategies

These strategies are listed per type (technology, use-cases) then they are factored.

#### 4.2.3.1 Insertion strategies based on new technologies

5G-TOURS is aiming to build new ideas and technologies around three main areas: broadcast, service and AI, as it is introduced in Section 3.

Referring to the 5G EVE architecture, it seems difficult to inject some new algorithms for orchestrations since the orchestration is not part of the experiment blueprint. 5G EVE orchestration is mainly based on MANO stack. Two approaches are implemented: OSM and ONAP. What could be injected is the preference of deployment of some VNFs. The preferences could cover a geographical area, specific EDGE cloud, specific hardware etc. This injection is feasible via the experiment blueprint/experimental portal.

Concerning the AI technology developed by 5G-TOURS, it exploits the monitoring data collected by 5G EVE platform. The NFVI information is assumed to provide the knowledge on the computational resources' capabilities (e.g., type of CPU, memory, data plane and accelerators) and availability (status and utilisation level). Building on such information and running the AI-based algorithms, the framework can then influence and optimise placement decisions made by the Virtualized Infrastructure Manager (VIM), while ensuring that resources allocation and SLAs are adhered too. Moreover, by using this information, we can further optimise resource utilisation by (i) enabling higher density for a given set of workloads under associated SLA, (ii) anticipating and reacting to changing loads in different slices and assisting the VIM in avoiding resource conflicts, and/or (iii) timely triggering of up/down scaling or in/out scaling of associated resources. Most of these objectives are feasible if the underlying NFVI is accessible (via SSH) and accept to receive policy inputs and share the infrastructure KPIs.

The broadcast technology developed in 5G-TOURS is initially aiming to enhance the capability of the CORE network by allowing specific functions to be able to treat and understand the broadcast flows. These functions 5G-Xcast User Plane Function and Control Plane Function and their revised version (MUF and MCF respectively) [56] could be uploaded to the VNF catalogue, and via the experiment blueprint they can be executed.

To summarise, Table 5 depicts the injection strategies per new technology:

**Table 5. 5G-TOURS technology and injection into 5G EVE possibilities**

| | 5G EVE Injection place | | | 5G-TOURS specific platform |
|---|---|---|---|---|
| 5G-TOURS Technology | Portal | Site | Location | |
| Broadcast | | | X | X |
| AI | | | | X |
| Service | X | | | X |

In the Table 5, the site refers to the 5G EVE available sites. The location refers to the place of the use-cases deployment. For example, the use-case for the wireless operating room will be located in CHU Rennes. This location is not part of 5G EVE, so, 5G-TOURS should add it in the location inventory of 5G EVE.

#### 4.2.3.2 Insertion strategies based on new use-cases

5G-TOURS is targeting 13 use-cases grouped into three clusters and introduced in Table 6. WP4 is in charge of the "Touristic city" cluster, WP5 is in charge of the "safe city" cluster, and WP6 is in charge of the "Mobility-efficient city" cluster.

**Table 6. Summary of the uses cases to be trialled by 5G-TOURS**

| Use case | Vertical customer | Slice type(s) | KPI requirements | Improvements provided | Vertical solutions |
|---|---|---|---|---|---|
| ***Touristic city (Turin)*** | | | | | |
| *Augmented tourism experience* | Museum | eMBB, URLLC, mMTC | Per-user data rate up to 500 Mb/s, latency < 10 ms | Improving visitor's experience | XR application (AR/VR/MR) |
| *Telepresence* | Museum | eMBB, URLLC | Latency < 10 ms | Remote museum visit | Robot & remote interface |
| *Robot-assisted museum guide* | Museum | URLLC | Latency < 10ms | Improved visitors' experience and safety | Robot |
| *High quality video services distribution* | TV broadcaster | eMBB | Per-user data rate of 25 Mb/s, several users/m$^2$ | Improved video user's experience | App for content / video distribution |
| *Distributed video production* | TV broadcaster | URLLC, eMBB | Latency < 5 ms Reliability > 99.99% | Concert by distributed orchestra | Media production backpack |
| ***Safe city (Rennes)*** | | | | | |
| *Remote health monitoring & emergency mgmt.* | Hospital | mMTC, URLLC | Several devices/m$^2$, reliability > 99.99% | Prompter safety reaction upon an anomaly | Wearables & patches for health monitoring |
| *Teleguidance for diagnostics and intervention support* | Hospital | URLLC, eMBB | Speeds above 100 Km/h, 2 Gb/s, latency < 10 ms, reliability > 99.999% | Saving lives through improved assistance | Remote treatment & diagnostics, smartglasses |
| *Wireless operating room* | Hospital | URLLC, eMBB | Latency < 5ms, reliability > 99.9999%, total data rate > 10Gb/s | Saving lives in the operating room | AR & Cobotic assisted surgery, smartglasses |
| *Optimal ambulance routing* | Hospital | mMTC | Speeds above 100 Km/h, thousands sensors/Km$^2$ | Faster journey to hospital | City sensors & navigation app |
| ***Mobility-efficient city (Athens)*** | | | | | |
| *Smart parking management* | Airport | mMTC | Density of 50,000 devices/Km$^2$ | Fast & personalised parking for drivers | Parking sensors & driver app |
| *Video-enhanced ground-based moving vehicles* | Airport | eMBB | Per-user data rate above 25 Mbps, speeds up to 100 Km/h | Improved airport logistics | Live video feeds application |
| *Emergency airport evacuation* | Airport / Security agency | URLLC, mMTC | Reliability > 99.99%, location accuracy ≤ 1m, several devices/m$^2$ | Safer emergency handling for travellers | Personalised evacuation application |
| *AR/VR-enhanced educational bus excursion* | School | eMBB | Per-user data rate up to 500 Mb/s, latency < 10 ms, speeds of 100 Km/h | Improving students' educational experience | AR/VR application |

All the use-cases, listed here, are being analysed to determine how they will be inserted on 5G EVE platform.

In each WP, the work is focused on the definition of the required VNFs, blueprints, networking needs etc. Once these parameters are defined, we can define the use-case insertion strategy.

At this stage of the analysis, the insertion strategies are presented in Table 7:

- Tourist city cluster (WP4 use-cases):
  - 5G EVE portal will be used to expose new "touristic" services.
- Safe city cluster (WP5 use-cases):

  - ONAP is the primary "insertion technology" for supporting new domains & services. This point means the need to declare the new site in the ONAP domain by updating its multi-cloud OpenStack to be able to instantiate VNFs in the right site;
  - 5GEVE portal is used to expose new "healthcare" services.
- Mobility-efficient city cluster (WP6 use-cases):
  - 5G EVE portal will be used to expose new "mobility" services.

**Table 7. Insertion strategies per use-cases cluster**

| *Use-cases Insertion point* | *Tourist city* | *Safe city* | *Mobility-efficient city* |
|---|---|---|---|
| *5G EVE portal* | Yes | Yes | Yes |
| *Orchestrator (ONAP, OSM)* | Use 5G EVE orchestrator. No update from 5G-TOURS is planned | Need to update ONAP domain to add new OpenStack cluster | Use 5G EVE orchestrator. No update from 5G-TOURS is planned |
| *Site* | Turin | Rennes | Athens |
| *Location* | Under analysis | Yes. Need to add new location to cover B-COM parking and the Operating Room in CHU Rennes | Under Analysis |

### 4.2.4 Technologies development

This step will take care of implementing the added components identified in Step 2 to make sure the variety of use cases in each node has the necessary infrastructure support. This task will be handled in conjunction with Task 7.1 dealing with the integration of the multiple and heterogeneous use cases with sometimes conflicting requirements on the Turin, Rennes and Athens infrastructure nodes available in this project. The integrated ecosystem targeted by this step will be necessarily fed by the implementation use-case-specific WPs to allow the integrated execution of trials in the three nodes.

# 5 Conclusion

This document illustrates the progress on the architecture and physical deployment in 5G-TOURS. The architecture has been designed to accommodate all the novel technologies developed during the project that will support and enhance the proposed use cases. In detail, this document describes:

- The requirements and overall architecture of 5G-TOURS, including a high-level description of every layer involved and an observation on security, which is included by design;
- The physical implementation progress on the three nodes involved: Torino, Rennes and Athens. A description of the current capabilities, including use case on-boarding and 5G EVE integration is also covered, alongside the plans for future enhancement;
- The progress on the development of 5G-TOURS technologies. These include i) Enhanced MANO; ii) AI Orchestration; iii) Broadcast support; iv) Service Layer and v) Enhancements to 5G RAN. The status in standardization fora is covered as well as the implementation and integration of the subjacent NFs or algorithms into 5G-TOURS architecture;
- The integration methodology into 5G EVE infrastructure, with the possible insertion points for the 5G-TOURS technology. The deliverable also covers the status in 5G EVE with special focus into the application portal and the interworking layer.

Two more deliverables of WP3 (D3.3 and D3.4) describing the architecture are planned during the project duration. The focus will shift to use case on-boarding, custom architectural solutions, and technology implementation in 5G EVE.

# Acknowledgment

This project has received funding from the EU H2020 research and innovation programme under Grant Agreement No. 856950.

# References


[1] 5G-EVE, "Website," [Online]. Available: https://www.5g-eve.eu/.

[2] GSA, "LTE, 5G & 3GPP IoT Chipset Report," June 2020. [Online].

[3] M. Fuentes, L. Carcel, Jose, C. Dietrich, L. Yu, E. Garro and V. Pauli, "5G New Radio Evaluation Against IMT-2020 Key Performance Indicators," *IEEE Access*, vol. 8, pp. 110880-110896, 2020.

[4] 5G Observatory, "5G Services. Announcements of commercial launches," March 2020. [Online].

[5] 5G EVE, "Collaboration with vertical use case projects – 5G EVE presentation at 5G PPP Technology Board," June 2020. [Online].

[6] M. Gramaglia, "5G-TOURS D3.1, Baseline architecture and deployment objectives," 2019. [Online].

[7] ONAP, "Open Network Automation Platform," [Online]. Available: https://www.onap.org/.

[8] O-RAN Alliance, "White paper, "O-RAN: Towards an Open and SmartRAN"," 2018. [Online].

[9] Y. Altevet and J. E. González, "5G-TOURS D2.1, Use cases, deployment and techno-economic requirements – high level description," 2019. [Online].

[10] R. Legouable, "5G-EVE D2.1, Initial detailed architectural and functional site facilities description," 2018. [Online].

[11] J. Garcia-Reinoso, "5G-EVE D4.2, First version of the experimental portal and service handbook," 2018. [Online].

[12] 3GPP, "TS 23.288: Architecture Enhancements for 5G System (5GS) to Support Network Data Analytics Services," 2019. [Online].

[13] 3GPP, "TS 28.533: Management and Orchestration of Networks and Network Slicing, Management and Orchestration Architecture," 2019. [Online].

[14] ETSI, "White Paper No. #32: Network Transformation; (Orchestration, Network and Service Management Framework)," 2019. [Online].

[15] ETSI NFV ISG, "NFV-MANO"Network functions virtualisation (NFV); management and orchestration)," 2014. [Online].

[16] ETSI, "Experiential Networked Intelligence (ENI)," [Online]. Available: https://www.etsi.org/technologies/experiential-networked-intelligence.

[17] R. Román, "5G EVE D1.3: 5G EVE end to end facility reference architecture for vertical industries and core application," 2019. [Online].

[18] b<>com, "Wireless Edge Factory," [Online]. Available: https://b-com.com/en/bcom-wireless-edge-factory.

[19] Flexible Netlab Platform, [Online]. Available: https://b-com.com/fr/bcom-flexible-netlab.

[20] G. Sacco, "5G-TOURS D4.1: Robotic, Media and Smart Cities Solutions for Touristic Cities," 2019. [Online].

[21] E. Cordonnier, "5G-TOURS D5.1: 5G-enabled solutions for safe cities," 2019. [Online].

[22] N. Papagiannopoulos, "5G-enabled solutions for mobility-efficient city," 2019. [Online].

[23] SONATA, "AGILE DEVELOPMENT, TESTING AND ORCHESTRATION OF SERVICES IN 5G VIRTUALIZED NETWORKS," [Online]. Available: https://www.sonata-nfv.eu/.

[24] 5G-TRANSFORMER, "5G Mobile Transport Platform for Vertical," [Online]. Available: http://5g-transformer.eu/.

[25] 5G-MONARCH, "5G Mobile Network Architecture for diverse services, use cases, and applications in 5G and beyond," [Online]. Available: https://5g-monarch.eu/.

[26] blueplanet, "NFV Orchestration. Intelligent and automated lifecycle management of virtualized services," [Online]. Available: https://www.blueplanet.com/products/nfv-orchestration.html.

[27] Nokia, "CloudBand. Adopt lean operations and increase business agility," [Online]. Available: https://www.nokia.com/networks/solutions/cloudband/.

[28] Ericsson, "Ericsson Orchestrator," [Online]. Available: https://www.ericsson.com/en/portfolio/digital-services/automated-network-operations/orchestration/ericsson-orchestrator.

[29] OSM, "Open Source MANO," [Online]. Available: https://osm.etsi.org/.

[30] G. Lavado, "OSM Service Assurance," 2019. [Online]. Available: https://osm-download.etsi.org/ftp/osm-6.0-six/8th-hackfest/presentations/8th%20OSM%20Hackfest%20-%20Session%208%20-%20OSM%20Service%20Assurance.pptx.pdf.

[31] Prometheus, [Online]. Available: https://prometheus.io/.

[32] F. Vicens, "AI Agents for OSM - Feature description," 11 06 2020. [Online]. Available: https://osm.etsi.org/gerrit/#/c/osm/Features/+/9063/1/AI+Agents+for+OSM.md.

[33] Acumos, “Acumos Marketplace,” 11 06 2020. [Online]. Available: https://marketplace.acumos.org/.

[34] OSM, "Release 6 Network Slice Primitives," [Online]. Available: https://osm.etsi.org/gerrit/#/c/osm/Features/+/7187/1/Release6/Network+Slice+primitives.md.

[35] ETSI, “Network Virtualization Research Challenges,” 2017. [Online].

[36] IETF, “RFC8172, Considerations for Benchmarking Virtual Network Functions and Their Infrastructure,” 2017. [Online].

[37] IETF, “RFC 8455, Terminology for Benchmarking SDN Controller Performance,” 2018. [Online].

[38] J. Mendes and X. Jiao, “Cellular Access Multi-tenancy through Small-cell Virtualization,” *Computer Communications,* no. 133, pp. 59-66, 2019.

[39] C. Zhang and P. Patras, “Deep Learning in Mobile and Wireless Networking: a Survey,” *IEEE Communications Surveys & Tutorials,* 2019.

[40] D. Bega, M. Gramaglia and M. Fiore, “DeepCog: Cognitive Network Management in Slices 5G Networks with Deep Learning,” in *IEEE INFOCOM*, Paris, 2019.

[41] Elastic Beats, "Lightweight data shippers," [Online]. Available: https://www.elastic.co/beats/.

[42] D. Gomez-Barquero, D. Navratil, S. Appleby and M. Stagg, “Point-to-Multipoint Communication Enablers for the Fifth Generation of Wireless Systems,” *IEEE Communication Standards Magazine,* vol. 2, no. 1, pp. 53-59, 2018.

[43] D. Gómez-Barquero, J. Giménez and R. Beutler, "3GPP Enhancements for Television Services: LTE-Based 5G Terrestrial Broadcast," in *Wiley Encyclopedia of Electrical and Electronics Engineering*, 2020.

[44] 3GPP, "TR 23.757: Study on architectural enhancements for 5G multicast-broadcast services," 2020. [Online].

[45] 3GPP, "TR 26.802: 5G Multimedia Streaming (5GMS); Multicast architecture," 2020. [Online].

[46] 3GPP, "RP-193248: New Work Item on NR support of Multicast and Broadcast Services," 2020. [Online].

[47] M. Säily, C. Barjau Estevan, J. Joan Gimenez, F. Tesema, W. Guo and D. Gomez-Barquero, “5G Radio Access Network Architecture for Terrestrial Broadcast Services,” *IEEE Transactions on Broadcasting,* vol. 66, no. 2, pp. 404-415, 2020.

[48] T. Tran, D. Navratil, P. Sanders, J. Hart, R. Odarchenko, C. Barjau Estevan and D. Gomez-Barquero, “Enabling Multicast and Broadcast in the 5G Core for Converged Fixed and Mobile Networks,” *IEEE Transactions on Broadcasting,* vol. 66, no. 2, pp. 428-439, 2020.

[49] O-RAN, "Architecture Description v1.0," 2020. [Online].

[50] O-RAN, "AI/ML Workflow Description and Requirements Version 1.0," 2019. [Online].

[51] O-RAN, "Operations and Maintenance Architecture Version 2.0," 2019. [Online].

[52] O-RAN, "Operations and Maintenance Interface Version 2.00," 2019. [Online].

[53] TM forum, "GB917, SLA Management Handbook, Release 3.1," [Online]. Available: https://www.tmforum.org/resources/best-practice/gb917-sla-management-handbook-release-3-1/.

[54] L. Luque, “5G EVE D3.2: Interworking Reference Model,” 2016. [Online].

[55] ETSI, “NFV-SOL 005, Network Functions Virtualisation (NFV) Release 2; Protocols and Data Models; RESTful protocols specification for the Os-Ma-nfvo Reference Point,” 2018. [Online].

[56] Tuan Tran, Ed., "Deliverable D4.1: Mobile Core Network," 5G-Xcast, June 2018.

[57] 3GPP, "Architectural enhancements for 5G multicast-broadcast services," 28 May 2019. [Online]. Available: http://www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_84/Docs/SP-190442.zip. [Accessed 5 September 2019].

[58] Qualcomm, "Snapdragon X60 5G Modem-RF System," 2020. [Online]. Available: https://www.qualcomm.com/products/snapdragon-x60-5g-modem.

[59] 5G EVE, "D4.2, First Version of the experimental portal and service handbook," 2019. [Online].

# Annex A: 5G EVE blueprints descriptors and example

## Blueprints descriptors preparation

In order to be able to use the different 5G EVE portal functionalities, the following steps must be taken:

- Two initial type of blueprints must be prepared i.e. the Vertical Service Blueprints (VSB) and Context Blueprints (CtxBs). The number of context blueprints needed depend on the number of contexts you want to consider in your experiment. For example, delay and background traffic, each having their own CtxBs.
    - A functional VSB can be found in this link[1].
    - Examples CtxBs can be found in these footnotes [2] [3].
- For all blueprints, the preferred format is yaml, however JSON is supported.
    - The following online tool can be used to convert your YAML files to JSON[4]
    - Once the yaml file has been converted to JSON using the above link, copy the JSON file from the browser by clicking anywhere in the JSON text, press ctrl+A to select all, followed by ctrl+c to copy and save the file on your device with the same name as before but this time with the .JSON extension.
    - At this point, for each of our blueprints and NSD we should be having two files, one in yaml and the other in JSON format.
    - The 5G EVE portal only accepts the JSON formats of the descriptors so the JSON files from this step will be the ones uploaded to the portal.
- Context Blueprints are accompanied with measurements tests that will be performed in the experiment.
- After preparing all the blueprint descriptors, a validation of the blueprint format is necessary. 5G EVE offers an online blueprint-validator tool. Detailed instructions on how to install and use the blueprint validator can be found at the following link [5].

## Prepare network service descriptors (nsd)

- For each of the blueprints in the previous step, the corresponding Network Service Descriptors in NFV-IFA014 format need to be prepared, this format has been adopted in the 5G EVE portal. More details about the NFV-IFA014 standard can be found in this document [6].
- After the preparation of the NSDs corresponding for each blueprints, one additional NSD is needed. This NSD is referred as the ***composite NSD***. The ***composite NSD*** is the final network service, which encompasses both the vertical service and context NSDs. It is the final NSD that will actually be executed by the 5G EVE portal. An example is included here [7].

[1] https://github.com/5GEVE/blueprint-yaml/blob/master/vsb/vsb_asti_agv/asti_agv_vsb_v2.yaml

[2] Delay_CtxB: https://github.com/5GEVE/blueprint-yaml/blob/master/ctx/ctx_delay/asti_ctx_delay.yaml

[3] Background_traffic_CtxB: https://github.com/5GEVE/blueprint-yaml/blob/master/ctx/ctx_bg_traffic/asti_ctx_bg_traffic.yaml

[4] https://www.json2yaml.com/convert-yaml-to-json

[5] https://github.com/5GEVE/blueprint-validator

[6] https://www.etsi.org/deliver/etsi_gs/NFV-IFA/001_099/014/02.01.01_60/gs_NFV-IFA014v020101p.pdf

[7] https://github.com/5GEVE/blueprint-yaml/blob/master/vsb/vsb_asti_agv/asti_agv_all_nsd.yaml