A Secure and Reusable Artificial Intelligence Platform for Edge Computing in Beyond 5G Networks

D5.1 Testing and Validation Methodology, Planning and Preparation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101015922

D5.1 Testing and Validation methodology, Planning and Preparation	
WP	WP5 – Use cases Integration, Validation and Benchmarking
Responsible partner	SPI
Version	V1.4
Editor(s)	Leonardo Goratti (SPI), Babak Mafakheri (SPI)
Author(s)	Leonardo Goratti (SPI), Babak Mafakheri (SPI), Antonino Albanese (ITL), Marco Beccari (ITL), Cristina Costa (FBK), Bruno Lepri (FBK), Nicola di Pietro (ATH), Marco Centenaro (ATH), Daniele Munaretto (ATH), Marco Marchetti (CRF), Giampiero Mastinu (POLIMI), Giorgio Previati (POLIMI), Miguel Rosa (AERO), Alberto Cristóbal Granda (AERO), Brendan McAuliffe (SRS), Estefania Coronado Calero (I2CAT), Javier Palomares (I2CAT), George Lentaris (ICCS), Stefano Secci (CNAM), Jérôme François (INRIA), Enric Pere Pages (ATOS), Stelios Koumoutzelis (8BELLS).
Reviewer(s)	Flávio Brito (EAB), George Avdikos (Eight Bells), Roberto Riggio (UNIVPM), Cristina Costa (FBK)
Deliverable Type	R
Dissemination Level	PU
Due date of delivery	31/12/2021
Submission date	31/01/2022

Version History				
Version	Date	Author	Partner	Description
V1	20/09/2021	Leonardo Goratti.	SPI	Table of content created
V1.1	13/10/2021	Leonardo Goratti.	SPI + All contributors	Consolidated table of content

V1.2	17/11/2021	Leonardo Goratti, Miguel Rosa, Javier Palomares, Estefania Coronado, Cristina Costa.	SPI contributions; AERO contributions; I2CAT contributions; FBK contributions	V1.2 created; Inputs for UC4 added; Inputs to UC3 in all related sections. AI@EDGE general architecture contributions
V1.3	23/11/2021	Leonardo Goratti, Marco Marchetti, Nicola di Pietro, George Lentaris.	SPI, CRF ATH, ICCS	New version created; Inputs to UC1; Inputs to UC4
V1.4	30/11/2021	Leonardo Goratti, Antonino Albanese.	SPI, ITL	New version created; Inputs to UC4
V1.5	06/12/2021	Babak Mafakheri, Leonardo Goratti, Marco Marchetti, Bruno Lepri, Cristina Costa, Giorgio Previati, Brendan McAuliffe, Antonino Albanese.	SPI, CRF, DFKI, ITL, POLIMI, SRS, AERO	New version created; Inputs to all Use cases
V1.6	09/12/2021	Leonardo Goratti, Babak Mafakheri.	SPI	New version created
V1.7	13/12/2021	Leonardo Goratti, Babak Mafakheri.	SPI	New version created; ready to be reviewed
V1.8	23/12/2021	Babak Mafakheri.	SPI	New version created; to apply the reviews
V1.9	10/01/2022	Babak Mafakheri, Miguel Rosa, Marco Marchetti, Daniel Reti.	SPI, AERO, CRF, DFKI	Final Version for editorial review
V2.0	18/01/2022	Babak Mafakher, Daniel Reti, Marco Marchetti, Miguel Rosa, Nicola di Pietro, Antonino Albanese, Giampiero Mastinu.	SPI, AERO, CRF, DFKI, ITL, POLIMI	Final Version
V2.1	28/01/2022	Irene Facchin.	FBK	Final review

Table of Contents

List of Tables

List of Figures

Glossary	
5G	5^{th} Generation of mobile communications
5GC	5G Core network
5G NSA	5G Non-Stand Alone
5G SA	5G Stand Alone
AERO	Aerotools
AGV	Automated Guided Vehicles
AI	Artificial Intelligence
AMF	Access and mobility Management Function
AUSF	AUthentication Server Function
BVLOS	Beyond Visual Line of Sight
CP	Control Plane
CPU	Central Processing Unit
CRF	Centro Ricerche Fiat
DFKI	Deutsches Forschungszentrum für Künstliche Intelligenz
HIL	Hardware in the Loop
HITL	Human in The Loop
IFEC	Inflight Entertainment and Connectivity
IIoT	Industrial Internet of Things
LBO	Local Break-Out
MEC	Multi-access Edge Computing
MP-TCP	Multi-Path Transmission Control Protocol
NFUC	Use case Network Function
NRT-RIC	Non Real-Time Radio Intelligent Controller

RAN	Radio Access Network
RT-RIC	Real-Time Radio Intelligent controller
SMF	Session Management Function
SPI	Safran Passenger Innovations
UC	Use case
UDM	Unified Data Management
UDR	Unified Data Repository
UE	User Equipment
UP	User Plane
UPF	User Plane Function
VIM	Virtualized Infrastructure Manager
VPN	Virtual Private Network

Executive Summary

Deliverable D5.1 (Testing and Validation Methodology, Planning and Preparation) focuses on the development of the activities that shall allow demonstrating the four AI@EDGE Use cases introduced in WP2 [1]. It introduces preliminary applications and per-use case integrated platforms (Milestone M26), as well as preliminary trial test plans (Milestone M27) which will progress toward addressing Objective 5 of the project within the framework of WP5. While the overview of each Use case is provided separately, we discuss also their connection with the ongoing AI@EDGE system-wide architecture design, as well as the mapping with the ongoing design of the Connect-Compute platform. Then, the required test bed implementation roadmap toward equipment and services development, integration, logistics and validation activities are presented on a per-Use case basis. This deliverable briefly discusses the risk analysis and the related mitigation actions envisaged at the current stage of the project. We finally emphasize that D5.1 shall provide the basis for a living document that will be continuously updated throughout subsequent installments of WP5 deliverables.

1 Introduction

This deliverable is about the preliminary plans for test bed development, integration, validation, logistics and tests of the four AI@EDGE Use cases that were introduced in WP2 and described in Deliverable D2.1. D5.1 starts providing the synopsis of each of the four Use cases linking also to other project activities, respectively ongoing in WP2 and WP4, regarding system architecture design and Connect-Compute Platform design, to create at the system level the necessary links with the project as a whole. To provide a self-contained and consistent handbook for the four Use cases, we also report on a preliminary risk analysis with associated mitigation actions. The core proposition of D5.1 is to provide the basis for a living document that will be updated as the project progresses to fulfil the project's overall Objective 5, and that shall be continuously revised in order to adapt to the contingencies that may arise during the course of the AI@EDGE project.

The four Use cases that leverage the AI@EDGE platform are summarized below.

Use case 1 (Virtual Validation of Vehicle Cooperative Perception) refers to a group of vehicles that communicate with each other to coordinate their actions. The data collected by these vehicles can be used by Artificial Intelligence Functions (AIFs) to build traffic scenarios that predict potential collisions.

Use case 2 (Secure and Resilient Orchestration of Large (I)IoT Networks) revolves around the various security and privacy functionalities of the 5G network for industrial IoT services in digital societies. This is because the increasing number of sensors and actors in an Industrial Internet of Things (IIoT) network environment has raised concerns about security threats. In this case, the AI capabilities of the network are being used to detect anomalies; thus, it supports mitigation and reorchestration of the Connect-Compute fabric.

Use case 3 (Edge AI Assisted Monitoring of Linear Infrastructures Using Drones in BVLOS Operation) focuses on monitoring the vast areas of roads network using drones equipped with AI@EDGE Connect-Compute platform. Its goal is to use the 5G capabilities that allow the transmission of images without having issues with the current technologies in communication and control.

Use case 4 (Smart Content & Data Curation for In-Flight Entertainment Services) revolves around an evolved Inflight Entertainment and Connectivity (IFEC) system that allows passengers to customize the content they receive on board. This system would allow passengers to select the exact content they want to consume onboard by employing Artificial Intelligent Functions (AIFs) to cure the data. For this purpose, UC4 deploys an edge-cloud infrastructure that relies on the AI@EDGE platform's objectives.

Throughout this document, each of the Use case leaders and partners discuss the sub-systems, their testbeds and development roadmaps, validation and test plans and the risk analysis. Thus, this document delivers the achievements of Milestones M26 (preliminary applications and per-use case integrated platforms), and Milestone M27 (preliminary trial test plan) of the AI@EDGE project.

This document is organized as follows. Section 2 is assigned to the Use case overview where each sub-system is introduced separately. The general AI@EDGE architecture, derived from D2.2, is reintroduced in Section 3, followed by the 5G system adaption of each Use case. In Section 4, the required testbeds for the test session preparation are listed and explained. The preliminary plan for the AI@EDGE integration with the Use cases is highlighted in Section 5, whereas the related activities with regard to the risk analysis are detailed in Section 6. Finally, Section 7 concludes the deliverable.

2 Use cases overview

This section is devoted to providing a concise synopsis of the four Use cases tackled by AI@EDGE, while more insights can be found in [1].

2.1 Use case 1 – Virtual validation of vehicle cooperative perception

The Use case 1 (UC1) reference scenario refers to several vehicles that exchange data with each other and through the network edge to coordinate their maneuvers. The exchanged data are useful to build traffic scenarios used by Artificial Intelligence Functions to predict potential collisions and dangers. Given the complexity and costs to support many vehicles in the real world [2] [3], the Use case 1 adopts emulation and simulation environments that are able to scale and perform exhaustive and reproducible tests. Considering different possible specific traffic scenarios, the roundabout is a particularly challenging situation where fluidity and safety are important and will be implemented in the Use case.

The Use case logic is based on Artificial Intelligence and Agents with a Reinforcement Learning approach. In the Use case context, each Agent is a digital twin that represents a vehicle. This approach will leverage the network and service automation features of the AI@EDGE Connect-Compute platform. In terms of 5G stack, the AI@EDGE platform 5G core will be interfaced with a 5G network emulator to allow testing a broader range of scenarios and configurations.

Cooperative Perception becomes more complex when dealing with mixed real and virtual traffic situations and to create such a scenario, a dynamic driving emulator operated by a real human driver will be interconnected with traffic simulators to design, implement, and test the digital twinning of a mix of real and emulated vehicles. The aim of the testbed is to create a geographically distributed Virtual Validation environment to support cooperative manoeuvres between vehicles collecting and sharing information from/to vehicles.

2.2 Use case 2 – Secure and resilient orchestration of large (I)IoT networks

In Use case 2 (UC2) certain security and privacy functionalities for multi-stakeholder 5G campus networks are evaluated and network accelerations are showcased. The increasing interconnectivity of actors and sensors in manufacturing is increasing the attack surface and thus making Industrial Internet of Things (IIoT) network environments susceptible. As a result, when evaluating novel Connect-Compute fabrics powered by 5G technologies, special care should be taken. In this Use case one aspect to be shown are AI capabilities in the network used for detecting anomalies including security threats, and hence support mitigation and reorchestration of the Connect-Compute fabric.

The UC2 scenario involves automated guided vehicles (AGV), as used for warehouse logistics, which are controlled and monitored over a 5G interface and a near edge server to provide security against cyber-attacks. This scenario assumes a multi-stakeholder environment, where different areas of an industrial site might be governed by separate stakeholders who share a 5G campus network, in order to showcase the importance of confidentiality and the advantages of federated learning. Federated machine learning is researched and applied in AI@EDGE to learn and make predictions about network behaviour, so the configuration can be improved, normal behaviour can be learned, and anomalies can be detected.

While the main functionality of this Use case will be the anomaly detection at the edge server with AI-based approaches, UC2 will also showcase accelerations of the application based on NetFPGA base network cards, as well as traffic breakout functionality, which allows to cache network data on edge servers and therefore improve speeds and reduce network traffic.

For the evaluation, a variety of cyberattacks will be implemented to test the anomaly detection performance, including man-in-the-middle, denial of service, data, and model poisoning attacks, to see if the anomaly is recognized.

2.3 Use case 3 – Edge AI assisted monitoring of linear infrastructure using drones in BVLOS

The Use case 3 (UC3) main application is monitoring large areas of roads network through the use of drones with the AI@EDGE Connect-Compute fabric as an embedded system, in order to use 5G capabilities to overcome the limitations of the current technology in communication and control (C2) and remote transmission of images.

The drone will be controlled in a BVLOS (Beyond Visual Line of Sight) mode through the 5G network to scan industrial infrastructures, to make corresponding 3D modelling, then to identify the different incidents that could exist and to send notifications to the drone operator alerting that an incident has been found. Meanwhile, the information (images, telemetry) is sent continuously to the central office in order to improve the drone's operator decision-making process.

2.4 Use case 4 – Smart content & data curation for in-flight entertainment services

Use case 4 (UC4) tackles an evolved Inflight Entertainment and Connectivity (IFEC) system on board civilian airplanes that can offer passengers different personalized packages of content produced by different content producers. An evolved IFEC system shall allow full content customization in such a way that passengers may even be allowed to bring their own content on board without any physical media. Simply the content is curated for the passenger that can thus carry it on board for a real ubiquitous experience.

If on the one hand such level of content personalization is far from trivial, an advanced concept exemplifies the possibility to make a rough content selection from a ground pool of contents. The selection shall be made by a system as automated as possible that relies on offline artificial intelligence training with historical data of different airlines, source/destination information and passengers' subscriptions. The refined online curation of the content takes place, instead, on board in the edge-cloud infrastructure relying on light-weight artificial intelligence content selection until a personalized bundle of media contents is offered to a passenger, to the extent possible. In the context of airplane infrastructure with limited resources, computationally light artificial intelligence approaches are crucial to make wise use of the network-compute fabric and reach fast convergence time in content curation. In this regard, Use case 4 targets to include also hardware acceleration to validate enhancements to the cabin edge-cloud infrastructure to carry out also computation-intensive artificial intelligence on board.

3 AI@EDGE architecture overview and mapping to Use cases

This section describes key design outputs like the AI@EDGE overall system design and the Connect-Compute platform that were developed in WP2 and WP4, respectively. For the sake of this deliverable, we are conscious that the key designs provided hereinafter are still a work in progress and therefore can be subjected to changes that shall be reflected in future instalments of WP5 deliverables.

3.1 AI@EDGE general architecture

In this section, we give a quick overview of the AI@EDGE general architecture. For a more detailed description of the general architecture, please refer to Deliverable D2.2 [4].

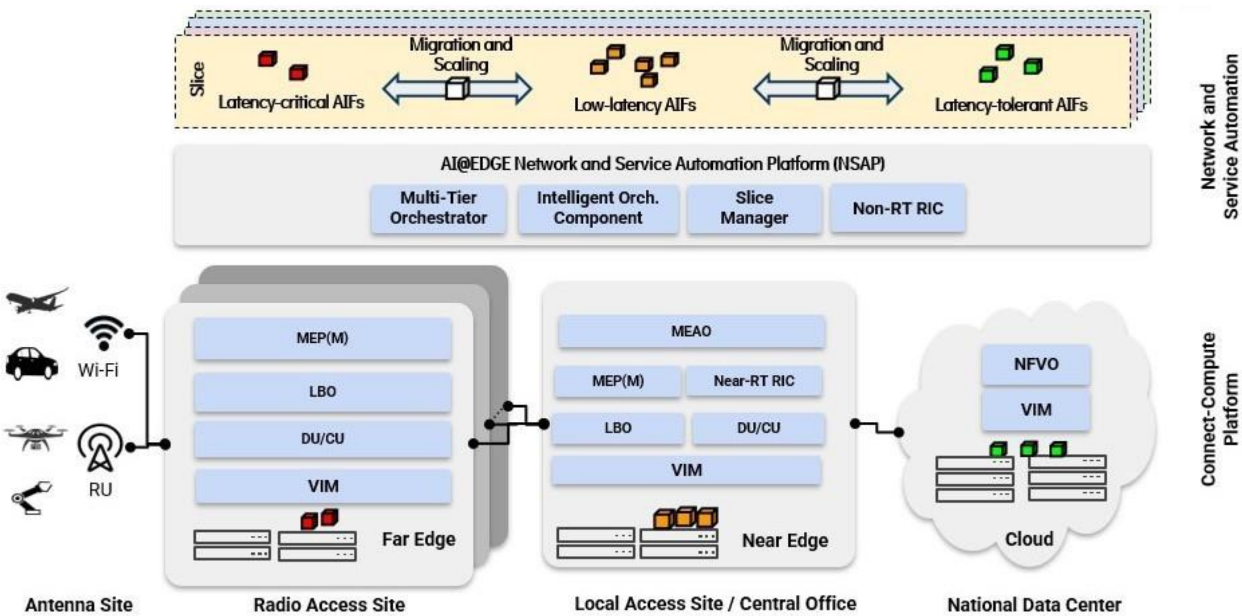

Figure 1 AI@EDGE system architecture.

The AI@EDGE system supports the orchestration and management of Artificial Intelligence Functions (AIFs). The AIFs represent an abstraction of software modules that implement Machine Learning (ML) models and expose various interfaces dedicated to (i) the input/output of data; and (ii) the AIF configuration. The AIFs implement the intelligence of the system and can support both network automation and user applications at the edge. AIFs are further described in D2.1 [1], which also details the AIFs reference model. Additionally, each UC will produce its own AIFs that will be specific for each scenario.

The AI@EDGE system general architecture is represented in Figure 1, where the location options of the cloud resources are depicted, namely Far Edge, Near Edge and Cloud. Furthermore, in Figure 1, the two main layers are highlighted: the Network and Service Automation layer (defined in detail in [5]) and the Connect Compute layer (defined in detail in [6]). AIFs can be deployed in both layers of the system, depending on which function they implement.

- The **Network automation layer** is where the network intelligence is placed and it is responsible for the system network automation. Automation loops can happen either on a local level or on a more global and extended level, depending on the scope [5].

- The **Connect compute platform** is responsible for providing the fabric for AIFs orchestration and management across various edge levels/locations. It is also responsible for assuring the connectivity between the different elements of the system, and for managing the available computing resources.

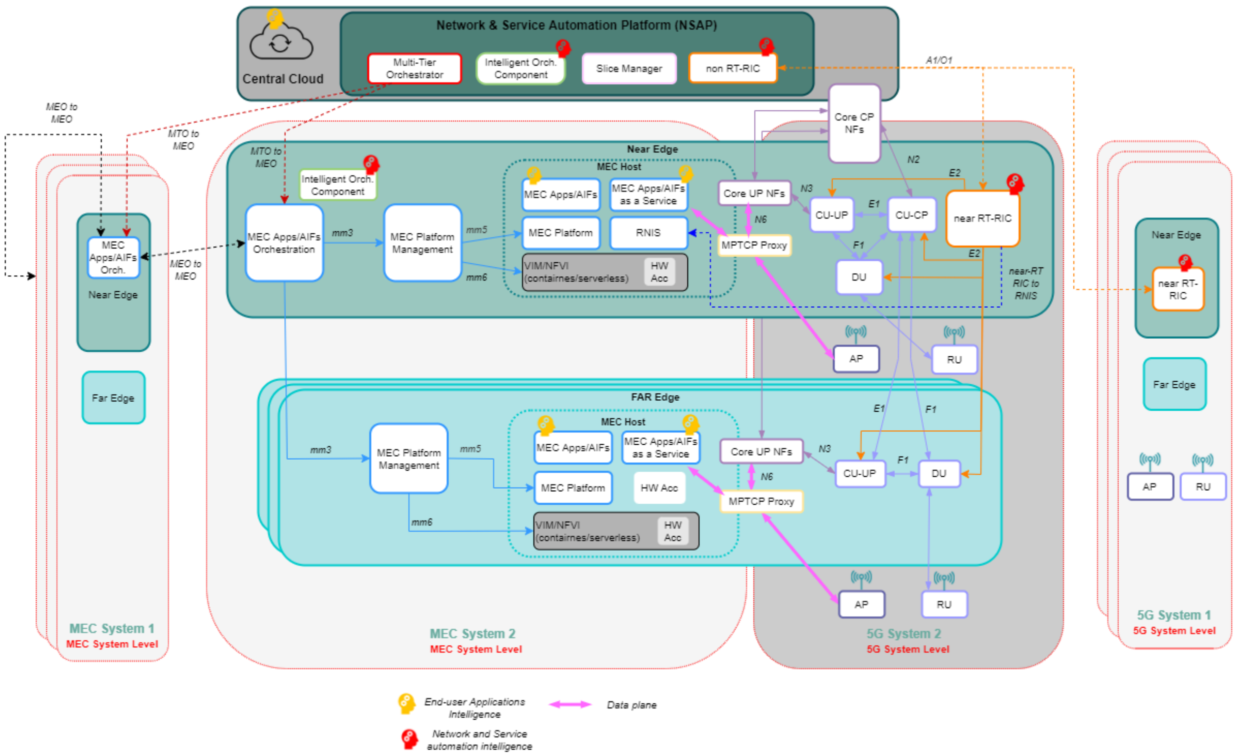

Figure 2 Draft system architecture of the AI@EDGE connected-compute platform.

Figure 2 depicts the Connect Compute Platform layer in detail, highlighting an option for the location of the platform components between the Central Office and the Radio Access Sites [4]. We report below some information about the Connect Compute Platform that is relevant to the current document. The interested reader should refer to D4.1 for more in-depth information.

The MEC system is a collection of MEC hosts and MEC management processes, which are necessary to run MEC applications, as per [7]. The MEC Apps/AIFs Orchestration element maintains an overall view of the complete MEC system. It is responsible for the placement of MEC Apps or AIFs when it is requested or needed. This decision can be based on rules or classic optimization methods or rely on the Intelligent Orchestration Component that uses AI/ML methods. The Orchestrator can be placed by the Near Edge or by another site as long as it has a direct connection with the Edge Sites (both Far Edge and Near Edge). The Orchestrator can be connected at most to one Near Edge and multiple Far Edges.

A MEC host is an entity that contains a MEC platform and the corresponding virtualization infrastructure, which provides compute, storage and network resources to MEC applications [7]. The MEC host is strategically placed by the edges of the network to provide computation and storage capabilities near the access network and provides, among other advantages, lower latency. To this aim, the 5G traffic is steered towards the MEC host where it can be processed (more details on the integration of the MEC host with the 5G infrastructure are given in the following paragraph). The MEC host can be therefore considered as an edge cloud able to host MEC applications and AIFs. An AIF can run on the MEC Host stand-alone or as a

MEC Service, providing services to other MEC Apps on the same or other MEC Hosts. The MEC Platform provides the functionalities required to run MEC applications and AIFs. It enables them to provide and consume MEC services, and that can provide itself with a number of MEC services, such as the RNIS service. The MEC host can also provide hardware acceleration services.

The virtualization infrastructure provided by the MEC hosts in AI@EDGE will be based on an extensible and open-source container orchestrator called Kubernetes [8]. This platform manages containerized workloads and services that facilitate configuration and automation processes. The MEC hosts will also deploy a FaaS Platform, supporting event-based, stateless, serverless functions.

Finally, the near-RT RIC will support AIFs relying on RAN data for network automation. The interested reader should refer to D4.1 for details on the near-RT RIC in-depth description.

3.2 5G system architecture for Use cases implementation

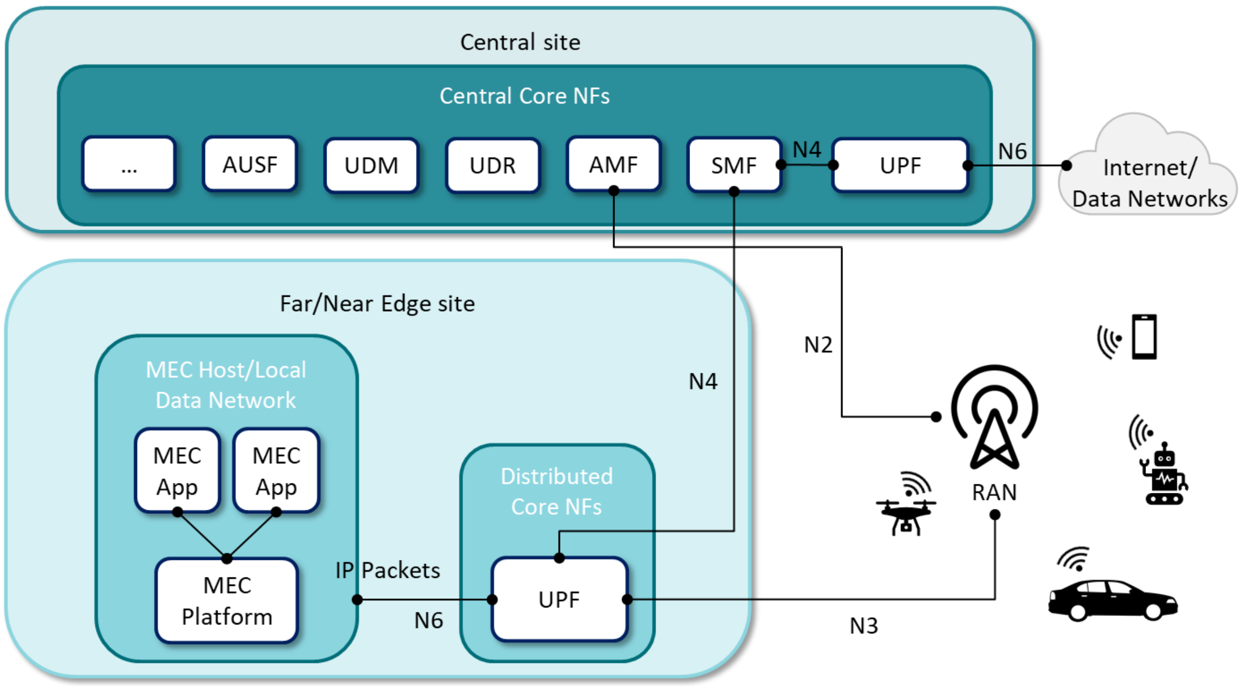

Figure 3 High-level reference distributed architecture of a 5G system, integrated with a MEC host.

Figure 3 provides a high-level representation of a 5G system architecture and its interconnection with other external architectural elements (MEC facilities, the Internet, other data networks, etc.). The purpose of the picture is to highlight a set of features, which will be part of the network deployments of the “connect” side of AI@EDGE’s platform in the project’s Use cases. First, the figure shows the network functions (NFs) that compose the 5G core network (5GC) [9]: the Access and Mobility Management Function (AMF), the Session Management Function (SMF), the User Plane Function (UPF), the Unified Data Management (UDM), the Authentication Server Function (AUSF), and the Unified Data Repository (UDR). Notice that this list is not exhaustive, and for representational purposes, we are only showing the minimum set of required core NFs for 5G connectivity. Furthermore, we are depicting the possibility of distributing NFs – which are virtualized – across different physical sites. In particular, Figure 3 shows a UPF deployed at an edge site and connected to the rest of the 5GC (the so-called control plane, CP) deployed at a central site. This will be an architectural choice highly leveraged by AI@EDGE because it allows to route data traffic directly from the User Equipment (UE) to the MEC facilities and the edge computing servers located in its proximity, and vice versa. Such a user-plane (UP) traffic local breakout enables low-latency communications and lightens the overall traffic over the backhaul of the network, ultimately with benefits to the quality of service that the 5G network can offer and to the users' quality of experience. Finally, Figure 3 explicitly reports the names of the standard 3GPP-defined reference points (N2, N3, N4, N6) [9] between the main represented elements.

Notice that Figure 3 is a reference upon which several variations can be built coherently, depending on specific architectural, functional, or performance requirements. For instance, an analogous system architecture can be adapted to scenarios with more than one edge site or in which the distribution of NFs is implemented differently.

In particular, Figure 4 represents the realization of the 5G system and edge site connectivity architecture that AI@EDGE is considering deploying to serve AI@EDGE's Use cases supported by ATH's 5GC (UC 1, 2, and 4). Namely, we are planning to set up a single 5G CP at FBK's premises that controls three independent UPs at POLIMI (UC1), DFKI (UC2), and SPI (UC4), possibly belonging to three different network slices. ATH's 5GC can natively support multiple edge sites, and such a deployment will showcase the flexibility and adaptability of 5G networks to simultaneously provide efficient connectivity services to separate scenarios, geographically distributed, and with different performance requirements. This matches, for instance, the needs of a national mobile network operator that provides dedicated user plane setups and edge nodes for the private use of specific vertical sectors. In this framework, UP traffic will be kept locally and directed towards the corresponding edge site, whereas only control signalling will travel towards or from the central CP via dedicated VPN connections. The full feasibility and implementability of such an architecture will be further investigated and confirmed at the beginning of the activities of tasks T5.2, T5.3, T5.5, from the project's month 13 on.

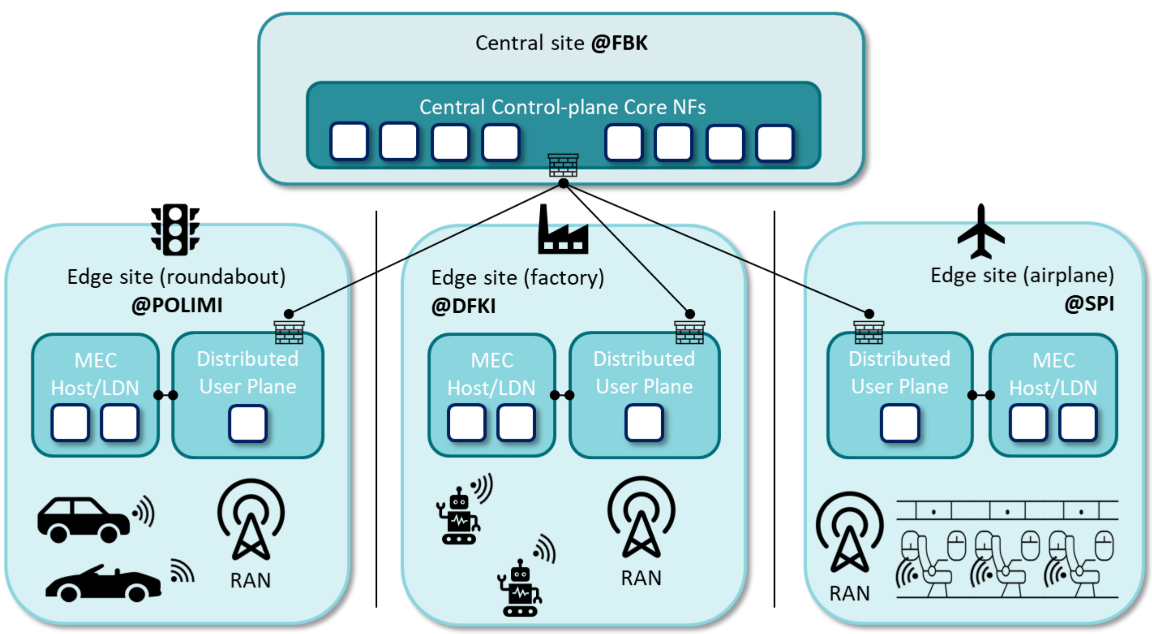

Figure 4 The 5G architecture with centralized control plane and distributed independent user planes that serve UC 1, 2, and 4.

The white little squares generically represent NFs and MEC Apps, in an analogous configuration of that in the previous figure.

As far as UC 3 is concerned, instead, it will be based on the 5G System deployed in the 5TONIC Laboratory, which will provide the required 5G Network Environment for the validation of the Use case. The 5G system and edge site connectivity architecture is depicted in Figure 5.

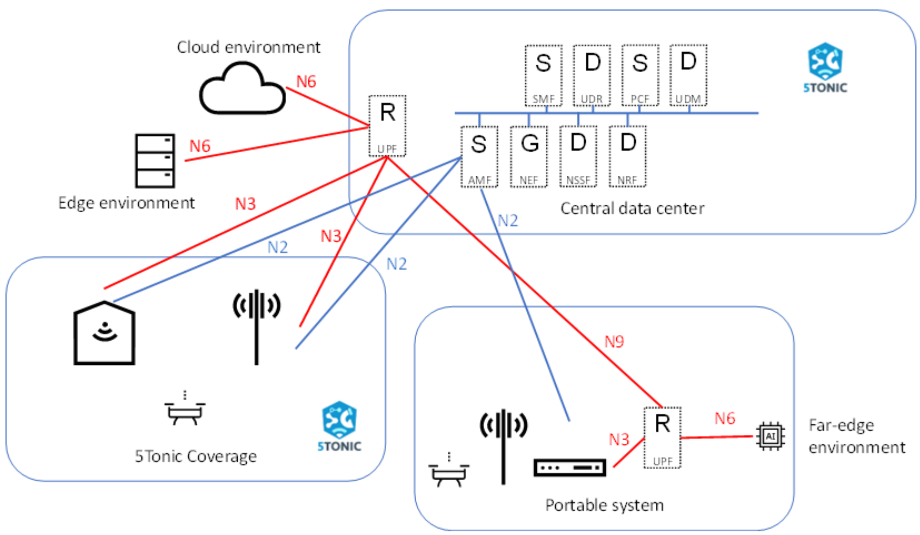

Figure 5 5Tonic system architecture.

The 5G System is composed of a central 5G Core, which provides control-plane network functions required for the 5G network: AMF, SMF, PCF, NRF, NSSF, UDM, and UDR. (Considering this list as an example of the functions potentially deployed in 5TONIC). This 5G core acts as the central core of an operator located at the datacentre, while edge sites and functions can be supported on several other testing spaces (further description of the different elements provided by 5TONIC is exposed later in this document).

With regard to Radio Access Network setup, 5TONIC provides several indoor and outdoor radio coverages, which are connected to the central 5G Core. The UPF is located close to the radio access and enables the possibility of deploying edge and central applications. In addition, 5TONIC provides a portable system that can be deployed in any location of Spain and provides 5G NR coverage as an extension of the 5TONIC network. The portable system can include an UPF and a far-edge environment for running applications (e.g., AI-based video processing).

Specifically, in the RAN setup, 5TONIC has different areas of coverage (both indoor and outdoor) to support a wide range of Use cases. The main experimental area (X3 Room) covers the following bands of LTE FDD: B7, B20, NBIot in B20 and LTE TDD B42; for 5G, it also has an NSA deployment with LTE as anchor in band 7 and NR TDD in band n78L (B43). As this 5G node is also configured to work as SA, both 5G capacities are supported by the network.

For indoor areas (Room 1S4), a small Ericsson antenna, called DOT, plus two additional antennas have been deployed, covering both technologies: LTE (B7) and 5G (n78L). So, this area and other rooms in the 5TONIC lab are covered with this NSA + SA solution. RAN Access Network in 5TONIC covers also outdoor, where an 8T8R antenna by Commscope has been deployed. It is normally configured as LTE but can be configured to work in 5G. Besides, this area is covered by the indoor NSA+SA setup located at X3 Room, with the Anchor LTE in B7 and NR in Band n78L. This 5G antenna installed in X3 is a 64T64R, a Massive MIMO antenna configurable to cover both Urban dense areas, such as Highrise buildings or Hotspots, as Rural areas using his spatial multiplexing capability.

3.2.1 5G system adaptation to Use case 1

UC1 embraces the general approaches provided by AI@EDGE regarding the development of the 5G system in a road environment. The UC will provide a completely simulated CRF environment to validate the Artificial Intelligence algorithms involved in the UC and the software and radio components of the telematic boxes. An edge node with decentralized user-plane core network functionalities will also be available at POLIMI, together with a driving simulator connected to the mobile network. The two systems, the one at CRF and the one at POLIMI, will share the simulation scenarios to evaluate algorithms with simulated vehicles that implement driving logic and a vehicle in which a human is directly involved.

The CRF site validation system will consist of a 5G access with core network emulation and will provide an automotive traffic simulation environment, on which various Cooperative Perception algorithmic techniques based on AI Reinforcement Learning will be evaluated. Once the Telematics box and the Cooperative Perception algorithms have been validated at CRF's validation site, they will be evaluated in the POLIMI environment, in which a real driver is involved. To be more consistent with respect to the project objectives, a better focus on the 5G Uu interface will be provided instead of the PC5 communication as previously described in the Use case deliverable "D2.1 Use cases, requirements, and preliminary system architecture".

The POLIMI edge node will host a UPF while the control plane will be located at the FBK headquarters, and a connection between the two sites will take place via a site-to-site Virtual Private Network (VPN). The edge node will also host the Cooperative Perception service that will interact with the driving simulator, which will be connected via a 5G Telematics box to a RAN emulation and to the AI@EDGE platform.

The Cooperative Perception will be implemented as a distributed agent system where each agent represents the digital twin of one vehicle. Both in CRF and POLIMI most of the agents will rely on Traffic simulation systems with the addition of agents that represents Telematics Boxes involved in the Use case. The Cooperative Perception will rely on the MEC provided by the AI@EDGE Platform.

The 5G connection could be validated only for the human-driven simulator and for the Telematic Box, all the data of the other simulated vehicles will be shared directly from the Worldsim [10] traffic simulator to the MEC App via ethernet cable. The following picture highlights in green the Use case specific components and in light blue the AI@EDGE components. Within the activities of the dedicated task, the chosen solution might be refined to improve the architecture or performance, according to specific needs or requirements that might arise.

The main Use case architecture and related components are depicted in Figure 6.

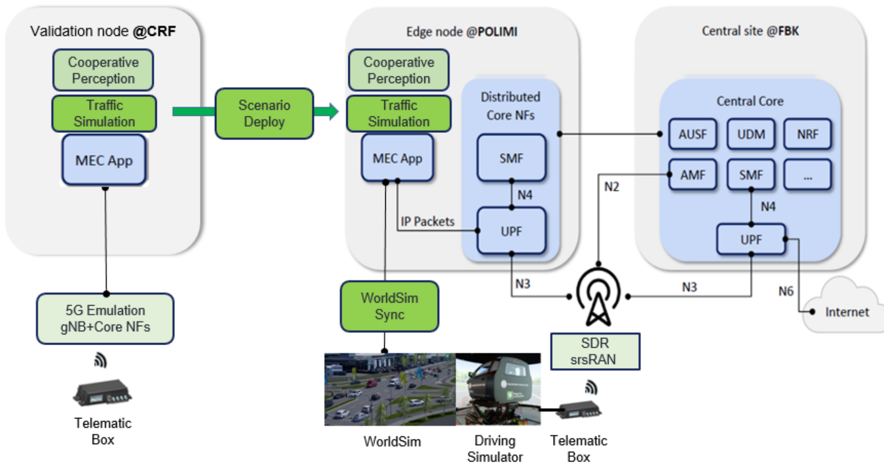

Figure 6 Use case 1 components.

3.2.2 5G system adaptation to Use case 2

UC2 will showcase the security and privacy aspects of 5G networks in the application of smart manufacturing environments. In the scenario, as depicted in Figure 7, multiple tenants are sharing a common 5G campus network, exploiting data streams for the anomaly detection AIF, while at the same time preserving the data privacy with federated machine learning.

For this demonstration, AGVs are controlled from the near edge. Different sensors are deployed to track the AGV and collect the position data on the near edge. All the devices are connected with the 5G network. An anomaly detection AI will run against the collected network and sensor data on the near edge servers (Anomaly detection AIF). The AI-based approach at the edge server aims for high detection accuracy with low latency.

FPGAs have the ability to be faster than general-purpose ICs, making them increasingly important for constructing 5G infrastructure elements. The fully functional 5G firewall based on NetFPGA [11] provided by CNAM, act as the first line of defence, speeding up the detection process by filtering all traffic for intrusions between eNodeB and Core.

A decentralized UPF will be provided by ATH and integrated into UC2 within an edge node component, with the advantage of allowing selected traffic to be locally offloaded for essential network edge applications without disrupting the existing network or risking network security. The Control Plane will be located at the FBK facilities and connected via a VPN tunnel to the User Plane at DFKI.

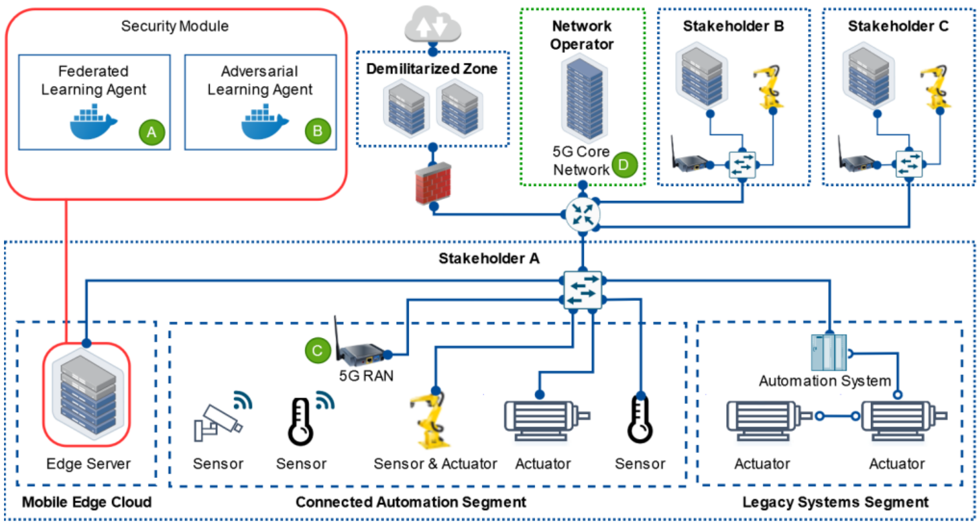

Figure 7 Generalized depiction of the Use case 2 scenario. Stakeholders A, B and D operate separate facilities and share one 5G core network. Each edge server has a security module.

3.2.3 5G system adaptation to Use case 3

UC3 will showcase the capability of AI-assisted 5G networks to control and operate drones in a BVLOS scenario, where reliability and very fluid data traffic are compulsory features in order to send telemetry data as well as image and video data with low latency.

The central core will be located at 5TONIC, where the AI@EDGE platform will be hosted. The Far Edge will be located at Aerotools and/or ATOS offices, where UP will be deployed to run AI functions dedicated to high-quality 3D modelling, and other tasks based on this 3D model such as visual analysis, labelling, alarm warnings and any function capable of enriching the operation. Part of this enriched information, such as warnings will be sent to the drone operator, who will use them to take decisions (to modify or keep the planned flight plan). The following figure shows the configuration and architecture of Use case 3 with specific components for the application in green colour and the AI@EDGE components in blue.

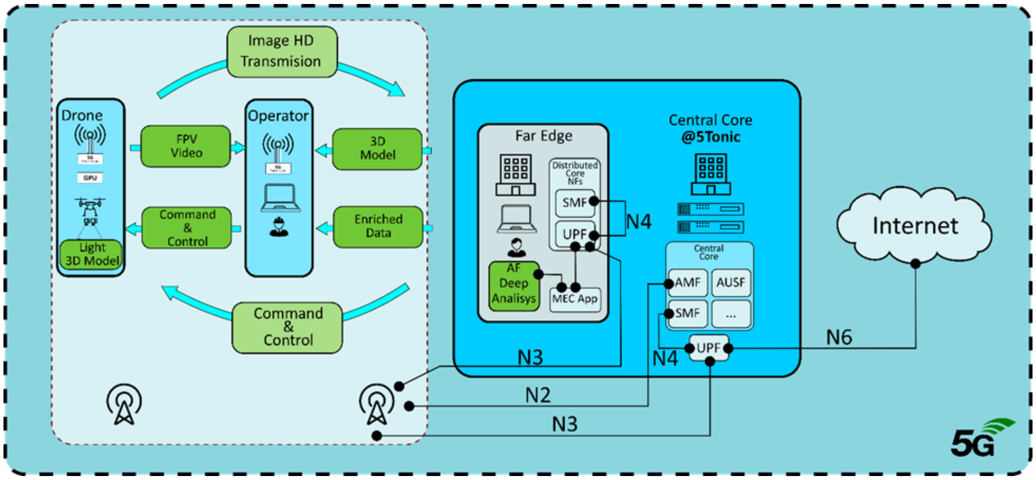

Figure 8 Use case 3 components.

3.2.4 5G system adaptation to Use case 4

The AI@EDGE Smart Content & Data Curation for In-Flight Entertainment Services, aka UC4, embraces the general approaches provided in Figure 3 and Figure 4 with regard to the development of the 5G system on-board an aircraft.

SPI shall host at its premises a UPF, while the control plane shall be located at the FBK office with a connection between the two that shall take place by means of a site-to-site VPN as illustrated in Figure 9. Specifically, during the flight time, an aircraft shall be envisaged as an edge-cloud connected to a ground data network through satcom technology, air-to-ground or a combination thereof; whereas on ground the aircraft can be connected to a dedicated (i.e., inside an airport) or to a commercial 4G/5G mobile network. Furthermore, the central core network (i.e., control plane) is assumed always on ground and connecting to an on board UPF for the purpose of implementing the local breakout (LBO). It must be mentioned that in the case of an airplane, the presence of a satellite connection in the Geostationary orbit, which is a de facto connectivity approach adopted by aviation, has to be taken into due account since 3GPP timers might fire before a UE association completes, or even the UPF association with the control plane (i.e., establish the N4 interface) can be completed successfully. Overall, this can turn into different problems that can cause malfunction of the 5G network on board. In general, services deployed on board, such as the content curation, shall be considered MEC Apps that service users requesting such particular service that does not require communication with a data network on the ground. Clearly, this introduces the advantages of saving on the bandwidth of the satellite or air-to-ground technologies, as well as greatly reducing service latency, thus leveraging on the typical advantages of MEC. Moreover, the integration of a multipath aggregation technology based on MPTCP proxy will be studied to allow UE usage of multiple access technologies for increased availability and bitrate.

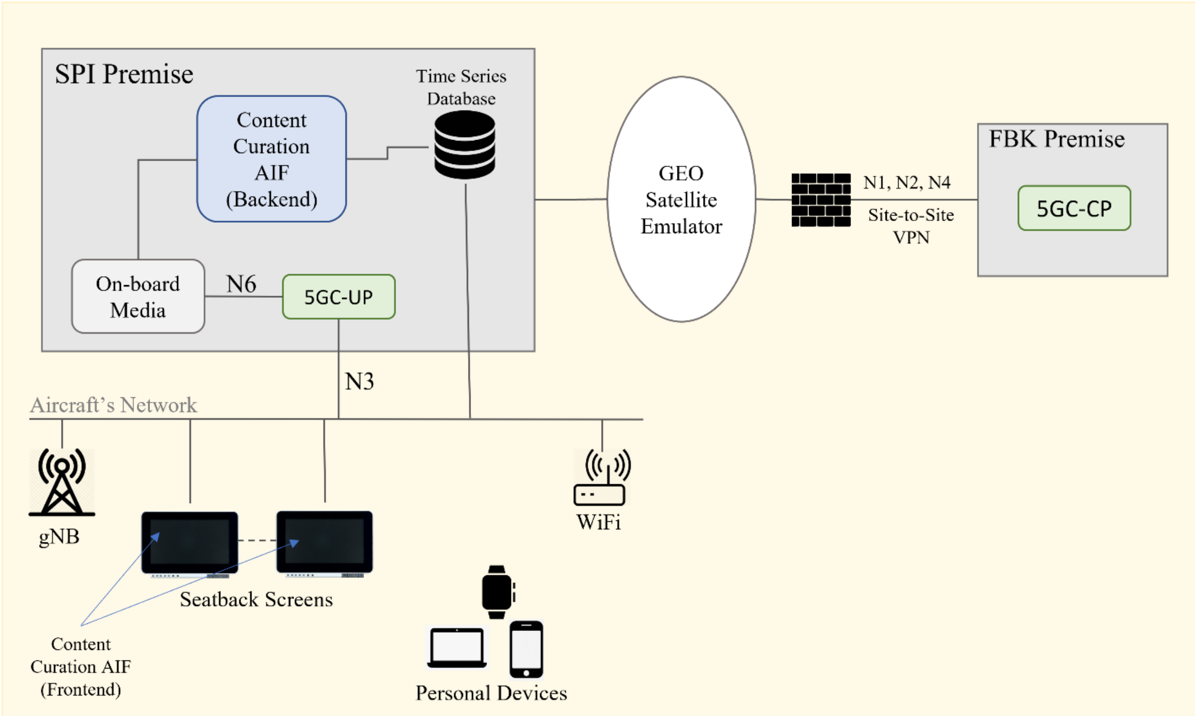

Figure 9 Use case 4 components.

4 Test bed status and roadmap

This section describes the roadmap that is currently envisaged for the four test beds that will be developed within AI@EDGE. Given the early stage at which the current plans are developed, needed adaptations shall be unveiled in future instalments of WP5 deliverables.

4.1 Use case 1 test bed

UC1 targets to develop two test bed facilities. The first facility relies on the infrastructure available at POLIMI in Milano site, where the driving simulator will be 5G connected and will send its vehicle dynamic to the network, in particular to an Edge Node on which a Cooperative Perception Algorithm will be executed. The POLIMI test bed facility is depicted in Figure 10.

Figure 10 POLIMI Driving Simulator.

In detail, the POLIMI site is based on the Driving Simulator and its Worldsim scenario environment. During the activities, a 5G telematic box will be integrated with the driving simulator to provide connectivity to the AI@EDGE platform. An SRS radio access point will be deployed to connect the driving simulator with the 5G network, whose control plane is hosted at FBK's premises and user plane within a local edge node (cf. also Figure 4). The edge node will also run the cooperative perception algorithm that works on a traffic simulation software (e.g. SUMO [12], MATSim [13]). The traffic simulation provides the simulation scenario for all simulated vehicles and the driving simulator will be part of the simulation using a mobile connected chain from the telematic box to the edge node through the RAN (in red in Figure 11). The traffic simulator will be also connected with the WorldSim driving simulator software to display simulated vehicles to the driver (in green in Figure 11).

A second facility will be the Validation Site available at CRF in Torino where a 5G emulator will test 5G enabled automotive Telematic Boxes and will provide a traffic simulation platform on which the Artificial Intelligent Agents Cooperative Perception distributed algorithm will be validated.

In detail, a complete 5G emulation system will be available and connected through RF cable with a Telematic Box under test. A GNSS signal emulator will be connected with the telematic box to provide satellite signals emulation and some On Board Devices (OBD) will be connected to provide eventual data from vehicle sensors. The Cooperative perception Algorithm will integrate a traffic simulation environment with simulated vehicles, the Telematics Box under test will represent the real vehicle and will be part of the traffic simulation because the vehicles dynamics data will be sent to the traffic simulation environment that will add the Telematics Box in the simulation (in red in Figure 12). The validation scenario can then be transferred or synchronized with the edge node at POLIMI.

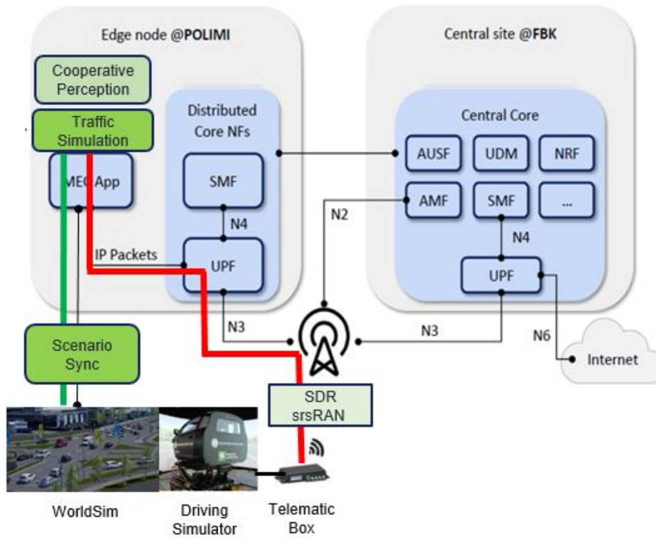

Figure 11 POLIMI Site.

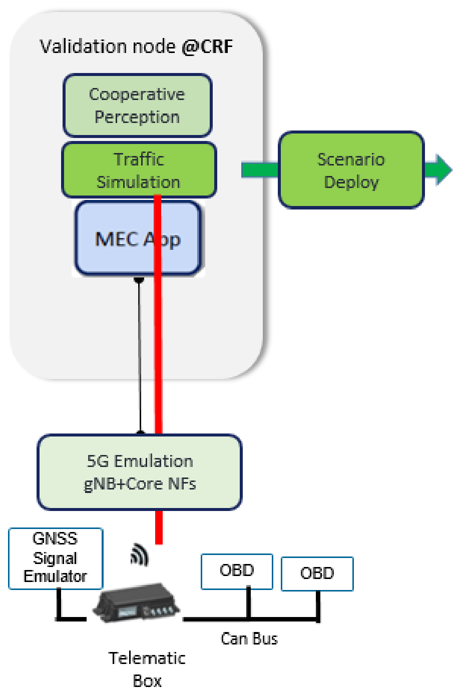

Figure 12 CRF Validation Site.

The Overall Test Bed plan grouped by Phases is detailed in the following planning depicted in the Figure 13. Details of Equipment (E), Service (S), Integration (I), Validation (V), Demonstrations and Logistics (D) activities are explained in the related chapters.

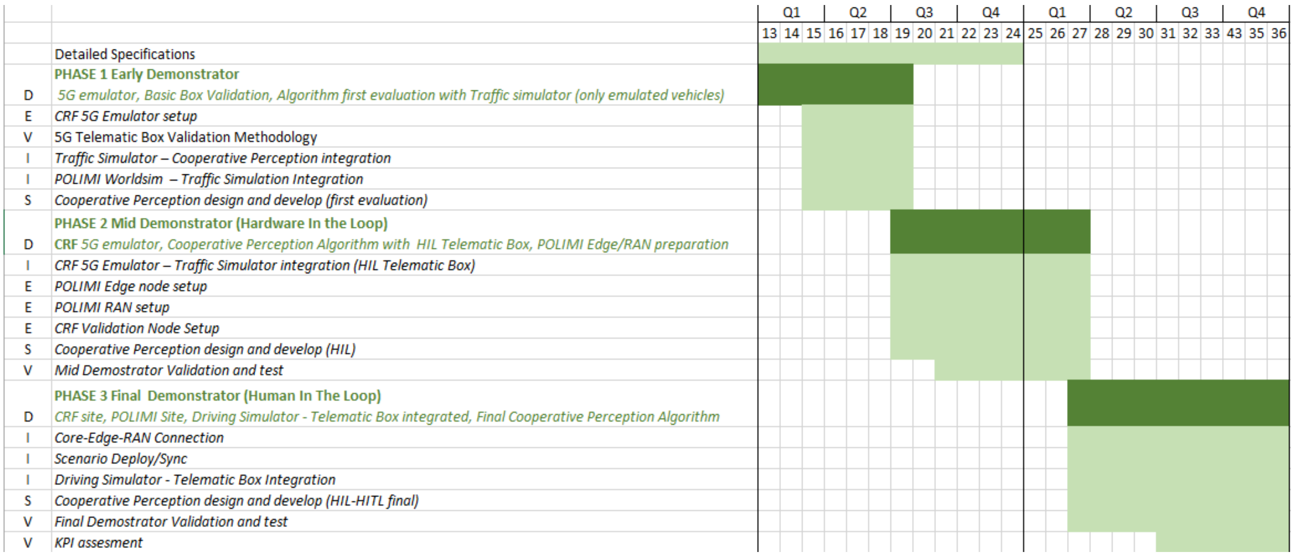

Figure 13 UC1 test bed development initial time plan.

4.1.1 Equipment development plan

The CRF validation site equipment will consist of the following elements.

- 5G Emulation system
  - Configured as Non-Stand Alone and Stand-Alone nodes
    - 5G-NR SA and NSA: DL 2CA, DL 4x4 MIMO, FDD and TDD
    - LTE / LTE-A: DL 3CA, DL 2x2 MIMO, FDD and TDD
  - Support 3GPP Rel. 15 and earlier
  - Support for different Country Codes (example: China, USA, Europe)
  - Ability to simulate communications up to 6GHz and openness in extending support for millimetre waves (28 GHz, 39 GHz)
  - Support for Mobility Handover in order to measure the availability of the service considering the different speeds of the car and the passage from one eNB to another,
- Control PC for Software to execute Validation and Test sequences,
- A local Edge Node emulation connected to the 5G Emulator to validate backend applications in a Hardware in the Loop mode.

The CRF emulation system is depicted in Figure 14 and components are described in the Table 1.

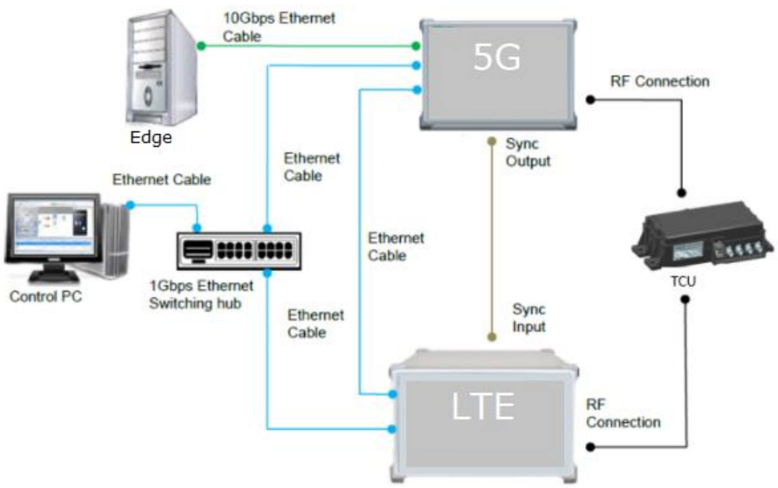

Figure 14 UC1 5G Emulation System.

Table 1 5G Emulation System Equipment.

Equipment Name	**Description**
5G Emulator	Base station simulator for NR
LTE Emulator	Base station simulator for LTE
TCU	Telematic Control Unit Under Test
1Gbps Ethernet Switching hub	Switching hub used to connect the following equipment • 5G emulator • LTE Emulator • Control PC
Control PC	PC on which the Validation and test software are to be installed
Server PC	PC with external server function installed. Required to test packet communication with an external server. A 10Gbps Ethernet adapter

For POLIMI's testbed facility, we are considering the option of a 5G RAN equipment relying on SRS's virtualized radio solutions that can run on a USRP X310 board for high performance in terms of time-frequency synchronization and high throughput. A different configuration may be chosen later, depending on the actual development of the activities.

At this early stage of the UC planning, ATH's 5GC functionalities at the edge node in POLIMI are envisaged to be deployed over a common-off-the-shelf Dell 240 server, with VMware as hypervisor and the following technical features:

- Intel Xeon CPU E-2146G 3.5 GHz 6C/12T,

- 32 GB RAM,
- 2 x 1 TB 7.2 K RPM SATA 6 Gbps 512n 3.5in Cabled Hard Drive,
- 4 x 1 Gb Ethernet.

The Control Plane (CP) functionalities, instead, will be deployed over a similar HW, located at FBK's premises, and shared for the purposes of UC1, 2, and 4 (cf. Section 3.2).
The POLIMI Driving Simulation Facility will consist of:

- Cable driven driving simulator DIM400,
- Worldsim traffic simulator environment,
- Cockpit with haptic HMI,
- CAN bus, LAN, Ethercat cockpit connections for 5G Node integration,
- Real time server with the following specifications
  - 2 x Intel Xeon Gold CPU E6144 3.5 GHz 8C/16T,
  - 48 GB RAM,
  - 4 TB Hard Drive,
  - 2 x 1 Gb Ethernet,
  - Concurrent Realtime Scheduler.

Preliminary Roadmap:
Phase 1 (to be completed by Q2 2022):

- First deployment of the CRF validation and test lab at Torino premises with 5G Emulator,
- Development of the traffic scenario and definition of the communication protocol between Traffic Simulator and Worldsim at Polimi.

Phase 2 (from Q3 2022 to Q1 2023)

- Deployment of the POLIMI Edge node with the distributed NFs and MEC app support,
- Deployment of the POLIMI RAN setup,
- Consolidation of the traffic scenario and traffic simulators setup,
- Experimental tests with data acquisition,
- Completion of the CRF Validation Node Setup with an Edge environment.

Phase 3 (From Q2 2023 to Q4 2023)

- Iteration of previous tests and validation of the UC1, analysis and KPI's verification.

4.1.2 Services development roadmap

FBK will develop solutions for the design and optimization of communication networks between self-driven vehicles (i.e., Reinforcement Learning (RL) agents) and human-driven vehicles agents. More specifically, we aim at coordinating the actions of a set of autonomous and controlled agents that coexist in a realistic environment. These agents will interact according to different levels of reciprocal integration and shared knowledge.

We will initially validate two extreme scenarios, centralized and distributed .

In the fully centralized scenario a central controller has access to perfect information on the environment. This information is distributed to each agent, which can then plan a well-informed action at each iteration. This ideal situation comes with high costs both in terms of resources and communication, and it will be useful to set a baseline for the following more realistic scenarios.

The second extreme scenario is a fully distributed system, where each agent interacts only with a local set of other agents. The notion of locality may be defined in terms of proximity (e.g., physically close agents that are within reach of the employed sensors), structural similarity (e.g., agents produced by the same vendor that may have facilitated communication), or logical similarity (e.g., agents operating with similar decision processes that may be more easily integrated).

On top of these two cases, we will explore and optimize heterogeneous scenarios better suited for real-world applications. Namely, we will model realistic communication latencies between agents of different groups, and introduce communication costs (e.g., in terms of delays or noise) to be taken into account when enlarging each agent's network. With these constraints in mind, we will design network optimization algorithms to obtain near-optimal cooperation, possibly including (expensive) queries to the central controller whenever needed. Different notions of optimality will be considered, namely maximized perception and maximized collective reward (e.g., traffic delay minimization).

The system will be developed using SUMO-RL [14], which is an interface to instantiate RL environments on top of the traffic simulator SUMO [12]. SUMO-RL provides ready to use implementations of state of the art RL algorithms, and is thus suitable for benchmarking purposes. The simulator SUMO, on the other hand, is a reference tool in traffic simulation, and provides physically accurate self-driving agents that will be used to mimic the human-driven agents.

ATH will provide UC1 with a fully virtualized 5GC that will encompass all the core networking functionalities required by the Use case. ATH's 5GC is entirely developed and tested at ATH's premises and is compliant with 3GPP standard specifications (cf. [9]). At the moment of writing this deliverable, ATH's 5GC is available as a 5G Standalone (SA) solution. Along the 24 months of activity of T5.2, whenever opportune, the provided functionalities may be further updated according to the evolution of the standard itself in the scope of the project's Use cases, and ATH will upgrade the 5GC software deployed for UC 1 with its available new releases.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022):

- First analysis of the Cooperative Perception with algorithmic options evaluation.

Phase 2 (from Q3 2022 to Q1 2023)

- Cooperative Perception design and development in the Validation environment at CRF site integrated with the Hardware in the Loop setup,
- Definition and recruitment of the panel of human drivers for UC validation, preliminary tests on drivers' acceptance,

Phase 3 (From Q2 2023 to Q4 2023)

- Final Cooperative Perception design and develop at the POLIMI site integrated with the Driving simulator.

4.2 Use case 2 test bed

The test bed for UC2 is located in a DFKI showroom in Kaiserslautern. The demonstrated application is based on Automated Guided Vehicle (AGV). The AGVs are controlled and guided over the edge server of the 5G network. The location data of the AGVs will be collected on an edge cloud and an AI-algorithm will perform anomaly detection on the aggregated sensor and network data of the AGVs.

Various positioning methods are used to track the position of the AGVs. This location data is sent to the edge server for monitoring and control. The AGV receives a stop signal from the edge cloud when a virtual or physical barrier is crossed. At the same time the available data is used for the detection of various attacks such as man-in-the-middle, denial of service and model poisoning attacks.

4.2.1 Equipment development roadmap

DFKI has set up 5G Network based on hardware from Druid and Airspan fr the core and RAN, respectively. This network is already set up but still has challenges that need to be issued. The localization hardware is already set up in the DFKI-Kaiserslautern showroom, as depicted in Figure 15. The AGV can be remotely controlled with a ROS interface over 5G network using a Rasberry Pi 4B bridge.

Figure 15 DFKI showroom with AGV, gNodeB and localization sensors mounted on trusses and 4 rolling stands carrying light barriers to span a cage.

For the tracking of the AGV position four different methods are considered as follows.

- Decawave DWM1001 System with RasPi 3B,
- MM Modem with Zotac mini PC,
- Allied Vision Industrial camera with NVIDIA mini PC,
- RevPi DIO light barrier.

CNAM and DFKI will work on the design of a programmable Smart-NICs (Network Interface Cards) for the near edge server in order to (i) satisfy real-time monitoring of network flows, (ii) guarantee the physical security of network link metrics then used for anomaly and intrusion detection, and (iii) exploit the reconfigurability of the Smart-NIC to mitigate attacks by blocking flows à la SDN.

CNAM will set up NetFPGA boards (represented in Figure 16) as Smart-NICs to compute the metrics for the anomaly detection AIF and intrusion detection algorithms. The boards will be programmed to be able to compute flow-level metrics and have them collected at dedicated integrated controllers by means of both API and ad-hoc south-bound interface. This effort is in coordination with Task 4.4.

The distribution of the Smart-NICs on edge servers and mini-PCs will be designed in order to support runtime traffic monitoring of the network infrastructure close to end-points, sensors and actuators, as well as 5G RAN and core elements. Figure 17 describes a possible deployment of the network switches for the interconnection of the UC2 testbed elements. The depicted switches will initially be deployed as standard L2 switches. In a later stage, in fall 2022, they will be replaced by three smart-NICs, two of which plugged to edge servers via the PCI slot. The third one would be connected using a serial connection to a mini-PC.

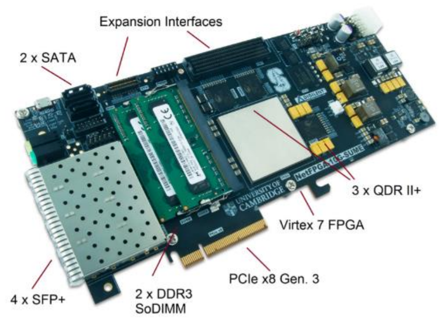

Figure 16 NetFPGA Sume Board.

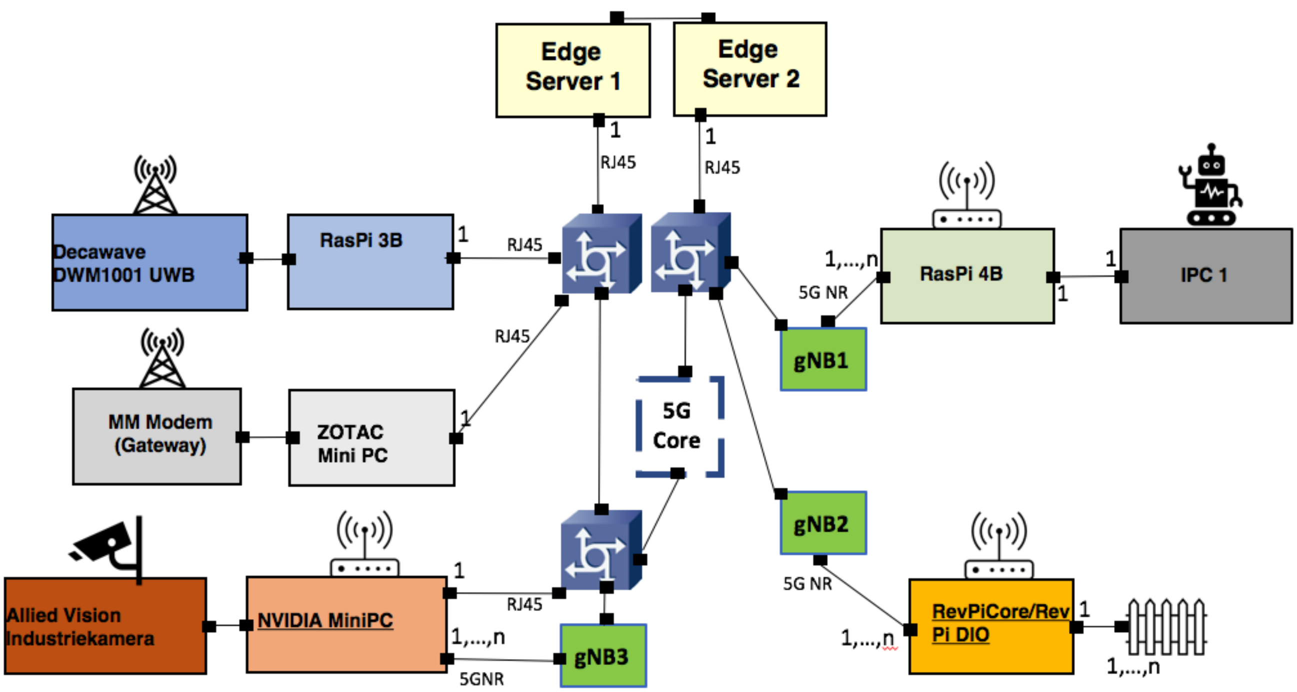

Figure 17 Preliminary network interconnect setting of UC2 testbed including three switches.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022)

- Stable 5G OpenRAN network running,
- Integration of sensors and start with data processing on the near edge
- Deployment of standard L2 switches

Phase 2 (from Q3 2022 to Q1 2023)

- Replacement of L2 Switches with Smart-NIC

Phase 3 (From Q2 2023 to Q4 2023)

- Improvement of stability and integration of software components

4.2.2 Services development roadmap

ATH will integrate its UPF for Mobile Edge Traffic Offload (Local Break Out) and integrate it as a Virtual Machine on the near edge server. This provides a cache for certain types of traffic and allows for the traffic to be broken out and provide data from the cache of the edge cloud. This caching helps ensuring low latency and saving bandwidth for specific applications such as video streaming. Therefore, specific traffic will be offloaded for key applications that are implemented at the network edge without impacting the existing network or breaking network security.

The anomaly detection AIF will be deployed in both centralized mode and federated learning mode on edge servers and mini-PCs. The set of metrics used by the anomaly detection AIF will be grouped in the following groups: CPU, RAM, network links (including those managed by Smart-NICS), gNodeB and 5GC, for both the physical layer and the container layer for the first two categories.

INRIA will coordinate with DFKI and CNAM, for the deployment of the anomaly detection AIF on the near edge, as well as on the evaluation of intrusion detection algorithms to partly be built upon the anomaly detection framework and metrics.

DFKI will aggregate and process the localization data on the near edge and create the cloud control for the AGV.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022)

- Preprocessing of localization data
- First iteration of anomaly detection AIF integration

Phase 2 (from Q3 2022 to Q1 2023)

- Stable demonstration of anomaly detection AIF
- Phase 3 (From Q2 2023 to Q4 2023) Integration of mobile edge traffic offload

4.3 Use case 3 test bed

The UC3 Testbed framework is composed of three environments that must work integrated and coordinated to emulate the required conditions for the development of equipment and services within the AI@EDGE platform. Throughout the developing process, the testbed will contribute to build the scenario for validating the technology. The three environments are identified as follows:

- The 5G network environment is provided by 5TONIC.
- The drone environment is provided by AEROTOOLS.
- The AI tool environment for running the automated incidents detection is provided by ATOS.

The 5G NETWORK ENVIRONMENT is built on top of 5TONIC laboratory based at the IMDEA NETWORKS building in Leganes (Madrid, Spain). Currently, the 5TONIC Data Centre (X1 room) has 24 racks to hold the equipment of all the 5TONIC members and for supporting communications, as it is shown in Figure 18. The required ventilation and power for those racks are provided, including electrical protection to guarantee uninterrupted operation.

Figure 18 5Tonic Data Center (X1).

There are indoor and outdoor 5G coverage areas for testing to be carried out, and the RAN Access Network is adapted to provide coverage and support to the different areas and testing cases. The principal experimental area is the room called X3 (depicted in Figure 19), that has a Radio deployment that covers the following bands of LTE FDD: B7, B20, NBIot in B20 and LTE TDD B42. For 5G, room X3 also has an NSA deployment with LTE as an anchor in band 7 and NR TDD in band n78L (B43). This 5G node is also configured to work as SA, so both 5G's capacities are supported. The 5G antenna deployed in X3 provides coverage both indoor and outdoor is a 64T64R. This Massive MIMO antenna is configurable to cover both Urban dense areas, such as Highrise buildings or Hotspots, as Rural areas using his spatial multiplexing capability.

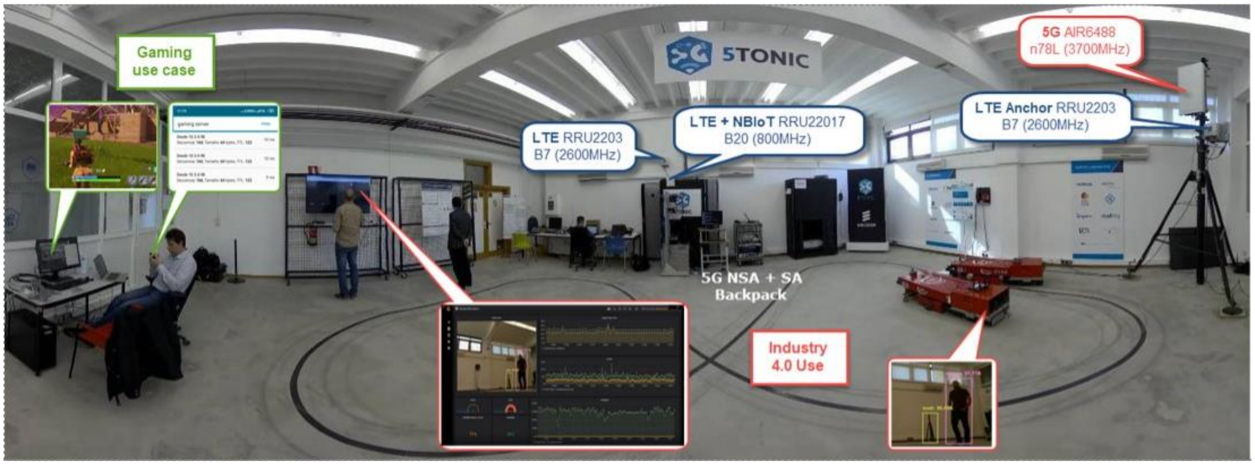

Figure 19 5Tonic main indoor experimentation area (YX3).

The outdoor area is shown in the Figure 20. Radio Access Network is provided by an 8T8R antenna by CommScope which is currently configured as LTE but that will be configured to work in 5G. This area is also supported by the indoor X3 NSA+SA setup, with the Anchor LTE in B7 and NR in Band n78L.

Figure 20 5Tonic outdoor experimentation area.

Additionally, other areas such as the room called 1S4 are covered by Radio Access Network, with the smallest Ericsson antenna called DOT and 2 different antennas that covers both technologies: LTE (B7) and 5G (n78L). So, this area and other rooms in the 5TONIC lab are covered with this NSA + SA solution, as shown in Figure 21.

Figure 21 5Tonic extra indoor experimentation areas.

The 5G System is composed of a central 5G Core, which provides all the control plane network functions required for the 5G network: AMF, SMF, PCF, Network Repository Function (NRF), Network Slice Selection Function (NSSF), UDM and the UDR. This 5G core acts as the central core of an operator and it is in the X1 datacenter of 5Tonic. Regarding the radio coverage, 5Tonic provides several indoor and outdoor radio coverages, which are connected to the central 5G Core. The UPF is located close to the radio access and enables the possibility of deploying edge and central applications.

In addition, 5Tonic provides a portable system that can be deployed in any location of Spain and provides 5G NR coverage as extension of the 5Tonic network. The portable system can include an UPF and a far-edge environment for running applications (e.g., AI-based video processing).

The **DRONE ENVIRONMENT** provides the required equipment and integration options for evaluating the new functionalities:

- Flying platforms of varied sizes and configuration with specific equipment for system integration and operations.
- Integrated systems (Navigation, C2, payload, data transfer) and computing devices (such as Raspberry PI3, Jetson Nano or similar).
- Stereoscopic cameras in dedicated stabilized gimbal to provide high quality footage.
- First Person View (FPV) camera to support drone´s operations when required by the operator.
- Operation of drone in BVLOS capacities, including communication and regulatory compliance.
- Communication devices based on a range of deployed networks.

Dedicated flying platforms and systems will be allocated to facilitate sequential integration and validation of new developments:

- The AT4-dev drone will be used as a small portable and adaptable platform for the initial tests in the first stage.
- As the development of AI@EDGE platform moves forward, a bigger and more powerful drone will be used, the AT6 model, that can embrace heavier and bigger devices while providing enough endurance for validating the operation.

Both drones will share the same configuration for most of the onboard systems (navigation, communication, payload management, etc.) to ease the transition between them. The functional scheme of this configuration is depicted in Figure 24, showing the functional concept of the integrated systems onboard related to the project.

Figure 22 AT4-dev drone for initial stages of integration.

Figure 23 AT6 drone for integration and validation.

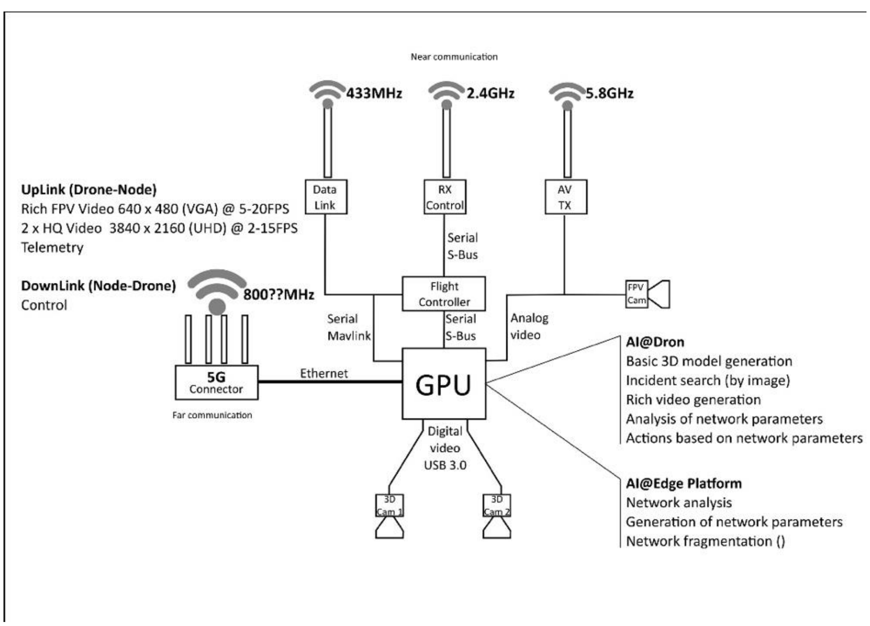

Figure 24 Functional scheme of the drone configuration.

The **AI TOOL ENVIRONMENT** for running the automated incidents detection tool is completing the test bed for UC3. Under this environment, several modules and a function based on Artificial Intelligence will be developed.

ATOS has developed a Computer Vision solution for the detection and classification of people and vehicles using real-time footage, that will serve as a basis for the development of an advanced solution to detect and locate incidents on roads.

This classification module has been created using a convolutional neural network, based on a MobileNet architecture and trained to detect up to 20 categories of "objects" like people, cars, trucks, etc. An example of a labelled image is shown in Figure 25. Due to the advances in object detection state of the art and the increasingly powerful edge computing devices that can be embedded on-board drones, new object detection architectures such as those based on YoloV5 that far exceed the accuracy of MobileNet models are intended to be used in the Use case.

The resulting module will be successfully run on a range of Edge Computing solutions using real-time footage coming from the drone camera.

Figure 25 Labelled images with identified objects by ATOS' classification module.

The detection of incidents related to the objects operating or moving within a changing and real environment is a challenging process that requires an innovative approach. In most of the cases, the training phase of a network is extremely difficult due to the high number of possible cases or the lack of skilled datasets for training. So, an innovative approach is to construct a 3D model of the monitored area and compare the recorded images to detect incidents and calculate their proper location.

These new modules and the function for detecting incidents will be developed during the working period for the UC3. They are named as follows:

- C2 Module,

- Object Detection Module,
- Video Module,
- SLAM Module,
- AIF Automated Incidents Detection Function.

4.3.1 Equipment development roadmap

With regard to equipment, 5TONIC will provide the necessary updates environment to operate in 5G communication network. With the current configuration at 5TONIC, 5G NSA and SA release 15 are supported, with coverage in 3.5 GHz band for the outdoor area and X3 room. The available radio bandwidth at 3.5 GHz band is 50 MHz. Nevertheless, by the beginning of 2022, due to a new regulation, the 3.5 GHz band in Spain will be re-ordered among the operators, and the available bandwidth for 5TONIC could rise to 100 MHz (*depending on Telefonica allowing the use of the full spectrum at 5TONIC) which would mean more performance and capacity. This re-ordering will cause a need for replacing some devices such as the street macro antenna at the site.

Connectivity and visibility of working devices will also be provided as well as the integration of the releases of the AI@EDGE platform during the process.

AERO will adapt the drone platforms according to the ongoing developments, starting from the AT4-dev drone with the systems required to test the different devices involved, up to the AT6 drone capable of integrating the whole set of developed systems:

- Communication and control,
- Payload (stereoscopic imagery),
- Navigation and control support (FPV signal),
- Video transfer,
- Enriching video unit.

With the aim of easing the development and integration process, AERO will work with a proprietary integration-platform to embed the different devices involved (shown in Figure 26).

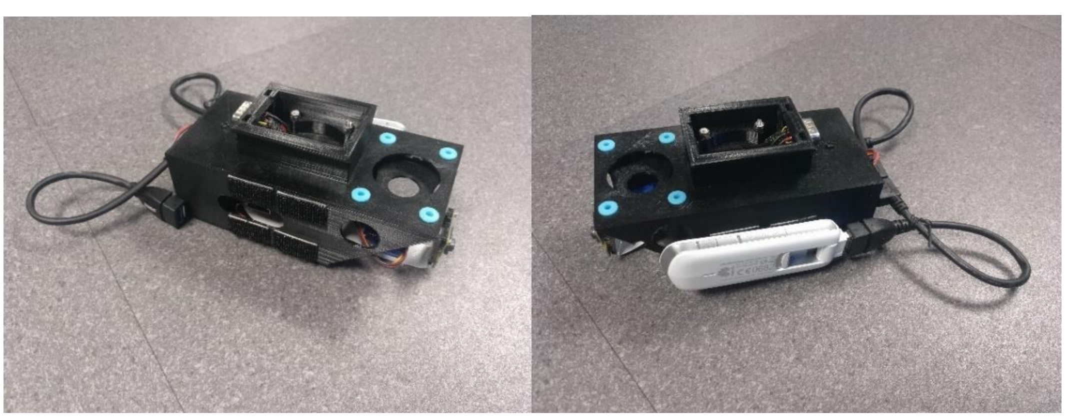

Figure 26 Different views of the integration platform that will support new devices.

This integration-platform can be connected to the drone systems (navigation, communication, etc, taking power from drone batteries and can also support computation devices or cameras for completing the integration process, as depicted in Figure 27 and Figure 28.

Figure 27 The integration platform onboard the drones to be used in the project.

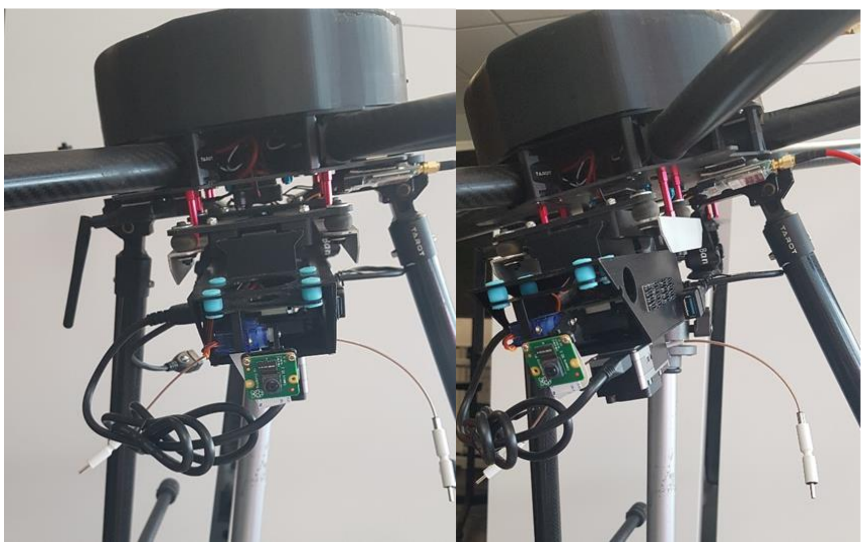

Figure 28 The integration platform onboard the AT6 drone with Raspberry Pi with integrated camera and video controller.

Preliminary roadmap

Phase 1 (to be completed by Q2 2022)

- Integration platform (drone & embedded systems) assembled and tested,
- 5G Network configuration tested,
- Systems and connections among 3 actors running for development.

Phase 2 (to be completed by Q1 2023)

- AT6 drone assembled and tested,
- Onboard systems integrated and tested,

- Adaptation of 5G Network in 5TONIC to specific requirements.

Phase 3 (to be completed in Q3 2023)

- Optimization of drone and embedded systems,
- Final configuration of 5G Network for testing.

4.3.2 Services development roadmap

ATOS will work on the development of services or software components included in the automated detection of incidents. This part has been planned with the above-mentioned approach, with different modules dealing with the specific needs of each part:

- **C2 Module, Command and control module:** This module will be responsible for receiving the drone control signals in Mavlink format in the ground station via the 5G network and sending them to the flight controller device. In addition, it will send the telemetry signals from ground station to the drone. It is important to highlight at this point that due to the criticality of this type of communication, it would be interesting to make use of a dedicated slice for this.
- The **Object Detection Module, Object detection and geo-positioning module**: ATOS will develop a Computer Vision solution for detecting and locating incidents in roads. This module will be based on a deep learning (DL) neural network, based on a YoloV5 architecture. The resulting module will be successfully run on any of the edge solutions using real-time footage coming from the drone camera. This tool will be deployed on the drone, will analyse the video acquired by the drone, identify the objects of interest, geo-referencing them and will send the events to be processed in the near edge by the Automated Incidents Detection AIF.
- The **Video streaming module**: The Use case considers two types of video streaming with different characteristics and functionalities. In both cases the video will be captured by the drone's on-board cameras and sent to the far edge where it will be analysed.
    - **Stream from FPV (First Person View) camera:** this stream is intended to provide the drone operator with real-time knowledge of the operating environment at all times. It is therefore necessary to prioritise transmission stability over transmission quality. This module will continuously interrogate the 5G network for bandwidth and availability parameters in order to reduce the quality of the video sent, if necessary, so that the broadcast is never interrupted.
    - **Stereoscopic camera stream:** this stream is intended for the generation of the 3D model of the infrastructure being inspected. It has to be of high quality and cannot be downgraded in case of network problems. In this case the system has to detect network saturation and initiate a process of recording the video to disk on board the drone and then transmit it when the network has sufficient capacity again or the connection is ideal again.
- **The SLAM module, SLAM simultaneously location and mapping module**: This module will calculate the position and orientation of the camera used to record the video that will be used as input to the 3D model. These data together with the data provided by the GPS of the drone and the stereoscopic video will be used by the Automated Incidents Detection AIF to generate the 3D model of the infrastructure being inspected.

Besides these modules to be integrated in the drone, a new function will also be developed and deployed on a Near Edge device, as closer as possible to the operating scenario.

- **The AIF for Automated incidents detection function**: This artificial intelligence function will be in charge of the automatic detection of incidents using the following procedure:
    - This AIF will be feed by the high-quality video stream and the telemetry data sent from the drone and the events sent by the object detection module
    - Using the video and the camera's position and orientation data, it will produce a high-precision 3D model of the infrastructure
    - By means of a data fusion layer, it will analyse the detected events and the 3D model, identifying possible incidences.

Preliminary roadmap

Phase 1 (to be completed by Q2 2022)

- Basics modules (C2, Object Detection and Video Streaming) installed and running.

Phase 2 (to be completed by Q1 2023)

- Basic modules iteration.
- SLAM & Automatic Detection Modules first iteration,

Phase 3 (to be completed by Q4 2023)

- Final iteration for all modules,
- Integration and testing.

4.4 Use case 4 test bed

UC4 targets to develop a testbed facility to demonstrate the concept of content curation for passengers on board civilian aircrafts relying on the infrastructure available at Safran Passenger Innovations Germany GmbH – SPI. Specifically, the cabin mock up, which reproduces a section of an A320 aircraft interior environment as shown in **Error! Reference source not found.** Partners of SPI for UC4 development, integration and testing activities are: SRS, ATH, CNAM, ITL, ICCS.

As mentioned in [1], SPI shall develop a test bed that includes two islands:

- A ground cloud infrastructure that is managed by means of Openstack VIM and
- An on-board cloud infrastructure, referred to as the Aero edge-cloud, that leverages a Kubernetes VIM; the edge-cloud has to be considered a 'Far Edge' as shown in Figure 2.

The two islands shall be separated by an emulated Geostationary satellite system that, preferably shall be implemented whereby the satcom emulator OpenSand [15], but in the minimum case it can be simulated with the Linux traffic control package for simple satellite link delay emulation. It is anyway planned that both islands shall be developed at the SPI premises in Wessling, near Munich.

The ground cloud infrastructure includes two Fujitsu Primergy RX200 S6 servers. Two of them mounting Intel Xeon X650 processor with 12 virtual CPU cores, CPU frequency 2.67 GHz, 32 GB RAM and 1 TB

HD internal storage, and two Fujitsu Primergy servers RX100 S7p mounting Intel Xeon E3-1230 V2 processor that has 8 virtual CPU cores, CPU frequency 3.30 GHz, 32 GB RAM and 1 TB HD internal storage. In addition, one Dell PowerEdge R6515 shall be added to the ground cloud.

The Aero edge-cloud test bed includes the following hardware components:

- 1 x SCU3 (system control unit) server, the third generation of aviation certified server manufactured by SPI with x86 processor architecture,
- RDU3 (removable display unit), the third generation of seatback screens manufactured by SPI that includes an ARM Cortex processors family and support for 4K video resolution,
- 1 x Supermicro server X10SDV-12C-TLN4F, which is a commercial-off-the shelf server with very compact form shape (i.e. similar to an SCU) with 12 CPU cores and 128 GB RAM,
- 1 x Fujitsu Primergy servers RX100 S7p for test bed management and not part of the Aero edge-cloud,
- 1 x TP-Link T2600G-52TS Ethernet Switch,
- 1 x SainSmart 16-Channel 12V Relay Module,
- RJ45 ethernet cables (i.e. 10 Gigabit Eth.) that provide overall connectivity to the test bed for functional purposes (i.e. to connect different machines).

The RAN that will be set up by SPI for UC4 shall include the following hardware components:

- USRP X310 for high performance in terms of time-frequency synchronization and high throughput,
- Jetway mother board with Intel Core i5-4300U and M.2 connected WiFi card, with the WiFi generation to that shall range from WiFi 4 to WiFi 6, depending on test bed evolutions,
- SPI manufactured RAVE WiFi access point.

The radio clients that will be integrated in the UC4 test bed shall include:

- Up to 20 x RDU3 that act also as clients for the passengers,
- 2 x 5G smartphones available from the off-the-shelf,
- WiFi dongles, for which the WiFi generation may change depending on market availability,
- 5G dongles will be further monitored in the mass market.

Figure 29 SPI A320 cabin mock up in Wessling (Munich – Germany).

4.4.1 Equipment development roadmap

The UC4 testbed currently developed at the SPI premise targets to reproduce a realistic IFEC infrastructure, deploying hardware that is currently delivered to many airlines that are customer of SPI.

The test bed that is being developed by SPI is shown in Figure 30 and Figure 31. Specifically, Figure 30 focuses on the RDU3 installation that is currently being set up to develop UC4 and where to integrate the AI@EDGE platform. It is worth noticing that each RDU3 contributes to the Aero edge-cloud since the seatback screens add to the pool of compute and storage resources, as well as they preserve the main functionality of rendering video catalogue and video contents to passengers. On the other hand, Figure 31 shows the other components of Aero edge-cloud that were illustrated. For example, the patch panel allows to finely connect multiple hardware components in the testbed, giving further possibility to develop different subnetworks (e.g. using VLANs). The SCU3 and the Supermicro servers also contribute to the Aero edge-cloud and once the system is fully operational, the content curation, 5GC network functions such as an UPF will be deployed, to name a few.

The RDU3 operating system (OS) is the embedded Linux PTXdist with docker engine installed. SPI has preliminarily identified this one to be the required set up to develop the Aero edge-cloud, while the SCU3 and the Supermicro will use an off-the-shelf Linux OS Ubuntu server 20.04 (Kernel version 5.4). We indeed remark that SPI will attempt to develop a particular type of edge-cloud that includes not only the servers, but also the seatback screens.

Hereinafter, we report the development roadmap for the UC4 partners.

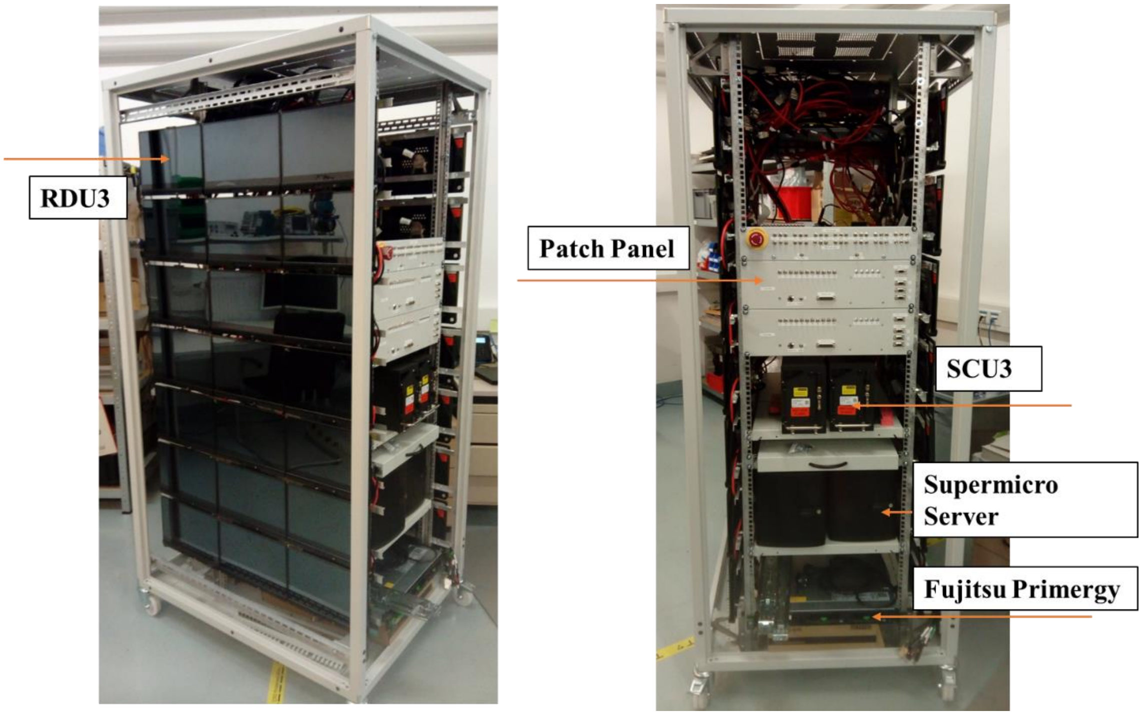

Figure 30 Use case 4 test rack overview.

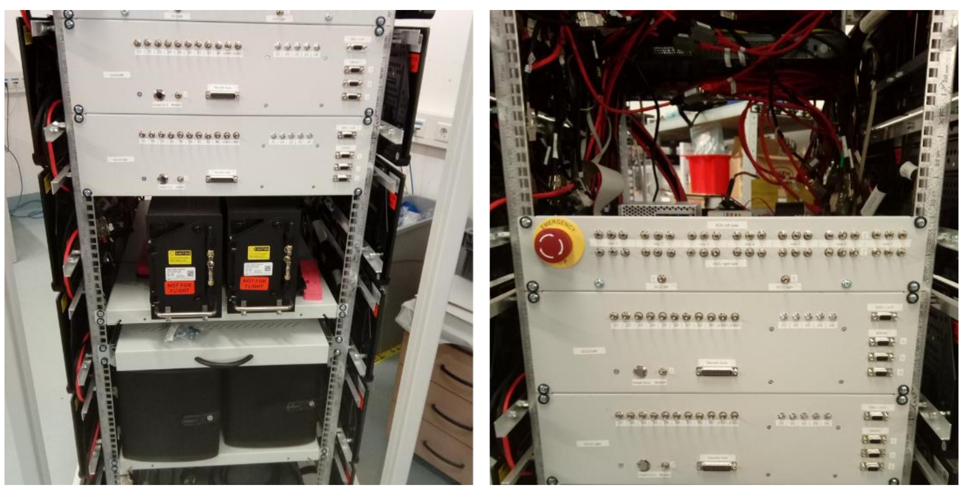

Figure 31 Use case 4 detailed view of the test rack.

At this early stage of the UC4 planning, ATH's 5GC functionalities (the UPF, cf. Section 3.2) within the on-board cloud infrastructure are envisaged to be deployed over a common-off-the-shelf Dell 240 server, with VMware as hypervisor and the following technical features:

- Intel Xeon CPU E-2146G 3.5 GHz 6C/12T,
- 32 GB RAM,
- 2 x 1 TB 7.2 K RPM SATA 6 Gbps 512n 3.5in Cabled Hard Drive,
- x 1 Gb Ethernet.

The CP functionalities, instead, will be deployed over a similar HW, located at FBK's premises, and shared for the purposes of UC 1, 2, and 4 (cf. Section 3.2).

In terms of HW acceleration equipment, ICCS can provide an acceleration card to demonstrate the computational power that can be available in such MEC scenario in the near future (for the current UC4 scenario, the accelerator can be viewed as an additive feature instead of a critical component of the demonstrator). Depending on final setup and performance specifications, the accelerator card can be installed either at the ground segment of UC4 (at FBK or ICCS premises via VPN connection), or at the on-board cluster. In the former case, the HW card will be installed at FBK/ICCS regardless of UC4 and the pending task will only be to define and perform the SW integration/connection to UC4 components. In the latter case, besides SW integration, the card should be installed to the on-board cluster by considering the following specifications (due to the availability of cards at ICCS):

- FPGA acceleration card for servers (Xilinx Alveo U200 or U280),
- PCIe interface, Gen3 x16 or Gen4 x8 (with extra network interfaces, 2x QSFP28),
- total power: up to 215W total electrical card load (with 150W fed from AUX connector),
- size: HxLxW=111.28x236x39.04-111.15x242x39.04, Weight=1066-1130g, passive cooling.

We note that the aforementioned cards have been successfully integrated and tested on a "supermicro" 1U server 1029GQ-TRT at ICCS premises. We also note that, alternatively, to avoid compatibility issues, a new card can be procured and integrated to the on-board cluster (the market cost ranges in 4-8 Keuro for November 2021).

In the UC4 scenario ITL will support the evaluation of a solution integrating a GPU (NVIDIA P4) as HW accelerator to increase performance when running the content curation AI in the aircraft environment. The evaluation will be performed initially at the ITL testbed and afterwards will be evaluated the possibility to install a GPU on the SPI testbed.

The industrial portable server at the ITL testbed is configured as follows:

- Mainboard with 2x PCI-E 3.0 x16, 2 PCI-E 3.0 x8,
- 2 x Intel® Xeon® Silver 4114, 10 Core, 2.2 – 3.0 GHz,
- 128GB DDR4 ECC memory,
- 1 x 256GB 2.5" SSD for OS (boot disk) removable,
- 1 x 1TeraByte 2.5" SSD, (storage) removable,
- GPU NVIDIA P4 (Low Profile, 75W, PCIe 16x),
- NIC Mellanox ConnectX-4 Lx EN NIC, 10GbE single-port SFP28,
- 2x USB3.0, 2x USB2.0, 2x LAN GbE.

The HW acceleration capabilities provided by the GPU may be used for supporting also video streaming and transcoding features dynamically adapting the video quality and the bandwidth used to the actual network conditions within the on-board infrastructure.

Preliminary Roadmap

Phase 1 (to be completed by Q1 2022)

- Complete assembling the Aero edge-cloud test bed (see Figure 30) and validate that hardware components are correctly set up,
- Complete installing the OS in each device and complete all configurations (ssh access, DNS, DHCP services, remote power on/off, etc.),
- Development of ITL test bed for video transcoding.

Phase 2 (to be completed in Q2 2022)

- Integration of Kubernetes in the Aero edge-cloud infrastructure,
- Contribute to set up the Layer-3 site-to-site VPN between the Aero edge-cloud and the FBK office in order to prepare for connecting the 5GC UP with the CP,
- Development of the test bed for the hardware acceleration card.

Phase 3 (to be completed in Q3 2022)

- Installation of the X310 SDR boards for 5G RAN,
- Installation of WiFi access
    - COTS WiFi access point
    - Aero certified WiFi access point,
- Test bed remote access by means of a VPN connection (i.e. Cisco Anyconnect client).

4.4.2 Services development roadmap

SPI shall develop several different services, some of which are enablers of the Aero edge-cloud as a whole, while some others provide the expected value-added services. The preliminary roadmap that is developed hereinafter hence does not include SPI developments, but also the ones of UC4 contributing partners. Generally, SPI will integrate in UC4 selected components of the AI@EDGE platform and building upon Kubernetes (e.g., Kubernetes version 1.21.2).

Since the Aero edge-cloud test bed that is being developed for UC4 is made of several components, SPI will make use of the SainSmart relay module to enable power control of the test bed from remote (i.e. electronically power on/off of the infrastructure as a whole, or a part thereof). SPI will create an independent subnetwork that includes all hardware components, RAN equipment and clients of the Aero edge-cloud, with the Fujitsu Primergy servers RX100 S7p that shall act as the test bed controller that provides also NAT service. This will be crucial to separate the Aero edge-cloud developed in UC4 from any other network segment and for which DNS and DHCP services will be set up. SPI will further deploy a software artefact (e.g., 5G-EmPOWER [16]) to enable a Near-RT RIC for the sake of 5G and WiFi RAN and, as mentioned, the Aero-edge cloud shall be separated also from other ground infrastructures whereby the emulation of a GEO satellite connection.

SPI shall deliver a value-added content curation AIF that is meant to provide content recommendation to passengers. For the sake of developing the content curation AIF, SPI will begin with analysing the owned IFE datasets collected over different flights of multiple airlines. The first step shall consist of pre-processing the dataset in order to provide full anonymity with respect to passengers, airlines, city pairs and time stamps, to name a few. Subsequently, a fraction of the anonymised dataset will be used on the one hand for applying machine learning offline, whereas the remaining part shall be used for online learning and decision making. The overall proposition consists of a content curation AIF that can take autonomous decisions in selecting content addressed to the passengers. The AIF shall be developed through docker containers and a serverless approach will be studied by SPI in the design. Currently, SPI envisages to split the content curation AIF into two parts: a frontend part that, as a container, runs in each seatback screen (or RDU3) and a backend part that runs in the Aero edge-cloud. Moreover, each seatback screen has to render the catalogue of selected contents for passenger's playback.

ATH will provide UC4 with a fully virtualized 5GC that will encompass all the core networking functionalities required by the Use case. ATH's 5GC is entirely developed and tested at ATH's premises and is compliant with 3GPP standard specifications (cf. [9]). At the moment of writing this deliverable, ATH's 5GC is available as a 5G Standalone (SA) solution. Along the 24 months of activity of T5.5, whenever opportune, the provided functionalities may be further updated according to the evolution of the standard itself in the scope of the project's Use cases, and ATH will upgrade the 5GC software deployed for UC 4 with its available new releases.

The HW accelerator of UC4 will implement a representative AIF of the final demonstrator to showcase the latency/throughput improvement that can be achieved in such use-cases. The function to be accelerated will be selected from the set of AIFs of UC4 based on function complexity and timing requirements. Representative AIFs for recommendation systems could be AI or ML, i.e., KNN, k-means, or CNN with autoencoder. After integration, the accelerated AIF will be able to be used on demand, either as part of the main user application or as a service.

ITL will provide the video streaming and transcoding application to allow the adaptation of the characteristics of the video streams according to the quality requested by each user, as well as to the network traffic conditions on board the aircraft. The application will also manage all data and configuration related to the users accessing the streaming services, by storing and maintaining updated the related data in the user database. The video transcoding feature can leverage the HW acceleration capabilities provided by GPUs.

SRS will provide UC4 with its 5G RAN via srsRAN, specifically srsENB and srsUE. Both srsENB and srsUE currently support 4G and 5G NSA. srsGNB will be made available in Q2 2022 and bring SA support to the RAN, in line with the srsRAN and project roadmaps. SA support for srsUE is expected at a later date. Both srseNB and srsUE are implemented in efficient and portable C/C++, the software supports a wide range of baseband hardware platforms including x86, ARM and PowerPC. srsENB supports functional split interfaces including splits 6, 7 and 8. Split 7 supports network deployment using commercial off-the-shelf ORAN RRU devices. The generic baseband I/Q interface allows srsUE to be used with most SDR front-ends including those based on Analog Devices and Lime Microsystems RFICs. Further modifications in line with the needs of UC4 may be implemented in an ad hoc fashion when needed, and when aligned with the roadmap of both the project and srsRAN.

About multi-connectivity, for the aggregation of multiple radio (WiFi, 4G and 5G) and wired (Ethernet) access technologies in UC4, the MPTCP proxy service is a network function that can be deployed as a physical network function, a virtualized network function or a container network function with the latter one preferred for deployment location the MEC host. The requirement is to be on the data plane path before reaching the application server. Different deployment modes will be studied, including positioning before, within and after the 5GC. The service consists in handling MPTCP sub-flow connections with the Ues (that have to be MPTCP capable), translating them to simple TCP connections with the application server. A TCP-level proxy under development at CNAM will be used for this purpose. It will integrate a predictive scheduler whose behaviour is to be influenced by runtime analysis of RAT measurements which, at the state being, are meant to be collected at the Near-RT RIC. The integration of the proxy can start as soon as the UC4 testbed becomes available and the design, improvement and demonstration of the different possible schedulers will be tested over T5.5 duration as detailed hereafter.

Preliminary Roadmap

Phase 1 (to be completed in Q1 2022)

- Output the first anonymized IFE dataset ready for use to develop the content curation AIF,
- Initial development of the video transcoding software (in lab development),
- Initial development of the MP-TCP proxy (in lab development).

Phase 2 (from Q2 2022 to Q3 2022)

- Development of 5G RAN SA (in lab development),
- Development of the content curation software (in lab development),
- Development of the hardware acceleration software (in lab development),
- Development of the video transcoding software,
- Development of the MP-TCP proxy.

Phase 3 (to be completed in Q4 2022)

- Maturation of the content curation software,
- Maturation of the video transcoding software,
- Maturation of the MP-TCP-proxy,
- Maturation of hardware acceleration software.

5 Integration, validation and test plans

In this section, the plans and roadmaps of each of the Use cases are discussed with some level of detail to provide an overview of the activities to be done in the following months towards the end of the AI@EDGE project.

5.1 Use case 1

This section reports on the preliminary plan for the integration activities, demonstrations and logistics that is currently envisaged for UC1.

5.1.1 Integration Roadmap

The 5G emulator at CRF premises will be deployed and tested together with some preliminary Telematic Box Validations (handover, authentication, etc.). In parallel, a first integration between the Cooperative Perception algorithms and the Traffic simulator will be performed. At this stage, only emulated vehicles will be considered. In this phase also a preliminary integration between the POLIMI Worldsim simulator and the traffic simulator will be performed.

When the Telematic Box is validated and the simulation environment ready, the next step will be the integration of the Telematic Box in the simulation environment. At this stage, real values (e.g., vehicle dynamics) collected from the emulated vehicles will be gathered and sent to the simulation environment by the Telematics box. The final integration will demonstrate the Cooperative Perception Algorithm that works both with emulated vehicles and the connected Telematic Box in the Validation environment. At POLIMI site the edge node will be prepared together with the RAN.

When POLIMI edge and RAN nodes are ready, it will be performed an integration between the Driving Simulator and the Telematics Box so the information gathered from the driving simulator will be sent to the Cooperative Perception Algorithm. At this stage also the FBK Central Site will be connected with the POLIMI Edge node.

The UC will rely for the MEC Orchestration on the Connect Compute platform elements deployed at FBK, which will act as the Central Office hosting the MEC Orchestrator. The Far Edge will be hosted at the main site of the UC. The FBK site will be connected to the main test site via VPN, in order to support the Orchestration of the AIFs on the MEC hosts at the far edge. As soon as the main site will be available, interoperability tests between the Far and Near Edge will start and will include connection availability; MEC host management; and AIF LCM (deployment, operation and termination).

Preliminary Plan

Phase 1 (to be completed by Q2 2022)

- Traffic Simulator – Cooperative Perception integration,
- Worldsim – Traffic Simulation Integration.

Phase 2 (from Q3 2022 to Q1 2023)

- 5G Emulator – Traffic Simulator integration (HIL).

Phase 3 (From Q2 2023 to Q4 2023)

- Edge-RAN Integration,

- Core-Edge Node Integration,
- Scenario Deploy from CRF Validation to POLIMI Edge Node,
- Driving Simulator – Telematic Box Integration.

5.1.2 Planned demonstrations and logistics

The demonstrator's plan takes an incremental approach of maturity as presented in the preliminary roadmap plan provided herein below.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022)

- Early demonstrator: it consists of the 5G emulator on CRF premises, a preliminary validation of the Telematic Box, first algorithm evaluation with a traffic simulator using only emulated vehicles.

Phase 2 (from Q3 2022 to Q1 2023)

- Mid Demonstrator (HIL): it consists of the 5G emulator on CRF premises connected with a Telematic Box added in the simulation environment, Cooperative Perception Algorithm that works both with emulated vehicles and the connected Telematic Box in the Validation environment.

Phase 3 (From Q2 2023 to Q4 2023)

- Final Demonstrator (HITL): it consists of the 5G emulator and validation system at CRF site, RAN and Edge at POLIMI connected with Telematic Box installed on the Driving Simulator, Algorithm completed with the Driving Simulator Agent that interacts with emulated ones. The algorithm validated on the validation site will be deployed and executed on the POLIMI site for the interaction with the driving simulator and taking into account Human behaviours.

5.1.3 Planned validation and test plans

At the end of each Phase validation and test will be performed to guarantee the correct integration between components, as well as end to end tests will be performed.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022): 5G Telematic Box Validation Methodology

- During this phase, a validation methodology will be defined to validate 5G Telematics Boxes.

Phase 2 (from Q3 2022 to Q1 2023): Mid Demonstrator Validation and test

- During this phase, the Telematic Box will be tested in the simulation environment and the correct data exchange will be tested. ATH's 5GC will be delivered to UC1 after internal functional and operational tests. Besides, the integration with the other elements of UC1's 5G system (Ues, RAN, applications) requires a series of tests that can be preliminarily divided into the following main categories: attach procedures, idle/connected state switch, detach procedures, tracking area update procedure, and traffic. These tests are planned to be carried out as soon as all the elements of the 5G system are functioning and available to the Use case, and indicatively within the first nine months of activity of T5.1. The results of such tests will be reported in the future deliverables of WP5.

Phase 3 (From Q2 2023 to Q4 2023): final demonstrator validation and test

- Driving Simulator and Telematics Box integration will be validated and will verify the data exchange between the driving simulator and the Cooperative Perception Algorithm.

5.2 Use case 2

This section reports on the preliminary plan for the integration activities, demonstrations and logistics that is currently envisaged for UC2.

5.2.1 Integration plan

The integration of the anomaly detection AIF is expected to be done at M18, once the UC2 testbed will be fully operational. Before this integration, preliminary testing will be done in a simulated/emulated ad-hoc environment, against UC2 testbed data collected in a reference scenario experiment.

The integration of the NetFPGA boards for the SmartNIC function and their exploitation for the anomaly detection AIF and intrusion detection algorithms is planned for M22-M25, with dedicated 1-week visits of CNAM staff at DFKI.

ATH edge node with user plane functionalities (UPF) for local breakout will need to be integrated into DFKIs near edge once around M26.

Preliminary Plan

Phase 1 (to be completed by Q2 2022)

- Integration of anomaly Detection AIF

Phase 2 (from Q3 2022 to Q4 2022)

- Integration of NetFPGA for SmartNIC

Phase 3 (From Q1 2023 to Q2 2023)

- ATH integration of edge node UPF for local breakout.

5.2.2 Planned demonstrations and logistics

Both the centralized anomaly detection AIF and the federated learning-based anomaly detection AIF will be demonstrated. The two modes will be compared in terms of learning time, detection time and attack detection precision time. We expect the centralized anomaly detection AIF to be demonstrated for M20, and the federated learning variant for M25.

Preliminary Plan

Phase 1 (to be completed by Q2 2022)

- Demonstration of AGV and localization sensors over edge cloud

Phase 2 (from Q3 2022 to Q4 2022)

- Demonstration of anomaly detection AIF

Phase 3 (From Q1 2023 to Q2 2023)

- Demonstration of federated learning based anomaly detection AIF.

5.2.3 Planned validation and test plans

As described in the D2.1 deliverable, recent technologies have been integrated within Industry 4.0 implementations, bringing new challenges in the cybersecurity area. In the next paragraph we introduce CyberBattleSim which is a potential solution that we can consider to generate attack scenarios is the context of Use case 2.

The standard way of implementing machine learning models that are dedicated to detect or mitigate attackers in the cyber domain is to rely on publicly available cybersecurity datasets for their training. By using state of the art machine learning algorithms, it is possible to reach an acceptable accuracy. Nevertheless, these static datasets are unlikely to help quickly and advanced autonomous decision making and response. For this purpose, Microsoft has released a simulated environment, called CyberBattleSim, based on the OpenAI gym library. CyberBattleSim "provides a way to build a highly abstract simulation of complexity of computer systems, making it possible to frame cybersecurity challenges in the context of reinforcement learning". Reinforcement learning is a machine learning technique with which autonomous agents learn how to conduct decision-making by interacting with their environment. This dynamic environment allows to train an autonomous agent to learn to take optimal actions to mitigate cyberattacks. Thus, this solution allows to create scenarios and insert custom autonomous (offensive or defensive) agents for training and testing.

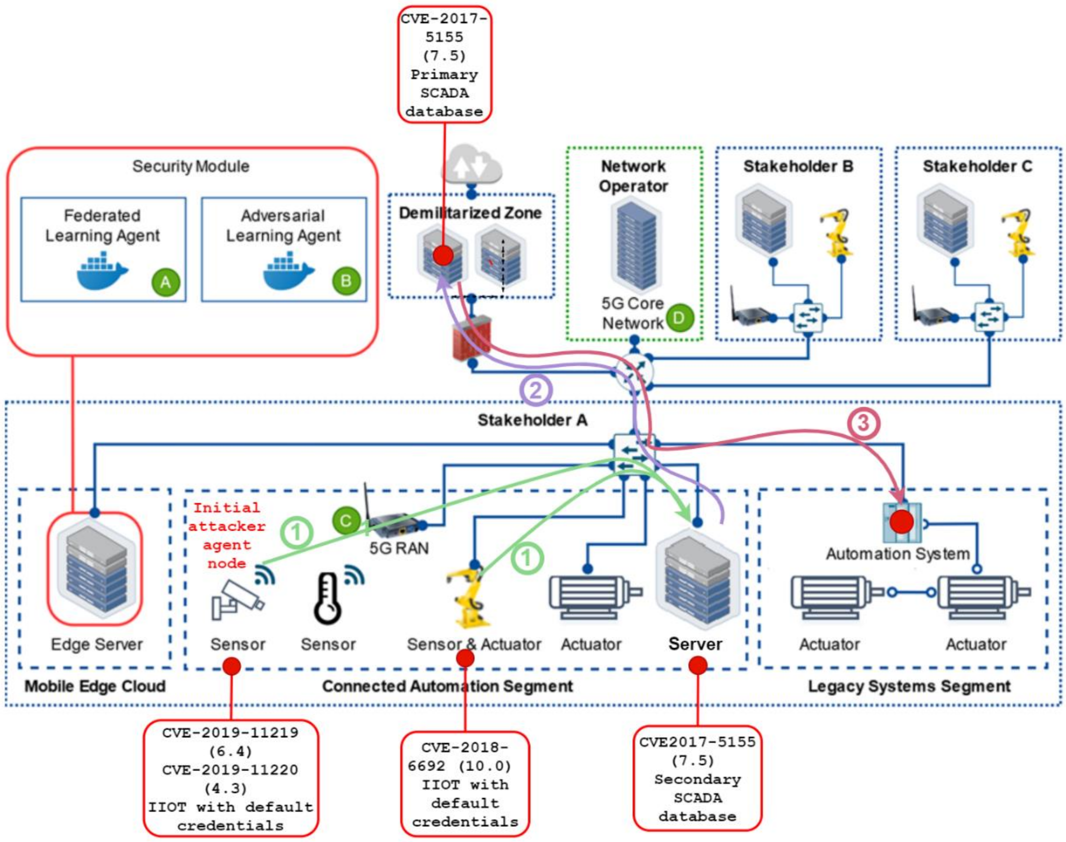

Figure 32 Example scenario with various vulnerabilities on different components in an industrial network.

The Figure 32 is an industrial network including IIoTs (Industrial IOT) (sensor, actuator, IP camera for monitoring, ...), traditional IT (not IOT systems) and SCADA equipment (automation system). The primary SCADA database is isolated from the whole network in a demilitarized zone by a firewall. The Figure shows the details of this network and the vulnerabilities associated with its components.

Using CyberBattleSim, a simulated attacker can exploit the predefined vulnerabilities to laterally move through the network. The goal is to own portions of the network by exploiting the discovered vulnerabilities. It also provides a way to train a defender agent that can learn by observing the industrial network behaviour and try to detect the presence of the attacker and mitigate the attack. Next, we provide an illustration of a potential attack simulated with the CyberBattleSim tool. Since it's a reinforcement learning based solution, we will use its terminology to describe the actions, the states and the reward of the agent. The attack scenario starts with a post-breach assumption where the two IIOT devices are compromised by an attacker exploiting one of the three vulnerabilities: CVE-2018-6692, CVE-2019-11219 or CVE- 2019-11220. These three vulnerabilities rely on the same underlying weakness, the use of known default credentials. The simulated attacker's goal is to maximize the cumulative reward by discovering and owning nodes in the network. The environment is partially observable: the agent, referring to the simulated attacker, is not able to observe all the nodes (devices and servers) and edges of the network (path to the DMZ and path to the automation system) in advance. Instead, the attacker takes actions to gradually explore the network from the currently owned nodes (a sensor). There are three types of actions an agent can perform: local attack, remote attack, and connecting to other nodes. In our scenario, the simulated attacker will combine this three types of actions to try to infect the automation system which is the node with the best reward. The CVE-2017-5155 targets a log database of a SCADA equipment using default credentials. Having taken control of the previously mentioned IIoTs and having gained access to the network, the attacker can exploit the weakness of this database, and thus compromise this device (1). If the device under the attacker's control is connected to the rest of the company's infrastructure, an in-depth takeover of the company's network and the IIoT network is possible, despite the presence of a firewall (2). After owning the primary SCADA database, the attacker proceeds with lateral movement to the automation system after steeling cached credential (3).

Preliminary Plan

Phase 1 (to be completed by Q2 2022)

- First basic operational UC2 testbed without hardware accelerators and AIFs.

Phase 2 (from Q3 2022 to Q4 2022)

- Demonstration of the centralized-mode anomaly detection AIF against injected attacks and failures.
- Demonstration of Smart-NIC-based attack detection and mitigation.

Phase 3 (From Q1 2023 to Q2 2023)

- Demonstration of the distributed-mode anomaly detection AIF against injected attacks and failures.

5.3 Use case 3

This section reports on the preliminary plan for the integration activities, demonstrations and logistics that is currently envisaged for UC3.

5.3.1 Integration plan

For the development and integration of the different functionalities described in this chapter, three phases have been identified and the process has been planned accordingly.

Preliminary Roadmap

Phase 1 (to be completed by Q2 2022)

This first phase is planned to prepare the coordination among the different systems and workflows. 5TONIC is to prepare development environment installing the required devices to emulate the working environment (drone + central office + drone operator) and to get drone control communication (C2) and video from the drone using the communication network currently operating in 5TONIC (4G, 5G NSA…). Additionally, the 5G SA Core will need an upgrade with the latest product enhancements.

The three developing actors will deploy a valuable DEVELOPMENT ENVIRONMENT connecting the 3 development nodes (AERO in Madrid, ATOS in Zaragoza and 5TONIC) to allow continuous and efficient development during this period. In parallel, ATOS and AERO will work in the development and integration of the modules and function for the operation. A dedicated channel for collaborative work will be kept during the whole process. AERO will provide the required drone platforms to integrate isolated systems (computing, communication) or cameras, as well as combined systems to be tested.

Phase 2 (from Q3 2022 to Q1 2023)

In this phase, the first iteration of the ongoing developments will be tested.

The results of the WP3 and WP4 and the 2022 version of the AI@EDGE platform released will be integrated to carry our connectivity and operation related to the range of scenarios stated in the Testcase table. 5TONIC will ensure for this phase that the 5G SA release 16 + Slicing is supported. The achievements on the modules and functions by ATOS and AERO will be also implemented and tested.

Phase 3 (from Q2 2023 to Q4 2023)

In the last phase, the second iteration of developments and the final version of the AI@EDGE platform will be deployed and tested. Validation tests will also be performed.

At the current state of planning UC3 activities, it is expected the possibility in 5TONIC to run a Dual Core (5G EPC + 5GC) that enables the option to experiment technologies interworking.

The results of the WP3 and WP4 and the 2023 version of the AI@EDGE platform will be integrated to carry out connectivity and operation related to the range of scenarios stated in the Test Case (Table 2). The development of the modules and functions by ATOS and AERO will be finished by then and therefore also implemented and tested. All developments shall be integrated in the test bed, and validation tests will be carried out.

5.3.2 Planned demonstrations and logistics

Phase 1 (from Q3 2022 and Q3 2023)

- Initial demonstration of 5GC and 5G SA RAN. Measurement of throughput, latency and packet loss metrics.

Phase 2 (from Q4 2022 to Q1 2023)

- Advanced demonstration of 4G/5G NSA/5G SA Rel. 16 in 2.6 GHz (4G), 3.5 GHz and 26 GHz bands.
- Advanced demonstration of 5GC and 5G SA RAN. Measurement of throughput, latency and packet loss metrics.

Phase 3 (to be completed in Q4 2023)

- Demonstration loss in 5G coverage, handed over to 4G
- Demonstration of UC3 test bed integrated AIFs, performance benchmark and collection of marked KPIs,
- Live demonstration or video recording of complete UC3 functionalities.

5.3.3 Planned validation and test plans

In order to have a consistent plan for testing at different stages of the development, a Test Cases table has been defined, contributing to improve the coordination among the partners as well. This Table will include the following fields that define the range of tests to be carried out.

Table 2 Use case 3 Test Cases table for detailed definition of testing procedures.

Test Case #	**Name of the test**
Slogan & Objective	Goals to be achieved with the test
Test Scenario (Pre-conditions)	Conditions to run the test
Expected Results (Post-Conditions)	Expected results according to KPI´s

General Time Plan (Validation Campaigns)	Scheduled period for the test
Test Sequence	Tasks for the test

For Phase 1, the proposed Test Cases will be run with the existing infrastructure and equipment. The validation will be performed mainly using 5G SA Rel. 15 in 3.5 GHz band. Measurement of throughput, latency and packet loss metrics.

In Phase 2, the edge applications will be deployed at 5TONIC Data Centre and ensured connectivity to them. For this purpose, a new antenna for the 26 GHz band could be installed to cover the outdoor and X3 areas. So, the tests can be run and benchmarked using the different technologies 4G/5G NSA/5G SA Rel. 16 in 2.6 GHz (4G), 3.5 GHz and 26 GHz bands.

In Phase 3, there will be the possibility to add the use of the 700 MHz band to the mix, for which a new antenna supporting this band will be required at 5TONIC. Additionally, it is expected to have the Dual Core available at 5TONIC, allowing to test the specific case of the operation not to be interrupted despite a loss in 5G coverage, as the service could be handed over to 4G.

5.4 Use case 4

This section reports on the preliminary plan for the integration activities, demonstrations and logistics that is currently envisaged for UC4.

5.4.1 Integration plan

The integration activities reported hereinafter shall be performed by SPI together with UC4 partners.

Preliminary Roadmap

Phase 1 (to be completed in Q3 2022)

- Integration of 5GC (i.e. CP and UP),
- Integration of 5G NSA,
- Integration of 5G UEs with programmable SIM cards (e.g. sysmocom),
- Integration of a Near-Real Time RIC.

Phase 2 (from Q4 2022 to Q1 2023)

- Initial integration of selected components of the AI@EDGE platform,
- Integration of the 5G SA RAN,
- Initial integration of the content curation AIF,
- Initial integration of video transcoding,
- Initial integration of MP-TCP proxy AIF,
- Initial integration of hardware acceleration.

Phase 3 (to be completed in Q2 2023)

- Complete integration of AI@EDGE platform components,
- Advanced integration of the content curation AIF,
- Advanced integration of the video transcoding,
- Advanced integration of MP-TCP proxy AIF,
- Advanced integration of hardware acceleration.

5.4.2 Planned demonstrations and logistics

We remark that UC4 demonstration activities will take place by means of the SPI developed and maintained Aero edge-cloud and in the SP owned A320 cabin mock up. Anyway, initial development and validation activities will take place in the lab at each individual UC4 participants.

Phase 1 (between Q4 2022 and Q2 2023)

- Demonstration of 5GC and 5G NSA RAN,
- Initial demonstration of 5GC and 5G SA RAN,
- Initial demonstration of content curation AIF,
- Initial demonstration of video transcoding,
- Initial demonstration of MP-TCP proxy AIF,
- Initial demonstration of hardware acceleration.

Phase 2 (to be completed in Q3 2023)

- Advanced demonstration of 5GC and 5G RAN,
- Advanced demonstration of content curation AIF,
- Advanced demonstration of video transcoding,
- Advanced demonstration of MP-TCP proxy AIF,
- Advanced demonstration of hardware acceleration.

Phase 3 (to be completed in Q4 2023)

- Demonstration of UC4 test bed integrated AIFs, performance benchmark and collection of relevant KPIs,
- Live demonstration or video recording of complete UC4 functionalities.

5.4.3 Planned validation and test plans

This section describes the validation and test plans that are envisaged for UC4 at the current state of development of the SPI test bed, as well as of the components under development at the UC4 partners. All the validation and test plans are summarized in Table 3.

Table 3 Use case 4 validation and test plans.

UC4 validation and test activities					
Test Case #	Slogan& Objective	Test Scenario(Pre-conditions)	Expected Results(Post-Conditions)	General Time Plan (Validation Campaigns)	Test Sequence
Aero Edge-cloud	Validation of infrastructure	SPI test bed	UC4 test bed ready for use	Q1 2022	Complete hardware set up, OS installation and scripts configuration
5G UE attach	Validation of end to end connectivity	SPI test bed and 5GC-CP at FBK	5G system health check	Q3 2022	5GC spin up, 5G RAN power on, 5G manual network selection
5G UE idle/connected state switch	Validation of 5G N1 interface	SPI test bed and 5GC-CP at FBK	5G system health check	Q3 2022	Validate the UE get connected
5GC tracking area updated	Validation of 5GC set up	SPI test bed and 5GC-Cp at FBK	5G system health check	Q3 2022	Validate AMF
5G Network throughput	Benchmark 5G connection	SPI test bed	5G system KPI measurement	Q3-Q4 2022	Validate maximum throughput is at least 200 Mbit/s
Content curation AIF	IFE dataset processing	SPI test bed	Dataset ready for use	Q1 2022	Base on SPI owned IDE datasets
Content curation AIF	Machine learning algorithm, software implementation and frontend-backend connection	SPI test bed	Software implementing a machine learning algorithm for analytics to the pre-processed IFE dataset and output rendered on the RDU3	Q1-Q4 2022	Initial selection of existing machine learning implementations (e.g. python based)
Video transcoding AIF	Validation of the first software release	ITL test bed	Software validation	Q4 2022	TBD

Hardware acceleration	Validation of the first software release	ICCS test bed	Software validation	Q3-Q4 2022	TBD
Multi-connectivity Throughput	Higher user access rate than with only 5G	AI@EDGE platform 2022 release	Aggregated Throughput increased by 50%.	Q4 2022	Tests with stationary conditions. Tests with changing radio conditions.
Multi-connectivity Reliability	Imperceptible RAT fault recovery	AI@EDGE platform 2023 release	Recovery time decreased to the 50 ms target.	Q4 2022	Tests with stationary conditions. Tests with changing radio conditions.

6 Risk analysis and mitigation actions

This section provides a risk analysis for the four testbeds, alongside related activity, associated risk level and mitigation action. It should be further specified that the tables presented hereinafter provide an initial assessment of the risks, which will be continuously updated as the project progress. The risk analysis for Use cases 1 – 4 are reported in Table 4, Table 5, Table 6, and Table 7, respectively.

6.1 Use case 1

This section provides a risk analysis for the development of UC1 to the extent that can be forecasted at the current stage of the project.

Table 4 Risk Analysis for Use case 1 and development of mitigation actions.

Test Bed stage	Activity	Risk	Risk level	Mitigation
Phase 1	Worldsim – Traffic Simulation Integration	The latency introduced by the integration makes Worldsim simulation slow.	Medium/ High	Some troubleshooting can be adopted based on the latency amount (reduce the number of simulated vehicles or the amount of sensor data)
Phase 1	5G Emulator Equipment provisioning	Chip Shortage and Purchasing process can slow the delivery and the 5G emulator setup	Medium/ High	Use of software emulators (e.g.: ns-3)
Phase 2	Telematics Box provisioning	Hardware and Software instability due to the prototypal box.	Medium/ High	Work closely with Tiers 1 to fix bugs and issues.
Phase 3	Driving Simulator – Telematic Box Integration	Risks on the absence or compatibility of the signals/information managed by both systems (Driving Simulator, telematic Box).	Medium/ High	Signal simulation or estimate some information.

6.2 Use case 2

This section provides a risk analysis for the development of UC2 to the extent that can be forecasted at the current stage of the project.

Table 5 Risk Analysis for Use case 2 and development of mitigation actions.

Test Bed stage	Activity	Risk	Risk level	Mitigation

Phase 1	5G OpenRAN set-up	The System will not work due to faults in the Equipment for RAN and Core.	Low	Support from the respective vendors. Fallback: use equipment from project partners
Phase 2	Anomaly detection on sensor data	Sensor data might not be sufficient for anomaly detection	Low	Look for different data sources in the testbed.
Phase 2	NetFPGA board integration	The PCI-based API reveals to be instable using the UC2 edge servers	Medium/ High	The serial connection will be used for board configuration. An ad-hoc SBI will be used for metrics collection.
Phase 2	Anomaly detection AIF integration	Learning time too high with the provided computing resources on-site	Low/Medium	A restriction on the set of metrics used by the LSTM cells will be applied.
Phase 3	Federated Anomaly Detection AIF	Results may show that merging distributed models might decrease the detection success.	Medium	Find a hybrid solution for merging models without decreasing accuracy.

6.3 Use case 3

This section provides a risk analysis for the development of UC3 to the extent that can be forecasted at the current stage of the project.

Table 6 Risk Analysis for Use case 3 and development of mitigation actions.

Test Bed stage	**Activity**	**Risk**	**Risk level**	**Mitigation**
Phase 1	Hardware Integration	Hardware: Weight and size of the physical device associated with the AI@EDGE platform to be integrated onboard allowing an optimal operation. Physical devices associated with the AI@EDGE platform have high power consumption, preventing acceptable flight time.	Low	If the physical device associated with the AI@EDGE platform cannot be integrated onboard (either by weight and size or power consumption) its functions could be, partially or totally, offloaded to a ground station.
Phase 3	Connection to 5G Network	Scenario: Stability of radio access connections linked to the status of 5G deployment in the selected area for the Use case.	Medium	Focus on AI@EDGE platform features that better suit the Use case and

				its technical requirements.
Phase 2	Testing of AI@EDGE functionalities	Access to full range of functionalities: Related to network deployment in the selected area: 5G connexion is NSA instead of SA.	Medium	If the 5G status on the Use case area is not optimal to efficiently perform the operation, repeaters to extend 5G coverage will be deployed.
Phase 1	Hardware Integration	Equipment on board: Interferences with other drone equipment (controller, GPS satellites, etc).	Low	Pre-checks and damping measures in case of detected interferences or different positioning of the equipment inside the drone.
Phase 3	Operation within 5G Network	Connection: Data corruption between the operator and the drone prevents proper control of the last one because of the state of 5G connection.	Medium	Automated procedures will be applied until the drone regains connection i.e., intelligent RTL (Return To Launch) function will be fully available.

6.4 Use case 4

This section provides a risk analysis for the development of UC4 to the extent that can be forecasted at the current stage of the project.

Table 7 Risk Analysis for Use case 4 and development of mitigation actions.

Test Bed stage	Activity	Risk	Risk level	Mitigation
Phase 1	Integration of the 5GC with the 5G RAN equipment	End-to-end interconnectivity failure that may delay the UC activities	Low	Proactively start this integration as soon as possible, leveraging the remote integration solutions that both ATH and SRS have available and can be exploited before the actual deployment of the radio access and

				core network components at UC4's site.
Phase 1	Integration of the MP-TCP proxy in the function chain	Too high forwarding delay or packet loss due to virtual switches for high bitrates	Medium	Deploy the MP-TCP proxy as a physical network function
Phase 2	Integration of AI@EDGE platform components	Delays in software development causes delays in UC4 developments	Low/ Medium	The basic Kubernetes installation will allow continuing with UC4 developments and tests
Phase 2	Evaluation of interworking between Near-RT RIC and MP-TCP proxy scheduler	gNodeB data processing too low at near-RT RIC to have actual prediction for the scheduler useful	Medium	Run the eNodeB and/or WiFi AP data measurement processing outside the Near-RT RIC, as an independent server
Phase 2	Content curation AIF	Software development, containerization require longer time that expected	Low/Medium	SPI will cooperate also with other SPI teams whenever needed to overcome technical hurdles
Phase 3	Integration of video transcoding in the SPI test bed	Integration requires longer time than planned due high complexity	Low	ITL and SPI respective competence will minimize this risk
Phase 3	Integration of hardware acceleration in the SPI test bed	Integration of acceleration with the content curation AIF turns more complex than expected	Medium/ High	SPI, ICCS and ITL will cooperate continuously to minimize this risk, including sending experts at the SPI premise
Phase 3	Performance of 5G network is below expectations	5GC-RAN-UE chain requires additional investigations and tests	Medium/ High	A trade-off between time to improve performance and acceptable performance to carry on with the Use case will be identified

7 Conclusions

Deliverable D5.1 reports on the planning, preparation and validation methodology for the demonstration of the four Use cases tackled by AI@EDGE. Throughout this document, each sub-system was described, the piloting methodology was illustrated and the KPIs were addressed. The deliverable is the first outcome of WP5 in which the AI@EDGE Connect-Compute platform design has been taken as an input to devise the overall methodology that shall be used on a per-Use case basis for the integration of the AI@EDGE components. Moreover, taking the KPIs defined in Task 2.3 as input, each Use case has defined, generally, the data and meta-data to be collected and evaluated in order to develop artificial intelligent functions, which are used to validate the AI@EDGE platform in real-life and simulated environments. In D5.1 we introduced also the testbeds status for all Use cases, their roadmap, integration, validation and test plans, as well as an initial risk analysis as the starting point for upcoming WP5 activities. Furthermore, the evaluation procedure, the test and validation schedule, and the related milestones have been defined in this deliverable considering that a continuous revision will be carried out by the Use case leaders and theirs partners in order to meet the planned deliveries. Finally, taking as an input the AI@EDGE platform developed in WP4 and the KPIs defined in WP2, this report paves the road for upcoming milestones regarding the initial, intermediate and final demonstration of the four Use cases.

References

[1] S. Secci and et al., “D2.1: Use cases, Requirements, and Preliminary System Architecture,” H2020 AI@EDGE, 2021.

[2] B. Mafakheri and e. al., "CarLink: A Real-Time V2X Validation Platform for a Safe and Sustainable Mobility.," 2021. [Online]. Available: https://www.techrxiv.org/articles/preprint/CarLink_A_Real-Time_V2X_Validation_Platform_for_a_Safe_and_Sustainable_Mobility/16922827.

[3] B. Mafakheri and e. al., "Optimizations for Hardware-in-the-Loop-Based V2X Validation Platforms," in *IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, Helsinki, Finland, 2021.

[4] M. Catalan-cid and et al., "D2.2: Preliminary assessment of system architecture, interfaces specifications, and techno-economic analysis," H2020 AI@EDGE, 2021.

[5] F. Brito and e. al., "D3.1 Initial report of systems and methods for AI@EDGE platform automation," H2020 AI@EDGE, 2021.

[6] C. Phung and e. al., "D4.1 Design and initial prototype of the AI@EDGE connect-compute platform," H2020 AI@EDGE, 2021.

[7] "Multi-access edge computing (MEC) framework and reference architecture. ETSI GS MEC V2.1.1," ETSI, Sophia Antipolis Cedex - FRANCE, 2019.

[8] "Kubernetes: an open-source container-orchestration system for automating computer application deployment, scaling, and management.," [Online]. Available: https://kubernetes.io/. [Accessed 2021].

[9] „“Technical Specification Group Services and System Aspects; System architecture for the 5G System (5GS); Stage 2 (Release 15), 3GPP TS 23.501 V15.12.0,“ 3GPP, 2020.

[10] "VI-WorldSim: A new, user-friendly, fully integrated graphic environment that accelerates vehicle development offline and on Driving Simulators," [Online]. Available: https://www.vi-grade.com/en/products/vi-worldsim/. [Accessed 2021].

[11] "NetFPGA: a platform to develop network hardware for reconfigurable and rapid prototyping," [Online]. Available: https://netfpga.org/. [Accessed 2021].

[12] "SUMO: Simulation of Urban MObility," [Online]. Available: https://github.com/eclipse/sumo. [Accessed 2021].

[13] "MATSim: An open-source framework for implementing large-scale agent-based transport simulations," [Online]. Available: https://www.matsim.org/. [Accessed 2021].

[14] "SUMO-RL: A simple interface to instantiate Reinforcement Learning environments with SUMO for Traffic Signal Control," [Online]. Available: https://github.com/LucasAlegre/sumo-rl. [Accessed 2021].

[15] "OpenSand, a user-friendly and efficient tool to emulate satellite communication systems," Thales Alenia Space, [Online]. Available: https://opensand.org/content/overview.php. [Accessed 2021].

[16] "5G-EmPOWER: a Mobile Edge Computing Operating System supporting lightweight virtualization and heterogeneous radio access technologies," [Online]. Available: https://5g-empower.io/. [Accessed 2021].