NEXT GENERATION INTERNET

Grant Agreement No.: 957216
Call: H2020-ICT-2018-2020

Topic: ICT-56-2020
Type of action: RIA

D4.5 Smart end-to-end INGENIOUS IoT System

Revision: v.1.0

Work package	WP4
Tasks	T4.1, T4.2 and T4.3
Due date	30/11/2022
Submission date	30/11/2022
Deliverable lead	TUD
Version	1.0
Authors	Anton Luca Robustelli (TEI), Christos Politis (SES), Eddy Higgins (iDR), Erin E. Seder (NXW), Francisco Javier Curieses Sanz (UPV), Giacomo Bernini (NXW), Gino Ciccone (TEI), Giuseppina Carpentieri (TEI), Ivo Bizon (TUD), Joe Cahill (iDR), Jose Costa Requena (CMC), Manuel Fuentes (5CMM), Miguel Cantero (5CMM), Nuria Molner (UPV), Pietro Piscione (NXW), Rania Rojbi (TUD), Roberto Bomfin (TUD), Shane Bunyan (iDR)
Reviewers	Carsten Weinhold (BI), Christos Politis (SES), Francisco Javier Curieses Sanz (UPV), Giacomo Bernini (NXW), Ivo Bizon (TUD), Jose Costa Requena (CMC), Michael Roitzsch (BI), Nuria Molner (UPV), Pietro Piscione (NXW), Tadeusz Puzniakowski (PJAKT)

Abstract	This deliverable describes the smart end-to-end IoT system proposed within iNGENIOUS. The system integrates under a central architecture the heterogeneous RAN landscape of the current and forthcoming IoT technologies fulfilling the requirements of different traffic classes. The deliverable is divided into three main parts. Namely, smart multi-technology IoT radio access network, smart IoT core network, and edge IoT orchestration and network slicing.
Keywords	Radio Access Network, Core Network, Network Slice Orchestration, 5G, Internet-of-Things

Document Revision History

Version	Date	Description of change	List of contributor(s)
V1.0	30/11/2022	EC version	See author list

Disclaimer

This iNGENIOUS D4.5 deliverable is not yet approved nor rejected, neither financially nor content-wise by the European Commission. The approval/rejection decision of work and resources will take place at the Final Review Meeting planned in July 2023, after the monitoring process involving experts has come to an end.

The information, documentation and figures available in this deliverable are written by the "Next-Generation IoT solutions for the universal supply chain" (iNGENIOUS) project's consortium under EC grant agreement 957216 and do not necessarily reflect the views of the European Commission.

The European Commission is not liable for any use that may be made of the information contained herein.

Copyright notice

Project co-funded by the European Commission in the H2020 Programme		
Nature of the deliverable:	**R**	
Dissemination Level		
PU	Public, fully open, e.g. web	✓
CL	Classified, information as referred to in Commission Decision 2001/844/EC	
CO	Confidential to iNGENIOUS project and Commission Services	

** R: Document, report (excluding the periodic and final reports)*

DEM: Demonstrator, pilot, prototype, plan designs

DEC: Websites, patents filing, press & media actions, videos, etc.

OTHER: Software, technical diagram, etc.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Executive Summary

The final development stage of the components that integrate to form the iNGENIOUS end-to-end smart IoT network system, are presented in this deliverable. These components, which were introduced in previous deliverables of WP4, are integrated under a central iNGENIOUS architecture, where the heterogeneous radio access network (RAN) landscape of the current IoT technologies can be combined to meet the requirements of different traffic classes. To validate and demonstrated the feasibility of the proposed architecture, the components are tested and integrated in several use cases related to the whole supply chain. This deliverable is divided into three main parts, which cover the three subcomponents of the iNGENIOUS IoT network layer. These are the smart multi-technology IoT radio access network, smart IoT core network, and edge IoT orchestration and network slicing.

The multi technology radio access network (RAN) components are firstly described in this document. A number of different techniques employed throughout the project are described, such as Flexible RAN, O-RAN, 5G, and Satellite communication links. Secondly, the core network is presented as the central component of the iNGENIOUS IoT network layer, and it interconnects the RAN with applications and services that consume and produce data at the end devices. Lastly, the orchestration and management framework integration with the other components is described. In addition, this document describes the relationship among the technical developments to the iNGENIOUS use cases followed by a summary of the exploitation potential.

After reading this document, the reader should be able to understand the individual components that were developed in WP4 as well as their position within the iNGENIOUS architecture.

Co-funded by the Horizon 2020
Framework Programme of the European Union

Table of Contents

Co-funded by the Horizon 2020
Framework Programme of the European Union

List of figures

Co-funded by the Horizon 2020 Framework Programme of the European Union

Co-funded by the Horizon 2020
Framework Programme of the European Union

List of tables

Co-funded by the Horizon 2020
Framework Programme of the European Union

Abbreviations

3GPP	Third Generation Partnership
5G	Fifth Generation
AD	Anomaly Detection
AGV	Automated Guided Vehicle
AI	Artificial Intelligence
CN	Core Network
DN	Data Network
eMBB	Enhanced Mobile Broadband
ESTI	European Telecommunications Standards Institute
GAD	Geographical Anomaly Detection
GBDT	Gradient Boosted Decision Tree
GFDM	Generalized Frequency Division Multiplexing
gNB	Next Generation Node B
GUI	Graphical User Interface
GW	Gateway
IP	Internet Protocol
IoT	Internet of Things
KPI	Key Performance Indicator
LCM	Slice Lifecycle Management
LSTM	Long Short-Term Memory
M2M	Machine-to-machine
MAC	Medium Access Control
MANO	Management and Orchestration
ML	Machine Learning
NB-IoT	Narrow Band IoT
NFV	Network Function Virtualization
NPN	Non-Public Network
NR	New Radio
NR-NTN	New Radio Non-Terrestrial Network
NSMF	Network Service Management Function
NSSMF	Network Slice Subnet Management Function
NST	Network Slice Template
NWDAF	Network Data Analytics Function
O-RAN	Open RAN
OFDM	Orthogonal Frequency Division Multiplexing
PDU	Protocol Data Unit
PHY	Physical layer
QP	Quality Prediction
QoE	Quality of Experience
QoS	Quality of Service

Co-funded by the Horizon 2020 Framework Programme of the European Union

RAN	Radio Access Network
RAT	Radio Access Technologies
RIC	RAN Intelligent Controller
RNN	Recurrent Neural Network
RT	Real Time
RTT	Round Trip Time
SDR	Software Defined Radio
SIM	Subscriber Identity Module
SMO	Service Management & Orchestrator
SoTA	State-of-the-art
TDMA	Time Division Multiple Access
TFT	Traffic Flow Template
TN	Transport Network
TS	Traffic Streeting
TSN	Time-Sensitive Networking
UE	User Equipment
UPF	User Plane Function
URLLC	Ultra Reliable Low Latency Communications
VLAN	Virtual Local Area Network
VPN	Virtual Private Network
VSMF	Vertical Service Management Function
xApps	External Applications

Co-funded by the Horizon 2020
Framework Programme of the European Union

1 Introduction

1.1 Objective of this Deliverable

The main goal of this deliverable is to describe the final iNGENIOUS end-to-end communication system starting with the lower layers, i.e., radio access network (RAN) technologies, going through the core network, and finally describing the management and orchestration system used to meet the requirements of IoT supply chain scenarios. Moreover, the employment of these techniques in the iNGENIOUS use cases is presented. The business potential and exploitation plans are summarized in the end of each chapter, and further detailed in deliverable D7.3.

This deliverable describes the final outcomes of iNGENIOUS related to the three aforementioned areas, which are part of the iNGENIOUS IoT Network architecture, highlighted in red, and depicted in Figure 1.

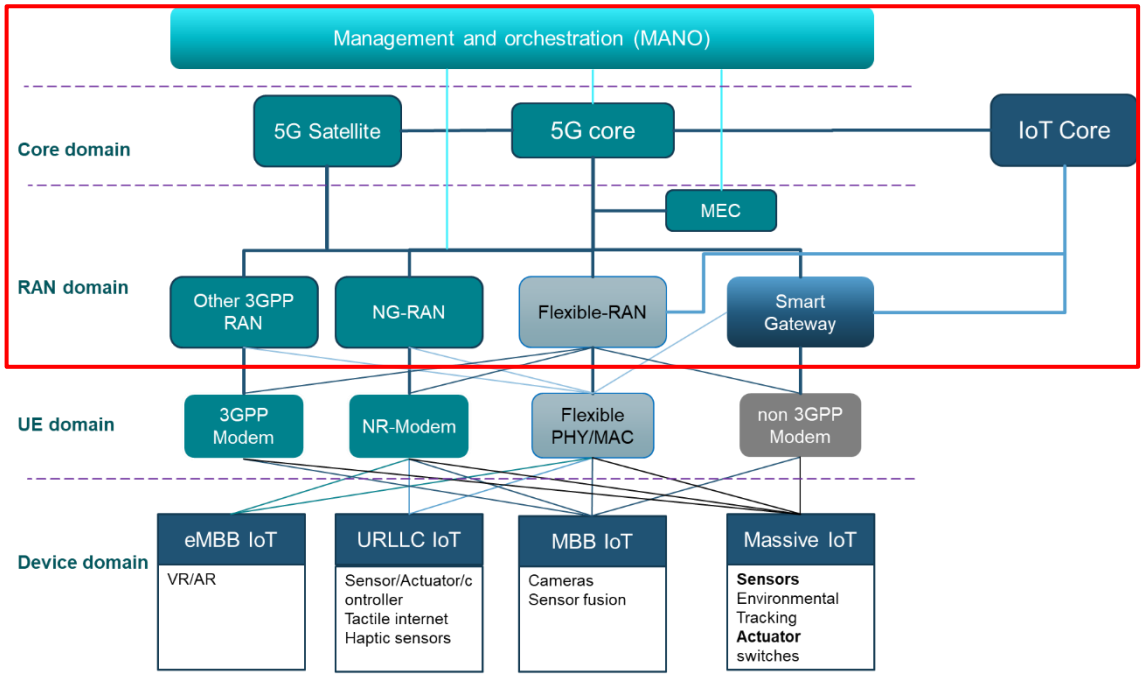


Figure 1: Illustration of the smart end-to-end IoT network layers with their corresponding interconnections.

1.2 Role of WP4 in iNGENIOUS

Using the iNGENIOUS architecture described in deliverable D2.4 [1] as reference, the work carried out within WP4 is focused on the Network layer. The Network layer is responsible for all the information exchange between the Things layer and the Data Management layer. Figure 2 illustrates the overall architecture highlighting the WP4 IoT Network layer.

Co-funded by the Horizon 2020 Framework Programme of the European Union

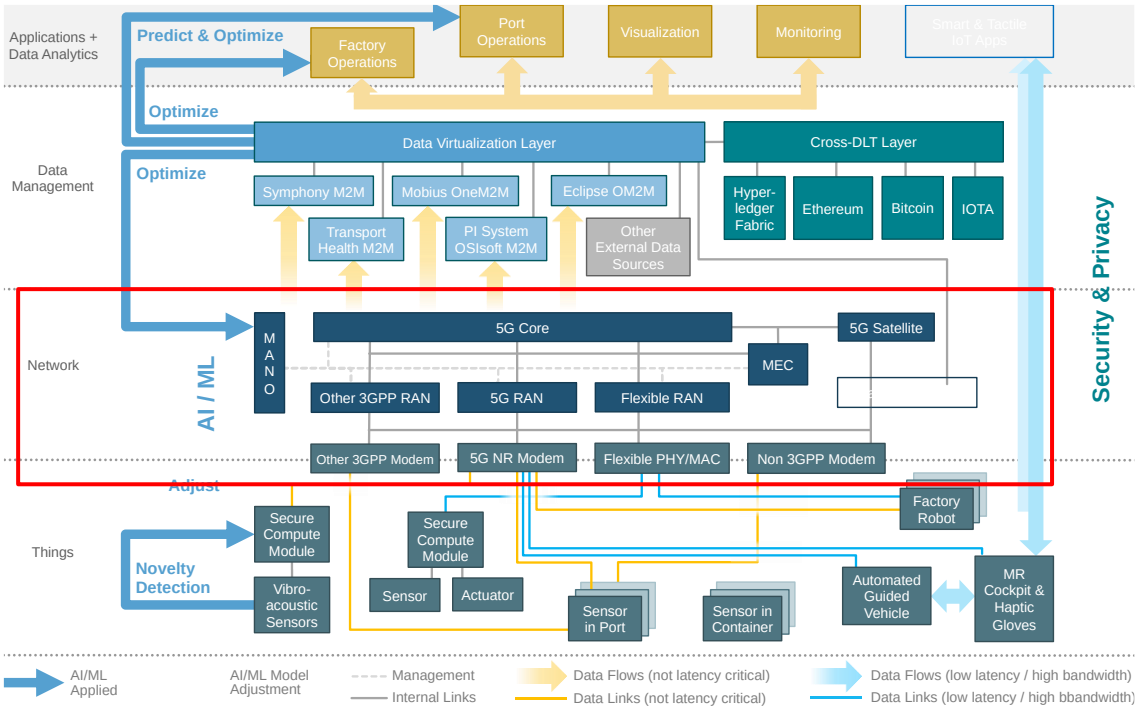


Figure 2: Complete iNGENIOUS architecture highlighting the components developed within WP4.

Due to the high diversity of communication devices and their associated network requirements, the iNGENIOUS Network layer must offer support to a range of bandwidth, latency, range, reliability, and energy-efficiency demands. Therefore, the iNGENIOUS network layer includes the state-of-the-art (SoTA) 5G core and MANO implementations, together with 3GPP compliant and non-3GPP radio technologies including terrestrial and satellite connections. One of the main goals of WP4 is to provide a network infrastructure that can be used by a diverse range of devices.

1.3 IoT Network Architecture

The main component of the iNGENIOUS IoT Network layer is the 5G core. It connects the 5G and Flexible RANs, the Smart IoT Gateway, and the satellite access networks. On top of the 5G core, the end-to-end network slice orchestration framework (MANO) offers overarching network and service orchestration capabilities that ensures that end-to-end performance requirements are met.

1.4 Role of WP4 in the Use Cases

All components developed within WP4 are present in different formats in the six iNGENIOUS use cases, as has been described in deliverable D4.1 [2]. In this document, at the end of each chapter, there is a paragraph describing the roles played by each component of the IoT Network layer in each use case. Table 1 presents the use case identification used throughout iNGENIOUS.

Co-funded by the Horizon 2020 Framework Programme of the European Union

Identifier	Use case definition
UC1	Automated Robots with Heterogeneous Networks – Factory UC
UC2	Transportation Platform Health Monitoring – Transport UC
UC3	Situational Understanding and Predictive Models in Smart Logistics – Port Entrance UC
UC4	Improve Driver's Safety with MR and Haptic Solutions - AGVs UC
UC5	Inter-Model Asset Tracking Via IoT and Satellite - Ship UC
UC6	Supply Chain Ecosystem Integration - DLT UC

Table 1: Use case naming identification.

1.5 Structure of the document

The document is structured as follows: Chapter 2 presents the Smart Multi-Technology IoT Radio Access Network that has been employed. Chapter 3 describes the Smart IoT network. Chapter 4 describes the end-to-end Network Slice Orchestration framework (MANO). Finally, Chapter 5 concludes the document.

Co-funded by the Horizon 2020 Framework Programme of the European Union

2 Smart Multi-Technology IoT Radio Access Networks

The iNGENIOUS architecture was designed to support multiple radio access technologies (RATs) to be used in the next-generation IoT networks. In the architecture detailed in deliverable D2.4 [1], several end devices with different requirements are presented e.g. enhanced mobile broadband (eMBB), ultra reliable low latency communications (URLLC), or Massive IoT. Therefore, *Task 4.1* within WP4 has developed and included innovations in multiple RATs, such as: Flexible PHY/MAC, employing a multiple access scheme to allow the dynamic resource allocation; AI/ML for RAN (O-RAN), researching the main functionalities of the RAN Intelligent controller and some use cases; Smart IoT Gateway, allowing routing and sorting of sensor data from network sensors; Satellite Connectivity, interconnecting the IoT data between Smart IoT Gateway and the cloud and researching direct access options for Satellite; and 5G New Radio (NR) modem, providing 5G connectivity to different AGVs or machines. These components introduce different performance criteria and requirements based on the connected devices. In this chapter, all development activities and their individual results are presented.

2.1 Smart RAN

2.1.1 Flexible PHY/MAC

In order to attain the requirements from a wide range of traffic patterns, the Flexible physical and medium access control (PHY/MAC) implementation used in the iNGENIOUS Factory use case employs a multiple access scheme that has been configured to allow dynamic resource allocation. This is achieved through a control channel, which allows the base station to have control over the radio resources, and to dynamically set the end user's configurations to target the requirements. The implemented flexible MAC design includes the multiple access schemes based on time-division multiple access (TDMA). At the PHY level, the implementation is based on the generalized frequency division multiplexing (GFDM) waveform, which is advantageous from the flexibility perspective (i.e., GFDM has extra configuration parameters when compared to orthogonal frequency division multiplexing (OFDM) (used in the 5G New Radio standard)) that can be tuned depending on the data type and user's needs.

The flexible PHY/MAC approach contributes to the development of smart networking starting from the physical layer. The performance can be further enhanced by the employment of data-driven techniques, such as machine learning, to assist the selection of the multiple configuration parameters. Moreover, integration of the PHY/MAC solution with the 5G network has been performed to enable 5G connectivity to the different UEs, and validate the PHY/MAC approach in a realistic 5G enabled industrial scenario. This has been done by integrating the PHY/MAC resource allocation with the 5G core and the management and orchestration (MANO) framework to setup and configure specialized end-to-end network slices tailored to the various mobile applications and UE requirements. This setup, built within TUD's testbed, is illustrated in Figure 3. Moreover, in this context, the integration of artificial intelligence and machine learning (AI/ML) techniques within the MANO

Co-funded by the Horizon 2020 Framework Programme of the European Union

framework (as described below) can help in achieving smarter NR networking and resource allocation at runtime.

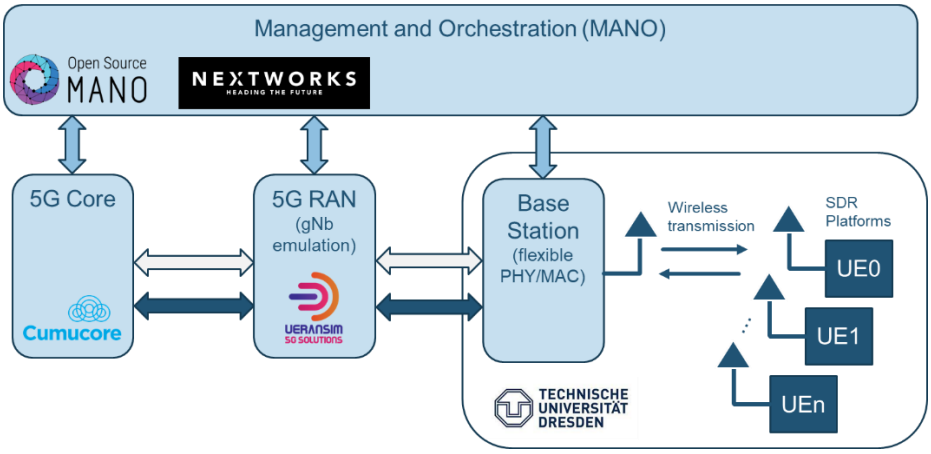


Figure 3: Setup at TUD's testbed used to demonstrate the integration between Flexible PHY/MAC, 5G core and MANO.

Further details about the Flexible PHY/MAC, used in iNGENIOUS, can be found in deliverables D4.1 [2] and D4.2 [3].

2.1.2 AI/ML for RAN/O-RAN

The O-RAN Alliance is committed to evolving radio access networks, with its core principles being intelligence and openness. With this purpose, several use cases where the automation, flexibility, and agility of the RAN are defined. In [4], the traffic steering (TS) use case is introduced as a widely used network solution to achieve optimal traffic distribution based on desired objectives. Due to the heterogenous nature of the iNGENIOUS network architecture in terms of capacity, throughput, and latency requirements of the different UEs, TS has been considered for further research. In order to achieve intelligent and proactive TS control, new components with AI/ML capacities as well as RAN control/guidance interfaces, have been tested.

To optimize resource management and increase the capabilities of the next-generation wireless networks, which is especially interesting in IoT networks, O-RAN has introduced new components in its architecture with AI/ML functionalities. The RAN Intelligent Controller (RIC) is designed to execute applications based on AI/ML, providing the possibility of managing the radio resources in real-time, through near-RT RIC, or in non-real time, through non-RT RIC. Moreover, the flexibility, interoperability, virtualization of the architecture, and open-source software make O-RAN an interesting research branch for the deployment of future 5G networks.

In order to improve the TS use case mentioned, the possible inclusion of functionalities of O-RAN architecture components within IoT radio access networks has been studied. In particular, the RIC and A1 interface, which interconnects and sends the policies set in service, management & orchestrator (SMO) to near-RT RIC, have been deployed and analyzed. As the TS use case is encompassed in the anomaly detection (AD) use case in O-RAN,

Co-funded by the Horizon 2020 Framework Programme of the European Union

this last use case has been studied making use of different external applications (xApps), located in the near-RT RIC, and the A1 interface.

The AD use case aims to detect and correct anomalies in the UEs connected to the network. Several AI/ML applications, called xApps, are executed in the near-RT RIC. The complete use case reported and explained in D4.2 [3] has been updated to the latest release of O-RAN including the updated versions of the xApps, that consider periodic updates of the generated models with new incoming data.

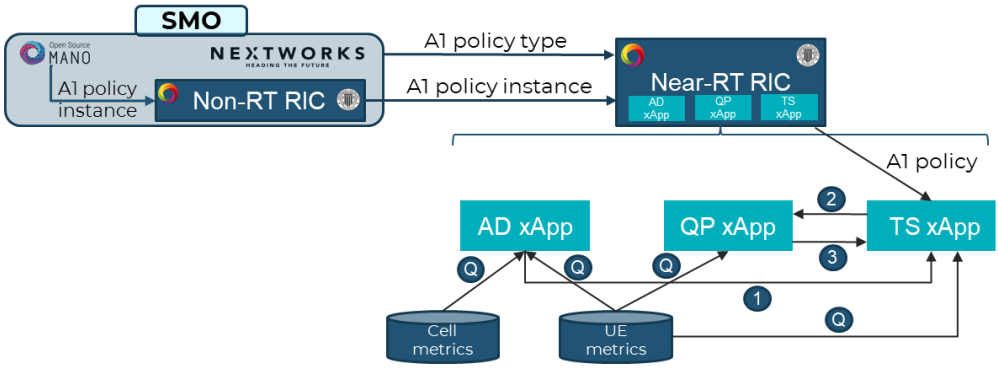


Figure 4. Architecture O-RAN and xApps deployed at UPV lab.

The final architecture employed to test the aforementioned o-RAN use cases is shown in Figure 4, where the MANO developed in the project, has been used as the SMO of the architecture. This integration allows MANO to create the A1 policy type, through an application programming interface (API) request, with several parameters and entries, such as thresholds, QoS or QoE policies, thus providing flexibility in policy setting. In this case, the A1 policy was instantiated with a threshold to establish a value that determines if one UE should change to another cell based on their throughput. The AD use case uses three xApps. First, the AD xApp finds anomalies in the UEs, e.g, throughput degradation, and sends this information to TS xApp (arrow number 1 in Figure 4), which determines if this UE will obtain better throughput in another cell. For this, TS xApp request a prediction throughput in the neighboring cells through the quality of prediction (QP) xApp (arrow number 2 in Figure 4), which is based on AI/ML and predicts the throughput in the different cells and returns this information to the TS xApp (arrow number 3 in Figure 4). Finally, TS xApp, based on the predicted throughput and the A1 policy, determines if it is necessary to initiate a handover of the UE. That is, if there is one neighboring cell whose predicted throughput is higher than the current throughput, in the percentage established for the A1 policy, TS xApp will send a handover request.

These three xApps consume the data from two databases, named cell metrics and UE metrics, through queries (Q). However, in a real scenario, these data are sent in real time by the gNB to the RIC. In addition, it is possible to add some entries in the A1 policy, such as UEid, latency, bandwidth, etc., to improve the network and have more parameters to adjust and optimize the RAN based on other key performance indicators (KPI) collected by the gNB.

Co-funded by the Horizon 2020 Framework Programme of the European Union

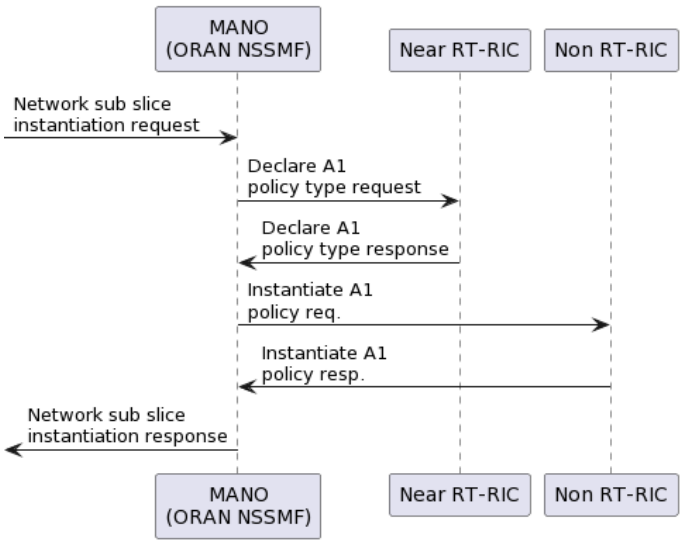


Figure 5: Sequence diagram of the interaction between MANO (O-RAN NSSMF) and O-RAN

Figure 5 presents a sequence diagram where the high-level interactions between MANO, particularly with the O-RAN network slice subnet management function (NSSMF) component, and O-RAN described as part of the end-to-end network slice provisioning process. The sub-slice belonging to the access part is provisioned as follows:

- If A1 policy is not defined, a request for declaring a new A1 policy type is sent to the near-RT RIC of O-RAN. The policy type contains the data model of the policy itself. For example, an A1 policy declaration contains the data structure of a QoS policy and its scope (values of each parameter of the policy), for providing a certain QoS of a given group of UEs.
- Then, the A1 policy is instantiated, meaning that the declared policy is applied. Using the above example, this means to numerically specify the QoS for a given set of UEs, specifying their international mobile subscriber identify (IMSI).
- Finally, the response of the A1 policy instantiation request is sent back for notifying MANO about the instantiation itself.

Hence, MANO manages all the declared and instantiated policies using the provided O-RAN A1 interface.

Open RAN architecture will be useful in a new communications paradigm, where there will be many different scenarios and many devices with different performance requirements, such as in the iNGENIOUS architecture. For this reason, the xApps development will help build more innovative components, use cases, and optimize the RAN in the upcoming years, building a flexible RAN and assigning adequate resources.

2.1.3 Exploitation Summary

Component Group: Flexible PHY/MAC and AI/ML for RAN/O-RAN

Co-funded by the Horizon 2020
Framework Programme of the European Union

- **TUD:** The Flexible PHY/MAC implementation in software defined radio (SDR) and its integration with 5G core and MANO allows for a unique testing platform where the interaction and coexistence of non-3GPP radio access technologies can be investigated, and further developed in the context of heterogeneous wireless IoT networks.

- **UPV:** The purpose O-RAN deployment is to analyze and learn about this new paradigm in RAN. That is, O-RAN deployment will be helpful in the research institute to open a new research area where academics and students can develop new applications (xApps), networks, or testbeds for many scenarios.

2.2 Next Generation RAN IoT

2.2.1 Smart IoT Gateway

The Smart IoT Gateway (GW) is the system element responsible for the appropriate routing and sorting of sensor data, coming from one or more sensor networks to higher layer data consolidation services and machine-to-machine (M2M) platforms. For performing these operations, the Smart IoT GW is able to interconnect multiple physical interfaces, as well as extracting and transforming messages as data traverses from one side to the other.

The Smart IoT GW exposes several physical and data-link interfaces to receive sensor data. Sensors can send messages to the Smart IoT GW either wirelessly (with technologies such as IEEE 802.11, LoRa, or Sigfox), or directly connected to the device (via Ethernet, I2C, or SPI). The Smart IoT GW will manage the routing and direct the received messages to the right output interface in a timely manner.

The Smart IoT Gateway ensures the connectivity for a vast number of heterogeneous IoT devices, by harmonizing different IoT technologies and application protocols and formatting data to be transferred across the network, terrestrial or satellite. The IoT interoperability enables the federation of different IoT platforms within heterogeneous domains, overcoming the compatibility issues between both standard and non-standard, proprietary and custom M2M solutions. The Smart IoT GW allows for improved optimization and prediction of the IoT data, which provides increased business value. Several technical tests, which are described in deliverable D6.1 [5], have been realized to verify the performance of the Smart IoT GW. Furthermore, a mid-term demo and a final demo were conducted at the port of Valencia during April 2022 and November 2022, where the performance of the Smart IoT GW was validated. Figure 6 and Figure 7 illustrate the performance of the Smart IoT GW through a custom dashboard. More information about the results can be found in the D6.3: “Final Evaluation and Validation”.

The Smart IoT GW Overview Dashboard in Figure 6 displays a histogram and the summary of the target routes selected for the messages sent by all related sensors. Furthermore, the latest round-trip time (RTT) ping values for the configured routes and general information about the gateway hardware is presented. The Container Overview dashboard in Figure 7 displays similar information regarding the message histogram as the previous Gateway Overview, however, only for the messages emitted by the specified

Co-funded by the Horizon 2020 Framework Programme of the European Union

container/sensor device. Furthermore, is depicts the measured temperature and humidity of the container as well as the battery life of the sensors.

Figure 6: Smart IoT GW Overview – Dashboard

Figure 7: Container Overview – Dashboard.

In Figure 2, the interaction of the Smart IoT GW in the iNGENIOUS architecture is also illustrated. Particularly, the Smart IoT GW interfaces with the sensors as it gathers and processes the data from the heterogeneous IoT devices. Then, the connectivity with the IoT cloud/Data center is obtained through satellite or terrestrial access network. Furthermore, the Smart IoT GW interfaces with the M2M platform. The primary use case is the communication between the Smart IoT GW and the SES Cloud Server via different types of

Co-funded by the Horizon 2020 Framework Programme of the European Union

physical network routes. The secondary use-case is the integration with external third-party systems, in which external software systems request data from the SES Cloud Server through standard application interface.

2.2.2 Satellite Connectivity

The role of satellite in the RAN IoT can be broadly considered in two areas: satellite backhaul and direct access to devices over satellite.

Satellite backhaul use cases are where the satellite network is used to backhaul the IoT traffic from IoT devices connected to a satellite edge node. This was illustrated earlier in the architecture overview provided in Figure 2. Satellite direct access is where IoT devices are connected directly to the satellite network. iNGENIOUS considers both cases to various degrees. From a demonstration and use case perspective, iNGENIOUS focused on the geostationary satellite constellations which is primarily used for satellite backhaul of IoT traffic, and therefore the use cases requirements, research and solutions are focused on the satellite backhaul configuration.

Satellite backhaul

In the Satellite backhaul setup, the IoT devices connect to an IoT gateway which uses the satellite backhaul to communicate with the core network. The IoT gateway optimizes the data before forwarding to the IoT cloud/data center via the satellite network as shown in Figure 8.

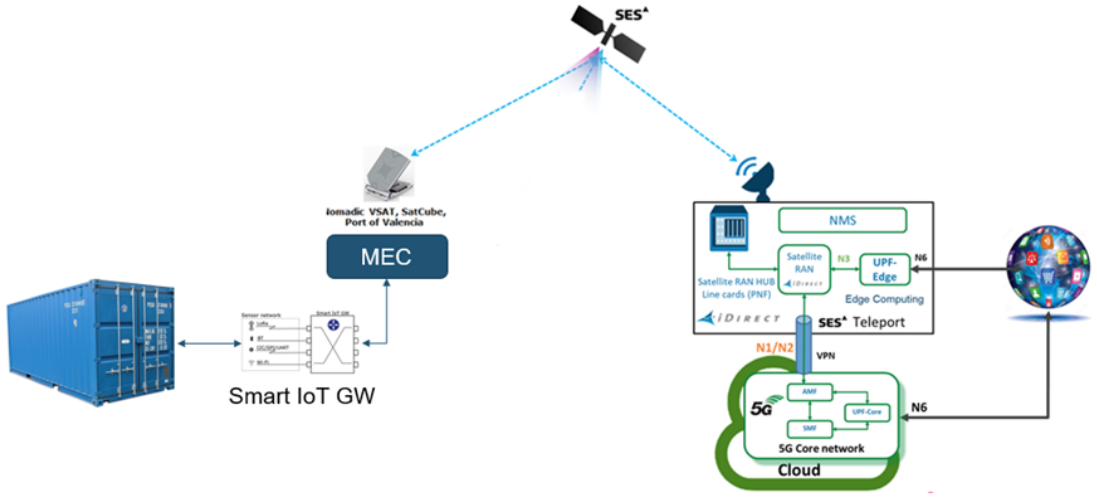


Figure 8: Satellite backhaul connectivity.

As illustrated in Figure 8, the satellite architecture also incorporates major concepts and components of the 5G architecture to provide the end-to-end satellite connectivity. The satellite architecture with 5G elements is explained in more detail in D4.2 (section 3.3) [3].

Satellite direct access

Co-funded by the Horizon 2020 Framework Programme of the European Union

Satellite direct access for IoT devices is where IoT devices connect directly to the satellite network, which in turn connects to the IoT cloud/data center. This allows delivery of the IoT content in a more efficient and cost-effective manner by utilizing a Direct-to-Satellite approach rather than using the satellite to backhaul IoT traffic.

Direct access of IoT devices over satellite can be categorized in the following three areas.

- **Non-3GPP IoT access** - Direct access of IoT devices over satellite using proprietary non-3GPP access technologies is already supported by many industry partners today. Within iNGENIOUS, iDR researched the use of their own proprietary access technologies to determine if they were suitable for connecting IoT devices over satellite.
- **3GPP 5G NR-NTN -** 5G New Radio Non-Terrestrial Network support is a new feature added in 3GPP Release 17. This offers the capability to use a standard 5G NR waveform over satellite links. This could offer new opportunities for both the direct access and satellite backhaul use cases.
- **3GPP NB-IoT NTN** - 3GPP have also made changes to NB-IoT to support NTN. These changes were studied in Release 16 and included in Release 17. In general, the changes are the same or similar to the changes outlined for 5G NR-NTN but tailored for NB-IoT.

There was no use cases requirement for direct access over satellite solutions within the iNGENIOUS project, however, research was carried out on all three areas mentioned above. Further details of the research on 3GPP 5G NR-NTN and 3GPP NB-IoT NTN are included in D3.2 [6] (section 3) and details of Non-3GPP IoT access (direct access) testing can be found in the D6.3 [7].

2.2.3 5G Modem

5G modems are devices that are used to connect specific vertical components, such as robots, sensors, cameras, or AGVs to the 5G network. 5G modems allow devices to be connected via Ethernet and to communicate with the 5G network wirelessly. In Figure 9 the integration of the 5G modem with ASTI's AGV is shown. This setup is used in the final Factory use case demo.

The modems provided by Fivecomm are a compact and versatile solution that permit, for example, to attach it to a robot or integrate into an AGV easily thanks to its reduced size. It is also very configurable via software, since it is possible to modify options normally not accessible in other commercial devices, such as services installation inside the modem or multiple VLAN management, making the modem more flexible for integration with different applications and adapting to their specific needs.

Within the context of iNGENIOUS, 5G modems are a part of the complete end-to-end system architecture. 5G modems have already been integrated and validated in several testbed facilities, including Cumucore premises and the UPV campus. Full results from the validation tests can be found in deliverable D4.2 [3], where latencies from 10 to 20 ms were obtained, and throughputs in the range of 100 to 500 Mbps depending on the testbed, enabling an effective communication for the iNGENIOUS use cases.

Co-funded by the Horizon 2020 Framework Programme of the European Union

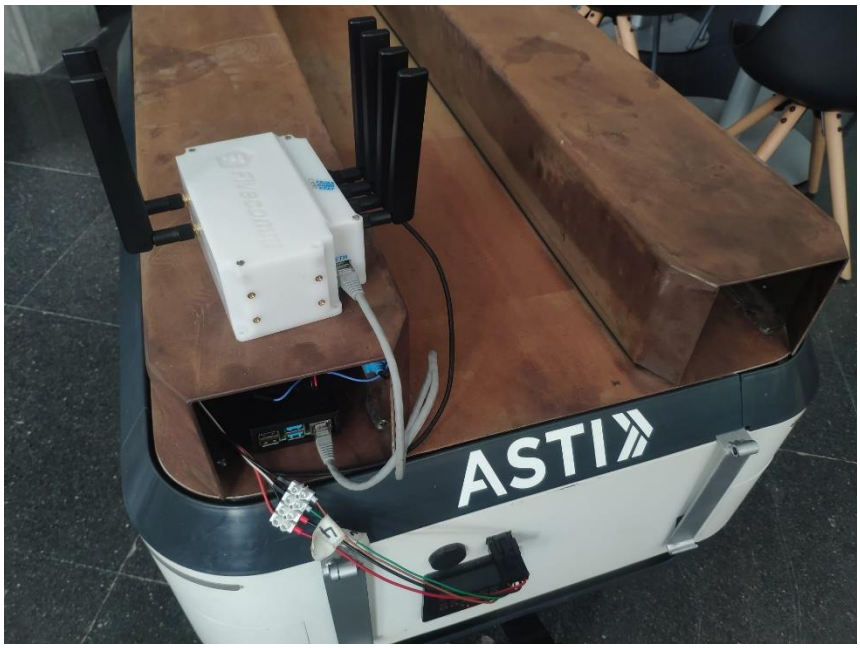


Figure 9: 5G modem integrated with ASTI's AGV. This setup has been employed in the Factory use case.

2.2.4 Business Potential and Exploitation Summary

Component Group: Smart IoT Gateway, Satellite Connectivity and 5G Modem

- **SES**: The Smart IoT GW can be used in several commercially attractive Massive IoT use cases and for multiple market verticals, such as Fixed Data, Aero, Maritime, Energy, Government, Agriculture, Cloud, and Video.

  For example, the Smart IoT Gateway can play an important role in solutions of disaster response where we need to re-establish communication (Internet, phone), to support the coordination efforts of humanitarian organizations in the field and to contribute to saving lives during humanitarian emergencies. The Smart IoT GW can help in the asset monitoring for equipment/sites. The edge-computing capabilities of the Smart IoT GW can be used to monitor the deployed equipment and can provide an alerting interface for the on-site personnel by sending notifications to the smartphones or other smart devices, in case there are issues with monitored devices.

  Furthermore, the Smart IoT GW can provide useful services to the agricultural sector. For example, Vines are especially exposed to the effects of climate change and to the spread of diseases. Providing the winegrowers with near-real time information of their vineyards will offer them the possibilities to adapt their management according to their needs. The availability of this information allows them for a more time and cost-efficient planning. The data from the sensors deployed in the vineyards can be processed by the Smart IoT Gateway that supports several IoT communications technologies and protocols, and intelligently routes the data to satellite or terrestrial backend networks.

- **iDR**: At the beginning of the iNGENIOUS project iDR aimed to research and develop use cases for indirect and direct access to IoT devices over a satellite network. Working closely with SES during the project while integrating their Smart IoT Gateway, has given iDR valuable information and requirements on

Co-funded by the Horizon 2020 Framework Programme of the European Union

how to enhance our existing satellite backhaul product offering for IoT networks. iDR will continue to work closely with SES on this after the project finishes.

Direct access for IoT over satellite is another area of research undertaken by iDR in iNGENIOUS and iDR expects this will also be an area of growth for satellite and IoT integration. The collaboration with other partners in the project will lead to future collaboration and closer integration of satellite and IoT. iDR identifies two categories of direct access or IoT devices over satellite namely, proprietary and 3GPP standard NB-IoT NTN based access. iDR have already invested in proprietary IoT access mechanisms and will continue to support and enhance our offerings in this area based on the learnings from the iNGENIOUS project research. Regarding 3GPP standard NB-IoT NTN based access, even during the lifetime of this project major strides forward have been made by 3GPP in this area. 3GPP Release 17 included NB-IoT NTN support and iDR plan to continue the research on NB-IoT NTN which started on this project.

- **5CMM**: The 5G modem is an innovative device for enabling NR technology use cases. It can be used in many verticals, having a high business potential. Its innovations such as the configurability and the reduced size make it a great option in the market. The exploitation roadmap focuses on providing the modem as a product for both investigation and commercial purposes.

2.3 Relation to Use Cases

The flexible PHY/MAC is employed as the RAN technology in the Factory use case. It wirelessly connects three UEs which required different requirements in terms of data rate and latency. The demonstration setup presented in the mid-term review is illustrated by Figure 10.

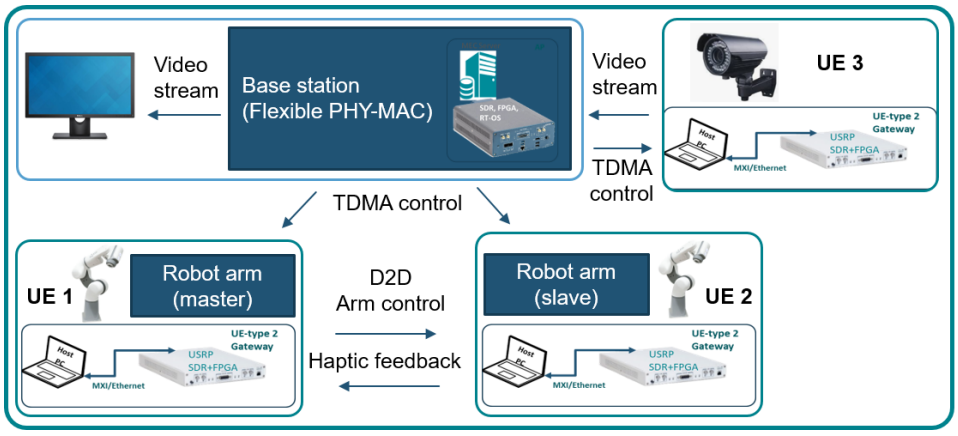


Figure 10: Factory use case mid-term demonstration at TUD's testbed. It highlights the Flexible PHY/MAC with multiple users.

The O-RAN components, RIC and xApps deployed by the UPV, can manage a network with different performance requirements on each UE, and improve handover, based on AI/ML applications, to improve the efficiency of the networks and QoS and QoE of the UEs. Therefore, this type of architecture can be helpful in the Factory use case with automated robots within heterogeneous networks, where AGVs and devices have different throughput requirements. For example, in a big factory with some gNBs, several AGVs are

Co-funded by the Horizon 2020 Framework Programme of the European Union

moving across the factory, and the signal received is not the same at all points, therefore, the throughput will be less than the requirements in some time slots. With this type of architecture, it is possible to define a threshold in which the AGV improves a handover to another neighboring cell with better throughput. Also, with an Open RAN architecture, it is possible to develop third party applications (such as beamforming or DSS (Dynamic Shared Spectrum) optimization, among others), that the client needs, opening up the network to a more interoperable and flexible architecture.

The Smart IoT GW is used in the Ship Use Case. This use case aims at providing E2E asset tracking via satellite backhaul from the IoT RAN to the corresponding data/control center, enabling real-time/periodic monitoring of predetermined parameters of shipping containers when they are sailing on the sea, while terrestrial IoT connectivity is provided when the ship approaches at the port. The shipping container is equipped with a certain number of heterogeneous IoT devices able to monitor the internal environment of the container (accelerometer, temperature, humidity) as well as to detect critical events (physical shocks, door opening). During the trip, the heterogeneous IoT devices send regular status updates and the Smart IoT GW gathers and processes the data and the connectivity with the IoT cloud/Data center is obtained through satellite or terrestrial access network.

For both the Transport and Ship use cases, the satellite network provides the backhaul connection between the edge node and cloud to ensure the edge node remains connected to the network even when there is no terrestrial or other means of connecting the edge node to the cloud. A good example of this is the real time monitor of the shipping container while at sea with no access to a terrestrial network. The satellite connectivity means the operator can continuously monitor the edge node and connected IoT devices which allows for active intervention on board or once the container reaches the shore.

The 5G modem has been integrated in several cases in iNGENIOUS including robots and AGVs, and both commercial and private networks have been tested. This component is used in the "*Automated Robots with Heterogeneous Networks*" use case for integrating with ASTI's AGV and controlling it in the context of the use case. In the final demo, the 5G modem will connect the AGV with a 5G private network deployed in ASTI's premises in Burgos. It is also used in *"Improve Driver's Safety with MR and Haptic Solutions"* use case, integrating it in ASTI's AGV, other robots and the cockpit where the haptic gloves are located.

Co-funded by the Horizon 2020
Framework Programme of the European Union

3 Smart IoT Core Network

The mobile core consists of all the network modules required to provide connectivity to end devices through the RAN. Thus, the core network provides the authentication and authorization functionalities based on the subscriber identity module (SIM) cards, but it also provides best effort connectivity to the mobile devices. The 5G core network has been extended with new features, such as 5G local area network (LAN) and network slicing. These new ingredients will facilitate the employment of 5G technology in industrial mobile infrastructure. The iNGENIOUS partners have designed and prototyped the 5GLAN functionality to demonstrate the integration of industrial IoT devices for both wired and wireless scenarios.

This section describes the functionality implemented in Cumucore 5G Core to support 3GPP specifications for deployment of 5G as part of IoT networks such as 5GLAN, network slicing and geographic anomaly detection.

3.1 5G Local Area Network

The 5G LAN-type service defined in 5G as part of the core network consists of a service over the 5G system offering private communication using IP and/or non-IP type communications. As part of the 5G LAN service it is possible to create a 5G virtual private network (VPN) capable of supporting 5G LAN-type service.

With access to the 5G LAN service, an IoT device or UE can communicate with any other device that is a member of the 5G VPN from any location where 5G service is available. Figure 11 presents an architecture interconnecting mobile devices with local area network (LAN) Data Networks (DN) with fixed devices.

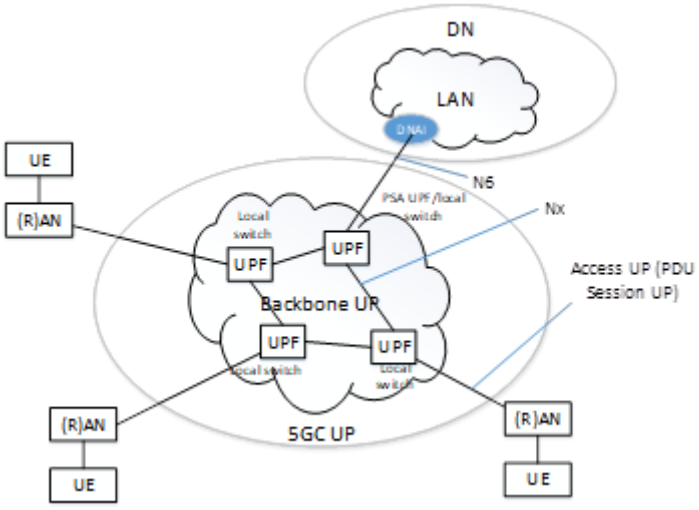


Figure 11: The user plane architecture of 5G-LAN communication framework (Ref. 3GPP TR 23.734) [8].

Thus, the 5G LAN-type service shall provide a mechanism for an authorized 5G VPN administrator to enable or disable a UE from accessing the 5G VPN, but also it is possible to remove a UE from a specific group of UEs of a private group communication. It is necessary that a 5G LAN-type service provide a mechanism to identify an authorized UE. Therefore, if UE-1 wants to establish

Co-funded by the Horizon 2020 Framework Programme of the European Union

private data communications with UE-2, it sends a request to the 3GPP network for an on-demand private data communication connection to UE-2. UE-1 can also indicate what type of data communication it wants (e.g. IP, Ethernet, or other). Consequently, the 5G network shall support the routing of non-IP packet (e.g., Ethernet frame) efficiently for private communication between UEs. Nevertheless, in this case UE1 is part of VPN-1 and cannot communicate with another UE that belongs to a different VPN. Thus, the 5G network shall enable the network operator to ensure UEs that belong to a different private group cannot send data to any or all of the UEs in a specific group.

Since each group of devices that are part of the 5G LAN service would have different requirements, the 5G network shall be able to provide the required quality of service (QoS) (e.g., reliability, latency, and bandwidth) for non-IP packet (e.g. Ethernet frame) for private communication between UEs. Moreover, with 5G LAN service the core network shall enable the network operator to support point-to-point addressing as well as multicast addressing between the different UEs in a private group. It is assumed that all UEs in a same private group use the same type of addresses (e.g. IP, Ethernet or other). The deployment of 5G LAN service should be easy to use and the core shall enable the network operator to create, manage, and remove private groups including their related functionalities (subscription data, routing and addressing functionality).

One of the key innovations of the 5G core that was not available in earlier releases of mobile networks is the support for direct communication based on Ethernet. Thus, the 3GPP 5GLAN support the fast routing, broadcast, support of virtual LANs, and Ethernet QoS classification. Ethernet frames are transported by the 5G network and routed to the correct destination 5G UE before being unpacked and forwarded to the correct Ethernet switch/device. The network must support the routing functionality based on Ethernet frame header information. The core should support Ethernet broadcast frames. Routing of Ethernet frames in the 5G system must be based on the outcome of spanning tree algorithm run by the Ethernet network being served. Moreover, as native support for Ethernet based communications, the Ethernet traffic flow classification must be based on Ethernet headers – Source and Destination MAC address, EtherType (including multiple EtherTypes in double-tagging), VLAN tags including VLAN ID and PCP, in addition to the existing fields used in Traffic Flow Templates (TFT). Thus, packet filtering and choice of 5QI should be based on Ethernet header information. The Ethernet transport service shall support traffic filtering and prioritization based on source and destination MAC addresses. The Ethernet transport service shall support traffic filtering and prioritization based on EtherType (including multiple EtherTypes in double-tagging)

In LAN networks, devices make use of discovery mechanism (e.g Bonjour, UPNP) to discover other devices online to be used and their characteristics. This discovery mechanism makes use of the multicast capabilities of the network. Therefore, it is important that 5G LAN support discovery mechanisms. Thus, on-demand establishment of a multicast communications within subset of UEs that are members of the 5G VPN, e.g. equipment A create a multicast on demand and B and C joins this multicast to receive A's multicast messages.

Co-funded by the Horizon 2020 Framework Programme of the European Union

As summary, connecting mobile with fixed devices natively have some requirements listed in this section which 3GPP has been addressing as part of the 5G-LAN service. iNGENIOUS project brings a prototype of the 5GLAN as part of Cumucore 5G core to bring a totally new communications feature to 5G where devices can communicate with each other natively over Ethernet and they can be part of secure private group.

3.1.1 Business Potential and Exploitation Summary

Component Group: 5Gcore – 5G LAN
• **CMC**: The 5G LAN allows the creation of a private group, i.e., a VPN, of devices that can communicate together as if they were physically connected. The mobile core with 5G LAN feature can be connected to a private LAN and the mobile devices will be able to connect to the devices connected to the LAN. CMC is differentiating from major mobile vendors by focusing in non-public networks (NPN) targeting private industrial wireless networks. Cumucore business potential consists in delivering innovative highly competitive product that consist of 5G core for industrial usage. Thus, Cumucore includes in the product features such as 5GLAN, TSN, and network slicing, which are strong requirements in industrial networks compared to consumer public networks**.**

3.2 Network Slicing

Network slicing has been defined by 3GPP to split physical resources for different services or applications. Network slicing transforms a shared physical infrastructure into a set of logical end-to-end networks each one designed with different requirements to be used by different business purposes. Network slicing is proposed to split network resources at RAN, transport and core network levels, and offer them to different customers that want to have their own dedicated network that attain specific requirements. Thus, in principle network slicing would make sense in public mobile networks that can provide service to enterprises, or verticals, that want to have their own slice of the public network.

In private networks the customer already has its own dedicated network resources where RAN, transport and core are not shared with other public users. Thus, in principle private networks do not need network slicing. However, the industrial or enterprise deployments of private networks require network slicing not only for offering network resources to other customers but to separate the resources and assign those to different group of devices. In industrial networks the variety of devices is larger than public networks with consumer devices. In industrial networks the devices are mostly machine-type, such as robots, video cameras, sensors, actuators, but also mobile devices. Moreover, some of these devices have different communications requirements. For example, remote-controlled drilling machines needs very low latency communications, while on the other hand, video camera needs higher bandwidth to exchange 4k video to the control room. This means that devices would have to be grouped based on the requirements and they need

Co-funded by the Horizon 2020 Framework Programme of the European Union

to be allocated to different resources in the RAN, transport and core network, which results in a different network slices within the private network.

Figure 12 illustrates at high level how network slicing provided by the MANO component can be used to separate, e.g., IoT traffic from consumer data, where each slice attains a different set of requirements. Cumucore has contributed to the iNGENIOUS project with 5G core that includes the network functions required for creating network slices. Moreover, the 5G core provides the interfaces such that MANO can be used to automatically create and efficiently orchestrate the network slice.

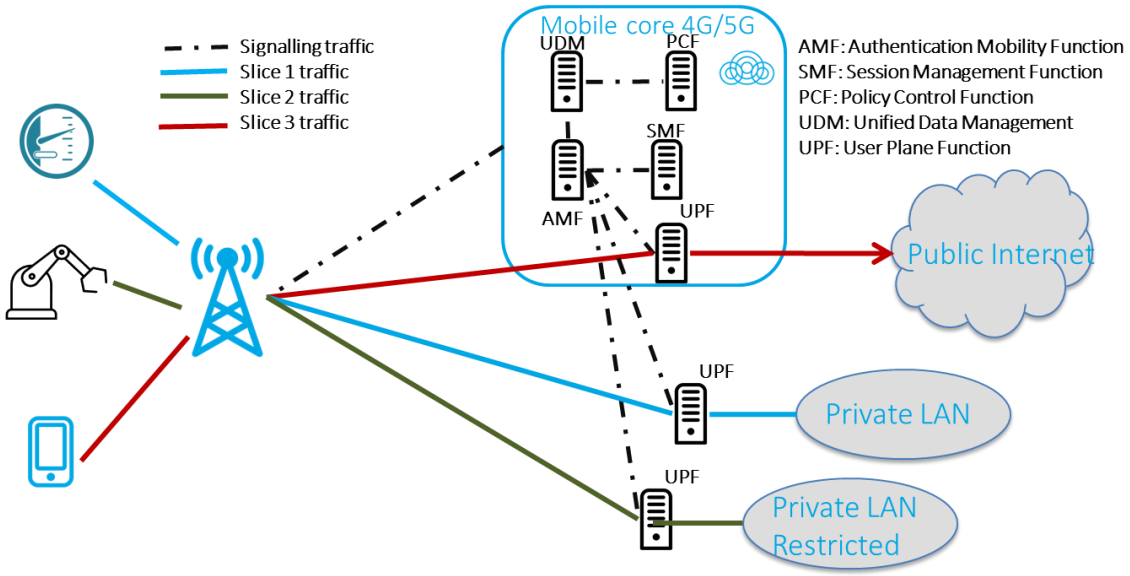


Figure 12: Illustration of traffic discrimination obtained with MANO integration.

3.2.1 Business Potential and Exploitation Summary

Component Group: 5Gcore – Network slicing

- **CMC**: Network slicing can be used in different contexts, and it can have different business models depending whether the network slicing is used in public or private mobile networks. In private networks, one model of network slicing consists of the usage to separate devices according to traffic requirements but also from security standpoint. In this scenario the network slices Is used by private mobile networks separate traffic from the devices accessing different Data Networks (DN). Each DN consist of fixed networks with some access restrictions. Another model in public networks consists of the usage of networks slices that could be leased to private customers or to other mobile operators that do not have sufficient coverage.

  A similar model within the context of private mobile network is to utilize the slices for separating services or applications. Thus, the traffic from the drilling machine would be routed to a separate data network where the servers controlling the drilling process are running. This traffic would be completely isolated from other to ensure high reliability and access only to that data network. Instead, the traffic from the video cameras would be connected to

Co-funded by the Horizon 2020 Framework Programme of the European Union

another data network where the video processing applications are running, and end users can visualize the video streams.

A typical smart factory floor consists of a large complex system and not every component of such a complex system will require the same level of security. Such differences are addressed by applying the concept of the security zone. According to ISA/IEC ISA2443 [9], each security zone defines a logical grouping of physical, information or application assets sharing a common security requirement [10]. Security conduits according to IEC 62443 [11] is defined as "Conduits are the special type of security zone that groups communications that can be logically organized into a grouping of information flows within and also external to a zone".

Therefore, network slicing can be exploited in different ways depending whether the slices are part of public network or private mobile network.

3.3 Geographical Anomaly Detection

The Geographical Anomaly Detection (GAD) functionality aims to automatically identify behaviors which deviate from some standard defined patterns, and which might therefore represent malicious activity by truck drivers or potential cyber-attacks able to corrupt the managed data position. This identification is carried out by analyzing the path followed by IoT devices mounted on trucks, which are moving inside the target area.

After a comparison of different algorithms and techniques, the density-based algorithms DBSCAN method was selected because it confirmed to be the most reliable approach in this scenario, where arbitrary shaped clusters with noise (outliers) had to be detected.

Based on the architectural definition activities available in D4.3 [12], the GAD component was integrated in the data flow pipeline of the Port Entrance use case.

The developed component is able to operate in a soft real time validation data flow chain. The real-time capability depends on the frequency of IoT data transmission (10 seconds).

In particular, the developed functionality periodically performs the following tasks with a configurable time-period within a range depending on several constrains (transmission frequency, computing power, and admitted latency):

- Receiving positioning data from IoT devices;
- Classification of the acquired dataset;
- Generation of classification results.
- In case of anomalies, generation of graphical results highlighting anomalous path sections (see Figure 13), where green dots identify valid path section, while red dots refer to invalid paths that were detected.

Co-funded by the Horizon 2020 Framework Programme of the European Union

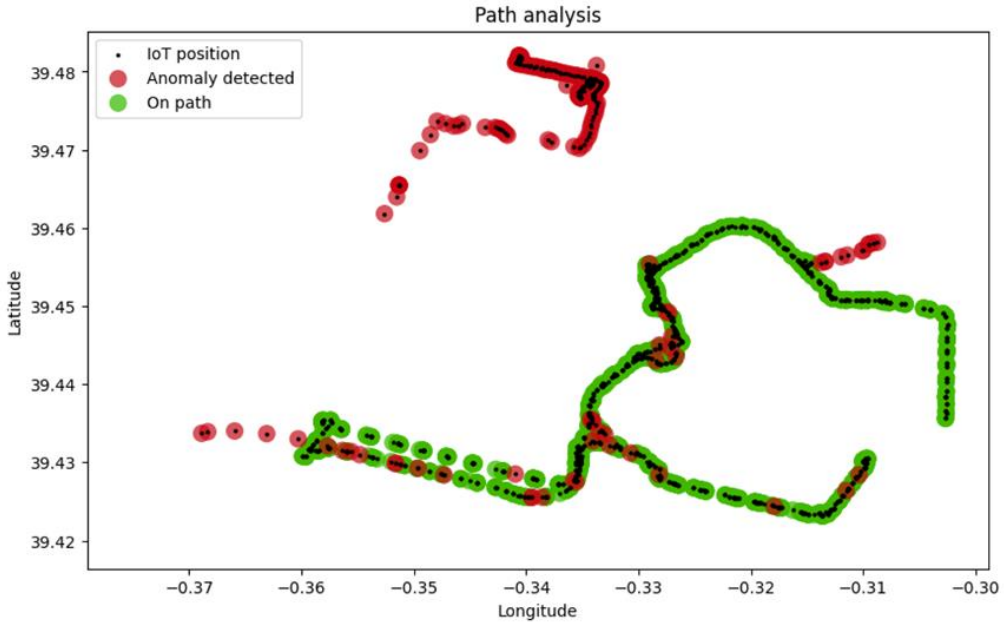


Figure 13: GAD graphical result.

An initial learning phase had to be executed using the admitted paths inside the target area. The data related to the admitted paths was generated through sampling of several routes followed by real trucks inside the target area.

Python v3.10 has been used to develop the GAD functionality. In particular, the Scikit-learn library [13] has been used for the implementation of the DBSCAN algorithm. Additional details are in deliverable D4.3 [12].

For security reasons (handling of private data, such as truck-id) the GAD module is integrated in the UPV Server (Smart Platform Services), as shown in Figure 14.

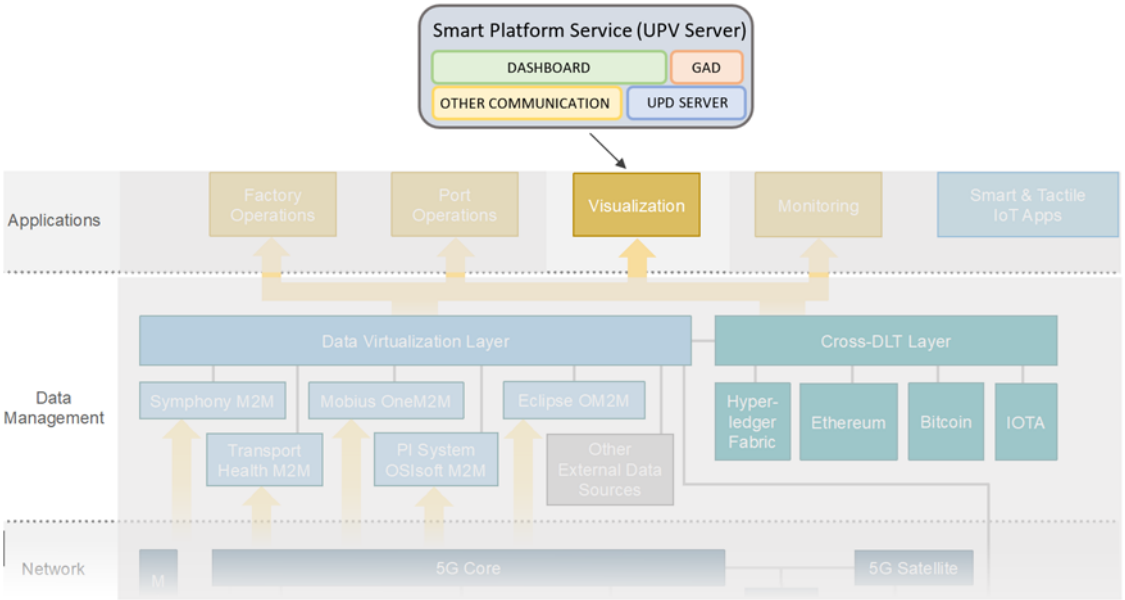


Figure 14: GAD integration in UPV Server.

Co-funded by the Horizon 2020 Framework Programme of the European Union

3.3.1 Business Potential and Exploitation Summary

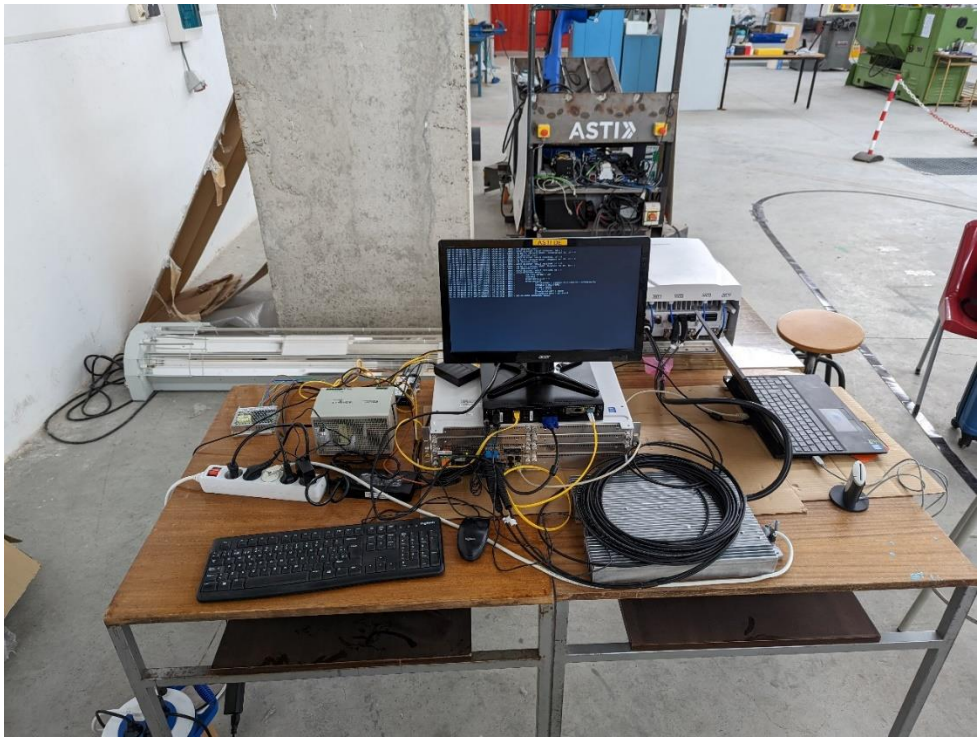
Component Group: GAD

- **TEI**: The GAD functionality has been designed with the aim to be easily adapted to all those end-user applications or services where real time tracking of moving devices is required. The machine learning module used in the anomaly detection phase can be used for several application scenarios. The training phase on specific admitted routes, and the tuning of the centroids dimension can be performed accordingly to the target scenario characteristics. In this way, the anomaly detection can be performed on every kind of conveyance, e.g., ships, trucks, planes, that are involved in the end-user supply chain.

  TEI plans to investigate how GAD can become part of future 5G services, also considering additional use cases that could benefit from this technology. For example, during a natural disaster, or terrorist attack, that is involving a specific geographical area, people connected to 5G network can receive an alarm message if they are to run across dangerous escape paths.

3.4Relation to the Use Cases

The 5GLAN and network slicing has been deployed in ASTI for creating seamless connectivity between fixed device that will connect with 5G modem integrated with an AGV. The objective is to control the AGV from the fixed device. Thus, network slicing was first used to create a slice named ASTI that is used to define the profile for all the devices that will connect to that slice. Furthermore, 5GLAN group was built using the ASTI slice to connect the fixed device in the LAN with the 5G modem in the AGV.



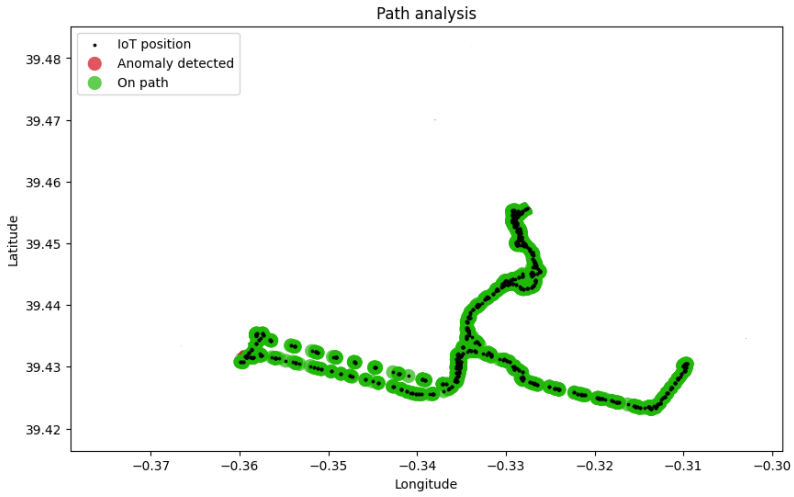
Co-funded by the Horizon 2020
Framework Programme of the European Union

Figure 15: Private industrial network with 5GLAN and network slicing.

The system was deployed and is operational for managing the AGV from the fixed device connected to the LAN.

Figure 16: AGV with 5G modem to be controlled from device connected to fixed LAN.

The GAD was used in the Port Entrance use case, where the GAD processed the tracking data coming from IoT device mounted on trucks entering Valencia port, detecting any geographic anomalies in real time, according to the availability of IoT data. The GAD produces and updates, as output, an image showing the evolving of the truck's path in the port of Valencia, using green dots as can be seen in Figure 17.



Co-funded by the Horizon 2020 Framework Programme of the European Union

Figure 17: example of GAD output in case of truck on path.

In case an anomaly is detected, the suffix "_ALARM" is added to the image name and the anomalies are highlighted via red dots as demonstrated in Figure 18.

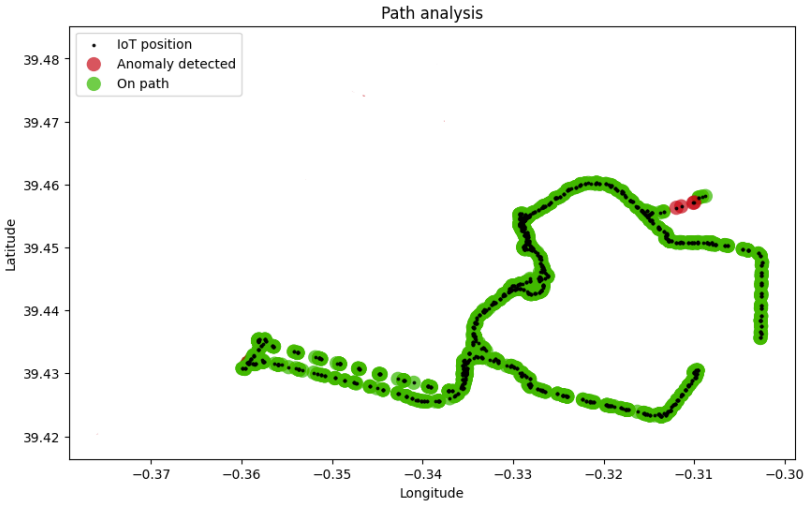


Figure 18: example of GAD output in case of anomalies detected.

The GAD's output is automatically taken in charge by the Smart Platform Services and showed on the dashboard in Figure 19.

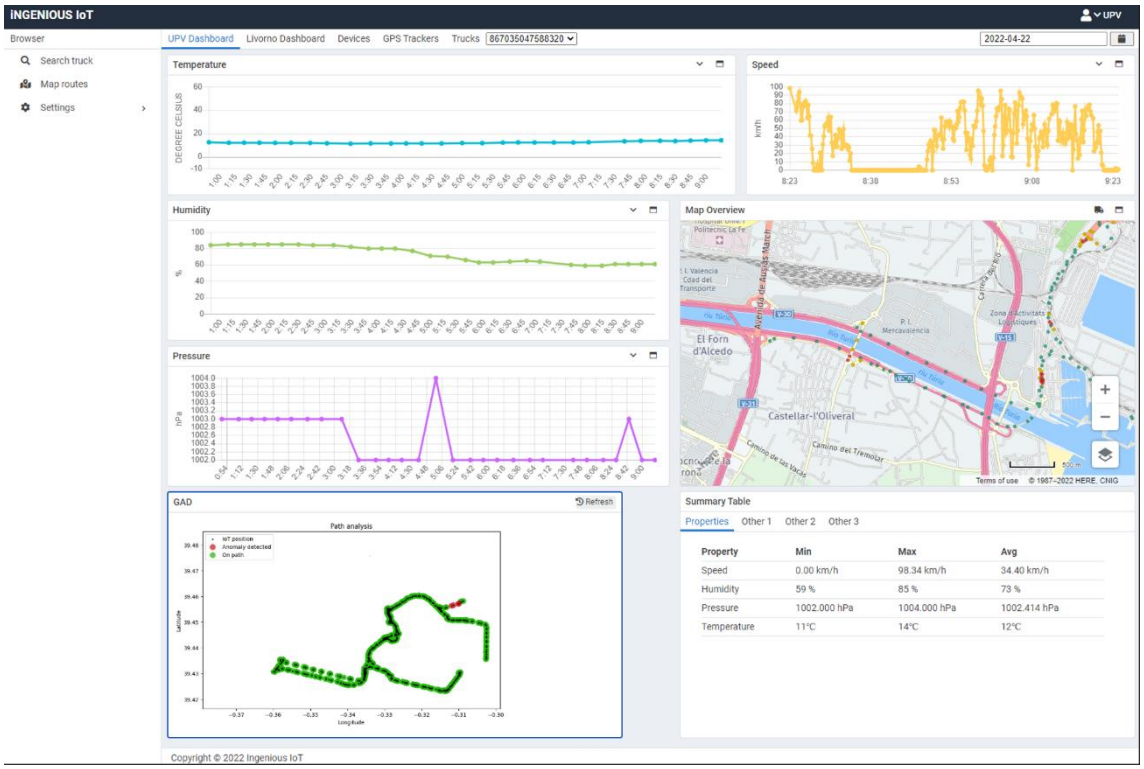


Figure 19: GAD output integrated in the Smart Platform Services' dashboard.

Co-funded by the Horizon 2020 Framework Programme of the European Union

4 Next Generation IoT Network Slice Orchestration

Within the Smart end-to-end INGENIOUS IoT System, a crucial role is played by the end-to-end network slice orchestration framework, which in some other deliverables has been identified with the terms Management and Orchestration (MANO) and cross-layer MANO. Substantially, the end-to-end network slice orchestration framework aims at providing automatic end-to-end network slicing capabilities to the vertical entities, regardless of the technologies implemented at the different layers of the network architecture.

This section concisely describes the state of the art of the orchestration system implemented and its functionalities are reported, focusing on the innovations and benefits brought to the iNGENIOUS architecture, and on the related business potentials that can be explored after the project lifetime. Finally, this section describes the relation between the NG-IoT network slice orchestration and the use cases within the project.

4.1 State-of-the-art and Reference Architectures

This section briefly describes the State of the Art and Reference architecture for the end-to-end network slice orchestration framework. In particular, after a brief description of this architecture, it is explained how its limitations are overcome. For further details about the state of the Art, reference the architecture of the NG-IOT Network Slice Orchestration and the innovations, refer to the corresponding specifications and to Deliverables D4.1 [2] and D4.4 [14], respectively.

The concept of network slicing is one of the pillars of the 5G network because offers a virtualized isolated network for each different service. This requires resource allocation and management in the end-to-end network where different network segments (radio, access, and core) could belong to different vendors. Moreover, in the domain of NG-IoT several UEs can be connected simultaneously, leading to an increased resource allocation demand. End-to-end network slice orchestration solutions come into place for addressing this kind of problem, orchestrating resources and (virtual) Network Functions (VNFs) for satisfying the high-level requirements of the services.

From a standardization perspective, 3GPP and ETSI have specified different standards in the domain of the management and orchestration of a 5G network. The 3GPP TS 28.500 specification [15] defines generalized mobile network management architecture, while the NFV-MANO architecture defines the MANO framework architecture [16]. The interaction between the 3GPP management system and the ETSI NFV MANO components is detailed by the 3GPP TS 28.528 specification [17], where the standard interfaces are detailed.

However, in the current network slice and service orchestration solutions some limits have been identified. The silo-based approach proposed implies lack of flexibilities in terms of management of the service and network slice,

Co-funded by the Horizon 2020
Framework Programme of the European Union

especially in the context of NG-IoT Time Sensitive Network (TSN) and Ultra Low Latency Requirements, where the network is continuously subject of changing on its status. Then, a full integration among the domain of 5G NR, NG-IoT and the edge computing is not reached yet. Moreover, the use of AI\ML technologies to fully support automation in the end-to-end network slice management, orchestration and operation is still not mature and enabling zero-touch network and service management. As already reported in the gap analysis study in D4.4 [14], the works related to AI\ML for network slice configuration make usage of simulated data, or the usage of specific datasets and only few of them collect data coming from the Network Functions.

4.2 Orchestration Functionalities

In this section, after a short explanation of the end-to-end network slice orchestration positioning in the iNGENIOUS architecture, the main high-level orchestration requirements are briefly described. Then, an explicit mapping between these requirements and the specific functionalities implemented by the end-to-end network slice orchestration framework is underlined. Moreover, early results as output of such implementation are reported. Finally, the implemented functionalities are reported in the use cases where the end-to-end network slice orchestration is involved in. For the main software technical details of the implementation of the end-to-end network slice orchestration framework refer to the deliverable D4.4 [14].

The end-to-end network slice orchestration component represents a crucial element within the network layer of the iNGENIOUS architecture depicted in Figure 20 (referred as MANO).

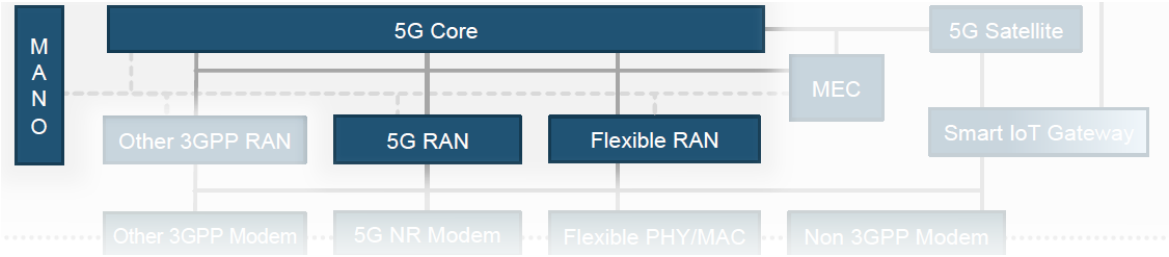


Figure 20: Network Layer of iNGENIOUS architecture.

For realizing its main functionalities, the end-to-end network slice orchestration interacts with other components belonging to the iNGENIOUS Network Layer. More details about these interactions are available in the deliverable D2.4 [1]. The purpose of these interactions is to perform the following high-level functionalities:

1. **Interoperability** with different interfaces, even not standardized.
2. **Resource abstraction** of different data models provided by different kinds of technologies and vendors.
3. **Modularity and flexibility** for supporting possible new technologies.
4. **Lifecycle management** of the end-to-end network slices.

Co-funded by the Horizon 2020 Framework Programme of the European Union

The abovementioned requirements are mapped into specific implemented functionalities of the end-to-end network slice orchestration framework within iNGENIOUS, whose high-level architecture is illustrated in Figure 21.

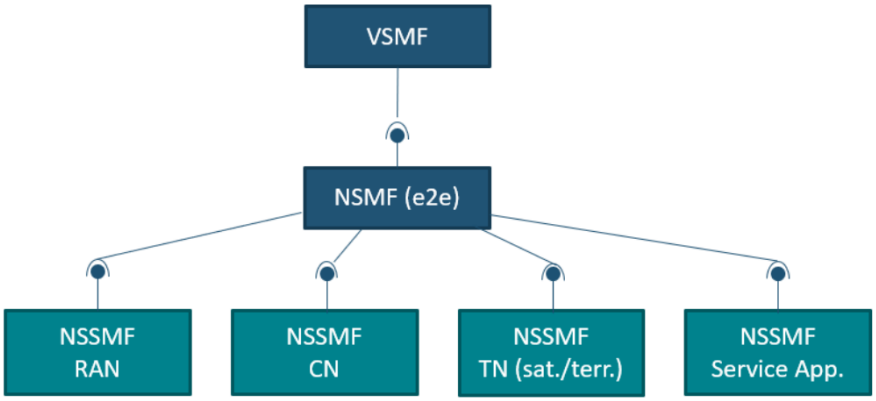


Figure 21: End-to-end network slice orchestration high-level architecture.

The end-to-end network slice orchestration high-level architecture is mainly composed of the Vertical Service Management Function (VSMF), the Network Slice Management Function (NSMF) and the Network Subnet Slice Management Functions (NSSMFs).

The VSMF is in charge of managing the end-to-end vertical services lifecycle. Relying on the high-level Information representing a vertical service, it interacts with the NSMF to request end-to-end network slice lifecycle management (LCM) operations. Then, each NSSMF is in charge of managing its own end-to-end network subnet slice, interacting with the corresponding network segment (RAN, Core Network (CN), Transport Network (TN), and so on), applying the needed configurations. Moreover, is possible to have an NSSMF for deploying the service applications.

Further technical details about the end-to-end network slice orchestration implementation are available in the deliverable D4.4 [14].

The four-high level functionalities earlier described, are mapped into the end-to-end network slice orchestration framework. In particular:

1. For what concerns **interoperability**, the end-to-end network slice orchestration can interact with the 5G core described in section 3 with (non-ETSI) mobile edge computing (MEC) and with different types of access networks, i.e., 5G RAN, Flexible RAN and other 3GPP RAN. From an operational point of view, this means having the end-to-end network slice orchestration implementing the protocol and logic for supporting the interactions with the aforementioned components. For example, the NSSMF RAN interacts with the RAN components, and the NSSMF CN with the Core Network.

2. The end-to-end network slice orchestration **abstracts** the information models at different levels. As has been described in deliverable D4.4 [14], the VSMF, NSMF, and NSSMF components are in charge of translating the high-level information model requirements in the specific interactions with the network segment of responsibility.

Co-funded by the Horizon 2020 Framework Programme of the European Union

3. **The modularity and the flexibility** for supporting different and possibly new technologies (standardized or not) have been implemented by using multiple NSSMFs. Each NSSMF is responsible of dealing with a specific component: NSSMF RAN implementing the logic for the interaction with the RAN component, NSSMF CN with the 5G core, and so on.

4. The end-to-end network slice orchestration implements the **end-to-end lifecycle management** functionality relying on the first three requirements: interoperability with different technologies, the model abstraction and modularity and flexibility. To this end, it is possible to automatically provision multiple end-to-end network slices on different network segments without the need of dealing with the technical details of each segment.

4.2.1 Software prototype and integration with other iNGENIOUS technologies

The design and implementation of the end-to-end network slice orchestration prototype have been produced and different integration activities with 5G core, O-RAN, and Flexible PHY/MAC have been carried out.

Specifically, the end-to-end network slice orchestration software prototype developed for iNGENIOUS follows the functional architecture design reported in D4.4 and highlighted in Figure 21, and includes:

- A VSMF prototype for the management of industrial IoT and supply chain vertical services
- An end-to-end NSMF prototype for the lifecycle management of network slices spanning RAN, edge and core network domains, In practice, it provides means and logic for automated provisioning of end-to-end network slices, including capabilities for runtime operation (i.e. logic for network slice scaling)
- An O-RAN NSSMF prototype (as reported in D4.4) for the enforcement of O-RAN Near Real-Time RIC configurations, with automated translation of RAN subnet slices performance requirements into O-RAN A1 policies
- An PHY/MAC RAN NSSMF prototype for the automated provisioning of RAN configuration on the Flexible PHY/MAC solution from TUD, with of RAN subnet slices performance requirements into PHY/MAC resource allocation
- A Core NSSMF for the management and runtime operation of the Cumucore 5GCore network slices, which supports automated deployment and configuration of the 5GCore NFs, as well as the dynamic provisioning of network slices following the Cumucore REST APIs. This NSSMF makes use of the ETSI OSM tool as NFV orchestrator
- A service application NSSMF for the management and runtime operation of application level virtual functions, which supports automated deployment and configuration of edge cloud-native applications (e.g. for the industrial IoT network slices and services). This

Co-funded by the Horizon 2020 Framework Programme of the European Union

NSSMF makes use of the ETSI OSM tool as NFV orchestrator to deploy VNFs on top of cloud-native virtual infrastructures.

Nextworks, as main contributor, is evaluating the opportunity to release the iNGENIOUS as opensource on top of its slicer tool. Available at: github.com/nextworks-it/slicer

Moreover, as part of the integration experiments mostly in the context of UC1, different interoperability tests (which will be fully reported as part of WP6 activities in D6.3) have been carried out between the end-to-end network slice orchestration framework and:

- 5G core for the LCM of core network slice, UEs and gNB configurations.
- O-RAN for the A1 policy management as reported in section 2.1.2.
- Flexible PHY/MAC for providing:
    - Slicing capability to a not-standard RAN.
    - 5G connectivity to real UEs relying on this RAN.

Figure 22 depicts a web graphical user interface (GUI) for the management of the end-to-end network slices. As part of task T4.3 activities, this web GUI has been designed and implemented to facilitate the network system administrator to manage the lifecycle of the end-to-end network slices.

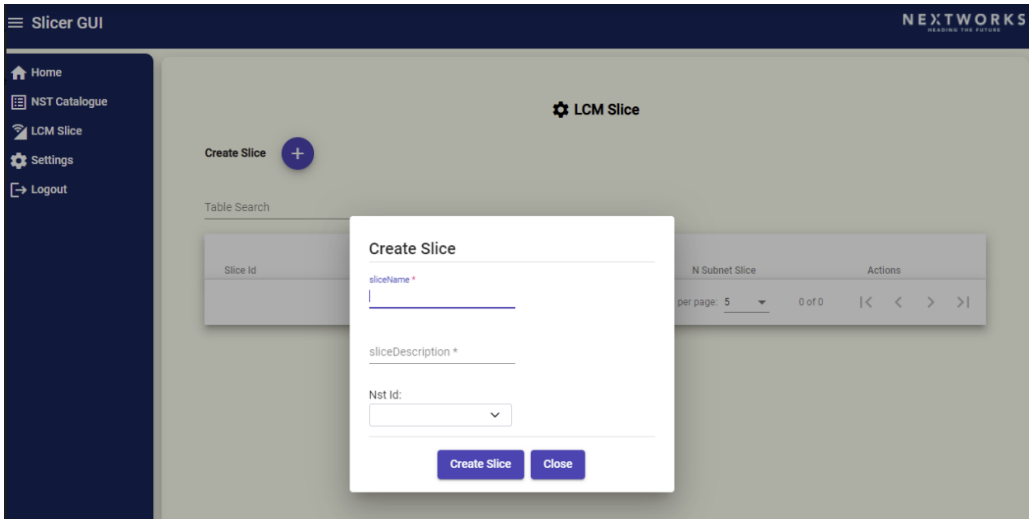


Figure 22: End-to-end network slice provisioning process.

Particularly in Figure 22, it is reported how the network system administrator fills the fields for instantiating a new network slice, specifying the name, the description and most importantly the network slice template (NST) identifier to refer to. This process can be repeated as many times is possible according to the resources available.

As part of the mid-term review demonstration, two end-to-end network slices have been provisioned using the UPV testbed relying on the Cumucore 5G core, and on an emulated gNB using the opensource UERANSIM software tool [18]. Finally, a couple of emulated UEs have been associated with them, as depicted in Figure 23.

Co-funded by the Horizon 2020 Framework Programme of the European Union

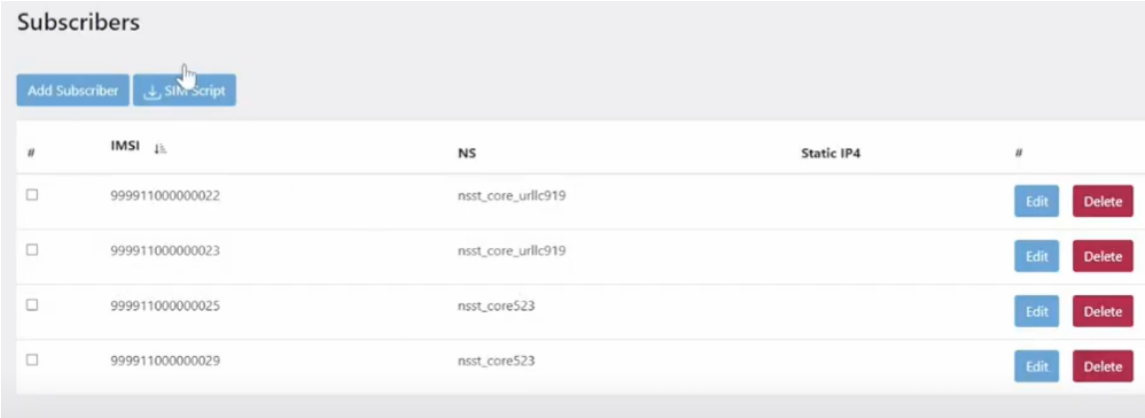


Figure 23: UEs associated to core sub-slices (Web GUI provided by Cumucore:

In this way, each couple of emulated UEs can run their service on top of the end-to-end network slice that is associated with it.

The described and implemented functionalities of the end-to-end network slice orchestration are directly related to the iNGENIOUS use cases as described in the section 4.3.

4.2.2 Automation capabilities through monitoring and AI/ML

As already mentioned in the previous section, one of the Innovation elements in orchestration framework, is the support of AI/ML technology. One challenge addressed by the inclusion of AI in the network management and orchestration framework is to avoid degradations in service caused by finite user plane function (UPF) resources while ensuring security, privacy and reducing data flow demands. In this way, an AI agent serves for the optimal auto-scaling of local UPFs placed at the network edge, thus supporting low latency communication services. The agent is responsible for inferring traffic patterns associated with local UPFs and relaying pre-emptive scaling operations to the end-to-end network slice orchestration in order to avoid service degradations associated with stressed UPF resources.

For feeding the AI/ML platform, which is built by an AI agent, it has been designed a monitoring platform for collecting, storing and managing Information related to network functions, as already initially described in deliverable D4.4 [14]. Figure 24 depicts the software components of the monitoring platform.

Co-funded by the Horizon 2020
Framework Programme of the European Union

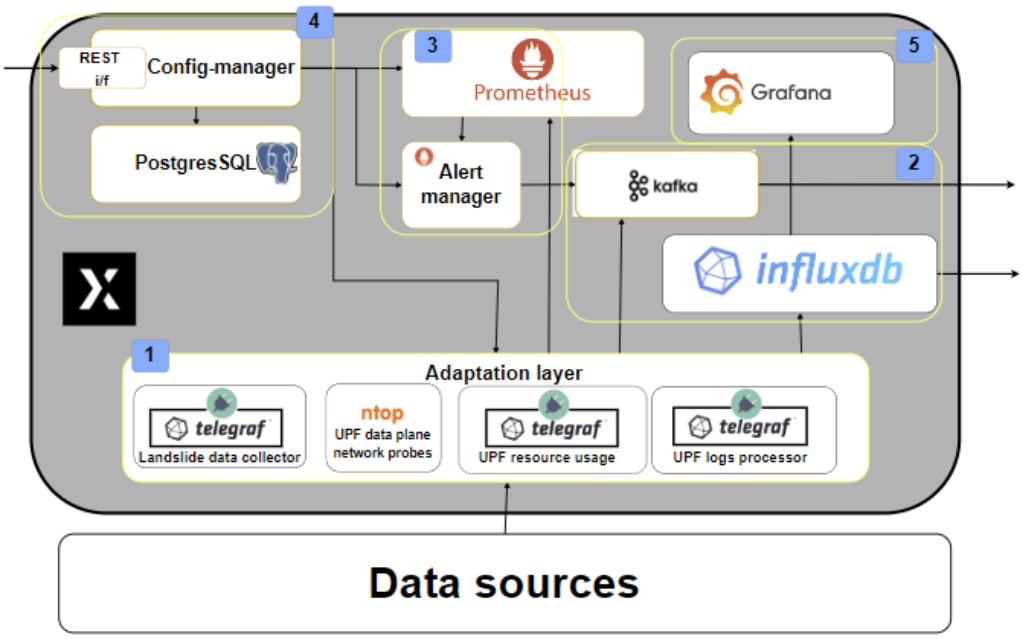


Figure 24: Monitoring platform main components and interactions.

The monitoring platform consists of four major groups of software components, each of them dedicated to a specific task.

1. The adaptation layer is in charge of collecting data from different sources using different plugins (e.g., telegraf and ntop).
2. Data are either stored in a database, using InfluxDB, or made available into a Kafka bus for "real-time" analysis.
3. It would be possible to manage some alerts based on threshold logic or aggregate the data themselves. Currently, this has not been adopted yet.
4. Data sources can be configured through the config manager and such configurations are permanently stored into a PostgreSQL database. Currently, default configurations are used.
5. The data can be shown on an open-source dashboard called Grafana.

The monitoring platform has two main data sources:

- The UPF, where data related to resource usage, protocol data unit (PDU) sessions and networking are collected.
- The Spirent landslide commercial tool, for customizing the stress test to be executed.

To collect meaningful data from the stress tests, the monitoring platform, 5GC and the Spirent landslide tool have been deployed into UPV premises as depicted in Figure 25.

Co-funded by the Horizon 2020 Framework Programme of the European Union

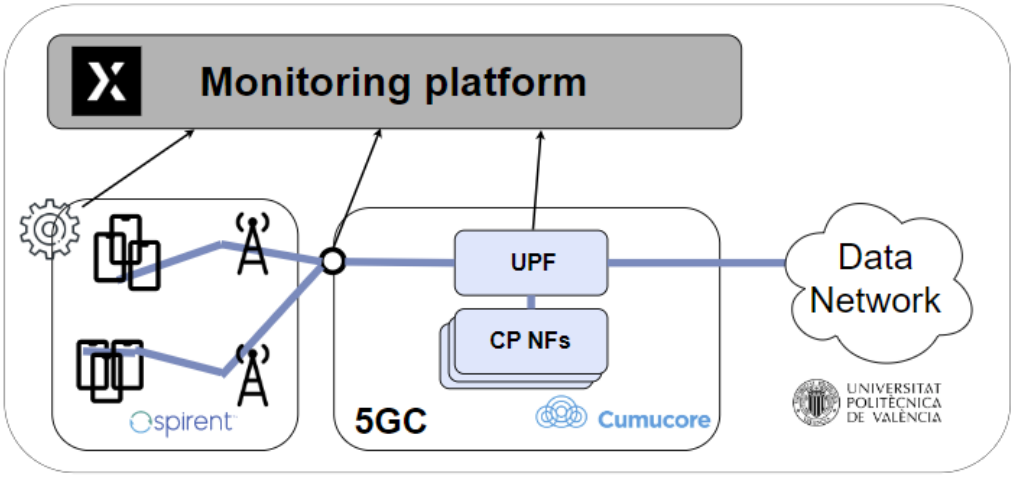


Figure 25: Integration and deployment of Monitoring Platform into UPV premises

Once properly configured the landslide tool with the input parameters, then one or more UEs connect to the 5GC and start to send a certain amount of traffic in a given time window. The number of UEs, the time window and the traffic amount are some of the parameters that can be configured from the tool.

For completeness, Table 2 describes the list of the parameters identified for performing the stress tests.

#	Parameter	Type	Description	Example
1	Duration of experiment(s)	Test parameter	How many seconds the experiment will last	300 sec
2	Iterations	Test parameter	Given the experiment, the number of repetitions	5
3	UEs count	Mobile subscriber	Total number of user equipments. These are simulated. In case of multiple gNB, these are equally distributed across them	100
4	PDU session per UE	Mobile subscriber	Number of active PDU sessions per UE	3
5	Activation rate	Mobile subscriber	Number of UEs that register each second to the gNB	4
6	Traffic starts when: all session established or as soon as session established	Client		Yes
7	Data start delay	Client	After how many seconds UE start sending traffic	2sec
8	gNB count	Client	Total number of gNB connected to the 5GC	5

Co-funded by the Horizon 2020 Framework Programme of the European Union

9	Type of traffic	Traffic params	TCP/UDP/ping/sctp/http/...	UDP
10	Bandwidth (Mbps)	Traffic params	How much traffic goes through the UEs	1Gbps
11	Burst traffic	Traffic params	Only for UDP	Yes
12	Tx/Rx ratio	Traffic params	Distribution of total traffic between source and transmitter	50%/50%
13	Packet size	Traffic params	Packet size in byte	64

Table 2: List of parameters of Landslide tool for stress testing the 5GC.

To be noted that the list could be subject to some minor changes. For each experiment a set of input parameters is needed and once run, data are collected from the monitoring platform. Then, these data are used for feeding the AI algorithm.

The high-level functional blocks of the AI agent framework, depicted in Figure 26, are the following:

- **Data Pre-processing**: The ingestion of the available monitoring data related to the UPF functionality. This data, collected by the Monitoring Platform, contains real-time data from the 5G infrastructure, and applications. The information can be related to specific UEs, (mobility, communication pattern, etc.), NFs, network slices, or the network as a whole. UPF load information available from the network data analytics functions (NWDAF), including CPU, memory, and disk usage, can be supplemented with user plane data like bandwidth, latency, packet loss, etc., as well as UE-related information (mobility, position, etc.) to get accurate predictions of future network conditions. The raw data is pre-processed to address issues such as missing values and unreal numbers.
- **Feature detection**: Using methods of dimensionality reduction, the many parameters collected by the monitoring platform are analyzed to achieve a dimensionality reduction on the set of data the AI agent uses for inference. The current methods being used are Principle Component Analysis and High Correlation Filter. Both are suitable candidates for the semi corelated nature of the input data.
- **Inference engine**: Prediction of future patterns of local traffic. Three models have been utilized for this functionality, each with a different approach to the challenge. The first model is a Long Short-Term Memory (LSTM) architecture that employs a Recurrent Neural Network (RNN). The second, is an eXtreme Gradient Boosting (XGBoost), which uses an optimized distributed gradient boosted decision tree (GBDT). The third, PROPHET, is an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects.
- **Scaling logic**: Once the inference block has produced a prediction on future traffic values and trends, the scaling logic determines if an action should be taken to scale the resources of the UPF. This information is then passed to the MANO for orchestration.

Co-funded by the Horizon 2020 Framework Programme of the European Union

- **Model performance reporting**: At runtime the AI agent also produces an evaluation of the model's performance by evaluating the previous time series prediction to the actual current status. Performance logic is used to determine if the current trained model is accurately performing to predict the traffic trends. When the model performance falls out of alignment, the AI model performance function triggers an evaluation of existing models, if no trained models provide enough accuracy or performance, a new training is triggered in the AI training function, which generates a new version of the model and deploys it in a new or existing AI agent.

The AI agent is designed to be packageable as cloud-native applications to be integrated into a cloud-native deployment environment for scalability. An example deployment scenario could be a hybrid edge/core cloud scenario, where dedicated AI functions management interfaces and workflows among the involved AI functions enable the computationally heavy model training to be moved to the core/cloud, together with an AI repository function to concentrate all training-related operations where computing resources can be easily available with reduced cost. The AI functions associated with inference can be deployed at the edge, and an AI orchestration function can be implemented as a dedicated AI Network Slice Subnet Management Function, responsible to instantiate and operate the AI functions at scale.

The cold start of an AI agent is addressed through a transfer learning approach combined with a continuous learning framework. Pretrained models for network flows are used as initial reference models when the agent is first deployed. The Performance logic module of the AI agent performs an evaluation of existing models, and if no trained models provide enough accuracy or performance, a new training is triggered before instantiating the UPF scaling logic. The amount of collected data and time window required before the instantiation depends on the achievable accuracy of the model training and the time-scale of the input. data trends.

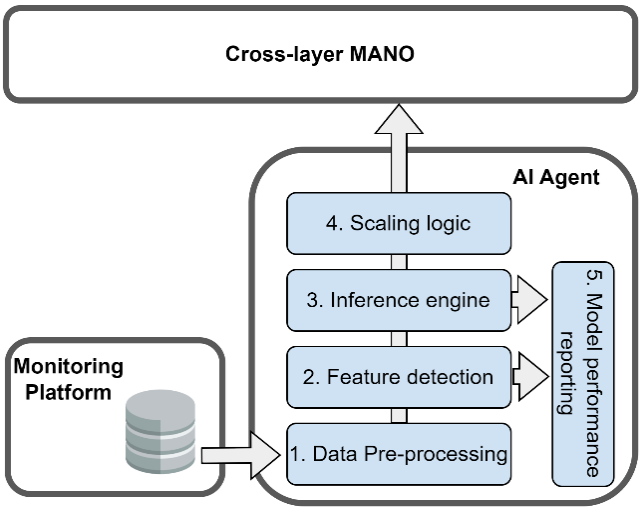


Figure 26: Functional blocks of the AI Agent.

Co-funded by the Horizon 2020 Framework Programme of the European Union

4.2.3 Business Potential and Exploitation Summary

Component Group: End-to-end network slice orchestration

- **NXW**: Nextworks, as main contributor of the iNGENIOUS end-to-end network slice orchestration framework, has identified some business potentials and exploitation opportunities for the outcomes and results reported in this chapter. Specifically, the iNGENIOUS network slice orchestration solution is perfectly aligned with the company strategy and interests, which look towards beyond 5G and 6G service orchestration frameworks, with native integration of in-network analytics, AI and ML capabilities and services. This provides concrete innovation and contribution to the existing company research-oriented network and service management portfolio. This is fundamental to be competitive in the telco consultancy market with cutting-edge beyond 5G and 6G ideas and solutions.

  Nextworks has recently identified the 5G non-public networks (NPNs) as a key topic to look at to exploit the company assets and knowledge (5G network management and industrial IoT areas). The iNGENIOUS results help the company in positioning itself as solution provider for enterprise and industrial IoT private networks through 5G in several scenarios, targeting pre-commercial pilots for industry 4.0, smart building and events, holiday resorts, hospitals, etc.

4.3Relation to the Use Cases

This section describes end-to-end network slice orchestration role in the Factory use case of iNGENIOUS, and how the implemented functionalities are exploited within the use case itself. In particular, at time of writing this deliverable, the final steps towards the integration between the end-to-end network slice orchestration and the Flexible PHY/MAC are ongoing, and they will be part of the final demonstration of Factory use case.

In an Industry 4.0 environment, different services (robot control, video streaming, and other services) can simultaneously run in the same physical environment relying on the same shared network infrastructure. Moreover, each of these services has its requirements in terms of reliability, latency, bandwidth, and so on. For Instance, the robot control service does not require a big amount of bandwidth, but on the other hand, it requires strict latency requirements. Oppositely, the video streaming service requires more bandwidth but no strict requirements on latency. This means to treat differently the types of traffic and then keeping them logically separated. Moreover, in this way could be possible also to provide priority to a certain type of traffic if needed from a certain type of services.

The logical separation of the network traffic on an end-to-end network is provided by the end-to-end network slices. Network slicing is one of the pillars of the 5G network, not only for providing logical separation of the traffic but also for providing the correct amount of resources needed for satisfying the service requirements. The end-to-end network slice orchestration plays a fundamental role in the network slicing because is responsible for the

Co-funded by the Horizon 2020 Framework Programme of the European Union

management of the lifecycle of the end-to-end network slices on top of a common and shared network.

Specifically, in the Factory use case, the end-to-end network is composed of a commercial and standard 5GC and experimental and not-standard access network. Additionally, the services that run on top of the common shared infrastructure are the haptic industrial robot control, video streaming and an emulated application with adjustable requirements. Each of these services has its own requirements that are directly mapped one-to-one with a specific end-to-end network slice with specific technical requirements.

The end-to-end network slice orchestration not only is responsible for the LCM and the provisioning of the end-to-end network, but also for translating the high-level requirements of these services into tailored actions to be executed towards the involved network segment.

In the Factory use case, the end-to-end network slice orchestration component is responsible for the orchestration of a 5G non-public network as depicted in Figure 27. The high-level functionalities described in the previous section that the end-to-end network slice orchestration realizes are:

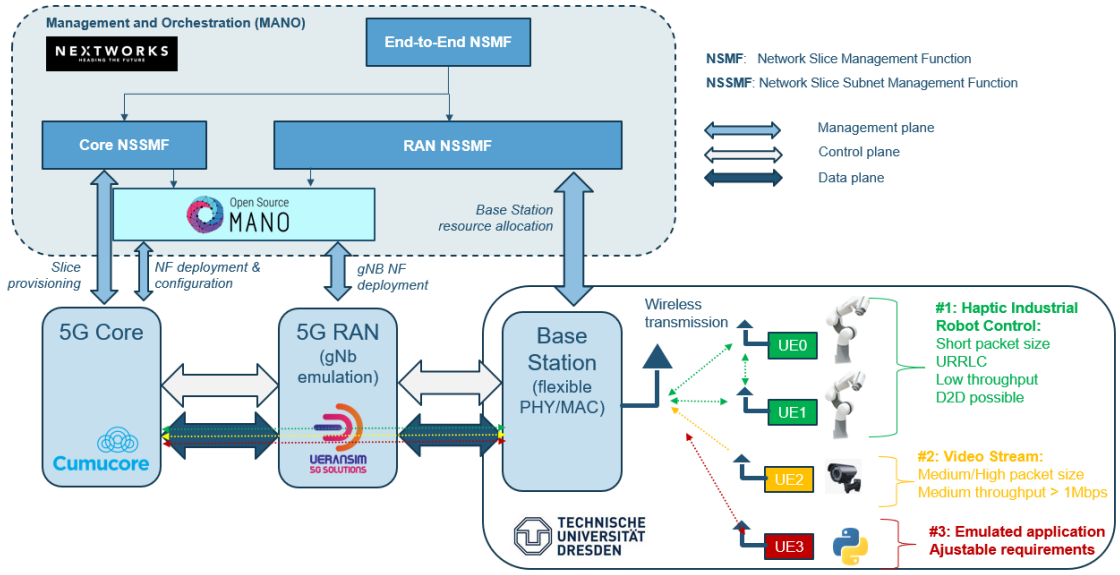


Figure 27: MANO role in Factory use case.

1. The interoperability with a standardized element, i.e., the 5G core and a non-3GPP element, i.e., the Flexible PHY/MAC. This is represented by the management plane depicted in the Figure 27. In particular, for the 5G core integration it has been designed, implemented and tested the Core NSSMF, which is capable of interacting with the 5G core for managing the core slices, the subscribers and gNBs associated to the slices. For the Flexible PHY\MAC, the RAN NSSMF provides connectivity and slicing capabilities. At time of the writing of this deliverable, the RAN NSSMF is under development and is planned to be completed for the final demonstration of Factory use case.

2. The information model representing the end-to-end network slices used in the Factory use case is the same that would be used in other use cases or scenarios. In particular, the translation process occurs at the NSMF and

Co-funded by the Horizon 2020 Framework Programme of the European Union

NSSMFs levels. The LCM slice requests are translated in the technical request towards the 5G core and the Flexible PHY/MAC by the Core NSSMF and RAN NSSMF, respectively.

3. The modularity and flexibility of the end-to-end network slice orchestration in the Factory use case are exploited by "plugging" an ad-hoc NSSMF for managing the resources in the Flexible PHY/MAC. A possible update of the Flexible PHY/MAC technology does not imply an update on the source code of the whole MANO component, but just a part of the RAN NSSMF.

To conclude, in the Factory use case, the end-to-end network slice orchestration plays a fundamental role because not only it provides a novel access network 5G connectivity, but also the automatic network slicing lifecycle management capabilities for the different running services in an Industry 4.0 environment.

For what concerns the status of the implementation, the end-to-end NSMF, and the core NSSMF have been designed, implemented, and tested. The RAN NSSMF has been partially implemented because integration activities are ongoing. The initial testing of these three components has been demonstrated during the mid-term review, where two different end-to-end network slices have been provisioned into a 5G partially simulated network, available into UPV premises.

Co-funded by the Horizon 2020
Framework Programme of the European Union

5 Conclusion

This document provides the reader with the final state of the technical work conducted within WP4 of the iNGENIOUS project, following the deliverables 4.1 [2], 4.2 [3], 4.3 [12], and 4.4 [14]. The roles and interactions among the technological components of the overall architecture are described to demonstrate the iNGENIOUS use cases. This document concludes WP4 and all its associated tasks.

The document was structured in three main sections where the innovations and achievements are described. The first section described the work related to RAN technologies, which considered the Flexible PHY/MAC and O-RAN designs targeting the wide range of requirements that appear when the entire supply chain system is considered. The RAN section also covered the integration of satellite connectivity within the RAN network, and the key role a Smart IoT gateway can play in translating and routing the data coming from the large variety of sensing communication protocols that can be on and off shore. The second section presents the 5G-LAN and its benefits when interconnecting among mobile and private networks in industrial scenarios. Finally, the third section describes the MANO framework, and the process for network slicing integration with the 5G core and AI/ML techniques. As it can be observed, the technologies investigated and developed within WP4 present characteristics that are key enablers of an efficient and smart end-to-end communication network that can address the entire supply chain.

Co-funded by the Horizon 2020
Framework Programme of the European Union

References

[1] iNGENIOUS Consortium, "D2.4 System and architecture integration," 2022.

[2] iNGENIOUS Consortium, "D4.1 Multi-technologies network for IoT," 2021.

[3] iNGENIOUS Consortium, "D4.2 Smart NR and NG-RAN IoT Designs," 2022.

[4] O-RAN Alliance , "O-RAN Use Cases and Deployment Scenarios Towards Open and Smart RAN (White paper)," 2020.

[5] iNGENIOUS Consortium, "D6.1 Initial planning for testbeds," 2021.

[6] iNGENIOUS Consortium, "D3.2 Proposals for next generation of connected IoT modules," 2022.

[7] iNGENIOUS Consortium, "D6.3 Final evaluation and validation," 2023.

[8] 3rd Generation Partnership Project (3GPP), "Study on enhancement of 5G System (5GS) for vertical and Local Area Network (LAN) services (Ref. 23.734)," 3GPP, 2019.

[9] International Society of Automation, "New ISA/IEC 62443 standard specifies security capabilities for control system components," [Online]. Available: https://www.isa.org/intech-home/2018/september-october/departments/new-standard-specifies-security-capabilities-for-c. [Accessed October 2022].

[10] 5G Alliance for Connected Industries and Automation, "Security Aspects of 5G for Industrial Networks," ZVEI – German Electrical and Electronic Manufacturers' Association & , Frankfurt am Main, Germany, 2021.

[11] International Eletroctechnical Commission (IEC), "Industrial communication networks – Network and system security – Part 2-1: Establishing an industrial automation and control system security program," 2011. [Online]. Available: https://webstore.iec.ch/preview/info_iec62443-2-1%7Bed1.0%7Den.pdf. [Accessed October 2022].

[12] iNGENIOUS Consortium, "D4.3 Core network automation design for 5G-IoT," 2022.

[13] P. e. al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research,* vol. 12, pp. 2825-2830, 2011.

[14] iNGENIOUS Consortium, "D4.4 Service orchestration at the edge," 2022.

[15] 3GPP, "TS 28.500 Management concepts, architecture and requirements for mobile networks that include virtualized network functions (Release 16)," July 2020.

[16] ETSI, "Network Functions Virtualisation (NFV) Release 2; Management and Orchestration; Architectural Framework Specification ETSI GS NFV 006," January 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/nfv/001_099/006/02.01.01_60/gs_nfv006v020101p.pdf. [Accessed October 2022].

Co-funded by the Horizon 2020 Framework Programme of the European Union

[17] 3GPP, "TS 28.528 Life Cycle Management (LCM) for mobile networks that include virtualized network functions; Stage 3 (Release 16)," July 2020.

[18] A. Güngör, "UERANSIM GitHub," [Online]. Available: https://github.com/aligungr/UERANSIM. [Accessed October 2022].

[19] O-RAN Alliance e.V., "Open Software for the RAN," September 2022. [Online]. Available: https://www.o-ran.org/software. [Accessed September 2022].

[20] iNGENIOUS Consortium, "D2.1 Use cases, KPIs and requirements," 2021.

Co-funded by the Horizon 2020 Framework Programme of the European Union