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

D5.2 Preliminary Validation and Use Case Benchmarking

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

D5.2 Preliminary Validation and Use Case Benchmarking	
WP	WP5 – Use cases integration, validation, and benchmarking
Responsible partner	DFKI
Version	V1.0
Editor(s)	Daniel Reti (DFKI)
Author(s)	Daniel Reti (DFKI), Antonino Albanese (ITL), Mario Paetta (CNAM), Stefano Secci (CNAM), Christina Costa (FBK), Omar Anser (INRIA), Nicola di Pietro (ATH), Marco Marchetti (CRF), Babak Mafakheri (SPI), Miguel Rosa (AERO), Mariangela Donato (CRF), Alessandro Roccasalva (CRF), Phung Ch-Dung (CNAM), Clavenna Antonella (ITL), Brendan McAuliffe (SRS)
Reviewer(s)	Javier Palomares, Estefania Coronado (I2Cat), Flávio Brito (EAB), George Avdikos (8Bells), Roberto Riggio (UNIVPM), Arfan Haider Wahla (FBK), Raman Kazhamiakin (FBK), Irene Facchin (FBK)
Deliverable Type	R
Dissemination Level	PU
Due date of delivery	30/06/2022
Submission date	29/07/2022

Version History				
Version	Date	Author	Partner	Description
0.0	15/04/2022	Daniel Reti, Antonino Albanese	DFKI, ITL	Table of Contents and editing assignment

0.1	06/05/2022	Mario Paetta, Stefano Secci	CNAM	Inputs to Section 4
0.2	09/05/2022	Christina Costa, Antonino Albanese	FBK, ITL	Inputs to Section 2
0.3	17/05/2022	Omar Anser, Nicola di Pietro, Marco Marchetti	INRIA ATH, CRF	Inputs to Section 3
0.4	18/05/2022	Babak Mafakheri	SPI	Inputs to Section 6
0.5	25/05/2022	Daniel Reti, Miguel Rosa, Marco Marchetti, Mariangela Donato, Alessandro Roccasalva Phung Ch-Dung, Babak Mafakheri, Nicola di Pietro	DFKI, CRF, AERO, CNAM, ATH, SPI	Inputs to Section 5 Update Sections 3, 4, 6
0.6	01/06/2022	Stefano Secci, Clavenna Antonella, Phung Ch-Dung, Marco Marchetti, Babak Mafakheri, Antonino Albanese, Mariangela Donato, Alessandro Roccasalva	CNAM, CRF, ITL, FBK, SPI	Update Sections 3, 4, 6
0.7	08/06/2022	Babak Mafakheri, Omar Anser	SPI	Update Section 6
0.8	14/06/2022	Daniel Reti, Albanese Antonino	DFKI, ITL	Update Section 4
0.9	17/06/2022	Cristina Costa, Brendan McAuliffe, Marco Marchetti, Mariangela Donato, Alessandro Roccasalva	FBK, SRS, CRF	Update Section 3
0.10	22/06/2022	Antonino Albanese, Cristina Costa, Phung Chi.Dung, Marco Marchetti, Miguel Rosa,	ITL, CNAM, FBK, DFKI, CRF, AERO, SPI	Finalize Sections 3, 4, 5, 6

		Babak Mafakheri, Daniel Reti, Antonella Clavenna, Stefano Secci, Mario Patetta, Alessandro Roccasalva		
0.11	23/06/2022	Daniel Reti	DFKI	Clean version – Ready for internal review
0.12	28/06/2022	Javier Palomares, Estefania Coronado	I2CAT	Internal review by I2Cat
0.13	05/07/2022	Flávio Brito	EAB	Internal review by EAB
0.14	06/07/2022	George Avdikos	8Bells	Internal review by 8Bells
0.15	08/07/2022	Roberto Riggio	UNIVPM	Review by the Technical Coordinator
0.16	20/07/2022	Arfan Haider Wahla, Raman Kazhamiakin	FBK	Final revision and addressing
1.0	28/07/2022	Daniel Reti	DFKI	Ready for the submission
	29/07/2022	Irene Facchin	FBK	Quality and final check

Table of Contents

List of Table

List of Figures

Glossary

4G	4th Generation of mobile communications
5G	5th 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
AIF	Artificial Intelligence Function
ALS	Alternating Least Square
AMF	Access and mobility Management Function
AMQP	Advanced Message Queuing Protocol
BVLOS	Beyond Visual Line of Sight
C2	Command and Control
CP	Control Plane
CPU	Central Processing Unit
CRF	Centro Ricerche Fiat
DFKI	Deutsches Forschungszentrum für Künstliche Intelligenz
DN	Data Network
DNN	Data Network Name
DUT	Device Under Test
FDD	Frequency-Division Duplexing
gNB	gNodeB (5G base station)

ICMP	Internet Control Message Protocol
IFEC	Inflight Entertainment and Connectivity
IIoT	Industrial Internet of Things
KPI	Key Performance Indicator
MEC	Multi-access Edge Computing
MEO	MEC Orchestrator
MIMO	Multiple-Input and Multiple-Output
MP-TCP	Multi-Path Transmission Control Protocol
NF	Network Function
NFUC	Use case Network Function
PLMN	Public Land Mobile Network
RAN	Radio Access Network
RL	Reinforcement Learning
S-NSSAI	Single Network Slice Selection Assistance Information
SD-WAN	Software-Defined Wide Area Network
SGD	Stochastic Gradient Descent
SGi	The Reference point between EPC and Public Data Network
SPI	Safran Passenger Innovations Germany GmbH
SVD	Singular Value Decomposition
TDD	Time-Division Duplexing
UC	Use case
UC1	Use Case 1

UC2	Use Case 2
UC3	Use Case 3
UC4	Use Case 4
UE	User Equipment
UP	User Plane
UPF	User Plane Function
V2N	Vehicle-to-Network
V2X	Vehicle-to-everything
VPN	Virtual Private Network

Executive Summary

This document represents the Deliverable D5.2 with the title “Preliminary Validation and Use Case Benchmarking”. It gives an overview on the status of the validation of the AI@EDGE use case demonstrators by M18. The use cases have been introduced and defined in the previous deliverables D2.1 and D5.1. In the document, an update of the AI@EDGE validation environment is given, including the reference architecture and the connect compute platform integration testbed. Then, for each of the use cases, the validation objectives and validation scenario are provided as well as the validation test cases, and the next validation steps. The latter are specific to each use case, but generally involve the integration of the AI@EDGE platform and improvement of the AI models for the Artificial Intelligence Functions (AIFs).

1. Introduction

In AI@EDGE a secure and reusable artificial intelligence platform for edge computing in beyond 5G networks is developed. Thereby a framework for closed-loop network automation is being developed to support flexible and programmable pipelines for secure and reusable AI models, as well as a connect compute platform to create end-to-end slices supporting a diverse range of AI-enabled network applications.

The AI@EDGE platform will be validated within the four use cases, which have been introduced and described in the Deliverable D2.1. In the Deliverable D5.1 the preliminary plans for the testbed development, integration, validation, and logistics for the AI@EDGE use cases have been elaborated.

In this Deliverable D5.2, the validation and benchmarking methodologies are described, and preliminary results are presented for the individual use cases. While some aspects of the AI@EDGE platform can already be found in the use cases, a complete integration has yet to be perused and will be documented in a future deliverable.

The first use case (UC1) is about the virtual validation of vehicle cooperative perception, where the AI@EDGE platform is used to support digital twinning and cooperative perception in the context of real and emulated vehicles. In UC2, the secure and resilient orchestration of a large (I)IoT network is showcased, where the AI@EDGE platform is used to operate an AI based intrusion detection system. The third use case (UC3) is about Edge AI assisted monitoring of linear infrastructures using drones in BVLOS operation. In this use case, the advantages of edge computing and the AI@EDGE connect compute platform are used for the monitoring of roads. In UC4, smart content & data curation for in-flight entertainment services, the delivery of curated content over 5G from an on-board edge cloud is showcased.

Throughout the remainder of this document, the validation of the use cases is described. In Section 2, the components of the AI@EDGE validation environment are summarised and the current status is described. From Section 3 to Section 6, the validation of the respective use cases is described. Hereby the validation objective, the validation scenario, the validation procedures and preliminary results are described. The validation procedures are described through test cases, where each test case describes a certain capability of the use case demonstrator. The structure of the test cases is presented in Table 1. In Section 7, this document is concluding the status for the validation status of the use cases and the next steps to be taken.

Table 1 Structure of test cases as introduced in the deliverable D5.1.

PHASE n	
Test case #n :	*Name of test case*
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 case*
Test Sequence	*Tasks for the test*

2. AI@EDGE validation environment

In this section, the recent state of the reference architecture for the use case implementation and validation is outlined. In addition, the connect compute platform integration testbed introduced in detail in D5.1 is summarized.

2.1 Reference architecture for the use cases implementation and validation

The general architecture of the AI@EDGE system is shown in Figure 1. The architecture comprises two main functional blocks: The Network and Service Automation layer and the Connect Compute Platform layer.

- The **Network and Service Automation layer** is where the network intelligence is placed, and it is responsible for the system network automation.
- 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.

The system is designed to host Artificial intelligence Functions (AIFs) and, depending on the function they implement, they can be deployed at both level of the system. In particular, some AIFs are used for network automation and optimization purposes, others are associated with applications and services. All system components, including AIFs, can leverage network and computing resources allocated in three different tiers, namely Far Edge, Near Edge and Cloud.

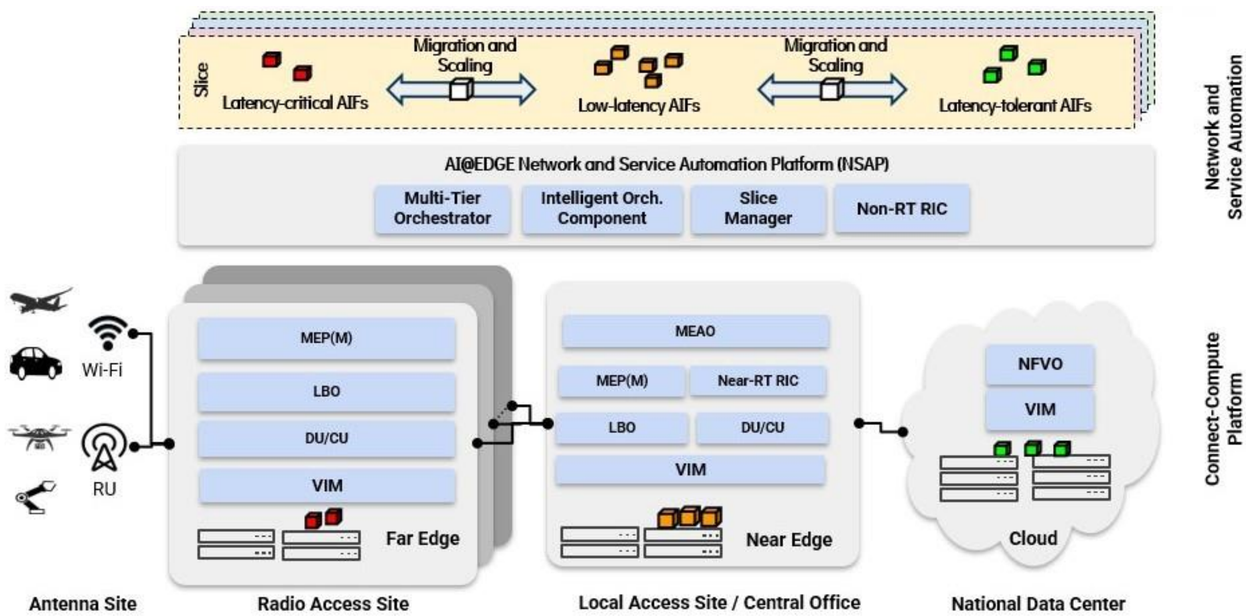

Figure 1 AI@EDGE architecture.

Use cases applications, including the related AIFs, due to the requirements constraints, will primarily leverage Far Edge and Near Edge nodes as execution environments. In fact, these nodes typically incorporate support for HW acceleration, also providing low latency features and the capability to process locally data when needed.

Figure 2 shows the functional diagram of the AI@EDGE architecture to be used as reference for mapping the HW and SW components relevant for each use case. Some components can be hosted in the cloud, others in the integration testbed at FBK and others at the use cases testbed sites (i.e., those located at the far and the near edge), as introduced and documented in D5.1 (AI@EDGE Deliverable D5.1, 2022).

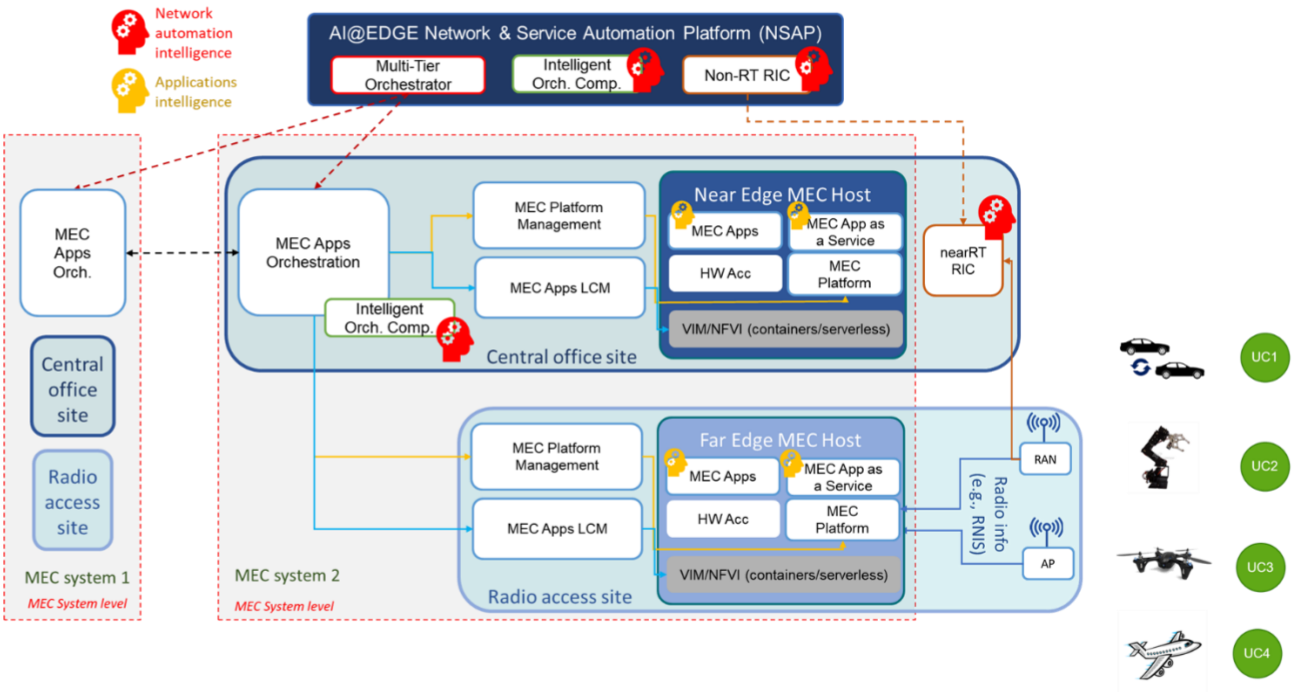

Figure 2 AI@EDGE functional architecture for mapping the HW and SW components relevant for the use cases.

2.2 Connect Compute Platform Integration testbed

The Connect Compute Platform Integration testbed hosts the main components of the system. As shown in Figure 3, the testbed is based on a Kubernetes cluster that spans both Near and Far Edge nodes.

The Near Edge host is connected to the SGi-2 port of the Athonet Core. It hosts the master node of the K8s cluster, together with the LightEdge[12] platform and with the Nuclio[3] platform. In addition to that, Near Edge node hosts applications and Serverless Functions. MEO platform is deployed in the testbed adjacent to Near Edge host. The Far Edge host is connected to the SGi-1 port of the Athonet Core (corresponding to the Edge UPF), and hosts Applications and Serverless Functions.

Near Edge and Far Edge are connected through a management network that allows the synchronization between the Kubernetes components.

[1] https://lightedge.io/

[2] Coronado, E., Yousaf, Z., & Riggio, R. (2020). LightEdge: mapping the evolution of multi-access edge computing in cellular networks. *IEEE Communications Magazine*, *58*(4), 24-30.

[3] https://nuclio.io/

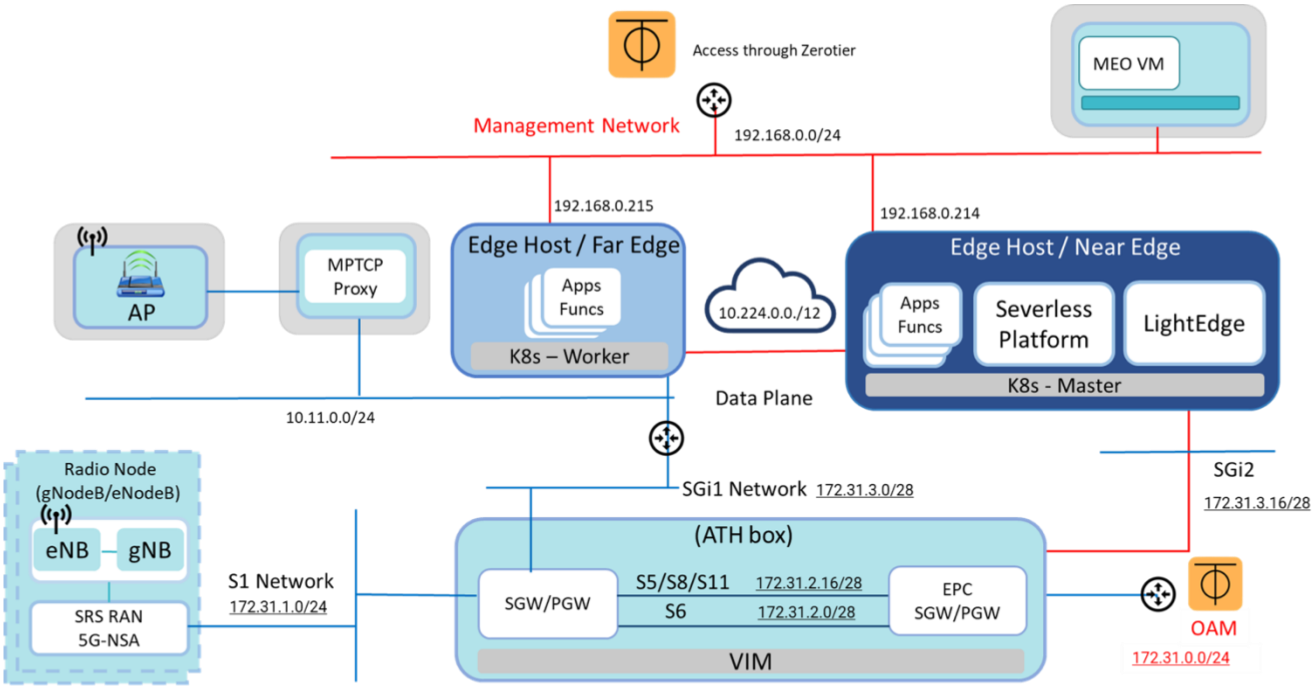

Figure 3 Integration Testbed Main Components.

The Testbed supports off-path multipath TCP experimentation with the deployment of Wi-fi AP and MPTCP proxy. The testbed is instrumented with metrics provided by Kubernetes and the 5G Core. Both the Kubernetes components and the Athonet Core emit metrics in Prometheus format. Which is structured in a plain text, designed in a way so both humans and machines can read it. A Prometheus Server or any suitable metrics scraper can be configured to periodically gather these metrics and make them available via time series database. In most cases metrics are available on the "*/metrics*" endpoint of the HTTP server. A Grafana server is also available for metric visualization.

Remote connection towards the testbed is allowed through a virtual network enabled by the ZeroTier service [1] This tool combines the capabilities of VPN and SD-WAN and emulates Layer 2 Ethernet with multipath, multicast, and bridging capabilities. The testbed uses two ZeroTier L2 networks, enabling different access points: (i) one through the management network, that allows the interaction with Kubernetes nodes and orchestration components; and (ii) one connected directly with the 5G Core.

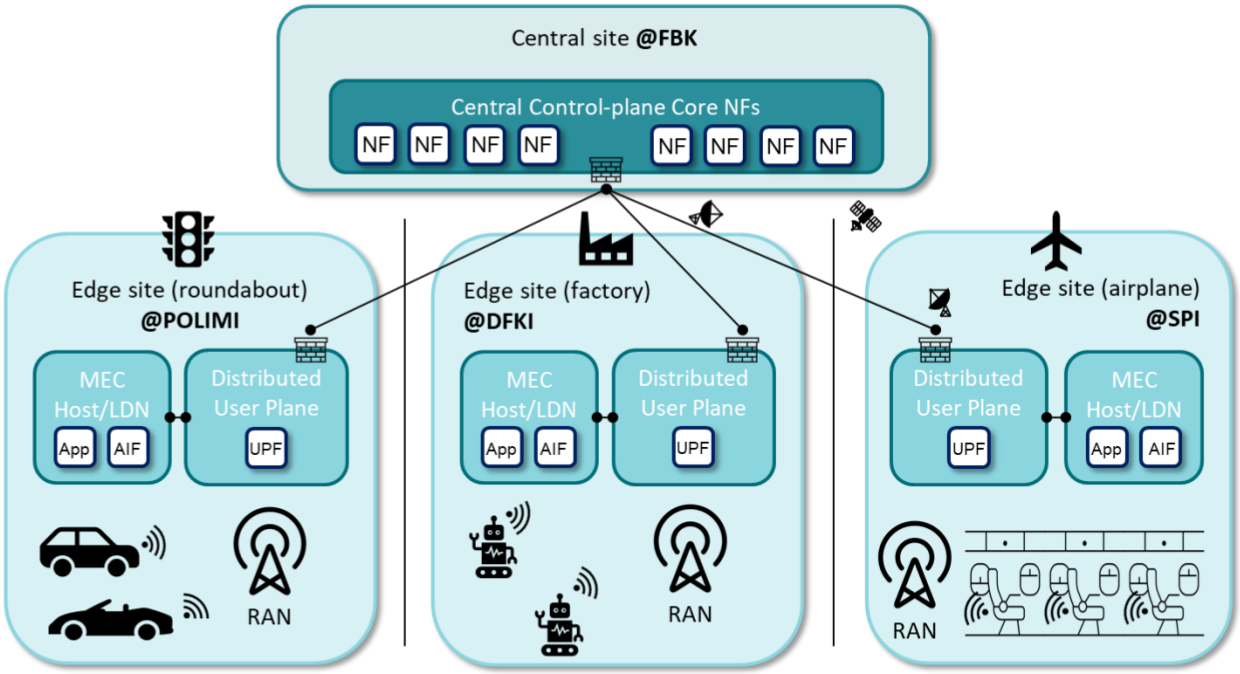

Figure 4 High-level architecture of the 5G network serving UC1, 2, and 4.

AI@EDGE's validation framework relies in each use case on 5G connectivity. As initially described in Section 3.2 of D5.1, Figure 4 shows a high-level representation of the 5G deployment specifically conceived for UC1, 2, and 4. The 5G architecture is split between a centralized control plane shared among the use cases and three distinct distributed user planes, each one exclusively dedicated to one of the involved use cases. Only control traffic is exchanged via VPN between geographically separated locations, whereas the on-premises UPF brings several benefits: (i) in UC1, the edge UPF guarantees low latencies and fast access for the connected UEs to the applications deployed over edge servers; (ii) in UC2, the on-site user plane answers the typical requirement of factory deployments to keep the data local, physically within the premises for privacy and security reasons; and (iii) in UC4, having a UPF on board of the airplane is the fundamental choice to allow an efficient user data exchange, with satisfactory quality of service that could not be supported by the satellite backhaul.

The work towards a full deployment of the proposed architecture has started and is still ongoing. It will be completed in the second half of the project's lifetime. Once all the 5G edge nodes with the distributed user planes will be deployed and remotely integrated with the central control plane, the end-to-end 5G connectivity will be tested for what concerns the UE attach procedures, their idle and connected states, the detach procedure, and finally the downlink and uplink traffic. We will report on such activities in D5.3.

The testbed devised for validating UC3 is relying on the 5TONIC environment to provide a 5G network capable of integrating AI@EDGE platform functionalities. The phase of dimensioning resources and designing the process for implementation has already started as scheduled and it will be completed during the second half of the project.

Figure 5 shows the proposed architecture of the AI@EDGE platform to UC3 testbed:

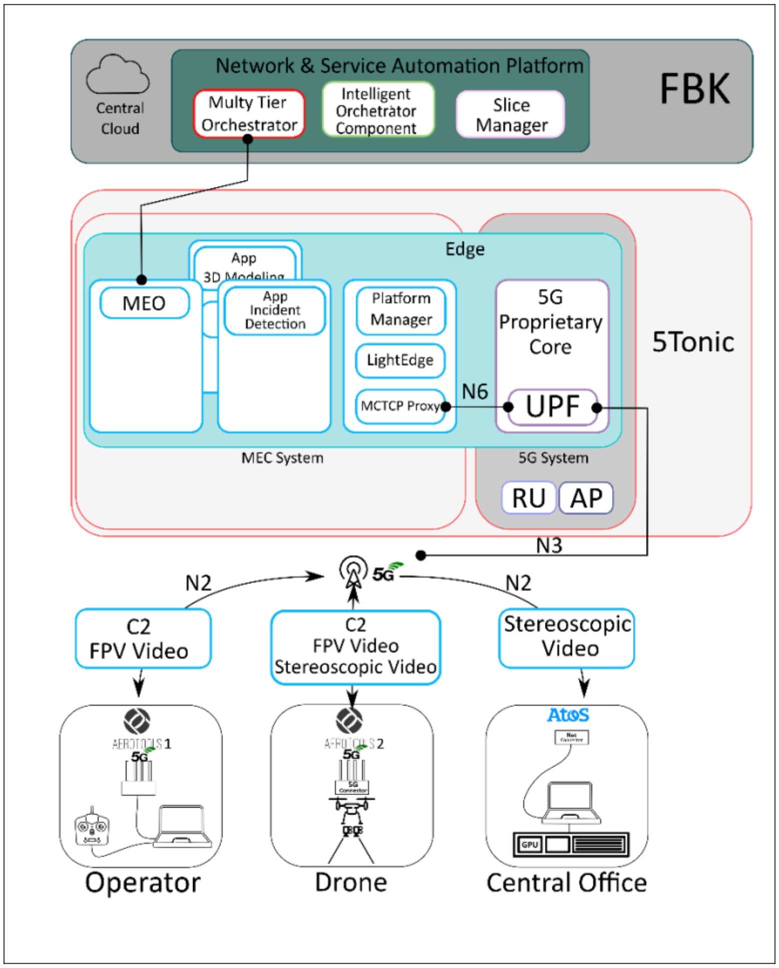

Figure 5 Proposed integration of AI@EDGE platform to UC3 testbed.

3. Use Case 1

The vehicle cooperative perception use case is based on several vehicles exchanging data related to their trajectories. Data is gathered at the network edge and is used to build a view of the surrounding environment that will be used by Artificial Intelligence Functions (AIFs), which will predict potential collisions and dangers. The end-to-end system development to demonstrate this use case is complex and expensive in the real world and for these reasons the use case adopts an emulation environment able to scale with such a complexity. The specific mobility scenario chosen is the roundabout, where fluidity and safety are of paramount importance.

The Cooperative perception emulation environment interconnects a dynamic driving emulator operated by a real human driver with a traffic simulator to design, implement, and test a mix of real and emulated vehicles. The optimization of communication between vehicles and between human-driven vehicles and autonomous vehicles is based on AI@EDGE platform. Each single autonomous vehicle is driven by a Reinforcement Learning (RL) agent. More specifically, the objective is using the Reinforcement Learning techniques for 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.

3.1 Validation objectives

The motivation of this use case is to increase the security and reduce the traffic congestion on a roundabout. The functionalities showcased in UC1 are the local traffic outbreak on the edge to increase availability, the extension of the driving simulator with a 5G connectivity and the intelligence added to vehicles to coordinate the manoeuvres.

The local execution of AI algorithms on vehicles with a direct vehicle to vehicle communications has some limitations regarding the possibility to identify and solve more complex traffic situations such as a roundabout. Traffic jams in roundabouts are also generated by better managing incoming and outgoing traffic flows and therefore having information on vehicles not only directly approaching the roundabout.

The introduction of edge computing nodes to offload the coordination functions between autonomous vehicles should allow a more complete solution regarding the traffic scenarios created in roundabouts and in its surrounding. From the algorithmic point of view, different approaches will be evaluated, starting from baselines of non-cooperative behaviour between vehicles up to global coordination scenarios.

The change of the communication prospective from “short-range” (802.11p, PC5) to “long-range” (5G) using MEC platforms adds some latency in Vehicle-to-Vehicle communication but allows for a wider communication between vehicles. In this context, the main validation objective is to assess the support of the MEC technology in this particular roundabout scenario.

The main challenges related to the scenario are:

- **Telematic Box and Driving Simulator integration:** Integration of a Telematics Box with the Driving Simulator to support the basic connectivity features and the correct operation.
- **Local breakout of traffic and offloading:** Capability to effectively support data collection via 5G from Telematics Box to edge servers to enable AIF processing of such data guaranteeing low latencies whenever needed.
- **Vehicle Coordination:** Capability to support the vehicles coordination considering the real driver reactions on the Driving Simulator.

Main KPIs related to those Challenges are:

- **Latency**: under 2000 ms, to receive Alerts messages to the driver (V2N communication).
- **Vehicle density**: 12000 vehicle/km^2 as expected number of simulated vehicles per a given area.
- **Positioning:** 1.5 m to deal with vehicle dynamics and movement.

The following table shows how each single KPI is relevant for a specific Use Case challenge.

Table 2 KPIs and respective challenges in UC1.

	Latency	Vehicle density	Positioning
Telematic Box and Driving Simulator integration	X	-	-
Local breakout of traffic and offloading	X	X	-
Vehicle Coordination	X	X	X

3.2 Validation scenario

The UC1 testbed facilities are two and depicted in the in Figure 6. The first facility relies on the infrastructure available at POLIMI in Milano, where the driving simulator will be 5G connected and will send dynamic to the 5G network, in particular to an Edge Node (far and/or near edge) on which a Cooperative Perception Algorithm will be executed. 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 algorithms will be validated.

- POLIMI Testbed: The first testbed is based on a driving simulator connected to the AI@EDGE platform through a 5G Telematics Box. The POLIMI driving simulator sends its dynamic data to an Edge Node on which a Cooperative Perception Algorithm will be executed.
- CRF Test Bed: The second Test Bed is available at CRF in Torino where a 5G emulator is used to test 5G enabled automotive Telematic Boxes. The Telematics Box sends data to the Cooperative Perception Algorithm, deployed on the Edge Node, through the 5G emulator.

CRF Testbed:

- MT8000A Network Emulator (Non-Stand Alone and Stand-Alone)
- 5G-NR SA and NSA: DL 2CA, DL 4x4 MIMO, FDD and TDD
- LTE / LTE-A: DL 3CA, DL 2x2 MIMO, FDD and TDD
- Ability to simulate communications up to 6GHz
- Control PC for Software to execute Validation and Test sequences.
- Local Edge Node with MEC Host for the execution of the AIF, the data collection server (AMQP Broker). On the local Edge is also running the traffic simulator that simulates the other vehicles.
- Telematics Box (DUT: Device Under Test) with the Uu and PC5 connection, V2X ITS stack and the data client to send data (AMQP Client)
- GNSS emulator

- Vector CANalyzer to support the internal Vehicle bus (CAN Bus)

Polimi Testbed:

- VI Grade Driving Simulator with its WorldSim scenarios simulator
- Telematics Box (DUT: Device Under Test) with the Uu and PC5 connection, V2X ITS stack and the data client to send data (AMQP Client)
- 5G RAN equipment with SRS's virtualized radio solution.
- ATH's 5GC functionalities at the edge node with the distributed User Plan
- Local Edge Node with MEC Host for the execution of the AIF, the data collection server (AMQP Broker) and the traffic simulator.

FBK Central Site:

- Central Control Plane

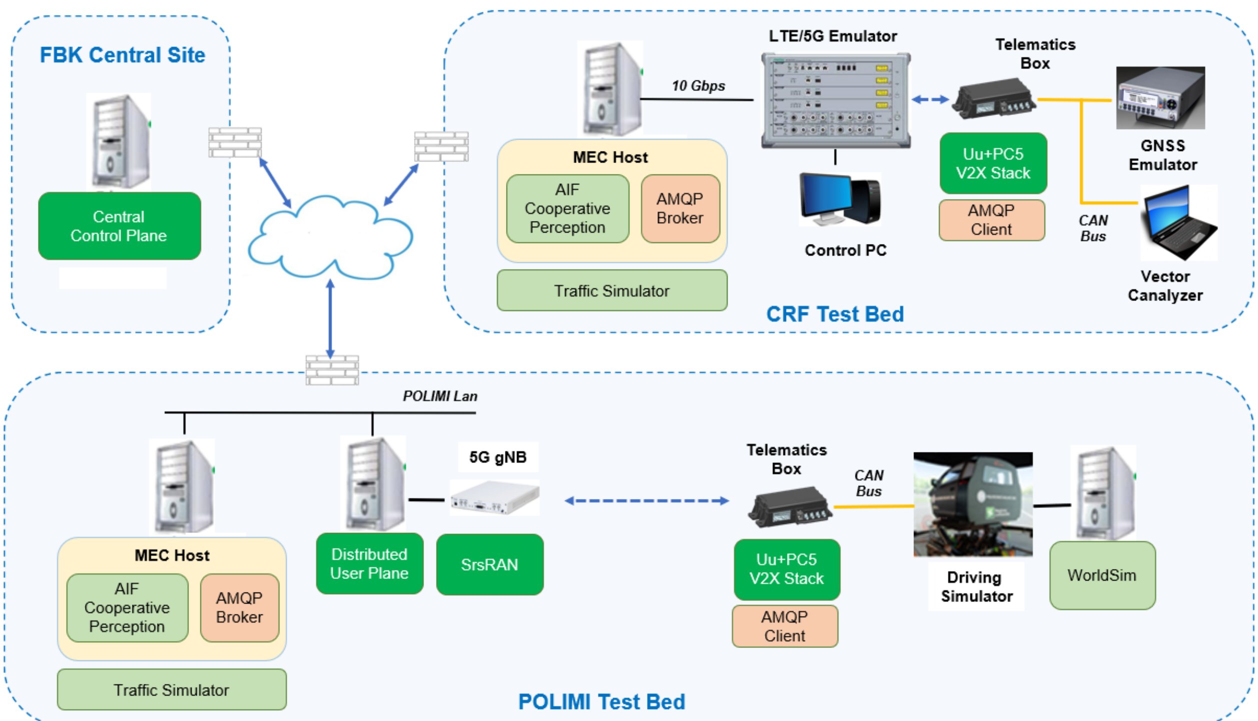

Figure 6 Use Case 1 Testbeds.

3.3 Validation procedures and preliminary results

In this chapter the procedure of the validation is given, separated into 4 test cases, where each test case is validating a separate aspect of the demonstrator.

3.3.1 Test case 1 – 5G Telematic Box Basic Validation

The connectivity tests that will be run in the CRF 5G emulation Lab are described in the following tables.

1 - Initial developments - Early Demonstrator	
Test case #1.1:	LTE to NR Handover (Non Standalone)
Slogan & Objective	• Handover between LTE and New Radio.
Test Scenario (Pre-conditions)	• LTE instance running on MT8000A. • 5G instance running on MT8000A. • Scenario with LTE cell and NR cell loaded.
Expected Results (Post-Conditions)	• DUT successfully performs handover to NR1 cell.
General Time Plan (Validation Campaigns)	• Q3 of 2022.
Test Sequence	• LTE1 and NR1 (SA) cell is available. • DUT registers to LTE1. • RRC connection is active. • Power of LTE1 cell decreases and NR1 increases. • Network performs cell handover to NR1 cell

PHASE 1 - Initial developments - Early Demonstrator	
Test case #1.2:	NR to NR Handover (Standalone)
Slogan & Objective	• Handover between New Radio and New Radio.
Test Scenario (Pre-conditions)	• 5G instance running on MT8000A • Scenario with 2 NR1 cell and NR2 cell loaded

Expected Results (Post-Conditions)	• DUT successfully performs handover to NR2 cell.
General Time Plan (Validation Campaigns)	• Q3 of 2022.
Test Sequence	• NR1 (SA) and NR2 (SA) cell is available. • DUT registers to NR1. • RRC connection is active. • Power of NR1 cell decreases and NR2 increases. • Network performs cell handover to NR2 cell. • DUT successfully performs handover to NR2 cell.

PHASE 1 - Initial developments - Early Demonstrator	
Test case #1.3:	IP data end to end- UDP (Standalone)
Slogan & Objective	• End to End IP Data connection from DUT to Edge Node.
Test Scenario (Pre-conditions)	• 5G instance running on MT8000A. • Scenario NR cell loaded. • Edge server connected to the MT8000A (10 Gbps connection).
Expected Results (Post-Conditions)	• DUT successfully receives UDP data from Edge server.
General Time Plan (Validation Campaigns)	• Q4 of 2022.

Test Sequence	• NR1 (SA) cell is available. • DUT registers to NR1 cell. • RRC connection is active. • Run iPerf on DUT side. • Trigger iPerf UDP data stream from edge server towards DUT. *(iPerf: https://sourceforge.net/projects/iperf2/)*

3.3.2 Test case 2 – V2X over 5G

PHASE 2 - Mid Demonstrator	
Test case #2.1:	V2X Messages from DUT to Edge server over 5G
Slogan & Objective	• Edge server receives CAM (Cooperative Awareness Messages) messages from DUT.
Test Scenario (Pre-conditions)	• 5G instance running on MT8000A. • Scenario NR cell loaded. • Edge server connected to the MT8000A (10 Gbps connection). • AMQP broker running on the edge server. • AMQP client running on DUT side.
Expected Results (Post-Conditions)	• CAM messages successfully received from Edge server.
General Time Plan (Validation Campaigns)	• Q1 of 2023.
Test Sequence	• NR1 (SA) cell is available. • DUT registers to NR1 cell. • RRC connection is active. • Data stream from DUT towards AMQP Broker.

PHASE 2 - Mid Demonstrator	
Test case #2.2:	V2X Messages from Edge server to DUT over 5G
Slogan & Objective	DUT receives DENM (Decentralised Environmental Notification Message) messages from the edge server.
Test Scenario (Pre-conditions)	• 5G instance running on MT8000A. • Scenario NR cell loaded. • Edge server connected to the MT8000A (10 Gbps connection). • AMQP broker running on the edge server. • AMQP client running on DUT side.
Expected Results (Post-Conditions)	• DENM messages successfully received from DUT.
General Time Plan (Validation Campaigns)	• Q1 of 2023.
Test Sequence	• NR1 (SA) cell is available. • DUT registers to NR1 cell. • RRC connection is active. • Data stream from AMQP Broker towards DUT.

3.3.3 Test case 3 – 5G Connectivity and Local Traffic Breakout

We describe here three connectivity tests that will be run as soon as the 5G deployment will be completed to validate the correct operations of the 5G network in this use case.

PHASE 2 - Mid Demonstrator	
Test case #3.1:	Connection between gNB and 5GC
Slogan & Objective	• Interface setup between gNB and 5GC.
Test Scenario (Pre-conditions)	• 5GC instance (remote control plane and edge user plane) running on servers or VMs. • 5GC configured with active license and running, gNB should be reachable through the network.

Expected Results (Post-Conditions)	• No connection errors. Log messages show gNB successfully attached to the AMF.
General Time Plan (Validation Campaigns)	• Q4 of 2022.
Test Sequence	• Configure the network interfaces and the CP, including all the related NFs. The system should show settings confirmation. • Set the IP address of the gNB in the whitelist of the 5GC's web interface. • Configure the N2 interface for interconnection between AMF and gNB. • Connect the gNB to the 5GC (AMF).

PHASE 2 - Mid Demonstrator	
Test case #3.2:	UE's attach to and detach from the 5G network
Slogan & Objective	• Check if UEs successfully attach to and detach from the correct PLMN and S-NSSAI.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC.
Expected Results (Post-Conditions)	• Log messages show UE successfully registered, attached and detached to the 5GC.
General Time Plan (Validation Campaigns)	• Q4 2022.

Test Sequence	• Configure the UE (virtual or physical) with the correct settings of PLMN, S-NSSAI and DNN. The system should show settings confirmation. • Register through the GUI the UE into the 5GC with SUPI identity. • Review the 5GC log messages related to the UE attachment. Verify that no error occurred. • Detach the UE from the 5GC.

PHASE 2 - Mid Demonstrator	
Test case #3.3:	Connectivity between UE and data network (DN)
Slogan & Objective	• Check uplink/downlink traffic between UE and DN through the 5GC (UPF), demonstrating the end-to-end connectivity between the connected devices and the edge servers.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC and attached to the 5GC.
Expected Results (Post-Conditions)	• Connectivity between UE and DN is operational. • iPerf shows uplink/downlink traffic. • ICMP messages are acknowledged.
General Time Plan (Validation Campaigns)	• Q4 2022.
Test Sequence	• Establish a new PDU session. Log messages should show the successful creation of UPF session. • Configure iPerf agents on the UE and in a reachable server of the DN. Verify that there are no registering errors.

	• Execute iPerf session or ping session. The test plan should start running. An iPerf or ping experiment will be started. • Review the 5GC log messages or check iPerf or ping results. There should be no errors, warning messages or dropped packets.

3.3.4 Test case 4 – Driving Simulator Integration

We describe here integration of driving simulator in the testbed set-up.

PHASE 1 - Initial developments - Early Demonstrator	
Test case #4.1:	CAN signal acquisition
Slogan & Objective	• Input/output signal acquisition through driving simulator CAN.
Test Scenario (Pre-conditions)	• WorldSim scenario running on driving simulator. • CAN bus configured and ready.
Expected Results (Post-Conditions)	• Successful acquisition of output signals from CAN of driving simulator. • Successful transfer of input signals through CAN of driving simulator.
General Time Plan (Validation Campaigns)	• Q3 of 2022.
Test Sequence	• Load reference driving scenario on driving simulator. • Start simulation. • Input/output signals on CAN. • Verify correct data transfer.

PHASE 1 - Initial developments - Early Demonstrator	
Test case #4.2:	WorldSim communication interface test

Slogan & Objective	• Input/output vehicle data in WorldSim.
Test Scenario (Pre-conditions)	• WorldSim scenario running. • Simulink interface model ready and running. • Communication established between WorldSim and Simulink.
Expected Results (Post Conditions)	• Successful acquisition of vehicle and sensor data from WorldSim to Simulink. • Successful control of vehicle dynamics from Simulink to WorldSim.
General Time Plan (Validation Campaigns)	• Q3 of 2022.
Test Sequence	• Load reference driving scenario on WorldSim. • Start Simulink model. • Enable communication between Simulink and WorldSim. • Start simulation. • Acquisition of vehicle and sensors signals. • Perform vehicle control from Simulink to WorldSim.

PHASE 1 - Initial developments - Early Demonstrator	
Test case #4.3:	SUMO communication interface test
Slogan & Objective	• Input/output vehicle data in SUMO.
Test Scenario (Pre-conditions)	• SUMO scenario running. • Matlab/Python interface model ready and running. • Communication established between SUMO and interface software.
Expected Results (Post-Conditions)	• Successful acquisition of vehicle data from SUMO. • Successful control of vehicle dynamics to SUMO.

General Time Plan (Validation Campaigns)	• Q4 of 2022.
Test Sequence	• Load reference traffic scenario on SUMO. • Start Matlab/Python model. • Enable communication between SUMO and interface software. • Start simulation. • Acquisition of vehicle data. • Perform vehicle control on SUMO.

PHASE 2 - Mid Demonstrator	
Test case #4.4:	SUMO – WorldSim integration test
Slogan & Objective	• Integration between SUMO and WorldSim.
Test Scenario (Pre-conditions)	• SUMO scenario running • WorldSim scenario running • Communication established between SUMO and WorldSim.
Expected Results (Post-Conditions)	• Successful data transfer between SUMO and WordlSim. • Successful data transfer between SUMO and WorldSim.
General Time Plan (Validation Campaigns)	• Q1 of 2023.
Test Sequence	• Load reference traffic scenario on SUMO. • Load reference driving scenario on WorldSim. • Enable communication between SUMO and WorldSim. • Start simulation. • Data tranfer between SUMO and WorldSim. • Check time synchronization between SUMO and WorldSim.

3.4 Next validation steps

There are several challenges to face in the roundabout scenario but in general in its mixed virtual-real environment.

The first challenge is related to the MEC, as an important component to support the seamless service management needed by the vehicle mobility. The AIF offloading management through the various MEC nodes, with low latency requirements, is fundamental in the mobility scenario and the orchestration and migration are very important enablers.

A second challenge is related to the AI functions. To develop vehicles which are autonomously choosing actions based on AI with Reinforcement Learning, it is important to define the external environment (traffic situations and maps) and to model the driver, in order to maximize the notion of cumulative reward identified for the vehicles.

The last challenge is related to the latency of the Human in the Loop system and the related Human reactions. For the use case success, it is fundamental evaluate the reactivity of the emulation and simulation environment in which a real driver is inserted, considering the addition of mobile communications and edge computing. In this Human in the Loop environment, it is important to take into account the integration of a real telematics box with a driving simulator, which is a completely new approach for driving simulators.

To summarize, the next validation steps needed to face the following challenges:

- Evaluate the MEC functions (Orchestration/Deploy/Migration) supporting the roundabout scenario.
- Compare different algorithms in respect to approaches for roundabout traffic optimization in a mixed real-virtual environment using the driving simulator and the AI@EDGE platform.
- Validate the latency target to allow the human driver to interact in real time with the simulation environment using the mobile communication channel and the AI@EDGE platform.
- Validate the integration of the telematic box into the electric and electronic architecture of the driving simulator. In this scenario the telematic box will act as a bridge between the driving simulator and the AI@EDGE platform.

4. Use Case 2

In use case 2 the security functionalities of AI@EDGE are showcased in an Industrial Internet of things (IIoT) scenario. The scenario consists of a 5G network and multiple stakeholders sharing the network. Therefore, the data privacy must be ensured.

4.1 Validation objectives

The motivation of this use case is the increasing demand for the protection of industrial IoT devices from attacks on availability and confidentiality. The availability of a MEC system for the project AIF supports low-latency model retraining and inference for the anomaly detection AIF, in the order of the ms. Low-latency retraining can increase model accuracy, and adaptive reorchestration of the UC2 platform stack to mitigate from attacks and anomalies. Moreover, given the short-range network between UE and application server in UC2 environment, with only few hops and few meters path, the deployment of Smart-NICs at MEC servers allows the integration of AI and monitoring acceleration functions in the embedded NetFPGA system, otherwise difficult to deploy because of limited computing resources at the far edge. The functionalities showcased in UC2 are the auto-configuration of hyperparameters for anomaly detection, the federated learning for anomaly detection, NetFPGA based intrusion detection and the local traffic outbreak on the edge to increase availability, as described in the following:

- **Intrusion detection with NetFPGA:** Capability to detect botnets that are unknown at the time of attack, as well as scan attacks and newly exploited vulnerabilities. Approach based on the integration of the Split and Merge algorithm at the state of the art in NetFPGA boards, developed by CNAM.
- **Auto configuration of HP for Intrusion detection:** The auto configuration of hyper parameters to adapt the pre-trained ML models to the deployed environment to improve the performance of intrusion detection from INRIA will be integrated
- **Anomaly Detection AIF:** Capability to detect any type of anomalies (attacks, malfunctioning, failures, etc) by monitoring the state of the connect-compute platform components (CPU, memory, disk, network states) at different layers and domains (containers, physical servers, UEs, radio front ends).
- **Local breakout of traffic and offloading (Athonet)**: Capability to effectively support data collection via 5G from connected devices to edge servers to enable AIF processing of such data, maintaining confidentiality and guaranteeing low latencies whenever needed.
- **UC Demonstrator:** In the demonstrator all the previously mentioned functionality will be integrated and tested in an environment based on a 5G Network, the connect-compute platform and robots and sensors for the application, as described in section 4.2.

4.2 Validation scenario

The use case two demonstrator aims to showcase and validate various security mechanisms in a 5G MEC scenario in the setting of an industrial site with a large network of industrial Internet of things. The components, as described in section 4.3, offer the functionality to detect known attacks based on anomaly detection with auto-configuration of hyper parameters, as well federated learning-based anomaly detection of unknown attacks and the split-and-merge based botnet detection. To showcase these security mechanisms, different attacks will be performed through a separate attacker device with access to the network. Namely, these attacks will be a botnet infection of the connected IoT devices, based on the Mirai botnet and credential brute force, and Metasploit based attacks, e.g., on the heart bleed vulnerability in TLS.

The physical implementation of the demonstrator is depicted in Figure 7, and consists of at least three AGVs, which are connected to the 5G RAN over a Raspberry Pi 4B. Also, different sensors will track the location of the AGVs, such as two radio wave-based localization, cameras, and light barriers. The RAN is connected to the 5G Core, from where data can be sent to the near edge with a MEC Host, hosting the Intrusion detection and anomaly detection.

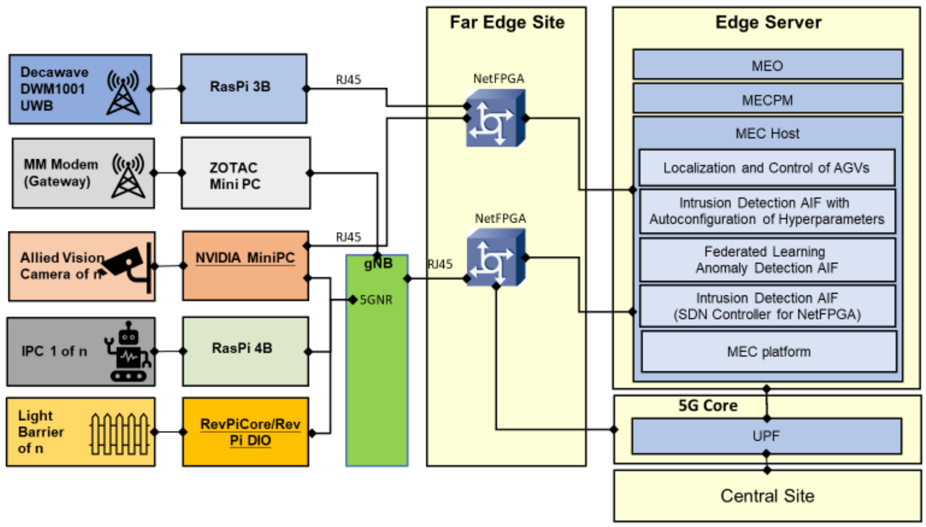

Figure 7 Network scheme of the use case demonstrator. On the left are the AGVs and sensors connected over radio and RJ45 to the far edge where NetFPGAs are located. Through the UPF on the 5G Core, data can be sent to the near edge, where the different AIF are located.

4.3 Validation procedures and preliminary results

In this chapter the procedure of the validation is given, separated into 4 test cases, where each test case is validating a separate aspect of the demonstrator.

4.3.1 Test case 1 – Intrusion Detection for Known Attacks

Test Case 1	Network Intrusion Detection based on machine learning for Known Attacks
Slogan & Objective	Benchmarking of the auto-configuration of a network intrusion detection system (NIDS) based on a machine learning algorithm.
Test Scenario (Pre-conditions)	• Demonstrator setting with AGV, localization, 5G network, Edge server • Vulnerable services on Edge server (weak credentials, HTTPS with heartbleed, etc.). • Run Attacks (DDoS, Botnet, Credential Brute Force, Metasploit exploits). • Build a knowledge base of prior experiences in attack detection to be able to predict the NIDS configuration. • Train and configure the NIDS offline and test it online. • Comparison to state-of-the-art techniques such as black box optimization solutions.

	• Comparison of the proposed solution with public dataset and testbench data from the demonstrator.
Expected Results (Post-Conditions)	• Metrics: time, memory, accuracy (different machine learning metrics), gain compared to traditional machine learning configuration approaches. • Accuracy > 97% and time delay in a few ms for inferring a configuration and for detecting an attack. • Qualitative description of degradation of service for 5G services and AGV.
General Time Plan (Validation Campaigns)	• **Q2 and Q3 Offline attack results.** • Q4 of 2022 Online attack: Attack and vulnerability implementation for test bench.
Test Sequence	• Generate a labeled intrusion detection dataset (IDD). • Build a meta-dataset from the IDD to infer a new configuration of the NIDS for an unseen dataset.

4.3.2 Test case 2 – Intrusion Detection for Unknown Attacks

Test case 2	Intrusion Detection for Unknown Attacks
Slogan & Objective	• Detect state-of-the art attacks that were unknown the time they happened without ad-hoc configuration of the attack profiles.
Test Scenario (Pre-conditions)	• Botnet installed in few x86 nodes letting it spread within the testbed. • Attack mitigation using NetFPGA Smart-NIC during botnet spreading. • Compared to alternative state of the art techniques such as Snort-based control.
Expected Results (Post-Conditions)	• Metrics: accuracy, detection time. • Accuracy > 97% and time delay in few ms for detection. • Qualitative description of service degradation for 5G services and AGV during the attack.
General Time Plan (Validation Campaigns)	• Q2 and Q3 Offline attack results. • Q4 of 2022 Online attack: Attack and vulnerability implementation for test bench.
Test Sequence	• Infect an NVIDIA MiniPC with Mirai malware. • Let the Mirai botnet propagate to other MiniPCs.

	• Use infected devices to start a DDoS attack. • Verify that the Split & Merge Aggregator is able to detect the attack and the spreading botnet, even before the attack takes place, and mitigate both threats by changing the routing rules of the NetFPGA SmartNICs.

4.3.3 Test case 3 – Anomaly Detection

Test case 3:	Anomaly Detection
Slogan & Objective	• Detect and characterize anomalous conditions generated by arbitrary events such as attacks, failures, misconfigurations.
Test Scenario (Pre-conditions)	• Inject anomalies in the infrastructure stack such as CPU overload, packet loss, link failures, attacks (both known and unknown from test cases 1 and 2). • Compare with standard monitoring systems alerts.
Expected Results (Post-Conditions)	• Visualization of anomalies by means of stack radiography. • Online characterization and root cause analysis of the anomalies.
General Time Plan (Validation Campaigns)	• Q4 of 2022: set up of probes on the stable UC2 testbed. • Q2 of 2023: anomaly injections and demonstration.
Test Sequence	• CPU overload injection. • Packet loss injection. • Link bandwidth decrease emulation. • Known attack emulation. • Unknown attack emulation. • Link failure emulation.

4.3.4 Test case 4 – 5G Connectivity and Local Traffic Breakout

We describe here three connectivity tests that will be run as soon as the 5G deployment will be completed to validate the correct operations of the 5G network in this use case.

Phase 1: Connection between gNB and 5GC	
Test case 4:	5G Connectivity and Local Traffic Breakout
Slogan & Objective	• Interface setup between gNB and 5GC.
Test Scenario (Pre-conditions)	• 5GC instance (remote control plane and edge user plane) running on servers or VMs. • 5GC configured with active license and running, gNB should be reachable through the network.

Expected Results (Post-Conditions)	• No connection errors. Log messages show gNB successfully attached to the AMF.
General Time Plan (Validation Campaigns)	• Q4 of 2022.
Test Sequence	• Configure the network interfaces and the CP, including all the related NFs. The system should show settings confirmation. • Set the IP address of the gNB in the whitelist of the 5GC's web interface. • Configure the N2 interface for interconnection between AMF and gNB. • Connect the gNB to the 5GC (AMF).

Phase 2: UE's attach to and detach from the 5G network	
Test case 4:	5G Connectivity and Local Traffic Breakout
Slogan & Objective	• Check if UEs successfully attach to and detach from the correct PLMN and S-NSSAI.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC.
Expected Results (Post-Conditions)	• Log messages show UE successfully registered, attached and detached to the 5GC.
General Time Plan (Validation Campaigns)	• Q4 2022.
Test Sequence	• Configure the UE (virtual or physical) with the correct settings of PLMN, S-NSSAI and DNN. The system should show settings confirmation. • Register through the GUI the UE into the 5GC with SUPI identity. • Review the 5GC log messages related to the UE attachment. Verify that no error occurred. • Detach the UE from the 5GC.

Phase 3: Connectivity between UE and data network (DN)	
Test case 4:	5G Connectivity and Local Traffic Breakout

Slogan & Objective	• Check uplink/downlink traffic between UE and DN through the 5GC (UPF), demonstrating the end-to-end connectivity between the connected devices and the edge servers.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC and attached to the 5GC.
Expected Results (Post-Conditions)	• Connectivity between UE and DN is operational. • iPerf shows uplink/downlink traffic. • ICMP messages are acknowledged.
General Time Plan (Validation Campaigns)	• Q4 2022.
Test Sequence	• Establish a new PDU session. Log messages should show the successful creation of the UPF session. • Configure iPerf agents on the UE and in a reachable server of the DN. Verify that there are no registering errors. • Execute iPerf session or ping session. The test plan should start running. An iPerf or ping experiment will be started. • Review the 5GC log messages or check iPerf or ping results. There should be no errors, warning messages or dropped packets.

4.4 Next validation steps

While defining the test cases for Use Case 2, the following next steps were identified:

Network Intrusion Detection System:

- Adapt the auto-configuration solution for online use.
- Determine when the configuration of the NIDS must be revised according to change in the dataset.
- Adapt the solution for unlabeled dataset.

Intrusion Detection for Unknown Attacks:

- Integrate Split & Merge Collaborative IDS in an SDN environment.
- Implement line-rate metric computation on NetFPGA SmartNICs.
- Split & Merge local anomaly detection assumes normal distribution of several traffic features. A Local Area Network such as the UC2 environment could defy these assumptions and potentially need adapting the Anomaly Score logic used by the algorithm.

5. Use Case 3

UC3 main application is monitoring large areas of roads network using drones in BVLOS (Beyond Visual Line of Sight) mode through the 5G network, a 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 monitoring application requires advanced functionalities to perform scanning and 3D modelling of the infrastructures, and then to locate identified incidents and send notifications to the drone operator. The dataflow (images, C2, telemetry) must run continuously to the central office in order to improve the drone's operator decision-making process.

Given the constrains in weight, energy consumption and other factors intrinsically related to drone´s performance, it is necessary to reduce systems onboard and therefore to move as much processes as possible out of the drone.

Leveraging the UC3 operation on MEC systems based on AI and Edge Computing assisted 5G networks, will provide the required support for optimal monitoring, accelerating computational and modelling processes, as well as improving reliability and range of operation.

5.1 Validation objectives

The main goals for this stage of the UC are focused on reaching an adequate integration scenario for the three areas of development identified and defined in D5.1 as the DEVELOPMENT ENVIRONMENTS., 5G NETWORK, DRONE and AI FUNCTIONS. The initial steps to integrate MEC functions of the AI@EDGE platform in 5TONIC ENVIRONMENT have started as well.

Therefore, the validation objectives are related to demonstrate firstly that an integrated framework has been achieved by connecting the Drone and the AI Function Environments within the 5TONIC 5G Network. Then to check if it is working successfully with the required dataflow. This scenario will allow the integration of new AI and Edge Computing functionalities generated by the project.

The test cases designed for this stage are to provide a proof of the successful integration as well as a reference dataset to demonstrate that the KPIs for the UC have been reached:

- 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…).
- Validation of the DEVELOPMENT ENVIRONMENT connecting the three development nodes (AERO in Madrid, ATOS in Zaragoza and 5TONIC) to allow continuous and efficient development during this period.
- Preliminary tests of the AI functions developed by ATOS and their integration in the workflow.
- Preparatory links with FBK site for the integration of AI@EDGE functions to be carried out throughout the second half of the project.

Main KPIs identified for the use case to be tested in this stage are::

- For the Drone operation:
    - the latency KPIs (initially set as 100ms the maximum end-to-end latency budget, composed of two components: Control Signal latency and Video processing latency);
    - and the reliability KPI (tentative metric) in terms of control signal packet loss.
- For the Anomaly Detection AIF, the KPI is the Mean Average Precision (mAP) with an Intersection over Union (IoU) equal to 0.5, . This target KPI for the AI@EDGE project, according to the dataset used for the project, will be mAP@.5 >= 0.6 (defining classes as identifiable items such as "persons" or "vehicles") - mAP@.5 refers to the mean average precision at an intersection over union value of 0.5.

5.2 Validation scenario

The **5G NETWORK ENVIRONMENT** is built on top of 5TONIC laboratory. In order to connect the other environments, dedicated VPNs have been set and tested for correct functioning to provide visibility and connectivity to all the systems involved.

The VPN generated by AERO for preliminary works is using MQTT to manage system information (device name and IP address). It is depicted in the following diagram:

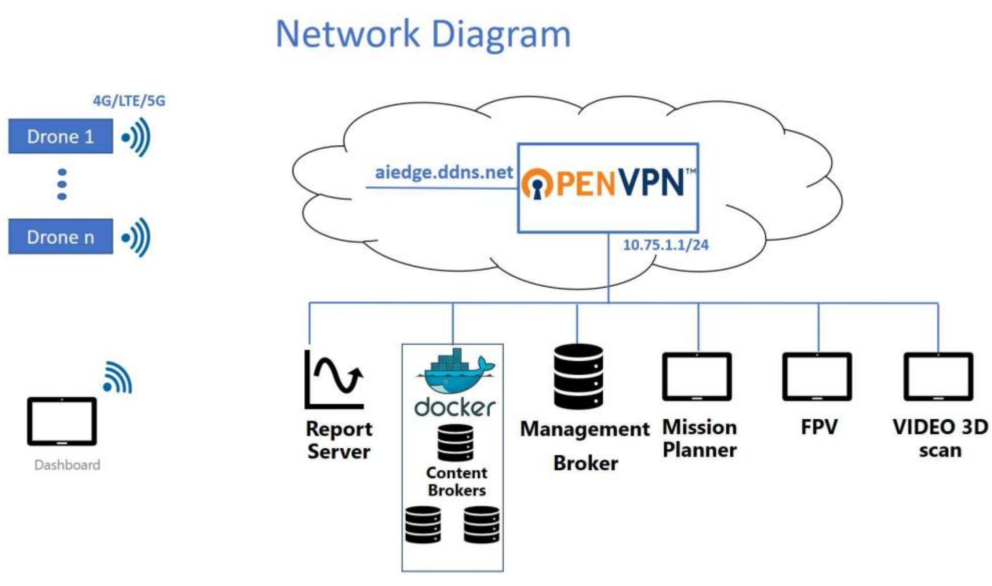

Figure 8 VPN connection for UC3 development process.

The 5TONIC VPN has been deployed recently according to the following diagram, correctly connecting the development environments.

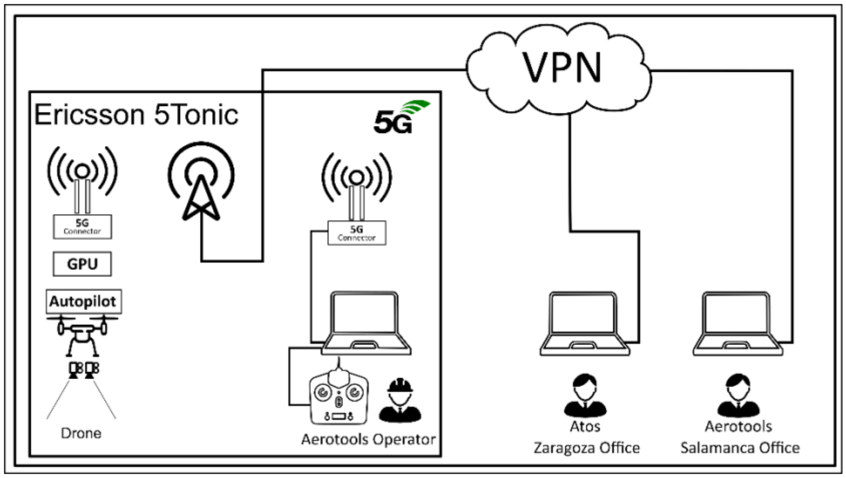

Figure 9 5TONIC VPN connecting UC3 development environments.

The **DRONE ENVIRONMENT** provides flying platform with specific equipment and integrated systems (Navigation, C2, payload, data transfer) and computing devices such as Raspberry PI3, Jetson Nano or similar, as well as stereoscopic cameras in dedicated stabilized gimbal to provide high quality footage and First Person View (FPV) camera to support drone´s operations when required by the operator. The following diagram shows the integrated devices onboard the drone:

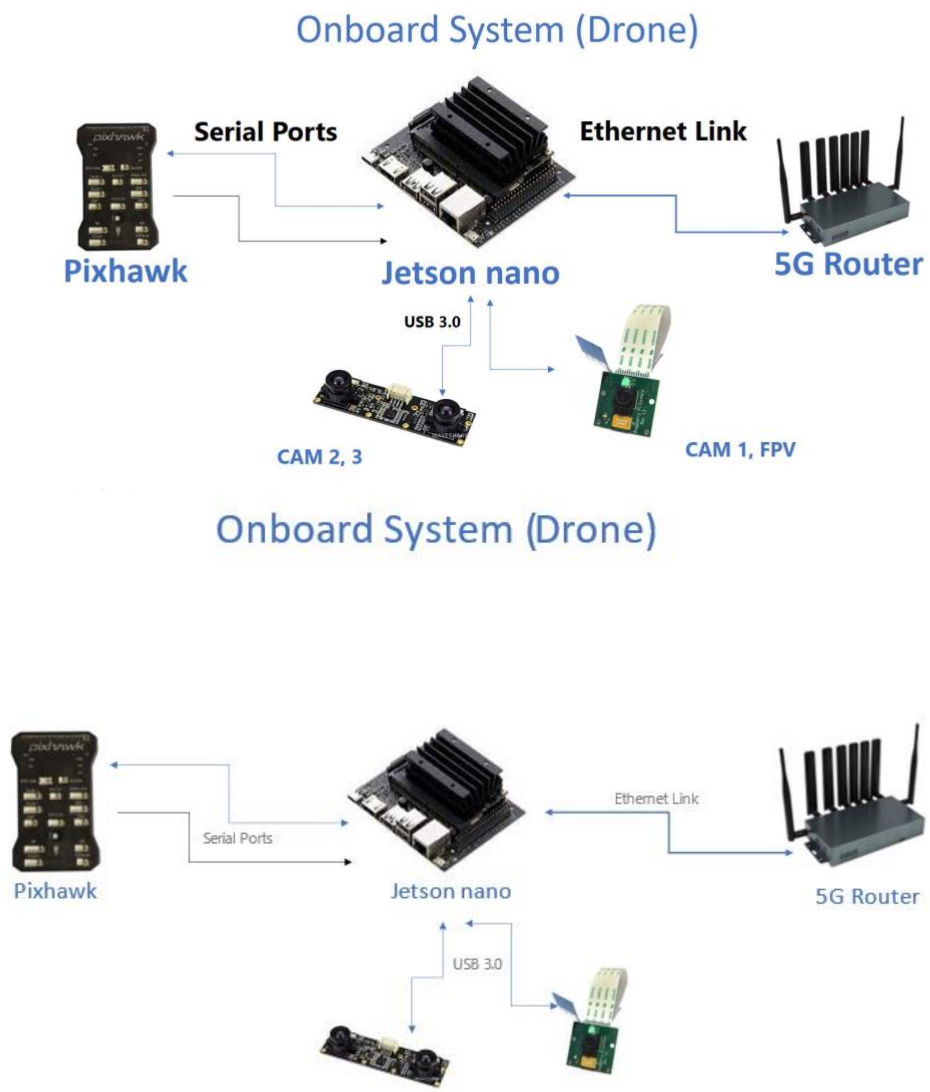

Figure 10 UC3 Drone Environment devices and connections.

- For a proper interpretation of the diagram, the Jetson nano is connected to the following devices:
    - Pixhawk using 2 serial ports:
        - Serial1: to be used by drone operator ground station (mission planner).
        - Serial2: to be used by Jetson nano to read GPS and orientation data.
- CAM1: to be used by Jetson nano to stream the FPV to drone operator ground station.
- CAM2-3: to be used by Jetson nano to push the images to the content broker.

And the dataflow designed for the integrated system is as follows:

- Pixhawk Serial1:
  - Shared by Jetson nano using ser2net software.
  - Mission Planner will use it in transparent mode, so will choose the IP/port to connect directly. This port is bidirectional so drone operator ground station can interact with the Pixhawk using mavlink.
- Pixhawk Serial2:
  - Jetson nano is reading GPS information and orientation + inertial data, using mavlink translator. This port is a one-way port, so it will only read from Pixhawk.
- CAM1 - FPV:
  - This video source needs to be streamed to drone operator ground station to be used as FPV (first person view).
  - Jetson nano is streaming it to drone operator ground station directly.
- CAM2 – CAM3:
  - This stereo CAM generates two video streams. Jetson nano is reading both in addition to position information from Pixhawk Serial 2. These images and the position information is packed and sent to content broker (RabbitMQ)
  - RabbitMQ stores CAM2/3 and positions information and serves to the consumer: Video 3D scan and Video detector.

The **AI TOOL ENVIRONMENT** for running the automated incidents detection tool is completing the testbed for UC3, developing several modules and functions based on Artificial Intelligence. In the context of this UC two AIF are being developed and tested: Anomaly Detection AIF and 3D Reconstruction AIF.

Anomaly Detection AIF: This artificial intelligence function is able to detect anomalies and locate them in the videos acquired by the drone. UC3 describes the inspection of critical infrastructure. This artificial intelligence function analyses the data sent by the drone and sends notifications and evidence to the pilot when an anomaly was detected, then the pilot may initiate the process of generating the 3D model of the incident scene for future evaluation using the 3D Reconstruction AIF. The implementation of this task involves the setup of several components which will pre-process the inputs and send them in the correct way to the AIF to perform the analysis. These components are:

- **On-board Data Server**: This component access to the cameras stream installed aboard the drone and the telemetry data bus (in MavLink format). It has a double function: to synchronise all data, since each telemetry data message and the cameras transmit at different frequencies and to emit this synchronised data at the selected frequency, in json format through RabbitMQ broker, feeding the AI function. This level of synchronism in the transmission of telemetry data and images is important for the proper functioning of anomaly detection AIF but it is fundamental in 3D reconstruction as we must ensure that the images sent and the position and orientation (GPS and Euler angles) of the cameras always correspond to the same instant of time.

- **RabbitMQ server**: The technology chosen to act as a messaging broker between the drone, the pilot and the different AIFs is RabbitMQ as it covers the following requirements:
    - It is fast enough, and it works with good metrics/monitoring.
    - Use standard protocols, publish/subscribe, request/response etc.
    - Completeness of messaging patterns.
    - Scales to 1 million messages per second.
    - Distributed.
    - JSON compliance.
    - This component works bidirectionally, it is the channel for sending configuration commands to the data server and events and notifications to the pilot and input data to the AIFs.

This AIF, responsible for anomaly detection and localization in video streams, use Detic as main detector and CLIP as visual-language model which will be the core of the search engine. When the user, the pilot in this case, wants to detect a concrete anomaly, he has to feed the AIF with a description using natural language, for instance: "*Overturned truck on the road*". CLIP will interpret the natural language search order done by the pilot and will identify where and when appear the anomaly described by this text in the sequence of images that the drone is sending. When this AIF detects an anomaly, it will send an event to the pilot through the message broker to trigger an order to initiate a 3D reconstruction if this is considered necessary.

3D Reconstruction AIF: This artificial intelligence function is in charge of generating the 3D model of the area where the anomaly has been detected. Once the pilot notices the detection event sent by the anomaly detection AIF, he shall, if deemed appropriate, initiate a circular flight (POI) around the point where the anomaly has been detected. At the same time the pilot will send through the RabbitMQ the request for video and telemetry acquisition to the on-board Data Server. The Data Server will start sending the necessary data to the 3D reconstruction AIF to start the process.

- **On-board Data Server:** Same functionality as in the previous AIF.
- **RabbitMQ server:** Same functionality as in the previous AIF.
- **Inertial Odometry & fast 3D model generator: The** 3D Reconstruction AIF needs images from the scene in different orientations to achieve a photorealistic result. The quality and the proper orientation of the images is the most important step in the process. Once the pilot sends the order to get the 3D model, the drone starts to fly around the area taking images and gathering the extrinsic parameters of the camera (position and orientation) in every instant of time. This component will perform a coarse reconstruction in real time aboard the drone, giving real time feedback to the pilot about the acquisition process. This allows the pilot to know whether the acquired data is enough, or it is necessary to continue with the flight, to get more images. Once the module identifies that the requirements of number and

quality of images acquired are met, it will send a notification to the system and the 3D modelling will be initiated by the AIF.

5.3 Validation procedures and preliminary results

For Phase 1 of the UC, 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 the 3.5 GHz band available at the 5TONIC network and measurements of throughput, latency and packet loss metrics will be performed.

The number of test cases will be 5:

- Initial integration to check visibility and connectivity among all the systems
- Indoor tests to check correct functioning within the indoor 5G network.
- Outdoor tests to check correct functioning in the outdoor 5G network where radio range is a relevant factor.

The different tests are shown in the following tables:

PHASE 1 - Initial set-up and developments	
Test Case #	**#0**
	Connection
Slogan & Objective	Testing of basic visibility and connectivity among all systems
Test Scenario (Pre-conditions)	Initial set-up for development
Expected Results (Post-conditions)	All systems connected and sending/receiving data
General Time Plan (Validation Campaigns)	Throughout first half of 2022 for testing initial set-up
Test Sequence	1. Procurement 2. Connecting 3. Visibility 4. Operation

Indoor tests:

PHASE 1 - Initial set-up and developments		
Test Case #	**#1**	**#2**
	Latency	**Reliability**
Slogan & Objective	Testing of communicationcommunication latency for drone control and video transfer.	Testing of reliability on C2 signals.
Test Scenario (Pre-conditions)	Initial set-up for development.	Initial set-up for development.
Expected Results (Post-Conditions)	C2 latency ≤ 50 ms Video latency ≤ 100 ms	C2 signal packet loss ≤ 1%.
General Time Plan (Validation Campaigns)	Throughout first half of 2022 for testing initial set-up.	Throughout first half of 2022 for testing initial set-up.
Test Sequence	1. Procurement 2. Connecting 3. Visibility 4. Operation	1. Procurement 2. Connecting 3. Visibility 4. Operation

Outdoor tests:

PHASE 1 - Initial set-up and developments		
Test Case #	**#3**	**#4**
	Latency	**Reliability**
Slogan & Objective	Outdoor Testing of communication latency for drone control and video transfer.	Outdoor Testing of reliability on C2 signals.

Test Scenario (Pre-conditions)	Outdoor conditions	Outdoor conditions
Expected Results (Post-Conditions)	C2 latency ≤ 50 ms Video latency ≤ 100 ms	C2 signal packet loss ≤ 1%
General Time Plan (Validation Campaigns)	Throughout first half of 2022 for testing outdoor scenario.	Throughout first half of 2022 for testing outdoor scenario.
Test Sequence	1. Procurement 2. Connecting 3. Visibility 4. Operation	1. Procurement 2. Connecting 3. Visibility 4. Operation

5.4 Next validation steps

Several challenges have been identified at this stage for the coming period.

The first challenge is operating the drone in outdoor scenario within the 5TONIC and AI@EDGE integrated environment, with adequate 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…).

The second is related to the AI functions for detection and location of incidents based on anomaly detection and 3D modelling. Development and integration of the AIFs in the Use Case workflow is a key step.

And, ultimately, the integration of AI@EDGE functions in the 5TONIC ENVIRONMENT to provide assisted monitoring operation.

6. Use Case 4

High-end connectivity has become an important issue for the airlines in the last couple of years, and many airlines nowadays offer passengers' personal devices to be connected to the cabin network besides the in-flight entertainment and connectivity (IFEC) systems. At the same time, airlines are increasingly investing in data analysis to provide the best experience for passengers with personalized services as well as to create operational efficiencies on the plane. Further, edge computing can bring several advantages to the aviation sector and aid to increase the performance of inflight networks. Since this technology is distributed which can bring computation, storage, and processing closer to the devices, it will avoid relying purely on remote infrastructures.

Within the framework of the Connect Compute platform of AI@EDGE, the UC4 is considered as the far edge located at the radio access site. SPI will also host at its premises user plane function (UPF) of the 5G core, with the control plane that will be located on the ground at FBK site. This allows the aircraft to be envisioned as an edge-cloud connected to the ground network through Satcom (or air-to-ground) technologies during flight time.

Initial UC4 objectives' timetable for month 18

Table 3 shows a summary of the activities for UC4 by month 18 which was discussed in more detail in D5.1 [2]. Here we briefly explain the works that have been done by each of the UC4 partners to fulfill the requirement of the use case for milestone MS5.3.

- **Equipment and development:**

  SPI has worked on the development and assembly of the Aero edge-cloud testbed. Following the initial time plan of the UC4, almost all the components of the testbed are set up correctly, including the required OS and other software for different components. The remote access through a VPN is already configured and we are able to access the testbed remotely.

  Moreover, SPI tested the integration of Kubernetes on the seatbacks and the servers. Initially, a small cluster is deployed for the test purposes, and the full integration was done in parallel with the final adjustments of the testbed, by the end of M18. More detail about the test rack and components are presented in Section 6.2.

  ATH has worked on the design of the 5G network that supports UC1, 2, and 4 (cf. Section 2.3), and has internally worked on the solutions to be deployed accordingly in Q3 and Q4 of 2022. Namely, concurrently with the activities of WP4, ATH has worked on the centralized control plane deployed at FBK, which will be upgraded to a 5GC control plane after M18; and has worked on the edge node solutions with user plane functionalities to be installed for UC4 at SPI's premises. The full implementation of the user plane functionalities at the SPI premises is, however, will be completed smoothly after M18.

- **Service development:**

  AIFs: Regarding the service deployment, SPI already provided the anonymized IFEC data set for the AI/ML applications. It is now being analyzed within the team and we are progressing to develop the first version of this AIF. Besides the recommendation system which was introduced in deliverable D2.1 [3] and discussed with some level of detail in D5.1, we may add a new AIF, called as RDU swapping, to the use case. Both applications are introduced and explained in Section 6.3.1.

  MPTCP: The first version of MPTCP-proxy has been developed and deployed as a container and is being tested at FBK. In the meantime, a laptop with MPTCP capability is being used for testing as a client. Integration of MPTCP into the RDU is underway.

  Video Streaming: The video streaming and transcoding application (VST) allows 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.

By M18 ITL will test the first version of the VST ready to run on a K8s based system located at its premises.

5G RAN: As outlined in D5.1, SRS will provide UC4 with its 5G RAN via srsRAN - specifically srsENB and srsUE. As of the latest release of srsRAN (srsRAN 22.04) both srsENB and srsUE support 5G SA. Currently srsRAN supports 5G SA mode in FDD bands (e.g. n3).

Specific to UC4, initial interoperability testing will focus on running a base implementation of srsRAN on the RDU3 hardware platform and the rest of the UC4 testbed. This will revolve around running srsUE on the RDU3 platform, and srsENB on the local servers, and maintaining a stable connection between these. Once this has been implemented and tested, any further UC4 specific modifications can be made to srsRAN.

As previously stated in D5.1, further modifications may be implemented in an ad hoc fashion when needed, and when aligned with the roadmap of both AI@EDGE and srsRAN.

Table 3 UC4 Objectives for M18

	Equipment development roadmap	**Services development roadmap**
Phase 1 **(To be completed by Q1 2022)**	• Complete assembling the Aero edge-cloud testbed and validate that hardware components are correctly set up. **(Completed)** • Complete installing the OS in each device and complete all configurations (ssh access, DNS, DHCP services, remote power on/off, etc.). **(Completed)** • Development of ITL testbed for video transcoding. **(Completed)**	• Output the first anonymized IFEC dataset ready for use to develop the content curation AIF. **(Completed)** • Initial development of the video transcoding software (in lab development). **(Completed)** • Initial development of the MPTCP and MPTCP proxy (in lab development). **(Completed)**
Phase 2 **(To be completed in Q2 2022)**	• Integration of Kubernetes in the Aero edge-cloud infrastructure. **(Completed)** • 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; 5G SA will be deployed after M18. **(Will be completed after M18)** • Development of the testbed for the hardware acceleration card. **(Integration with the test rack is ongoing)**	• Development of 5G RAN NSA **(In lab development is ongoing)**. • Development of the content curation software **(Initial version is completed).** • Development of the hardware acceleration software (in lab development). – **(Integration with the UC4 test rack will be investigated after M18.)** • Development of Video Transcoding. **(In lab deployment is completed)**

		• Development of the MP-TCP proxy. **(Integration with the test rack is ongoing)**

6.1 Validation objectives

Relying on the advantages of edge computing, the MEC in UC4 of the AI@EDGE is beneficial to improving the performance and connectivity of the inflight networks. On the one hand, the MEC-enabled 5G system for an aircraft offers real-time, low-latency, and high-bandwidth access to the local data network located inside the aircraft, which saves on the bandwidth of the satellite or air-to-ground technologies. On the other hand, by employing such a technology for the IFEC systems, there could be a more powerful network in terms of computation, processing, and storage due to the aggregation of all physical devices (RDUs, and SCUs) onboard a civilian aircraft. Such a network will have better resilience to a link failure compared to the traditional centralized networks. Thus, taking advantages of both AI/ML algorithms and MEC technology, UC4 of AI@EDGE project targets an evolved IFEC system that is able to offer passengers with different personalized packages of content produced by different content producers; and to predict any possible failure of the seat back screens based on the environmental conditions.

Following the main objectives of the UC4 that were already introduced in D2.1 and elaborated in D5.1, the main focus of this use-case is the implementation of an aero edge-cloud network which includes seat backs (RDU3), aero-certified and COTS servers. To this end, four test-cases are in progress as the use-case validation:

- Demonstration of Artificial Intelligence Functions (AIFs) as the MEC apps within an aero edge-cloud network. Two possible AIFs for this use case are elaborated in Section 6.3.1.
- Deployment of MP-TCP to increase the performance of existing interfaces between the content and the end-user, including WiFi, cellular (4G/5G), and wired technologies. Integration of LiFi is another possibility that is under investigation for UC4. This test-case is discussed in Section 6.3.2.
- Providing an application for the video streaming and transcoding with the capabilities to adapt the video streaming based on the users' request. More details about the video transcoding application are discussed in Section 6.3.3.
- 5G connectivity and local traffic breakout validation which is a series of tests that will be run as soon as the 5G network (centralized control plane and distributed user plane) is fully deployed, in the second half of the project's lifetime. These tests will validate the operativity of the 5G setup, including the correct reachability of the edge servers by the connected devices. These tests are functional and precede the final end-to-end performance tests of the use case.

6.2 Validation scenario

Figure 11 illustrates the whole architecture of use case 4, smart content & data curation for in-flight entertainment services. It represents the realization of the 5G systems and edge site connectivity that is considered for this use case of the AI@EDGE. Moreover, it is considered as the far edge located at the radio access site with the capability of communication with the upper-level orchestration within the framework of AI@EDGE.

Besides two possible radio access technologies (WiFi and LiFi), 5G network will be adapted into the system in order to provide simultaneous and efficient connectivity services to the IFEC systems as well as to the passengers on board. In this architecture, as is shown in Figure 4 and in Figure 11, the user plane function of the 5G core will be kept locally and the user plane traffic will be routed to the local data network at the edge site. Whereas the control signaling will be passed toward the central core network, located on the ground, through a satellite communication link. The services deployed on board in this configuration (discussed in Section 6.3), such as content curation and RDU swapping (i.e., AIFs), are served as the MEC apps which are isolated from a data network on the ground. Leveraging on this architecture and the advantages of the MEC, there would be a saving on the bandwidth of satellite communication as well as a reduction of the service latency. Besides this, and in order to further increase the capacity of the radio access network, the integration of the multipath TCP is under development (detailed in section 6.3.2). More details on the Hardware and software implementation/integration are discussed below.

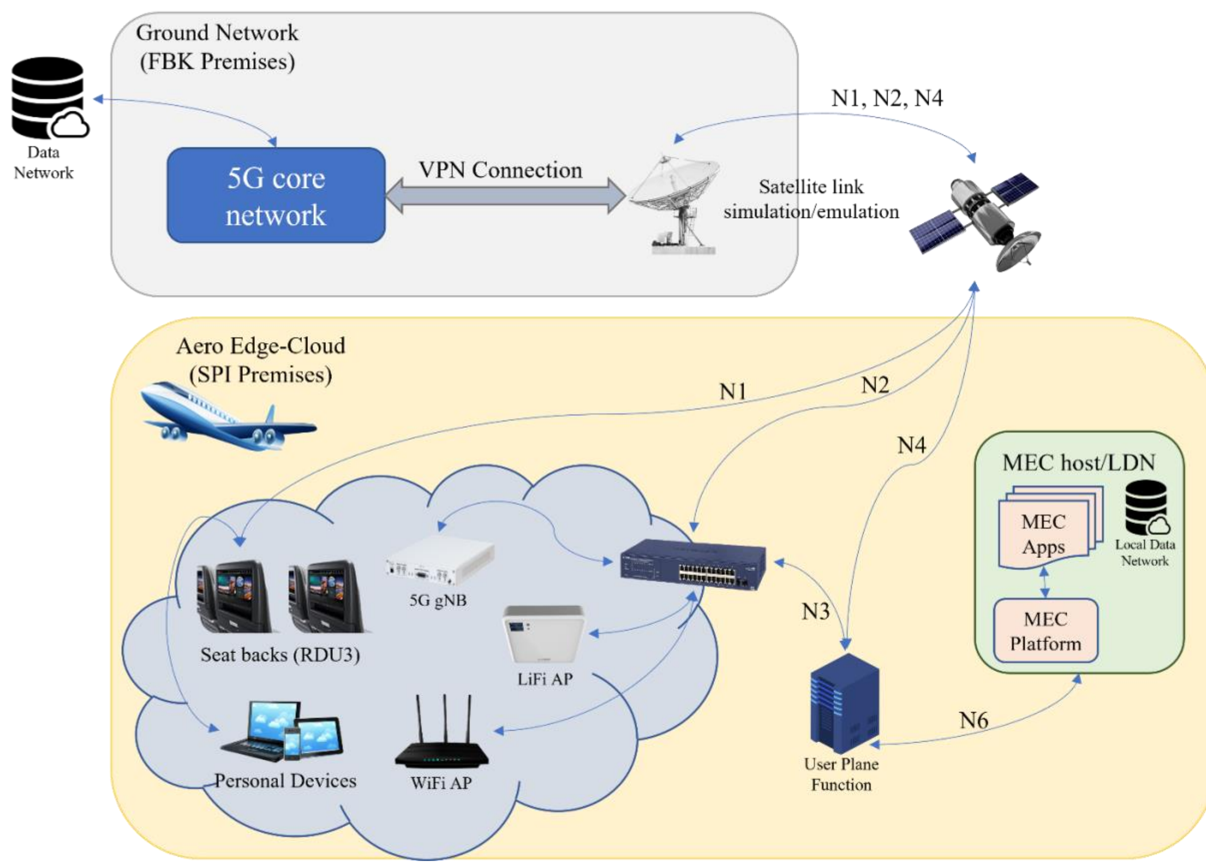

Figure 11 Use case 4 overall architecture.

1. 5GC - UPF

As anticipated in D5.1 and recalled in Figure 4, the 5GC deployment that serves this use case will consist of a centralized 5G control plane deployed at FBK and of a distributed on-site user plane. The 5GC hardware and software are provided by ATH. As reported in D5.1, the edge core network functionalities are planned 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 functionalities will be deployed over a similar HW shared for the purposes of UC1, 2, and 4 (cf. Section 2.3).

Such hardware choices will be finalized at the moment of the actual network deployment, in Q3 and Q4 of 2022.

1) 5G RAN SW

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. The generic baseband I/Q interface allows srsUE to be used with most SDR frontends including those based on Analog Devices and Lime Microsystems RFICs.

As previously mentioned srsUE will be run on the RDU3s, to provide a 5G SA connection between them and the srsENB instance running on a COTS server compute platform. Both srsUE and srsENB will require SDR front ends to be connected to the computer hardware.

2) Test rack (HW/SW)

The test rack that is assigned to the AI@EDGE project contains 21 RDU3, 1 SCU, and 1 supermicro, as well as other supporting hardware. The main components of the test rack, RDUs and SCU, are aeronautically certified products of SPI, and they are being used in commercial aircraft services. Each of the RDUs is a capable computing platform that aircraft passengers use for in-flight entertainment such as watching movies, listening to music, ordering meals, or shopping. The RDU models in the AI@EDGE rack have a resolution of 3840x2160 pixels (4K) with a 13.3” screen size. The SCU is used as media servers in the aircraft to handle certain networking related tasks as well as make media content available to passengers. They are usually located at the front of the aircraft with other computing equipment.

Figure 12 A shot of AI@EDGE-UC4 test rack.

As illustrated in Figure 12, the components of the use case 4 have been mounted on a movable hardware rack that has a small but enough form factor to be moved between different rooms for project needs or demonstration purposes.

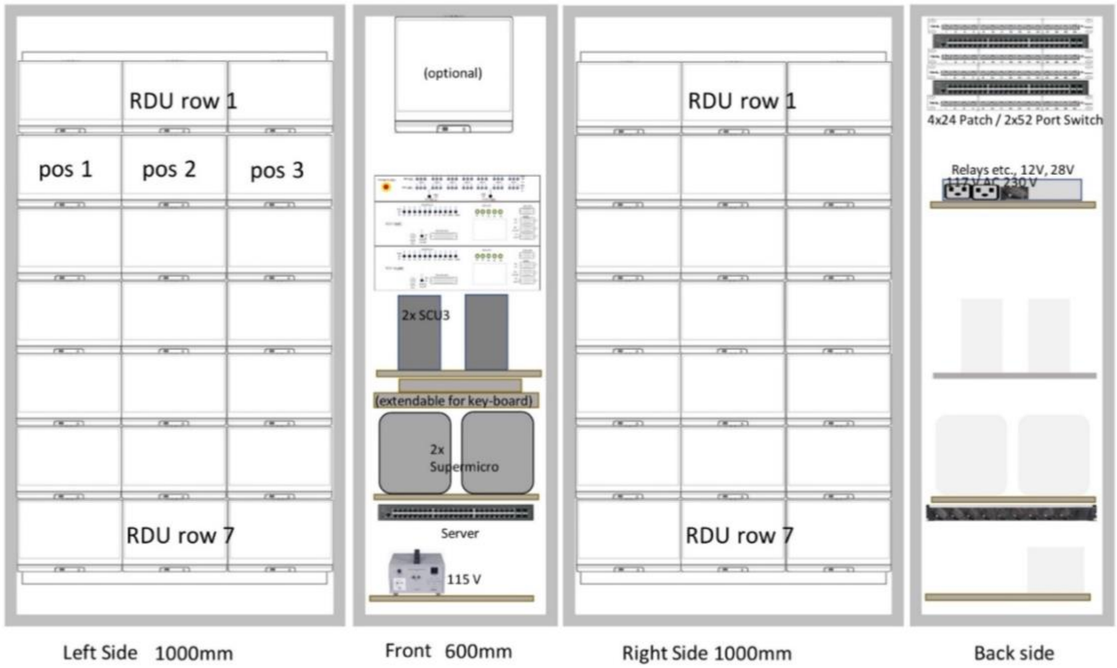

Figure 13 Physical layout of the test rack.

Figure 13 shows an overview of the whole layout of all four sides of the rack. Of course, not all the sides are assigned to the use case 4 of the AI@EDGE project, but part of it (the left side) is specifically for this project while some other components could be shared with the other SPI activities.

It has a front panel with switches that allows us to control power supply to all devices, as well as several connectors that can simulate inputs from the aircraft to the SCUs. The power switches can also be put in an automatic mode which means that the configuration is left to the relay boards which can be controlled remotely by a Raspberry Pi inside the test rack. This has the advantage of being able to turn on/off each RDU separately without physical presence at the test rack, which means experiments and tests can be executed remotely by turning on the required devices before the test and turning them off afterwards, instead of leaving all devices running indefinitely. This also facilitates the integration of other software that is being deployed by the other partners. Some tests also result in a misconfiguration that requires a power reset, which can also be done remotely. Ultimately the relay boards enable significantly lower power consumption and better remote operability.

The different devices of the rack require different power supply parameters. This is the reason for the multitude of voltages that are mentioned in the figure with 28V for the RDUs, 12V for fans and the relay boards, 115V for the SCUs and 230V for the servers and switches.

The RDU devices are mounted on the sides of the rack which provides physical access to all the displays for interaction through the touchscreen interface. They are all connected to network switches which gives us the option to control them remotely and lets them communicate with other devices of the network.

The operating system of the RDUs that are in service in commercial aircraft is tailored to inflight entertainment purposes. It has an intuitive touch screen interface for navigating and interacting with the media content and is configured with the necessary security measures in place that prevent misuse of the system. These properties are not ideal for a development and testing platform where the needs are more focused on supporting unusual networking configurations or external hardware that is not part of the IFEC product family. For this reason, the operating system on the RDUs has been modified from the flight-ready production system to a more custom development system with support for running Docker containers. This functionality can be used for simpler development of software for different purposes such as running ML/AI applications, video streaming and transcoding, or the networking software needed for testing new project implementations.

6.3 Validation procedures and preliminary results

In this chapter the procedure of the validation is given, separated into 4 test cases, where each test case is validating a separate aspect of the demonstrator.

6.3.1 Test case 1: AIFs development

Developments & Integration	
Test case # 1.1:	Recommendation System
Slogan & Objective	• Suggesting relevant content to the passengers on board.
Test Scenario (Pre-conditions)	• Analysis of IFEC dataset to develop a model-based collaborative filtering for the content curation (More details are described below).
Expected Results (Post-Conditions)	• To develop an ML-based application for the content curation for on board passengers
General Time Plan (Validation Campaigns)	• First complete version of recommendation system - Q4 2022 • Integration of recommendation system into the aero edge-cloud – Q1/Q2 2023
Test Sequence	1. Preparation of anonymized IFEC dataset for the content curation. 2. In-lab development of content curation software. 3. Integration of the application with the aero edge/cloud infrastructure.

Initial developments of the test case:

SPI uses the library "Surprise" that contains machine learning algorithms which predict "ratings" of unrated songs, similar to the one shown in Table 4.

Table 4 Example of table with unrated items.

	Song 1	Song 2	Song 3	Song 4
User A	4	2	?	5
User B	1	3	2	?
User C	2	?	5	3
User D	?	5	2	3

The dataset extracted from our database does not contain explicit rating given by passengers, but we have created a score which reflects how much a user likes a song based on the duration of the songs played. We used different methods to estimate this score. (Binary rating based on a threshold, rating from 1 to 5, like the one shown in Table 4).

We created user ids and song ids by combining different columns (flight Id, bundle Id, LRU Id, seat number). Then some algorithms were tested and their performance with the Precision@k and Recall@k metrics were evaluated. The precision@k ond recall@k are meant to evaluate the recommendations. The precision@k is the proportion of the recommended relevant items in the top-k sets, while the recall@k is the proportion of relevant items in the top-k recommendations. However, the suggested model has some drawbacks. It only predicts ratings for the users who are in the training set. To consider the inputs from a new user we would need to build a dynamic system.

To build our collaborative model-based recommendation system, we performed the following tasks:

- Extract datasets with relevant features from our database by executing SQL queries
- Load the datasets on a Jupyter Notebook and perform data preprocessing tasks (remove duplicates, null values if necessary)
- Create the required features for the training set
- Test different algorithms
- Evaluate the performance of the different algorithms with some metrics (Precision@k and Recall@k)
- Identification of the relevant features in our database to extract datasets with the required features for the training set.

The library Surprise contains built-in algorithms that take as input a dataset with three columns: User Id, Item Id, and Rating, as shown in Figure 14.

	UserId	SongId	rating
0	2676548917	110844390	4
1	2676548917	110844391	4
2	2676548917	110844392	3
3	2676548917	110844393	4
4	2676548917	110844394	1
...	...	...	...
1621066	3030630946	244981713	4
1621067	3030630946	24497970	1
1621068	3030630946	24436231	3
1621069	3030630946	24436231	2
1621070	3030630946	24436231	2

1621071 rows × 3 columns

Figure 14 Surprise dataset object.

The user Id is identified by combination of several other items including the flight Id, the bundle Id, the LRU Id and the location. Applying this method, the user Id is independent of any specific person, which preserve the privacy issues, instead, it represents a particular seat in a particular flight.

It is assumed that on a particular flight there is one user per seat. The Media Id has been combined with the track number to create the song Id. The rating has been then calculated by making the difference between two log dates of two successive actions. This difference corresponds to the time spent by a user on a particular item (e.g, song). A binary rating is tested, where 0 is assigned if the listening time is less than 20s and 1 otherwise. Moreover, a rating from 1 to 5 with different thresholds is elaborated.

In order to train the existing dataset and make a prediction, the following algorithms have been used: ALS, SGD, SVD, SVD ++, CoClustering and NormalPredictor. After the training phase, a prediction for the existing users can be generated.

- **Evaluation of the performance of the algorithms**

The dataset is split into a train set and a test set. The true value is compared to the predicted value to evaluate the model. The cross-validation method has been used with the RMSE (Root Mean Squared Error) metric to evaluate the quality of the model, but we cannot use these results to compare different methods of rating. A binary rating is more likely to give a lower RMSE than a rating from 1 to 5 because the scale is smaller. Therefore, it is preferable to use the precision@k or recall@k.

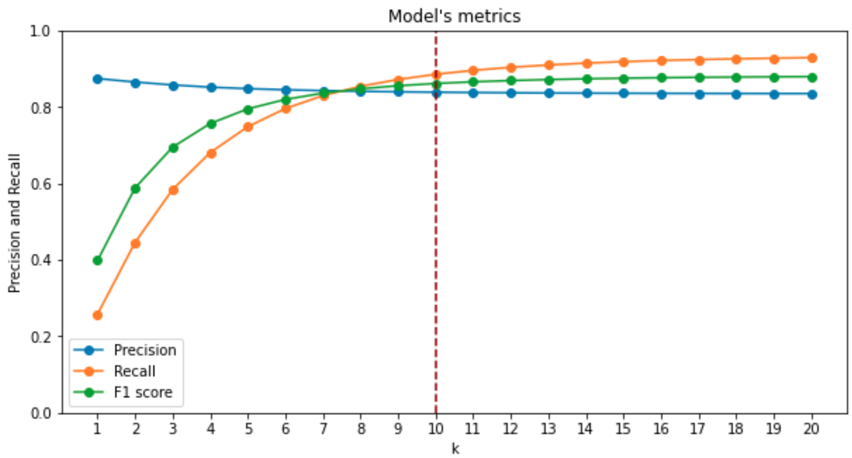

Figure 15 Precision@k and Recall@k.

Among the possible algorithms introduces earlier, we used the ALS algorithm to generate recommendations due to its lower execution time. The scores for precision and recall were almost similar for all the algorithms tested.

- **Drawback of the collaborative filtering model-based algorithm**

The collaborative filtering model-based algorithms from the library Surprise only works for users who are in the training set. Therefore, it does not work for a new user. To include the inputs from a new user, we need to build a dynamic system.

- **Popularity based model for each airline**

The popularity-based model is a simple model where the most popular movies are recommended. We will measure the popularity of a movie with three components: the watching ratio calculated from the PAX logs, the IMDB ratings, and the release date of the movie. Then we will sort the movies and recommend those with the highest popularity. The dataset extracted from our database contains logs for a time range of one month (logs between 1st of November 2021 and 1st of December 2021).

To calculate the watching ratio, we calculated the watching time by using the records of the following events:

- VideoPlayback: records when the passenger starts the video,
- VideoPlaybackStop: records when the passenger stops the video,
- PlaybackComplete: records when the video plays to completion.

We, then, sorted these records by user Id and log date and grouped them by user Id and we made the difference between two successive records within each group defined by one user Id. Moreover, to calculate the ratio, we used the column "Runtime". It is worth noticing since the runtime is not available for some of the videos in our database, we must use the IMDB dataset to extract the runtime.

As we do not have the IMDB Id in our database, we developed an algorithm which matches titles in the airlines' dataset with the titles from the IMDB dataset. The algorithm compares strings with a certain tolerance and applies it to the titles of the movies.

Finally, the airline's dataset will be merged with the IMDB dataset to extract the IMDB ratings and the release date of the movies.

Developments & Integration	
Test case # 1.2:	RDU Swapping
Slogan & Objective	• To gain knowledge and to build expertise in predictive maintenance of IFEC systems, • To analyze the energy efficiency of the RAVE IFEC system to contribute to greener IFEC systems, • To predict a possible failure of a seatback (RDU3).
Test Scenario (Pre-conditions)	• Validation and analysis of datasets for development of an application to predict possible RDU failure to be replaced (More details are described below).
Expected Results (Post-Conditions)	• To develop an ML-based application to predict when an RDU has to be replaced based on the on board environmental conditions.
General Time Plan (Validation Campaigns)	• First complete version of RDU swapping - Q4 2022 • Integration of RDU swapping into the aero edge-cloud – Q1/Q2 2023
Test Sequence	1. Preprocess the dataset to make it suitable for further machine learning experiments, 2. ML experiments and Results, 3. Deploy the model, Design GUI and Complete the application.

Initial developments of the test case:

- **Dataset Overview:**

The available data is from LogBitDataset, which contains various logging information of Removable Display Units (RDUs) of multiple flights, and according to the practical requirement, the prediction of 'whether an RDU should be swapped or not' has been proposed as the use case. Machine learning methods shall be implemented to tackle this use case.

First of all, it is worth to mention that the dataset was exported from LogBitDataset. There are 27 features (columns) in the dataset (an example is demonstrated in Figure 16), and they are: FlightId, ACDInfoId, SwVersion, AirlineCode, TailNumber, FlightStartTime, FlightEndTime, DocumentId, Location, HwPN, HwSN, Etc, min2002Temp, avg2002Temp, max2002Temp, resetCount, NormalPowerOff, minLinkState, maxLinkState, AvgVoltage, MaxVoltage, MinVoltage, voltageTripCount, AvgLCDCurrent, SwDate, Swapped and FlightDuration. (Described in Table 5).

FlightId, FlightStartTime, FlightEndTime and DocumentId are removed before experiments. The data types of these features are mainly categorical (textual), time(date), and numeric.

Table 5 Dataset features.

Feature Name	Description
FlightId	FlightId for flights, each FlightId represents a unique flight
ACDInfoId	Aircraft Configuration Database ID to identify and load specific Rave configuration
SwVersion	Current software version of the RDU
AirlineCode	Code for the airline companies, each airline has a unique code
TailNumber	Aircraft registration number
FlightStartTime	The real-world time the flight started
FlightEndTime	The real-world time the flight ended
DocumentId	Represents a bundle of logs offloaded from one (or sometimes more flights because of extended maintenance window)
Location	The seat number of the RDU within the aircraft
HwPN	Hardware part number identifies the hardware version of the Line Replaceable Unit
HwSN	Hardware Serial number, a unique identifier for each RDU
Etc	Elapsed time count – the accumulated seconds that an RDU is powered on
min2002Temp	Minimum temperature recorded on that RDU during the flight
avg2002Temp	average temperature of an RDU during the flight
max2002Temp	maximum temperature recorded on that RDU during the flight
resetCount	The number of times the RDU was reset during the flight
NormalPowerOff	The number of Normal powers off an RDU had during the flight
minLinkState	The min LinkState columns signify is the connection on any port was up or down during the flight.
maxLinkState	The max LinkState columns signify is the connection on any port was up or down during the flight.
AvgVoltage	average voltage of an RDU during the flight
MaxVoltage	maximum voltage recorded on that RDU during the flight
MinVoltage	minimum voltage recorded on that RDU during the flight
voltageTripCount	The total number of times the voltage violated the threshold during the flight
AvgLCDCurrent	average LCD backlight current recorded on that RDU during the flight
SwDate	Last software update date
Swapped	Whether an RDU has been swapped or not
FlightDuration	The duration of the flight in seconds

Here are some examples:

SwDate: '4/10/2020 00:35:29', '4/10/2020 01:58:59'…. time series values

TailNumber: 'JA213A', 'JA218A', 'JA135A'…. categorical values

max2002Temp: 54.0, 64.5, 58.0, 62.0…. numeric values

Swapped: 0, 1

	FlightId	ACDInfoId	SwVersion	AirlineCode	TailNumber	FlightStartTime	FlightEndTime	DocumentId	Location	HwPN	HwSN	Etc	min2002Temp	avg2002Temp
29	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	6J	00-5120-02	137589	1713540	55	60.1875
30	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	36H	00-5120-02	137280	1399980	48.5	54.14285714
31	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	37A	00-5120-02	137566	1692900	49	54.35714286
32	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	11K	00-5120-02	137274	1403340	44.5	48.16666667
33	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	21K	00-5120-02	137380	1394460	49.5	58.94736842
34	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	18C	00-5120-02	137320	1404720	53	55.75
35	1778990	132735	41.07.03	ANA	JA150A	12/7/2020 8:14	12/7/2020 9:10	c54f6911b5caad292a1755b3d677aeb2	7A	00-5120-02	137420	1400520	54.5	58.375

Figure 16 Screenshot of several rows of the partial features in the dataset.

There are 15,618,913 rows in the dataset and after preprocessing procedure, 15,420,437 has a value of 0 in Swapped and 54,369 of value 1. '0' indicates the RDU doesn't need swapping, and '1' indicates it does. Based on the target label situation ('0' and '1'), we conduct the ML experiment. Before starting the ML experiments, data preprocessing is necessary. We use Numpy, Pandas and scikit-learn libraries to fulfill these tasks.

- **ML experiments and Results:**

We choose H2O (after comparing with several other platforms) as the ML platform to perform the experiments. In the preprocessed dataset, there are 22 features: ACDInfoId, SwVersion, AirlineCode, TailNumber, Location, HwPN, HwSN, Etc, min2002Temp, avg2002Temp, max2002Temp, resetCount, NormalPowerOff, minLinkState, maxLinkState, AvgVoltage, MaxVoltage, MinVoltage, voltageTripCount, AvgLCDCurrent, SwDate, Swapped and FlightDuration. After evaluating the weights of each feature in the dataset, we start from 12 features: ACDInfoId, AirlineCode, TailNumber, FlightDuration, EndOfLogETC, rawResetCount, Swapped, SwVersion, SwDate, HwPN, HwSN and NormalPowerOff.

We keep 20% of the dataset as the testing dataset before applying the balancing method. The remaining 80% data is used as the training dataset. After the 1st round of experiments on 12 features, we rank the importance of each feature in the experiment and furtherly keep the top 6 important features and conduct the 2nd round of experiment. These 6 features are: ACDInfoId, AirlineCode, TailNumber FlightDuration, EndOfLogETC and HwSN.

The results of all experiments are filled in a table ((Figure 17). H2O has two different platforms: R Cluster in Jupyter Notebook version and Flow Web version. Both have almost the same functions and R version is Coding based and Flow is GUI based.

			F12 Flow F1 score		F12 R Cluster F1 score		F6 Flow F1 score		F6(set1) R Cluster F1 score		F6(set2) R Cluster F1 score		f12 manual smote		f6 manual smote	
max model	max secs	class_sampling_factors	running time	f1	running time	f1	running time	f1	running time	f1	running time	f1	running time	f1	running time	f1
10	3600	-	01:07:40.722	0.6558	27.75147	0.719257	-	-	27.81709	0.825863	-	-	-	-	-	-
10	3600	0.005, 1	-	-	28.84222	0.822216	-	-	26.6119	0.813837	-	-	-	-	-	-
10	3600	0.01, 1.	-	-	27.85497	0.856315	-	-	26.34569	0.840250	-	-	-	-	-	-
10	3600	0.05, 1.	-	-	27.67202	0.877397	-	-	25.02766	0.869197	-	-	-	-	-	-
10	3600	0.08, 1.	-	-	28.19353	0.880690	-	-	26.20694	0.872235	-	-	-	-	-	-
10	3600	0.1, 1.	-	-	27.46146	0.881575	-	-	25.97841	0.872877	-	-	-	-	-	-
10	1800	0.1, 1.	-	-	23.09487	0.881714	-	-	20.00941	0.869600	-	-	-	-	-	-
10	1800	0.1, 1.	-	-	20.57505	0.880222	-	-	19.86418	0.871244	-	-	-	-	-	-
10	1800	0.1, 1.	-	-	20.58567	0.879992	-	-	20.08418	0.871696	-	-	-	-	-	-
10	1800	0.1, 1	-	-	20.63322	0.878529	-	-	18.9968	0.872588	-	-	-	-	-	-
5	1800	0.1, 1.	00:43:49.350	0.8751	13.91566	0.881707	-	-	12.99551	0.871747	-	-	-	-	-	-
5	3600	0.005, 1	00:30:53.98	0.8098	15.93664	0.832334	-	-	16.06741	0.834435	-	-	-	-	-	-
6	3600	0.01, 1.	00:44:02.212	0.8400	18.62383	0.856843	-	-	17.67502	0.849531	-	-	-	-	-	-
7	3600	0.01, 1.	00:28:11.967	0.8452	21.60625	0.85181	-	-	19.85162	0.848593	-	-	-	-	-	-
7	1800	0.01, 1	00:28:11.573	0.8287	17.18833	0.842121	-	-	15.94013	0.837871	-	-	-	-	-	-
6	1800	0.01, 1	00:45:30.399	0.8413	15.70891	0.850656	-	-	14.6892	0.842451	-	-	-	-	-	-
6	1800	0.1, 1.	02:06:16	0,8741	15.61098	0.880660	00:39:21	0.8550	14.51756	0.865921	29.38895	0.848813	3.331574	0.828337	3.024485	0.802769
6	1800	0.1, 1	00:38:11	0.8757	22.71082	0.869760	00:32:22	0.8515	14.7184	0.866071	14.77576	0.868992	3.007836	0.826434	2.890934	0.808491
6	1800	0.1 ,1	00:39:14	0.8614	15.2184	0.878220	00:35:24	0.8664	14.2568	0.873577	14.77971	0.870286	2.97319	0.826772	2.973289	0.807628
6	1800	0.1 ,1	00:30:25	0.8709	14.36851	0.880452	00:44:31	0.8657	14.39157	0.869480	14.02871	0.870873	2.885386	0.830999	2.946818	0.804097
6	1800	0.1, 1	00:27:35	0.8654	15.41938	0.881411	00:39:11	0.8667	14.68452	0.870218	14.64521	0.870028	2.925042	0.834366	2.862598	0.810667
				0.868		**0.878**		**0.861**		**0.869**		**0.866**		**0.829**		**0.807**

Figure 17 Experiments Results.

Figure 17 shows the performances of the experiments with different parameter settings and feature sets. Since the dataset is imbalanced, two main methods are adopted to deal with the imbalance problem: SMOTE (Synthetic Minority Over-sampling Technique: a method to deal with imbalance dataset) and built-in auto-balancing method provided by the platforms, which are also included in the Figure 2. When running experiments on H2O, there are many parameters that can be manually configured and here we present 3 important parameters: Max model (maximum number of models tried), Max secs (maximum seconds that the models run in total) and class_sampling_factors (the ratio majority class decreases, and minority class increases, and this is parameter of the built-in auto balancing method).

The results are close to each other, and after comprehensive consideration (from timewise and score), The model generated from 12 features dataset running on H2O with R cluster can provide the satisfying result (last row in table with bold font is the average result of previous 5-row experiments: 0.878).

Among the implemented models, the DRF (Distributed Random Forest) algorithm-based model averagely outperforms others based on our results.

6.3.2 Test case 2: Multi-path TCP and MPTCP proxy

PHASE 1 - Initial developments	
Test case # 2:	MPTCP
Slogan & Objective	• To test multi-path aggregation at the transport layer in multi-connectivity multi-RAT scenarios.

Test Scenario (Pre-conditions)	• Test the possibility of increasing TCP throughput by using all links at the same time. • Test the ability to keep the end-to-end TCP connection when one of the links is lost or degraded.
Expected Results (Post-Conditions)	• Increase TCP throughput to roughly the sum of all available links. • The end-to-end TCP connection still works when a link fails or degrades.
General Time Plan (Validation Campaigns)	• Complete version of RDU MPTCP – Q3 2022 • Setup and test - Q4 2022
Test Sequence	1. Test with two links equivalents in terms of performance. 2. Test with two links with different delay or bandwidth. 3. Test in case one of the links is failed. 4. Test in case one of the links is degraded. 5. Test in case one of the links is restored after failing/degraded.

Initial developments of the test case:

So far, an initial implementation of MPTCP is deployed on a single laptop at SPI with collaboration of CNAM. In this phase of MPTCP deployment, the laptop uses 2 interfaces (WiFi and LTE) and download data directly from an MPTCP-enabled server located at CNAM.

The green curve of Figure 18 shows the aggregate data rate where we have MPTCP enabled on both client (at SPI) and server (at CNAM); whereas the red curve represents the normal WiFi data rate and the blue is dedicated to the LTE performance. From this initial test, we can see not only higher throughput when having an MPTCP-enabled configuration, but also a continues connection even if one interface fails (The LTE interface were manually dropped for some seconds).

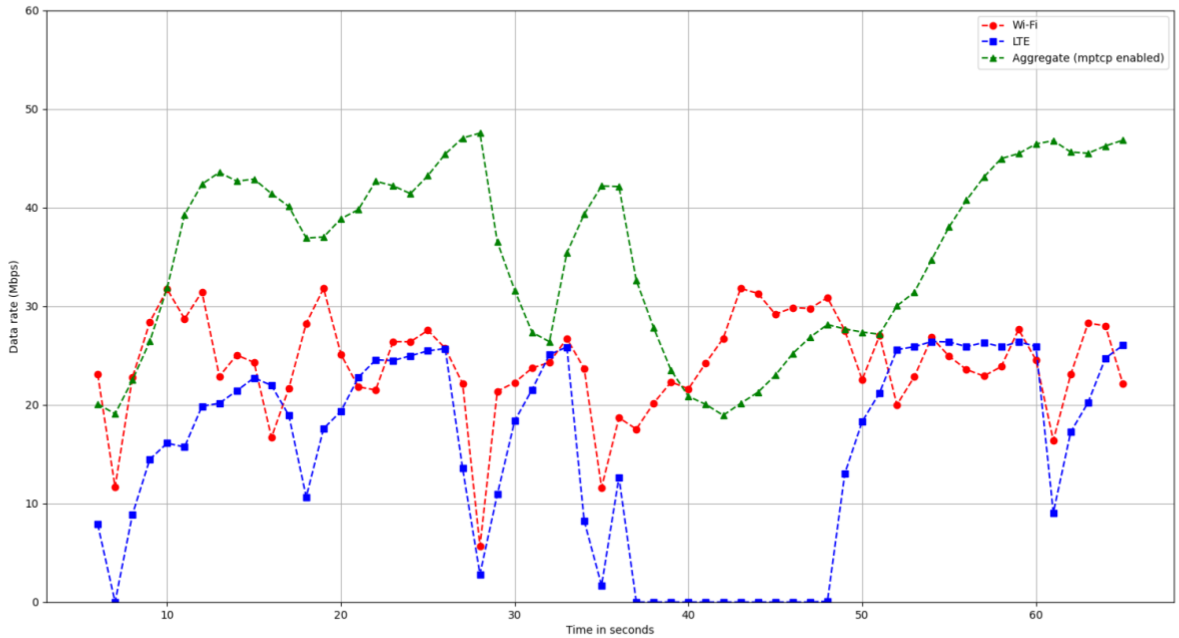

Figure 18 Initial MPTCP test results.

6.3.3 Test case 3: Video Streaming and Transcoding

PHASE 1 - Initial developments	
Test case # 3:	Video Streaming and transcoding
Slogan & Objective	• Allowing the user to select a video content from a local database and stream it on the user adapting the video to his terminal device capabilities. • To test performance/energy efficiency increasing using HW accelerators (GPU).
Test Scenario (Pre-conditions)	Please see Figure 19.
Expected Results (Post-Conditions)	• The selected video is streamed to the user equipment. • Performance increase in transcoding: 5x using a GPU instead the CPU.
General Time Plan (Validation Campaigns)	• Q2/2022: First trials at ITL local testbed. • Q4/2022: integration in the SPI testbed.

Test Sequence	1. User registration and authentication 2. Check that the user can access the application for selecting a video among those stored locally at the MEC server. 3. Select video using the EU device. 4. Check the quality of the video streamed at the UE.

Initial developments of the test case:

The diagram and the physical testbed at ITL premises used for developing and testing the video streaming and transcoding service is depicted in Figure 19.

The application is deployed on the K8s clusters and the user equipment is connected to the system.

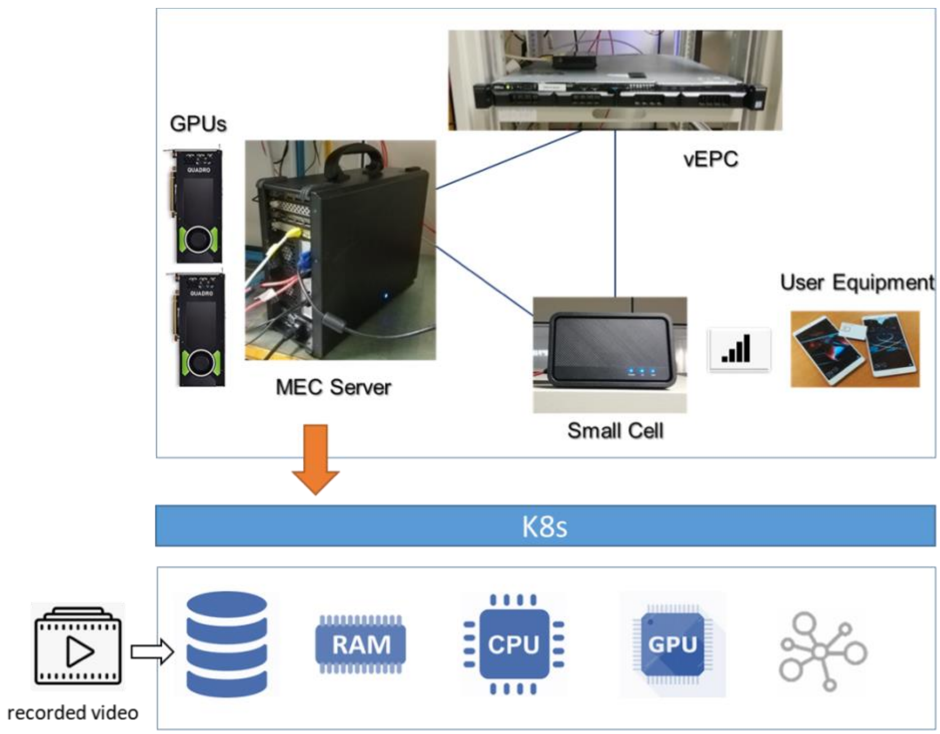

Figure 19 Video streaming and transcoding test scenario.

6.3.4 Test case 4: 5G Connectivity and Local Traffic Breakout

We describe here three connectivity tests that will be run as soon as the 5G deployment will be completed to validate the correct operations of the 5G network in this use case.

PHASE 1 - Initial developments

Test case # 4.1:	Connection between gNB and 5GC
Slogan & Objective	Interface setup between gNB and 5GC.
Test Scenario (Pre-conditions)	• 5GC instance (remote control plane and edge user plane) running on servers or VMs. • 5GC configured with active license and running, gNB should be reachable through the network.
Expected Results (Post-Conditions)	• No connection errors. Log messages show gNB successfully attached to the AMF.
General Time Plan (Validation Campaigns)	• Q4 of 2022.
Test Sequence	1. Configure the network interfaces and the CP, including all the related NFs. The system should show settings confirmation. 2. Set the IP address of the gNB in the whitelist of the 5GC's web interface. 3. Configure the N2 interface for interconnection between AMF and gNB. 4. Connect the gNB to the 5GC (AMF).

PHASE 1 - Initial developments	
Test case # 4.2:	UE's attach to and detach from the 5G network
Slogan & Objective	Check if UEs successfully attach to and detach from the correct PLMN and S-NSSAI.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC.

Expected Results (Post-Conditions)	Log messages show UE successfully registered, attached and detached to the 5GC.
General Time Plan (Validation Campaigns)	• Q4 of 2022.
Test Sequence	1. Configure the UE (virtual or physical) with the correct settings of PLMN, S-NSSAI and DNN. The system should show settings confirmation. 2. Register through the GUI the UE into the 5GC with SUPI identity. 3. Review the 5GC log messages related to the UE attachment. Verify that no error occurred. 4. Detach the UE from the 5GC.

PHASE 1 - Initial developments	
Test case # 4.3:	Connectivity between UE and data network (DN)
Slogan & Objective	Check uplink/downlink traffic between UE and DN through the 5GC (UPF), demonstrating the end-to-end connectivity between the connected devices and the edge servers.
Test Scenario (Pre-conditions)	• 5GC (remote control plane and edge user plane) running on servers or VMs and connected to a gNB. • 5GC configured, gNB reachable and interconnected to the 5GC AMF. • UE connected to the same gNB. UE must be pre-provisioned into the 5GC and attached to the 5GC
Expected Results (Post-Conditions)	• Connectivity between UE and DN is operational. • iPerf shows uplink/downlink traffic. • ICMP messages are acknowledged.
General Time Plan (Validation Campaigns)	• Q4 of 2022.

Test Sequence	1. Establish a new PDU session. Log messages should show the successful creation of UPF session. 2. Configure iPerf agents on the UE and in a reachable server of the DN. Verify that there are no registering errors. 3. Execute iPerf session or ping session. The test plan should start running. An iPerf or ping experiment will be started. 4. Review the 5GC log messages or check iPerf or ping results. There should be no errors, warning messages or dropped packets.

6.4 Next validation steps

To sum up the UC4 of the AI@EDGE platform, following the time plane of UC4 from D5.1, the works are in line with the initial plans for M18. There were some challenges regarding the delivery of some hardware which are required for the 5G RAN implementation. This delay, however, is not affecting dramatically the progress of the UC4. After M18, the works will mostly go toward the integration direction; in which different applications, components, and test cases will be integrated into the main testbed, while the integration with the rest of AI@EDGE platform will be also investigated and will be deployed.

7. Conclusions

In this deliverable D5.2, the preliminary validation strategies for the use cases at month 18 of the AI@ EDGE project were presented and preliminary benchmarks were given. For each of the four use cases the objectives and scenarios were defined, including the software and hardware requirements. While this deliverable focuses on the validation of the individual modules, future steps will involve the integration of the MEC features and the overall AI@EDGE platform, and on refining the AI models for the respective AIFs. In addition, the test cases for the respective modules within the use cases demonstrators and a roadmap for the tests were defined. This document represents the state of the development in M18 of the AI@EDGE project.

For UC1, the features to be integrated and validated are the offloading of network traffic at the edge, the latency advantage of the AI@EDGE platform and the telemetric box as an interface between the digital twin and the AI@EDGE platform. Similarly, UC2 will use the edge platform and leverage the MEC system for low latency retraining of the anomaly detection AIF and accelerate the monitoring with the deployment of Smart-NICs at the MEC Servers in addition to the traffic offloading functionality. Further steps for UC2 also involve the improvement of the AIFs, by including online detection, and adapt the AIFs to the use case environment. UC3 will improve the AIFs as well and integrate AI@EDGE function to provide assisted monitoring operation. In the UC4, the further steps are the integration of all modules into the main testbed and the AI@EDGE platform.

All use cases are showing good progression and are well planned. The updated plan for the integration of the AI@EDGE validation environment and the connect compute platform with the use cases was also given. Further steps for each use case have been identified.

References

(n.d.). From ZeroTier – Global Area Networking: https://www.zerotier.com/

AI@EDGE Deliverable D2.1. (2022). From https://aiatedge.eu/wp-content/uploads/2021/09/AI@EDGE_D2.1_Use-cases-requirements-and-preliminary-system-architecture_v1.0-1.pdf

AI@EDGE Deliverable D5.1. (2022). From https://aiatedge.eu/wp-content/uploads/2022/03/AI@EDGE_D5.1_Testing-and-Validation-Methodology-Planning-and-Preparation_v2.0.pdf