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

D5.3 Final 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

D5.3 Final validation and use-case benchmarking	
WP	WP5 – Use cases integration, validation, and benchmarking
Responsible partner	Italtel S.p.A. (ITL)
Version	1.0
Editor	Antonino Albanese (ITL)
Author(s)	Antonino Albanese (ITL), Nicola di Pietro (ATH), Babak Mafakheri (SPI), Giorgio Previati (POLIMI), Elena Campi (POLIMI), Alessandro Roccasalva (CRF), Miguel Rosa (AERO), Daniel Reti (DFKI), Raman Kazhamiakin (FBK), Phung Chi Dung (CNAM), Javier Palomares (i2CAT).
Reviewer(s)	Neiva Linder (EAB) Roberto Riggio (UNIVPM)
Deliverable Type	R
Dissemination Level	PU
Due date of delivery	31/12/2023
Submission date	20/12/2023

Version History				
Version	Date	Author	Partner	Description
0.1	06/09/2023	Antonino Albanese	ITL	ToC and assignment
0.2	06/10/2023	Nicola di Pietro (ATH), Babak Mafakheri (SPI), Giorgio Previati (POLIMI), Elena Campi (POLIMI), Alessandro Roccasalva (CRF), Miguel Rosa (AERO), Daniel Reti (DFKI)	ATH SPI POLIMI CRF AERO DFKI	First round of contributions – update of Sections 2, 3, 4, 5, 6
0.3	11/10/2023	Antonino Albanese	ITL	Clean version for the light review
0.4	19/10/2023	Javier Palomares	i2CAT	Update the content in section 2.1 (Reference architecture)
0.5	28/11/2023	Babak Mafakheri	SPI	Update Section 6 – UC4
0.6	29/11/2023	Lorenzo Uccello Miguel Rosa	POLIMI AERO	Update Section 3 – UC1 Update Section 5 - UC3
0.7	04/12/2023	Daniel Reti	DFKI	Update Section 4 – UC2
0.8	06/12/2023	Antonino Albanese	ITL	Clean version for the review
0.9	13/12/2023	Neiva Linder Roberto Riggio	EAB UNIVPM	Review
1.0	14/12/2023	Antonino Albanese	ITL	Final Version for the submission
1.1	15/12/2023	Irene Facchin	FBK	Final check and submission

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

Table of Contents

List of Figure

List of Table

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
DDoS	Distributed Denial-of-Service
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)
GPRS	General Packet Radio Service
GTP	GPRS Tunnelling protocol
ICMP	Internet Control Message Protocol
IFEC	Inflight Entertainment and Connectivity
IIoT	Industrial Internet of Things
IMDB	Internet Movie Database
KPI	Key Performance Indicator
LEO	Low Earth Orbit
MEC	Multi-access Edge Computing
MEO	MEC Orchestrator
MIMO	Multiple-Input and Multiple-Output
MPTCP	Multi-Path Transmission Control Protocol
NF	Network Function
NFUC	Use case Network Function
ODBC	Open Database Connectivity
PLMN	Public Land Mobile Network
RAN	Radio Access Network
RDU	Removable Display Units
RL	Reinforcement Learning
S-NSSAI	Single Network Slice Selection Assistance Information
SCU	System Control Unit

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
SQL	Structured Query Language
SVD	Singular Value Decomposition
SUMO	Simulation of Urban MObility
SUPI	Subscription Permanent Identifier
SYRROCA	System Radiography and Root Cause Analysis
TBM	Telematic Box
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
xApp	eXtended Application

Executive Summary

The deliverable D5.3 “Final validation and use case benchmarking “, describes the validation objectives, validation scenarios, validation strategies, and planning of the projects use cases and reports results and KPIs of the four use cases applications deployed on the respective testbeds into which the AI@EDGE platform has been integrated. To meet the use cases requirements, the related applications were designed and developed to leverage the project's technological enablers.

Due to the heterogeneity of the four use cases, each testbed has been equipped with dedicated HW and a distinct end-to-end 5G solution. For what concern the SW, the same platform, including the RAN part, has been used, although with different configurations.

The availability of the document allows to declare the achievement of the milestone MS5.6 (Final validation completed) expected for M36.

In summary, this deliverable reports contributions and results from tasks 5.1, 5.2, 5.3, 5.4, and 5.5 that dealt with the integration, validation and benchmarking of the project's use cases.

In particular:

- Section 2 reports the reference architecture for the use cases implementation and validation including the Network and Service Automation (NSAP) layer and the Connect-Compute Platform (CCP) layer, also describing the final layout and configuration of the integration testbed used as a reference for the deployment of the entire platform. In addition, it reports the 5G network infrastructure architecture implemented in the four testbeds for providing the communication and Edge computing framework.
- Sections 3, 4, 5, and 6 are dedicated to the description of the validation objectives, validation scenario, validation procedures, and validation results of the four use cases. 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 independent of the specific use case in order to provide a homogeneous vision to the reader. Finally, the validation results are presented on a test case basis. Each Section ends describing benefits and impact of using the AI@EDGE platform for enhancing each use case while also highlighting possible improvements after the end of the project.
- Section 7 reports the conclusions.

1. Introduction

The deliverable D5.3 reports the results of the final trials carried out in the testbeds where the four project use cases were designed, implemented, tested, and validated. The initial testbeds have been updated by integrating several HW and SW components related to the 5G radio, Edge, and Core functionalities, as well as integrating the Connect-Compute Platform. The deliverable includes contributions and results of tasks 5.1, 5.2, 5.3, 5.4, and 5.5.

Figure 1 shows the interaction between the technical work packages (WP2, WP3, and WP4) and the integration work package (WP5) which defined the evaluation methodology and KPIs measurement for benchmarking the AI based applications associated with the four use cases when running on the AI@EDGE integrated platform. Furthermore, the connect-compute platform was deployed on all the testbeds also integrating a complete 5G network which includes, in each testbed, several HW and SW components related to the radio, Edge, and core functionalities. Finally, the use cases applications were designed and developed to leverage the project's technological enablers to achieve the expected KPIs. Achieving this result required a very complex system integration process driven by the specific requirements, constraints, and needs of each testbed, mainly due to the heterogeneity of the four use cases.

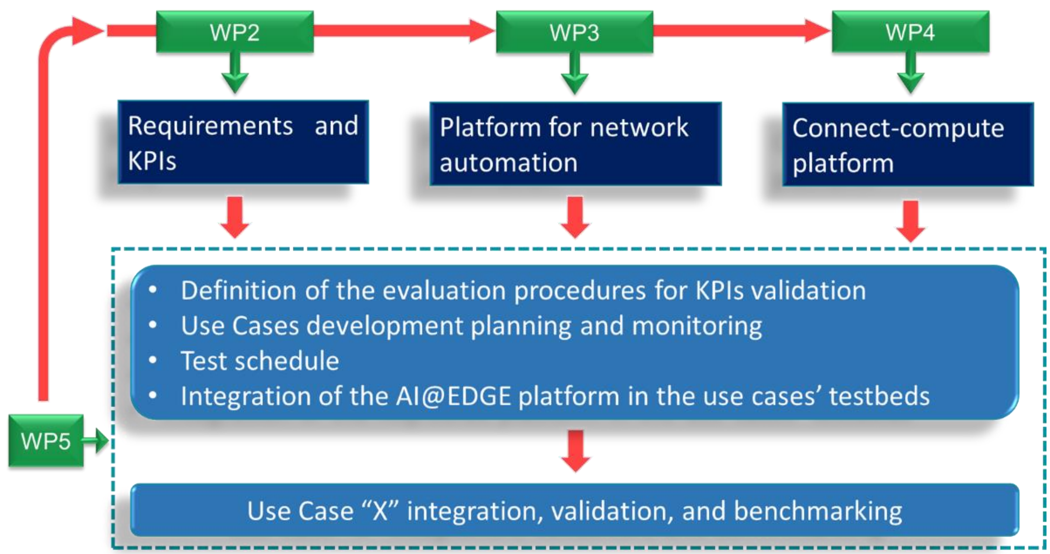


Figure 1 Work Packages interaction

The first use case (UC1) is about the virtual validation of vehicle cooperative perception, where the AI@EDGE platform is used to support 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.

Figure 2 summarizes the four use cases.

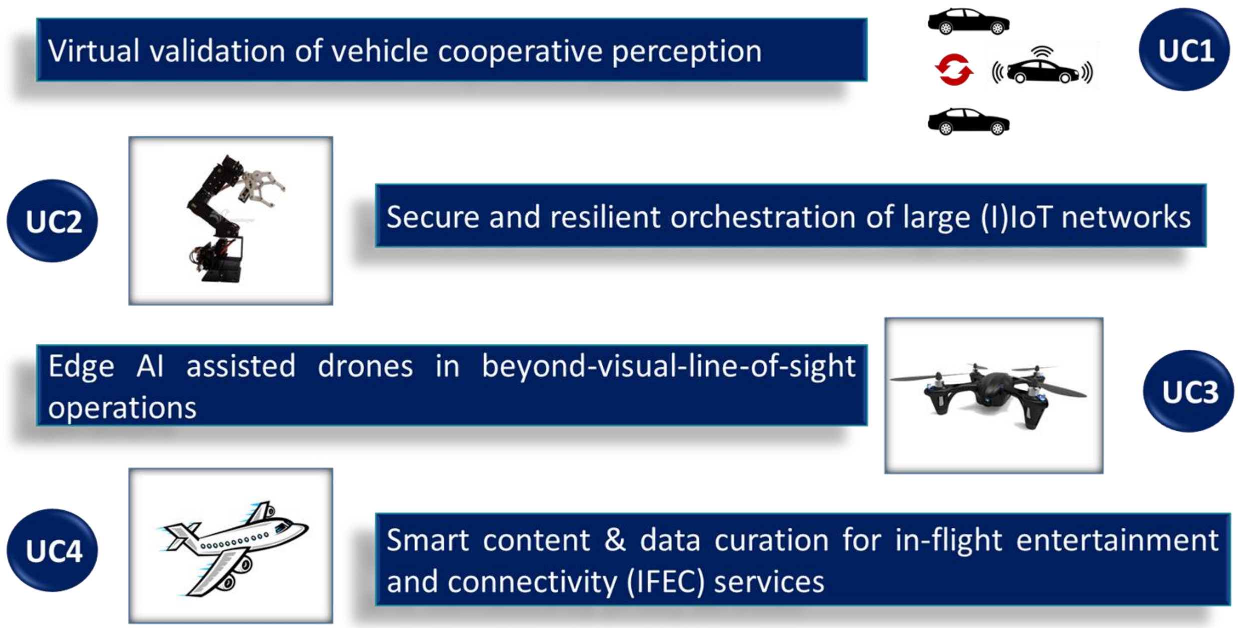


Figure 2 AI@EDGE Use Cases

2. AI@EDGE Validation environment

This section reports the final reference architecture for the use cases implementation and validation is reported, as have been defined in the deliverable D2.3 [3]. In addition, it describes the final layout and configuration of the integration testbed used for the deployment and testing of the entire AI@EDGE platform.

2.1.Reference architecture

The functional architecture of the AI@EDGE system is shown in Figure 3, serving as a reference for mapping the relevant hardware and software components relevant for each use case. The architecture consists of two primary operational sections: (i) the **Network and Service Automation (NSAP) layer**, which is where network intelligence is concentrated, responsible for the management of the system network automation; and (ii) the **Connect-Compute Platform (CCP) layer,** which takes charge of orchestrating and managing Artificial Intelligence Functions (AIFs) across different edge levels/locations. Moreover, it ensures the seamless connectivity between various system elements and oversees the management of its computational and hardware resources.

The system is structured to accommodate AIFs, which, depending on their specific functions, can be deployed at any level within the system. Certain AIFs are employed for tasks related to network automation and optimization, while others are linked to applications and services. Every component within the system, encompassing both the modules constituting the system's architecture and the AIFs, can utilize network and computing resources distributed across the three different tiers, namely Far Edge, Near Edge and Cloud.

With respect to the state of the architecture described in D5.2 [2], the following additional work was conducted towards improving and/or implementing the proposed solutions of AI@EDGE:

- The definition of the communication service bus for interconnecting the various NSAP modules.
- The enhancements made to the Non-Real-Time RAN Intelligent Controller (non-RT RIC), which is the key element to implementing non-RT intelligent closed loop automation related to the 5G System at both the NSAP and 5G System Platform management levels.
- The definition of the Slice Manager, which provides control over the lifecycle of network slices, allowing for Create, Read, Update, and Delete (CRUD) operations on slice instances.
- The implementation of the Data Pipeline, defined in project Deliverable 3.2, [4], which is responsible for delivering up-to-date and relevant data to AIFs. Given that the IOC executes multiple AI/ML models, it also relies on the Data Pipeline for data. Additionally, the Data Pipeline encompasses the management of the AI/ML models' lifecycle, including instantiation, updating, and replacement. Its components include the Data Collector, Data Processor, Data Repository, Model Manager, and Model Repository. The Data Collector acquires data from various sources necessary for model training. The Data Processor is responsible for cleaning, filtering, and preparing data for use by the AIFs. For data that may take time to obtain, the processor can store it in the Data Repository, allowing for its reusability by multiple models. The architecture also enables the lifecycle management of AI models through the Model Manager. This component evaluates the performance of one or more models, either periodically or in response to events specified in the AIF descriptor and monitors them accordingly. If the performance falls short of expectations, the Model Manager initiates an update, which is stored in the Model Repository along with associated metadata. This setup facilitates the reuse of models by different AIFs.

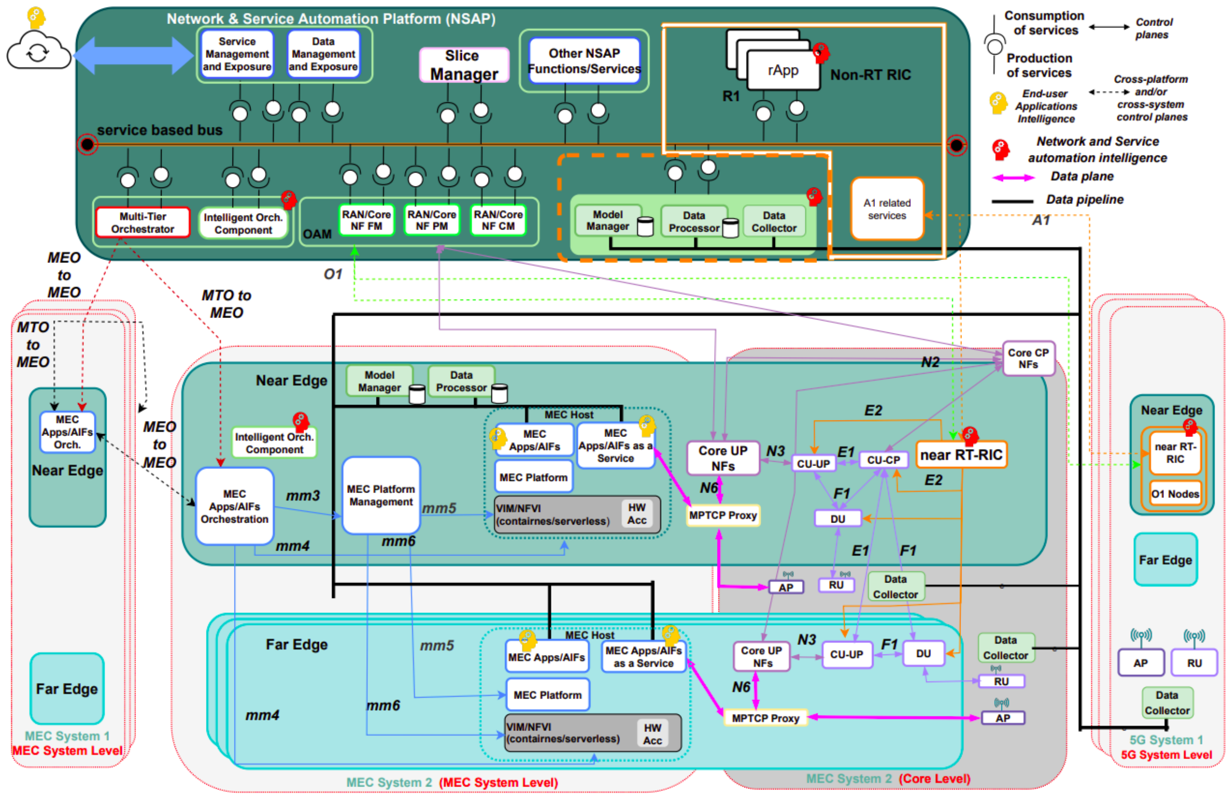


Figure 3 AI@EDGE functional architecture

2.2. Integration testbed

The integration testbed hosts the main components of the system. Specifically, it contains the instance of the Connect-Compute Platform deployed in a MEC system, an instantiation of the NSAP layer with the corresponding software modules, and the relevant networking modules for the 5G scenarios. The testbed is based on a Kubernetes cluster that spans both Near and Far Edge nodes, while the NSAP part is deployed in the Cloud VMs with the dedicated Kubernetes cluster to host its modules.

The description of the testbed architecture is presented in Figure 4. The integration between Connect-Compute Platform in MEC and the MTO component in NSAP is performed via publish/subscribe mechanism exposed with RabbitMQ[1]. This allows for integrating multiple MEC systems, such as acceleration testbed (hosted in ICCS). The metrics of all the MEC systems and the underlying components

[1] Details at: https://www.rabbitmq.com/

are collected through local Prometheus2 instances at MEC, and then further pushed to NSAP, to the Thanos3 metrics aggregator for further usage of the MTO and IOC for the orchestration purposes.

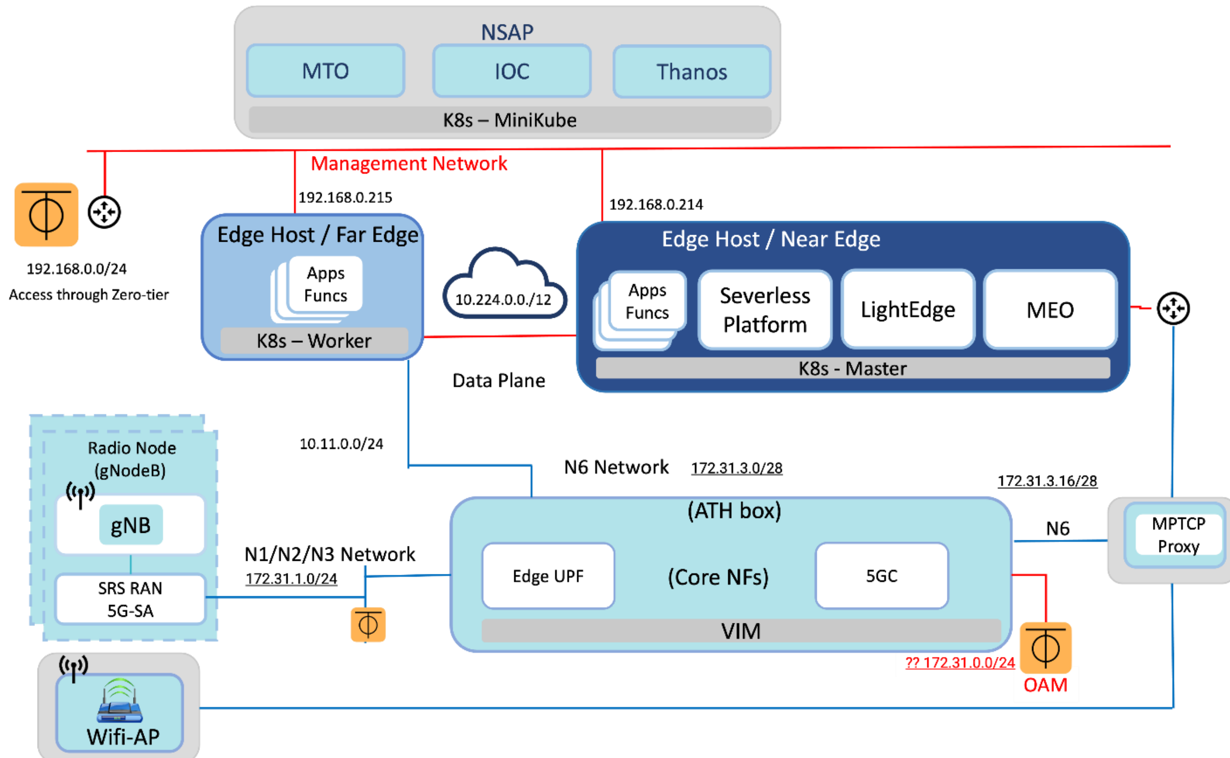


Figure 4 Integration testbed architecture

Remote connection towards the MEC part of the testbed is allowed through a ZeroTier4-enabled virtual network. 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. As for the NSAP part of the testbed, the connection is performed via SSH connectivity as the corresponding infrastructure is hosted in private Cloud (MS Azure5) and is exposed to Internet.

[2] Please see: https://prometheus.io/
[3] Please see: https://thanos.io/
[4] Please see: https://www.zerotier.com/
[5] Please refer to: https://azure.microsoft.com/en-us

As already described in Deliverable D5.2 [2] , the Near Edge host of the integrated testbed is connected to the SGi-2 port of the Athonet Core and hosts the master node of the K8s cluster, together with the LightEdge6 platform and with the Serverless platform (based on Nuclio7). The Far Edge host is connected to the SGi-1 port of the Athonet Core (corresponding to the Edge UPF). The updates to the 5G infrastructure with respect to the preliminary model of the integrated testbed described in D5.2 are presented in the following section.

2.3. 5G infrastructure

We described in Section 3.2 of D5.1, [1], and Section 2.2 of D5.2, [2], the 5G network architecture that supports AI@EDGE's UC1, UC2, and UC4. Such use cases were characterized by a common shared control plane of the 5G core network, deployed within WP4's integration testbed (at FBK's premises) and distributed dedicated user planes. Such architecture was slightly updated after the completion of D5.2, and its final version is depicted in Figure 5.

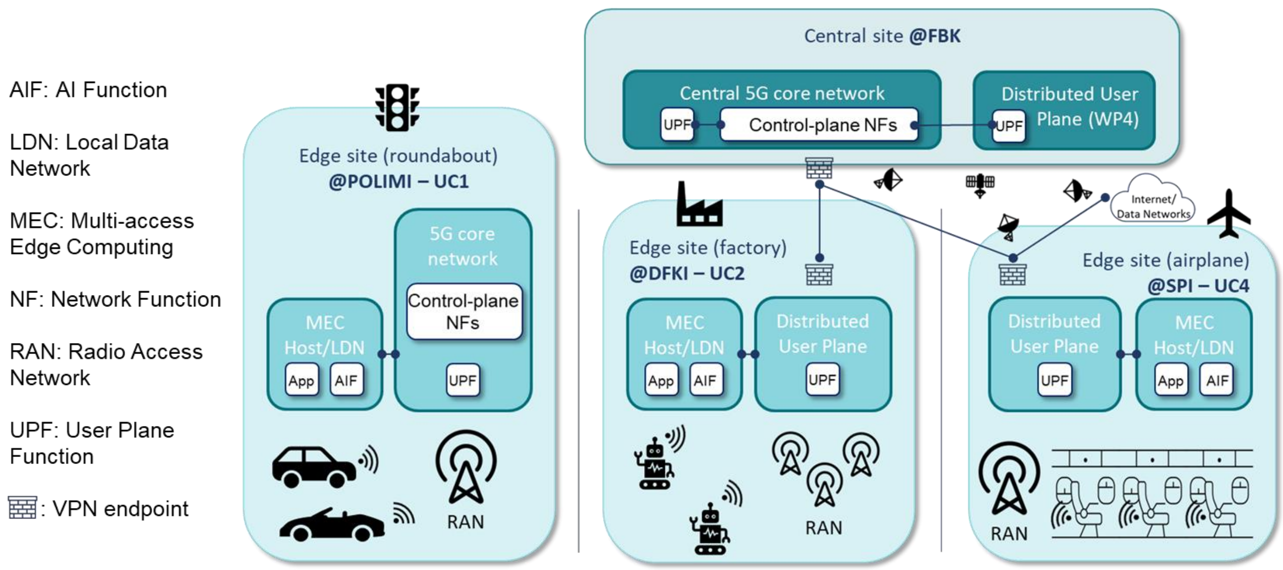


Figure 5 High-level architecture of the 5G network

Namely, the consortium decided to separate UC1 from the other use cases for security reasons, related to the network isolation requirements imposed by the utilization of POLIMI's driving simulator. In practice, the architectural features of the 5G network of UC2 and UC4 remain the same described in previous deliverables, whereas for UC1 we proceeded to deploy at POLIMI a standalone 5G network with dedicated core network, fully devoted to the use case's operation. Such a deployment choice comes without

[6] Please see: https://lightedge.io/
[7] Please see: https://nuclio.io/

drawbacks: the network requirements identified in the initial phases of the project for the effective operations of each use case are still met, and the goal of validating an advanced multi-site network architecture is still achieved with the network deployment that interconnects the integration testbed, UC2, and UC4.

The testbed devised for validating UC3 is relying on the 5TONIC[8] environment to provide a 5G network where AI@EDGE platform functionalities have been integrated. The preliminary architecture described in Section 2.2 of D5.2, [2], has been updated and its final version is depicted in Figure 6.

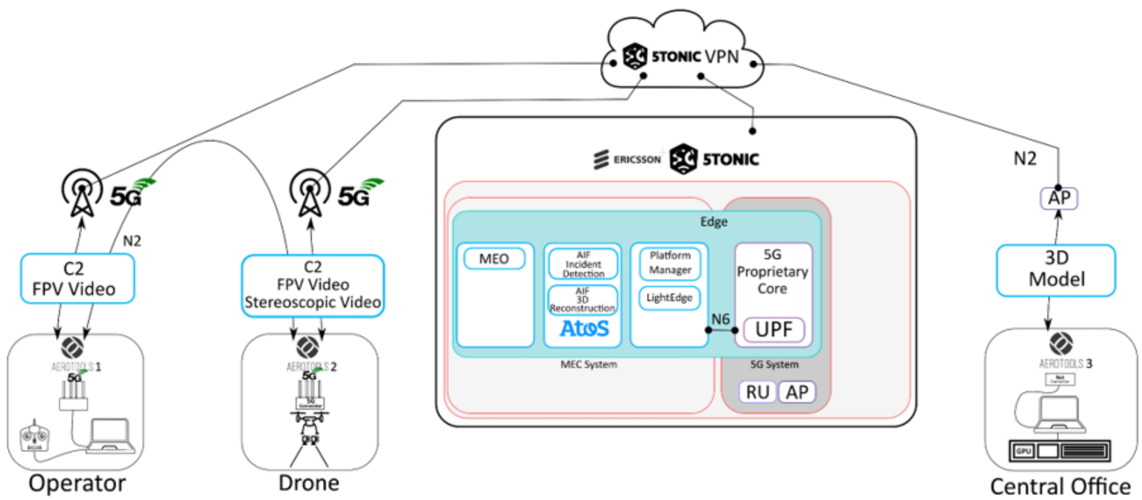


Figure 6 Final UC3 testbed architecture

3. Use Case 1: Virtual validation of vehicle cooperative perception

The use case is developed with the aim of allowing Cooperative Connected and Automated Vehicles (CAVs) to safely navigate into a roundabout. Vehicles exchange data related to their trajectories (position and velocities). Data are gathered at the network edge and used by the Artificial Intelligence Function (AIF) to increase traffic fluidity while avoiding potential collisions. Additionally, the AIF provides a reasonable ride comfort to passengers (avoid sudden breakings or accelerations of vehicles). The end-to-end system developed to demonstrate this use case is complex and expensive in the real world. So, a virtual reality environment has been chosen and adopted. In such an environment a real human driver is included. The real human driver operating in a driving simulator[9] allows to obtain feedback on his/her perception of driving safety, fluidity, and comfort. Both a subjective and objective feedback on perception can be

[8] Please refer to: https://www.5tonic.org/

[9] Please see: www.drismi.polimi.it

obtained. This feedback is of paramount importance to obtain a trustable evaluation of the results of the use case.

The virtual reality environment is composed by the dynamic driving simulator of Politecnico di Milano and the traffic simulator SUMO[10]. This combination was obtained -according to our knowledge- for the very first time in the world. The test of a mix of one really driven vehicle and many emulated vehicles has been made. The communication between the emulated vehicles and the human-driven vehicle is based on AI@EDGE platform. Each single autonomous vehicle is driven by a Reinforcement Learning (RL) policy. More specifically, the objective is using the RL techniques for coordinating the actions of a set of cooperative and controlled agents that coexist in a realistic environment. These agents will interact with each other and with human driven vehicles. One agent is driven by a real human driver in the driving simulator, the others are emulated in the traffic environment.

3.1. Validation objectives

The motivation of this use case is threefold. While running into a roundabout, we aim to increase safety, reduce the traffic congestion *and pollution*, and provide riding comfort. The functionality 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 Artificial Intelligence (AI) coordinating the manoeuvres of the automated vehicles.

The local execution of AI algorithms on vehicles with a direct Vehicle-To-Vehicle (V2V) communication has some limitations regarding the possibility to identify and solve more complex traffic situations such as a roundabout. V2V is typically based on Dedicated Short-Range Communications (DRSC), which has strong limitations on bandwidth and communication speed, restricting the capacity of vehicles to comprehensively understand the environment around them. These boundaries are even more crucial considering complex scenarios, such as the roundabout, and assessing the ever-growing number of sensors on automated vehicles and the resulting increased data sent. For this reason, a promising solution in this filed is the combination of V2V and Vehicle-to-Infrastructure (V2I) communication architectures, such as the one proposed through the presence of a Connect and Compute Platform.

The introduction of edge computing nodes guarantees the offloading of the coordination functions between autonomous vehicles, but exploits the features of 5G networks, ensuring the quality of service required in this context.

The change of communication perspective from “short-range” (802.11p, PC5) to “long-range” (5G) using MEC platforms adds some latency in V2V 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 roundabout scenario.

[10] Please see: https://www.eclipse.org/sumo/

We exploit V2NN2V (vehicle-to-network-network to-vehicle paradigm). This is simpler and cheaper with respect to V2V paradigm. It also allows for complex scenarios and technologies, such as CCAM (Cooperative Connected and Automated Mobility). V2NN2V is driven by the AI@EDGE AIF. This could not be obtained with the conventional V2V paradigm, which is only based on vehicular communications and does not consider any infrastructure. Furthermore, considering only V2V communications and their low-data availability, the cooperation is milder and not able to produce relevant results like the ones enabled by a policy in charge of routing the whole traffic, acting as a regulator.

The main challenges related to the scenario are:

- **Virtual reality setup, merging of:**
    - the real human driver
    - the automated vehicles (cooperative agents driving autonomous vehicles)
    - the vehicles driven by human driver models.
- **Telematic box and driving simulator integration:**
    - Integration of a telematic 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 telematic box to edge servers to enable AIF processing such data guaranteeing low latencies whenever needed.
- **Vehicle coordination:**
    - Capability to support the vehicles coordination considering the real driver reactions on the driving simulator.
- **Objective and subjective feedback from the real human driver:**
    - Check that the AIF provides an acceptable and sustainable management of the traffic into a roundabout.

Main KPIs related to those challenges are:

- **Latency:** under 160ms, to allow Connected and Automated Vehicles to safely navigate the roundabout.
- **Vehicle density:** 12000 vehicle/km as expected number of simulated vehicles per a given area.
- **Positioning:** 1.5m to deal with vehicle dynamics and movement.

Table 1 shows how each single KPI is relevant for a specific use case challenge.

Table 1 Use case 1 challenges and KPIs

	Latency	Vehicle density	Positioning
Virtual reality setup	X	X	X
Telematic Box and Driving Simulator integration	X	-	-
Local breakout of traffic and offloading	X	X	-
Vehicle Coordination	X	X	X
Objective and subjective feedback	X	X	X

3.2. Validation scenario

The aim of the validation is to check that the AIF of AI@EDGE works for the selected scenario, i.e., the traffic into a roundabout. The very final validation refers to the positive or negative feedback of the real human driver that drives into the roundabout. The objective KPIs that are related to the mentioned feedback by the human driver are Latency, Vehicle Density, and Positioning.

The UC1 testbed facilities are two. The first facility relies on the infrastructure available at POLIMI, where the driving simulator will be 5G connected and will send vehicle kinematics data (position, velocity acceleration) 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 enabling automotive telematic boxes and will provide a traffic simulation platform on which the artificial intelligent agents cooperative perception distributed algorithm will be validated.

POLIMI testbed: the first testbed is based on a driving simulator connected to the AI@EDGE platform through a 5G telematic box. The driving simulator sends its dynamic data to an edge node on which a cooperative perception algorithm is executed. The testbed configuration is depicted in Figure 7. It includes the following:

- VI-Grade driving simulator with its VI-WorldSim[11] scenario simulator.

[11] Please see: https://www.vi-grade.com/en/products/vi-worldsim/

- Telematic box (DUT: Device Under Test) with the Uu and PC5 connection, V2X ITS stack and the data client to send data (AMQP[12] client).
- 5G RAN equipment Ettus X310[13] with SRS's virtualized radio solution (srsRAN Project[14]).
- Athonet's full 5G network (core/edge).
- Local edge node hosting the Connect-Compute Platform and the NSAP, for the execution of the AIF, the data collection server (AMQP broker) and the traffic simulator.

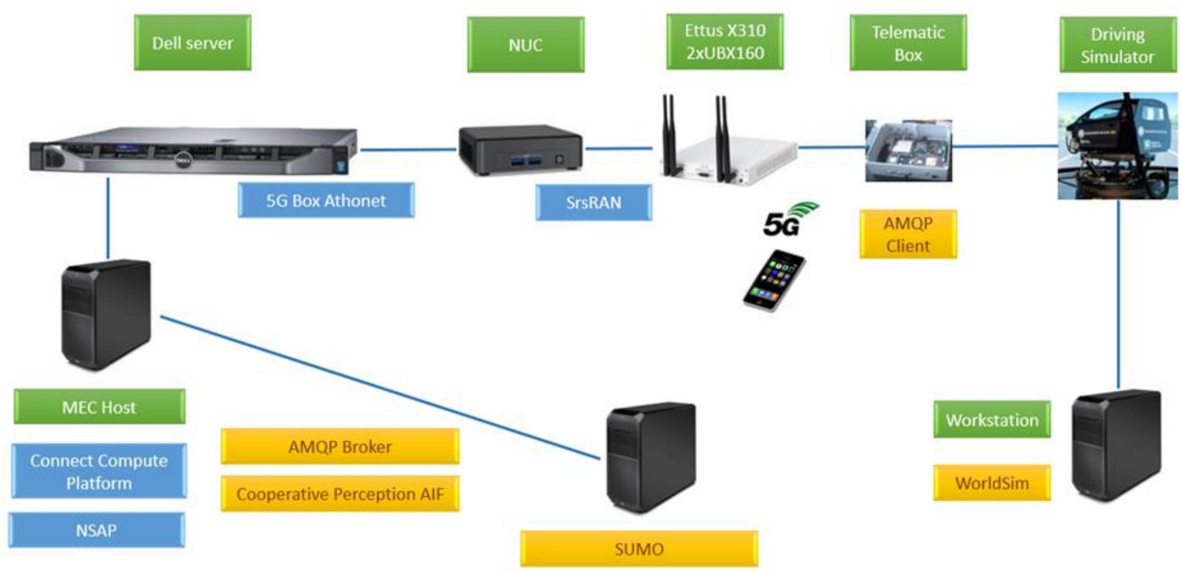


Figure 7 POLIMI testbed configuration

CRF testbed: the second testbed is available in CRF in Torino where a 5G emulator is used to test 5G enabled automotive telematic boxes. The telematic box sends data to the cooperative perception algorithm, deployed on the edge node, through the 5G emulator.

The testbed configuration is depicted in Figure 8. It includes the following:

- MT8000A Network Emulator (Non-Stand Alone and Stand-Alone).
- 5G-NR SA and NSA: DL 2CA, DL 4x4 MIMO, FDD and TDD.

[12] Please see: https://www.amqp.org/
[13] Please see: https://www.ettus.com/all-products/x310-kit/
[14] Available at: https://www.srslte.com/5g

- 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).
- Global navigation satellite system (GNSS) emulator.
- Vector CANalyzer to support the internal Vehicle bus (CAN Bus).

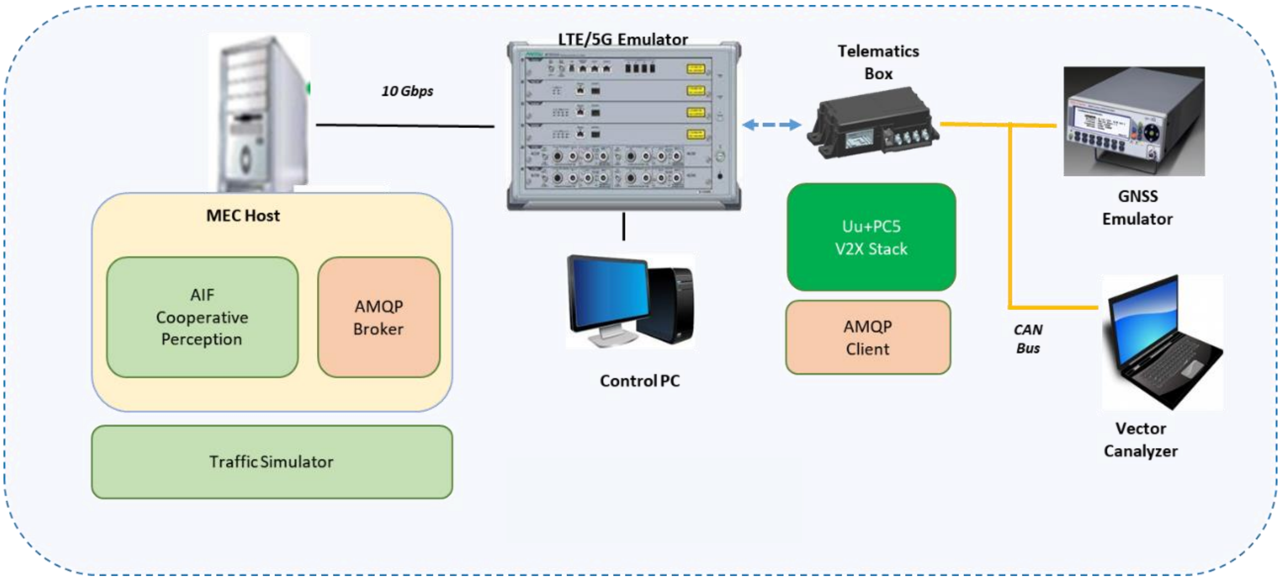


Figure 8 CRF testbed configuration

3.3. Validation procedures

Tests are carried out considering a panel of drivers navigating through the roundabout. The traffic is constituted by both automated vehicles and human-driven vehicles. For participants to have a better understanding of the situation, they repeat the manoeuvres entering the roundabout from each of its legs. For each leg, drivers are asked to compare two different percentages of automated vehicles present in the scenario, namely 20% and 80%. The subject is not aware of the predetermined order in which the two traffic configurations are presented for all the legs of the roundabouts. For each tester, the order is maintained throughout the testing process, while can be different for different testers. After the test, drivers are asked to fill a questionnaire about their perceptions. They are asked to focus on their safety feeling and on their impressions on traffic smoothness.

3.3.1 Test case 1 – Driving Simulator and AIF integration

Test case #1.1:	**AIF implementation in the final simulation environment**
Slogan & Objective	Validation of the new policy operation in the final simulation environment.
Test Scenario (Pre-conditions)	• New WorldSim scenario implemented reproducing the real-world environment. • SUMO configuration and network files correctly compiled. • Established stable connection between AIF, Driving Simulator and SUMO co-simulation, employing UDP and TCP communication protocols.
Expected Results (Post-Conditions)	AIF properly implemented in the testbed and capable of driving automated vehicles.
General Time Plan (Validation Campaigns)	Q3 of 2023
Test Sequence	• Identify the geometric and visual details of the final real simulation environment. • Create the simulation environments in WorldSim and SUMO. • SUMO scenario validated against measured traffic data. • Verify the connection between the terminals representing AIF and SUMO, via TCP protocol. • Verify connection between the Driving Simulator and SUMO, via UDP protocol. • Introduce and test the new policy

Test case #1.2:	**Policy behaviour modifications**
Slogan & Objective	Study of policy behaviour with other vehicles and human in the loop; comparison of the policy with the previous one.
Test Scenario (Pre-conditions)	• Established stable connection between AIF, Driving Simulator and SUMO co-simulation. • New policy correctly implemented in the testbed.
Expected Results (Post-Conditions)	Defining differences in policy behaviour in relation to newly identified objectives, such as reducing consumption and optimizing passenger comfort. The objective is to achieve improved policy behaviour by training in a more intricate and challenging environment.

General Time Plan (Validation Campaigns)	Q3 of 2023
Test Sequence	• Verify the connection between AIF, SUMO and the Driving Simulator. • Test the new policy in the simulation environment with other automated and non-automated vehicles and a human participant. • Compare the performance of the new policy with the previous one

Test case #1.3:	**Human perception analysis**
Slogan & Objective	Analysis of human perception of automated and connected vehicles in the simulation
Test Scenario (Pre-conditions)	• Established stable connection between AIF, Driving Simulator, and SUMO co-simulation. • Captured the behaviour details of the policy used. • Defined a precise structure for test and questionnaire to collect results.
Expected Results (Post-Conditions)	• Detailed analysis of policy behaviour. • Identification of necessary changes, both to the simulation environment and to the policy's objective function. • Different results with respect to users' abilities and experience in the simulator.
General Time Plan (Validation Campaigns)	Q4 of 2023
Test Sequence	• Introduce human in the loop. • Define a sequence of tests and a questionnaire capable of getting accurate and complete indications from participants. • Considered several participants with different experience and driving skills.

Test case #1.4:	**Comfort and safety analysis**
Slogan & Objective	Detailed analysis of a passenger's perception of comfort and safety inside the AIF-guided simulator.
Test Scenario (Pre-conditions)	• Chosen an automated vehicle to be used to control the driving simulator. • Utilization of VI-CarRealTime with the SUMO simulator in order to define all the required data to control the driving simulator on the basis of the motion of the chosen automated vehicle.

	• Established a connection between SUMO and the driving simulator allowing the two co-simulations to be initialized in parallel.
Expected Results (Post-Conditions)	• Detailed analysis of policy behaviour, in terms of comfort and safety perceived by the passenger. • Verification of limit values in terms of accelerations and jerks used to train the policy.
General Time Plan (Validation Campaigns)	Q4 of 2023
Test Sequence	• Choose the automated vehicles to be replicated. • Implement the required procedure between SUMO and VI-CarRealTime to obtain the required data to control the driving simulator motion accordingly to the motion of the chosen automated vehicle. • Repeat the experiment with different human beings as passengers of the automated vehicle to evaluate the perceived comfort and safety.

Test case #1.5:	**Latency analysis**
Slogan & Objective	Introduction of the latency in both communication directions with the edge node where the AIF is running. Verification of the ability of the policy to adapt to different latency values.
Test Scenario (Pre-conditions)	• Testbed completely functioning. • Implemented a modifiable communication latency in both communication directions with the edge node where the AIF is running.
Expected Results (Post-Conditions)	Definition of the maximum communication latency that allows a proper operation of the policy.
General Time Plan (Validation Campaigns)	Q4 of 2023
Test Sequence	• Set up the testbed, including the modifiable communication latency. • Organize several test sessions by increasing the communication latency.

3.3.2 Test case 2 – V2X over 5G

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

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)	• Q2 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)	• Q3 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

The following three connectivity tests 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)	• Q3 of 2023.
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 of 2023.

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[15] shows uplink/downlink traffic. • ICMP messages are acknowledged
General Time Plan (Validation Campaigns)	• Q4 of 2023.
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.

[15] Please see: https://iperf.fr/

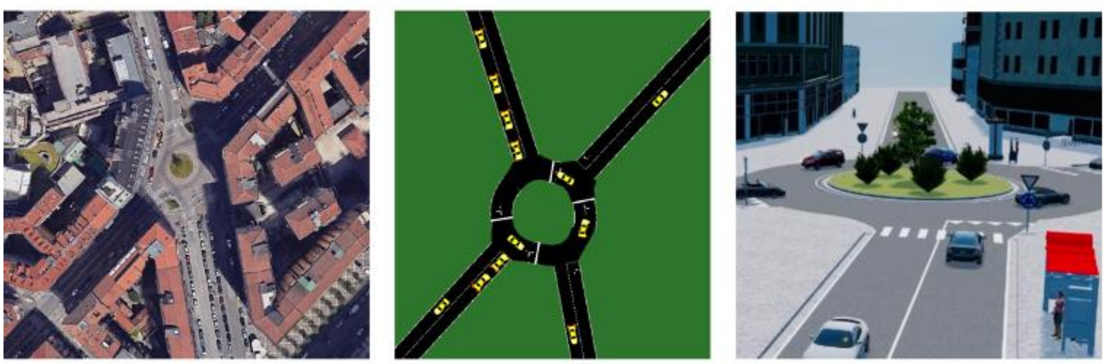
3.4. Validation results

3.4.1 Test case 1 results – Driving Simulator and AIF integration

Test case #1.1: AIF implementation in the final simulation environment

The final roundabout is inspired by a real-world one in Milan, Italy. Figure 9 presents how the roundabout looks in the real world, while panels B and C show the simulation environments of SUMO and VI-WorldSim. It is a four-leg mini-roundabout, with medium-high traffic and, therefore, posing a challenging environment for the AV policy.

A calibration procedure was conducted with the aim of replicating the number of vehicles approaching the intersection and their positions during the simulation. Firstly, measurements were taken for the maximum queue length, upstream and downstream flows for each leg, considering road vehicles, pedestrians, and bicycles on the actual roundabout. This process was repeated for six consecutive time slots, each lasting 10 minutes. Subsequently, the results of these measurements were compared with simulations conducted in SUMO to calibrate the most relevant parameters that define the traffic conditions in the considered scenario. Table 2 reports the calibrated parameters of the roundabout model.



A) Actual roundabout in Milan B) SUMO's simulation environment C) VI-WorldSim's simulation environment

Figure 9 Final simulation environment and SUMO's and VI-WorldSim's simulation environments

Table 2 SUMO's IDM calibrated parameters

Calibrated parameter	Value
JmCrossingGap *(minimum distance between the vehicle and the pedestrian that is heading toward the point of conflict of its trajectory with that of the vehicle)*	1,3545
JmTimegapMinor *(minimum time interval for a vehicle to enter an intersection where it does not have the right-of-way, before a vehicle with right-of-way)*	1,7792
Impatience *(driver's intent to obstruct a vehicle with the right of way)*	0,1182

Acceleration *(maximum acceleration for the selected vehicle type)*	1,7634
Deceleration *(maximum deceleration for the selected vehicle type)*	4,2939
Tau *(minimum time interval between consecutive vehicles)*	1,3472
ActionStepLength *(driver reaction time)*	0,5050

The two simulation environments, SUMO and VI-WorldSim, have been connected by a UDC connection via lan cable. The communication frequency is 200Hz and the delay between the two environment is less than 5ms. The communication protocols have been successfully tested, proving the complete implementation of the digital twin of the real scenario.

Test case #1.2 and #1.3: Policy behaviour modifications and Human perception analysis

PRELIMINARY TESTS

The final simulation environment has been used to test the final policy running the Cooperative, Connected and Automated vehicles (CCAVs) in the network. Firstly, preliminary tests have been performed on a restricted group of drivers. A panel of ten participants has been selected for the tests. The participants were chosen from individuals without previous experience with driving simulators. The panel consists of 5 females and 5 males, aged between 22 and 33 years, with driving experience ranging from 1 to 15 years. Before the test, each participant was given instructions on how to operate the driving simulator and signed an informed consent form. Additionally, each participant spent about ten minutes driving in a simple motorway scenario to become familiar with the driving simulator before the actual test. A team of psychometrics experts have guided the tests with humans in the loop. The final policy has been analysed, considering both qualitative and quantitative results.

Table 3, Table 4, and Table 5 show the qualitative results represented by the answers to the questionnaire filled out by human drivers.

Table 3 Answers to the first question of the survey

Regarding the traffic smoothness, which of the following statements do you agree with the most?	Number of answers
Traffic in the scenario with 20% of AVs was definitely safer than in the scenario with 80% AVs	1
Traffic in the scenario with 20% of AVs was partially safer than in the scenario with 80% AVs	3
Traffic in the scenario with 20% of AVs was partially less safe than in the scenario with 80% AVs	4

Traffic in the scenario with 20% of AVs was definitely less safe than in the scenario with 80% AVs	1
I did not perceive differences	1

Table 4 Answers to the second question of the survey

Regarding safety perception, which of the following statements do you agree with the most?	Number of answers
Traffic with 20% of CCAVs was definitely safer	0
Traffic with 20% of CCAVs was partially safer	2
Traffic with 20% of CCAVs was partially less safe	2
Traffic with 20% of CCAVs was definitely less safe	6
I did not perceive differences	0

Table 5 Answers to the third question of the survey

Globally, which of the two scenarios did you prefer?	Number of answers
I definitely preferred the scenario with 20% CCAVs	0
I partially preferred the scenario with 20% CCAVs	3
I partially preferred the scenario with 80% CCAVs	3
I definitely preferred the scenario with 80% CCAVs	4
I cannot say which scenario I preferred	0

Table 6 and Table 7 present the quantitative analysis of the final policy, considering fuel consumption and traffic smoothness. With regards to fuel consumption, the worst-performing and best-performing vehicles are used as normalising factors, generating a score between 0, lower fuel consumption and 1, worst performance, for each vehicle. The traffic smoothness is represented by the crossing time and number of vehicles that completed their path. Both quantities are computed as function of the percentage of CCAVs. Crossing time is defined as the time interval between departure and arrival for every vehicle.

Table 6 Normalized consumption and emission scores given a penetration rate of CCAVs, considering a simulation of 3600 seconds

% CCAVs	CCAVs		Human driven vehicles (HD)			
	Consumption	Emission	Consumption	Emission	#CCAVs	#HD
0	-	-	0.74	0.69	0	1540
20	0.61	0.56	0.64	0.58	308	1232
80	0.46	0.38	0.49	0.44	1232	308
100	0.43	0.36	-	-	1540	0

Table 7 Crossing time and number of vehicles that completed their path as a function of the percentage of CCAVs, considering a simulation of 100 seconds

	0% CCAVs	20% CCAVs	80% CCAVs
Average crossing time [s]	56.26	54.49	49.01
Maximum crossing time [s]	87.53	83.32	79.66
N. vehicles [-]	35	39	41
Reduction of crossing time	Ref.	3.15%	12.88%

Finally, it's worth noting that the policy used for the tests is the result of a refinement of the policy behaviour driven by the feedback of test drivers. In order to ensuring the real-world acceptability of such technology, also the driving comfort of vehicle occupants is considered. To this end, lateral and longitudinal jerks' constraints have been included in the policy, with defined thresholds rooted in experimental data and on the test drivers feedback. This approach resulted in the development of a comprehensive and functional motion planning algorithm.

According to the qualitative results, 80% of the participants perceived the scenario with 80% CCAVs to be safer. Additionally, 70% of the participants preferred the scenario with 80% CCAVs. Moreover, as the number of CCAVs increased, both CCAVs and HDs reduced their fuel consumption and emissions on average, and the average crossing time decreased.

FINAL TESTS

Final tests have been performed on a broader group of drivers. They are intended to obtain a complete understanding of the human perception of Cooperative, Connected and Automated Vehicles (CCAVs).

Forty volunteers (50% male; mean age = 23.6, SD = 2.26; average years of licensed driving experience = 5.24, SD = 2.86) were recruited and asked to drive along two scenarios, each with four replications. Before starting the driving scenarios, participants were asked to answer a Likert-type questionnaire developed to assess participants' disposition towards autonomous vehicles and the prospect of driving alongside autonomous vehicles (AVs). At the end of the driving session, participants were asked to answer the same questions in order to assess the impact of the driving experience on their willingness to share roads with autonomous vehicles. The Wilcoxon match-pairs test was used as a non-parametric analogue of paired sample t-tests in paired comparisons involving ordinal data; r was used as an effect size in Wilcoxon test. The r value varies from 0 to 1; r values ranging from 0.10 to 0.3 indicate small effect, from 0.30 to 0.5 a moderate effect, and r values larger than 0.5 suggests large effect size. Rank-mean consistency analysis (i.e., Wilcoxon matched-pairs test) results are summarized in Table 8.

Table 8 Participants' Disposition towards Autonomous Vehicles and the Prospect of Driving alongside Autonomous Vehicles: Pre-Driving and Post-Driving Comparisons

Items content	**Before Driving**		**After Driving**		**Comparisons**	
	Mdn	***SD***	***Mdn***	***SD***	Z_W	***r***
Disposition to share roads with AVs	3.00	1.02	3.00	0.92	0.18	.03
Disposition towards driving an AV	2.00	0.98	2.00	0.98	0.20	.03
Differences between sharing roads with human drivers and AVs	3.00	0. 89	3.00	0.99	-1.62	.26
Importance of humans keeping control of driving	3.00	0.88	2.00	1.10	1.95	.31
Larger predictability of AVs vs. human drivers	2.00	0.88	2.50	0.98	-2.74^{**}	.43
Possibility of positive interactions with AVs	3.00	1.07	3.00	1.03	1.24	.20

Note. Scores ranges from 1 to 5; AV: autonomous vehicle; ZW: Wilcoxon matched-pairs z value; r: Effect size measure for Wilcoxon matched-pairs test.
*** p <.01*

As it can be observed in Table 8 , after the driving session, participants considered AVs significantly and moderately more predictable than human drivers; notably, this is the only comparison that reached statistical significance. Thus, it seems to indicate that the driving sessions in which participants had the opportunity to drive alongside autonomous vehicles did not impact negatively on participants disposition towards AVs. During each driving session (i.e., 8 scenarios), participants were asked to identify the percentage of AVs (i.e., 20% vs. 80%). On average participants were able to correctly identify 3.68 (i.e., 46.0%; *SD* = 1.32) scenarios out of 8 total driving scenarios. This result suggests that the perceived differences between scenarios with 20% AVs and 80% AVs was limited; moreover, the number of correctly identified driving scenarios significantly discriminated the scenario with a predominance of human drivers (20% AVs, 80%

human drivers: $M = 2.21$, $Mdn = 2.00$, $SD = 0.81$), from the scenario with a predominance of AVs (80% AVs, 20% human drivers: $M = 1.47$, $Mdn = 1.00$, $SD = 1.05$), Wilcoxon matched-pairs z value = -2.79, p <.01, $r = -.44$. These findings seemed to suggest that no significant differences between the different scenarios were perceived by participants, and that the driving session seemed to foster participants' perception of AVs as more predictable than human drivers, which may, in turn, promote the willingness to drive alongside AVs.

Test case #1.4: Comfort and safety analysis

The dynamic driving simulator allows the subjective evaluation of the comfort perceived by a real human passenger in a vehicle driven by the policy. The passenger is seated in the driving seat, but the vehicle is driven by the policy to simulate a passenger on a CCAV. Although the old policy was efficient in optimizing fuel consumption and traffic flow, the levels of acceleration reached by the vehicle were not suitable to be used in a real environment. Lateral acceleration reached values higher than 1g in both directions, making the vehicle greatly uncomfortable, unsafe and even unrealistic.

The new policy has been trained to limit accelerations and jerks below limit values obtained from experimental tests, corresponding to a lateral acceleration below 0.43 g, a lateral jerk below 1.18 m/s^3, and a longitudinal jerk below 2.9 m/s^3. In this case, the vehicle reaches lower values of lateral acceleration, remaining below 0.5 g, and satisfies the limits on jerks. This solution has been tested with real human drivers. Passengers' feedback reported as satisfactory the perceived comfort provided by the policy. Referring to the time required to navigate the roundabout, the limits on the lateral acceleration reduce the performance of the vehicle, and an increment of about 3 seconds in terms of crossing time can be observed with respect to the old policy. Figure 10 shows the results in terms of longitudinal and lateral acceleration for the old and the new policy.

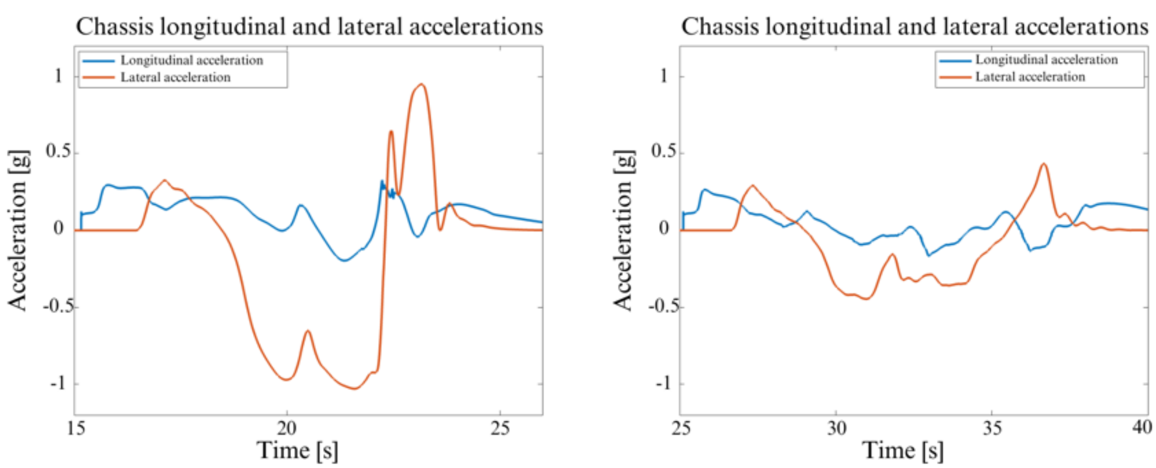


Figure 10 On the left, lateral and longitudinal accelerations during the roundabout crossing of a CCAV with the old policy. On the right, lateral and longitudinal accelerations during the roundabout crossing of a CCAV with the new policy

Test case #1.5: Latency analysis

The policy interaction with the simulation environment has been modified considering the possibility of tuning the delay with which it sends its state to the AI@EDGE CCP and receives the observation of the environment. Specifically, two buffers have been implemented to create a queue of observations and actions, respectively. In this way, it is therefore possible to introduce a simulated delay which is an integer multiple of the simulation step.

To define the real delay, the TBM connection with the AI@EDGE CCP has been tested. A ping has been used to evaluate the communication delay since it carries an amount of data comparable with the real one given by the CAN bus on the vehicle, and it ensures proper time synchronization in reading time intervals. The ping has been used for a total time of 15 minutes to obtain several measurement points and a statistically relevant population of data.

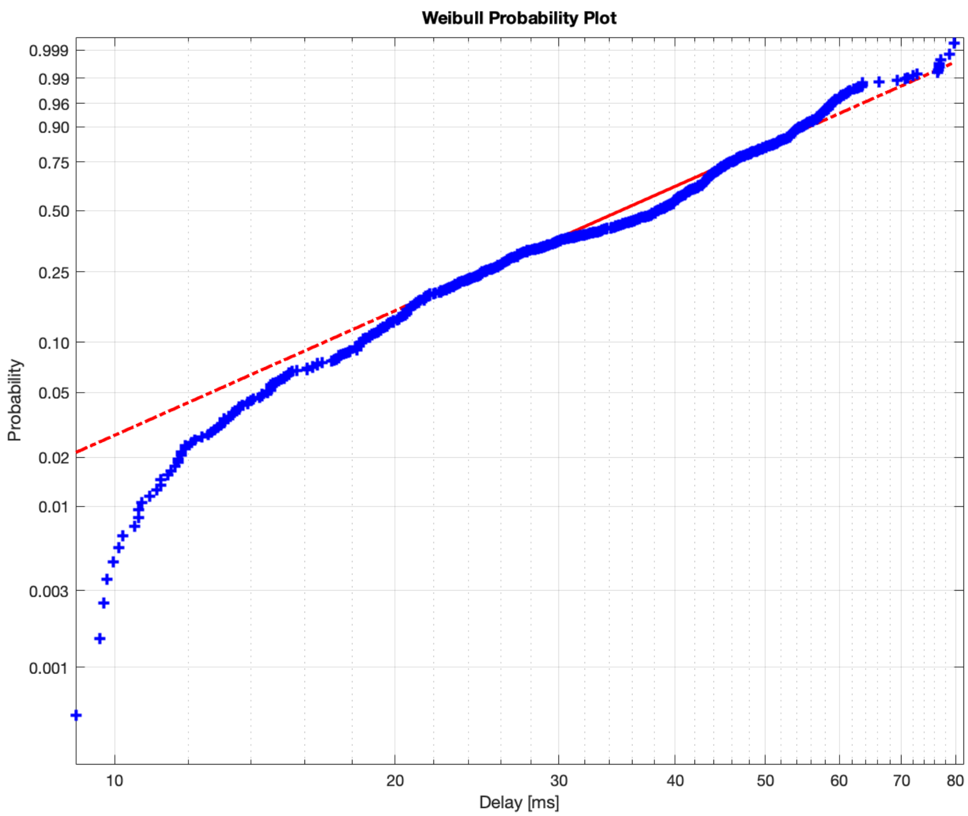


Figure 11 Delay measurement data function of their probability, considering a Weibull probability distribution as reference in red

As shown by Figure 11, the data exhibit a behaviour comparable to that of a Weibull probability distribution. In Table 9, the characteristics of this distribution are collected, together with its percentiles.

Table 9 Weibull probability distribution parameters and percentiles

Parameter	Value	Unit of measure
Weibull distribution parameter #1	41.134	–
Weibull distribution parameter #2	2.9016	–
2.5% percentile	11.587	ms
50% percentile	36.253	ms
97.5% percentile	64.501	ms

Considering what has been achieved, the communication delay which is used for the experimental test is equal to 36.253 ms. This value is real for the ego vehicle driven by the human driving in the dynamic driving simulator and simulated for all other vehicles in SUMO.

3.4.2 Test case 2 results – V2X over 5G

Test case #2.1: V2X Messages from DUT to Edge server over 5G

V2X message from DUT to EDGE server are handled by an AMQP client with the encoding and decoding locally in addition to the sending and receiving CAM messages. A unidirectional data (in the form of CAM packets) flow must be obtained in the policy from the TBM via AMQP broker. Consequently, a client which also can decode the CAM packets arriving from TBM via broker, would be needed to facilitate the data collection at the policy side. The AMQP client along with the V2X decoder interfaces are being dockized and integrated with other applications. Then an end-to-end data communication testing from TBM to Edge server has been conducted. In which the generated CAM packets at V2X stack has been sent by TBM to the EDGE, and they are received. The CAM packets received at the EDGE being forwarded to the AMQP broker and then the Data subscription has been performed as at policy side to use the information about the vehicle positioning, speed, and acceleration.

The test setup on 5G emulation bench with a data connection to a local host server is depicted in Figure 12 and the V2X message exchange by the TBM is shown in Figure 13, where AMQP client receives first the CAMs from V2X stack and then it forwards to the local host server. For which the communication delays are given in Table 9. The main objective of this test case was to understand the correct functioning of the TBM with an emulated 5G network in a lab environment and to get the correct set of parameters from the 5G connectivity point of view. Those parameters then are being used for the connectivity in the POLIMI lab for real connection and results are given in the Test case 1.5.

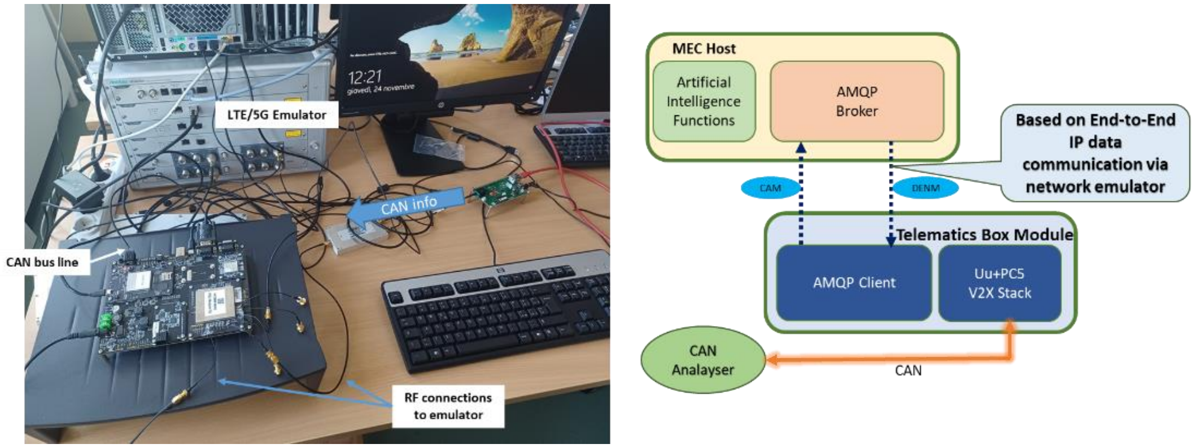


Figure 12 Test setup on Virtual Validation bench at CRF site where the CAN data are simulated by CAN analyser

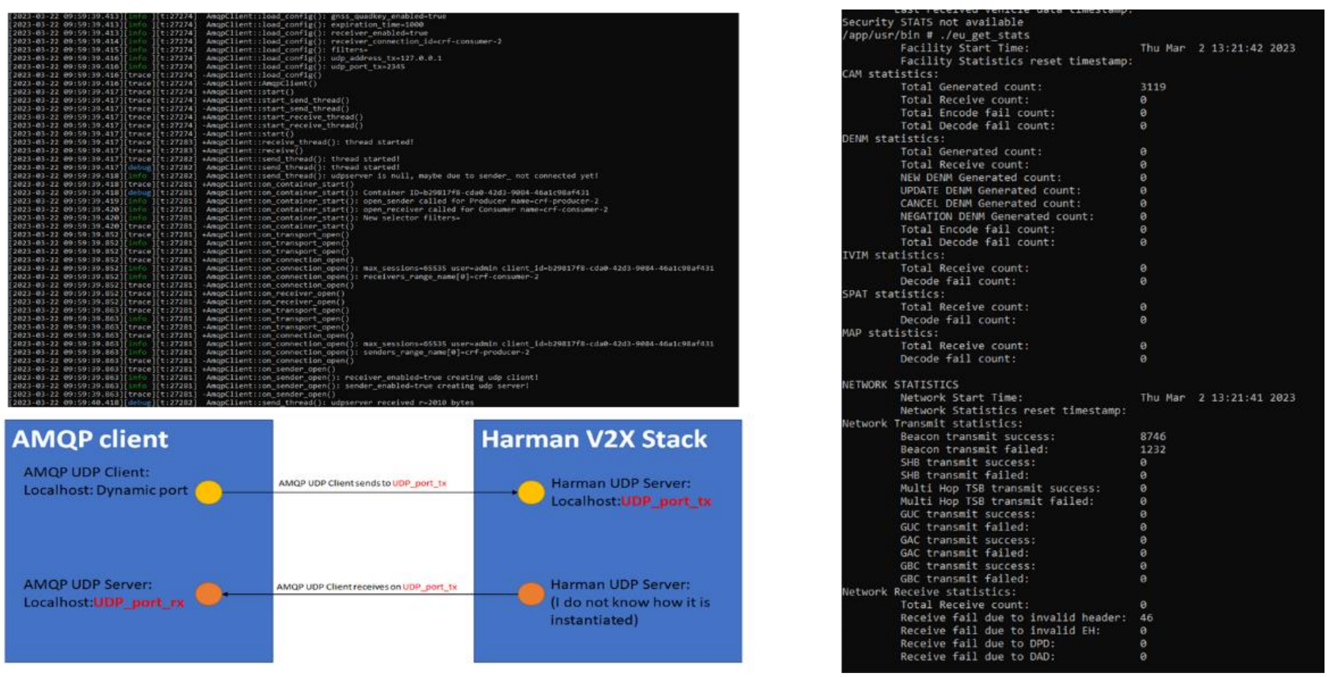


Figure 13 V2X message exchange between the TBM and the server

Test case #2.2: V2X Messages from Edge server to DUT over 5G

The same AMQP Client-Server based method is also used to process V2X messages from EDGE to DUT. Under the V2X architecture, a TBM or TCU device often receives CAM messages from nearby vehicles as well as other messages from warning traffic stations and roadside devices that complies with ITS-G5 standards. Most of these messages fall into the V2V and I2V categories, and the UC1 context limits them to V2N2V. As a result, a direct V2X message from EDGE to DUT is now outside the purview of this use case. Nevertheless, the TBM that is being examined in this case is able to receive those messages from the EDGE, and local decoding can also be done at the V2X stack that is installed on the device.

3.4.3 Test case 3 results – 5G Connectivity and Local Traffic Breakout

Test case #3.1: Connection between gNB and 5GC

No connection errors. Log messages show gNB successfully attached to the AMF. Control plane messages between the RAN and the 5GC are correctly exchanged.

Test case #3.2: UE's attach to and detach from the 5G network

Log messages show UE successfully registered, attached, and detached to the 5GC.

Test case #3.3: Connectivity between UE and data network (DN)

Connectivity between UE and DN is operational.

iPerf[16] shows a maximum of 50Mbit/s in download and 5Mbit/s in upload throughput over band n78 (40-MHz bandwidth).

ICMP messages are acknowledged.

3.5. Final remarks

AI@EDGE CCP has proven to be crucial to orchestrate the traffic in a complex scenario such as the roundabout. The most important developments introduced by the project include:

- The definition of a new communication protocol, called V2N-N2V. This protocol shows better performance than others so far used, of which V2V and V2I are examples. Unlike V2V protocols, it is capable of handling large amounts of data exchanged between vehicles and the CCP present in the centre of the roundabout. This allows for better management of the intersection and the development of truly cooperative policies. In contrast to V2I protocols, in this case the infrastructure is very simple, allowing a reduction in associated costs.
- The implemented communication architecture is completely independent of the vehicle type and, therefore, the manufacturer. It uses data that each vehicle is capable of generating and provides easily applicable indications regardless of the vehicle model. This made it possible to develop a system that is extremely adaptable to both different intersections and traffic conditions.

[16] Please see: https://iperf.fr/

- 5G mobile communication turns out to be essential for the proper functioning of the system employed. Indeed, thanks to it, it is possible to achieve lower values of communication latency and ensure the proper functioning of the policy. 4G technology, aversely, would have a much higher latency, and autonomous vehicles would suffer from a high mismatch between received and real-time information.
- The driving simulator at the DriSMi laboratory of the Politecnico di Milano proved to be instrumental in achieving the key results described in this document. Especially for a technology still in an early stage of research, a state-of-the-art simulator allowed rapid advancement and the acquisition of a database composed of both those obtained from the vehicle and those measured directly on humans, thus being able to objectify human response.
- Another key point to describe the importance of the AI@EDGE project turns out to be the enrichment of the literature on the subject and the interaction it had with policymakers. Indeed, several papers have been produced using the preliminary results obtained so far and presented at various conferences. One of these is EARPA (European Automotive Research Partners Association), during which it was possible to update European policymakers on the progress of the project and its future development.

A much larger test campaign was conducted at the DriSMi laboratory of the Politecnico di Milano in early December. In section 3.4, a portion of the final results related to subjective perception of the simulation environment is presented. Before the final review, they will be enriched with those derived from sensors mounted on drivers, aiming to objectify human perception. These final tests and results will ensure new possibilities for interaction with policymakers and OEMs in the future. A new paper discussing the project in its entirety will also be produced.

4. Use Case 2: Secure and resilient orchestration of large (I)IoT networks

This use case is showcasing a smart manufacturing scenario with a 5G campus network. The objectives are to showcase the security and privacy aspects of the AI@EDGE Connect-Compute Platform leveraging the utilisation of AIFs (anomaly and intrusion detection), of NetFPGA[17], and federated learning to detect intrusions and anomalies in the network.

4.1. Validation objectives

The following functionalities are to be showcased in this use case:

- **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

[17] Please see: https://netfpga.org/

integration of the Split and Merge algorithm at the state of the art in NetFPGA boards, developed by CNAM.

- **Auto configuration of Hyper Parameters 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**: 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

As described in D5.2, in the validation scenario different security mechanisms will be showcased by simulating an attack through the Mirai botnet[18]. The demonstrator setup was extended with an IT Zone, which simulates and enterprise network environment. In the scenario an employee desktop from the IT Zone is infiltrating the malware, e.g., by opening a malicious email attachment. The botnet will then propagate to the Manufacturing Zone and compromise the IoT cameras using default credentials. The IoT cameras will then attack the 5G infrastructure with a DDoS attack. Ideally all security AIFs should be able to detect the attack on different levels, such as TCP flow (Autoconfiguration of Hyperparameters), IP metadata (NetFPGA) and network patterns (Federated Learning Anomaly Detection AIF).

[18] Please see: https://github.com/jgamblin/Mirai-Source-Code

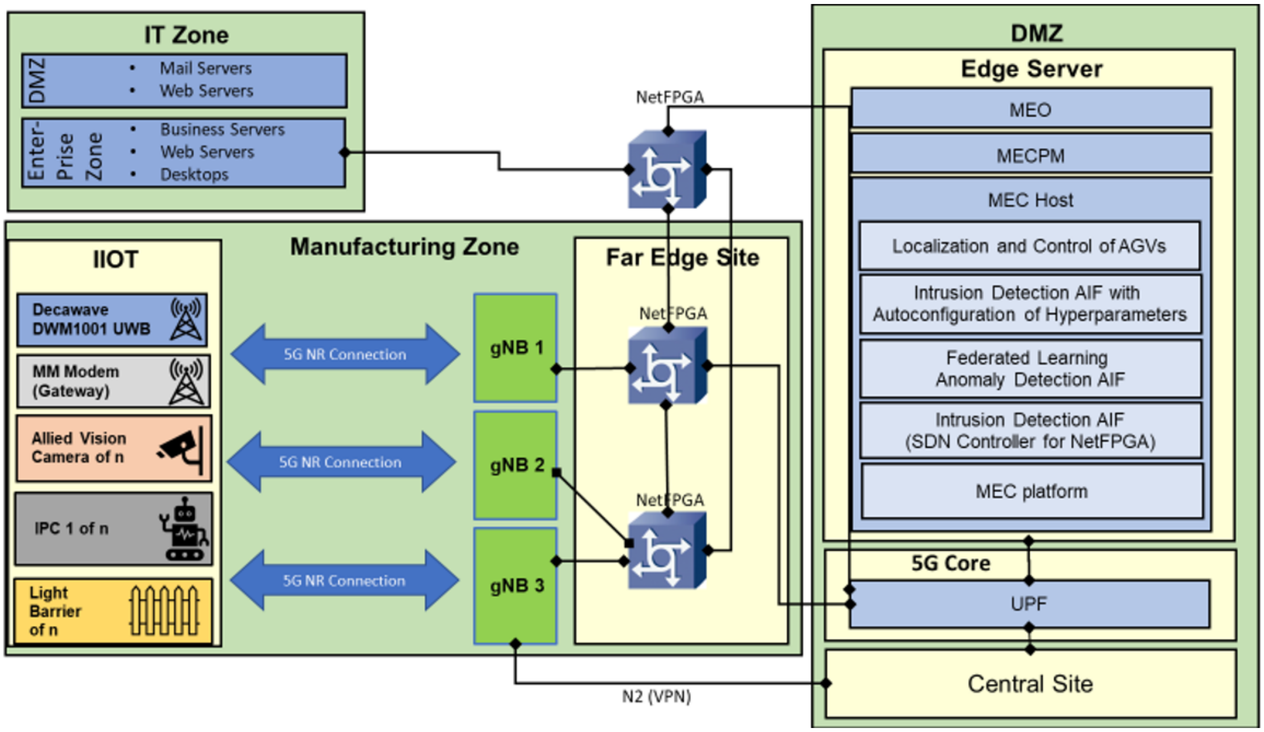


Figure 14 Components of the UC2 Testbed

The Use Case 2 testbed is depicted in Figure 14. It is split into three zones:

- the IT-Zone.
- the Demilitarized Zone (DMZ).
- the Manufacturing Zone.

In the IT Zone, typical building blocks of an enterprise network are to be found such as mail servers and employee desktops. In the Manufacturing Zone the IoT and the RAN can be found, whereby the IoT devices are connected to the network over a 5G connection. In the DMZ is the Edge Server, the 5G Core and the Central Site. The NetFPGAs connect the gNBs with the UPF and thereby analyze the metadata of the GTP traffic packets. In the Edge Server the MEO, the MECPM and the MEC host are located, where the showcased AIF are hosted. Namely the Localization and Control of AGVs, the Intrusion Detection AIF with Autoconfiguration of Hyperparameters, the Federated Learning Anomaly Detection AIF and the Intrusion Detection AIF for the NetFPGAs. The UPF is connected to the control plane of the core on the remote central cite and thus showcasing the distributed aspects of the AI@EDGE platform.

4.3. Validation procedures

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)	• Run DDoS attack with Mira Botnet. • 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 Mirai botnet.
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)	• Q4 of 2023
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. • Train on CIC IDS[19] Dataset and reconfigure for collected attack traces

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)	• Classify network traffic and compare anomaly score to baseline from literature of Split & Merge algorithm
Expected Results (Post-Conditions)	• Qualitative description and comparison to other literature • Accuracy not possible to measure because real world data is not labeled.
General Time Plan (Validation Campaigns)	• Q4 2023

[19] Available at: https://www.unb.ca/cic/datasets/index.html

Test Sequence	• 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 2024
Test Sequence	• CPU overload injection. • Packet loss injection. • Link bandwidths decrease emulation. • Known attack emulation. • Unknown attack emulation. • Link failure emulation.

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

Phase 1: Connection between gNB and 5GC	
Test case 4.1:	**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)	• Q2 of 2023.
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 4.2:	**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)	• Q2 2023.
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 4.3:	**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)	• Q2 2023.
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. Validation results

4.4.1 Test case 1 results – Intrusion Detection for Known Attacks

In Test Case 1, we employed meta-learning for the auto-configuration of Network Intrusion Detection Systems (NIDS), with detailed methodology outlined in the deliverable D4.2 [5]. This section expands upon that foundation, presenting in-depth experiments and advanced results using the IDS2017[20] and IDS2018[21] datasets. It's important to note that both datasets are segmented into daily intervals, each encompassing a distinct type of attack, spanning a total of 17 days.

Figure 15 (an infinite value due to division by zero is marked as 100% on day 13) illustrates the daily performance variations of Meta-learning (MtL) in comparison to Hyperparameter Optimization (HPO). In this setup, the number of HPO iterations is capped at the time MtL takes to infer a configuration. Performance difference is quantified as (p(MtL) - p(HPO))/p(HPO). Generally, the performance variation is minor across most days. However, there are notable deviations: a decrease in precision by -10.89% on Day 8 and -28.05% on Day 9. Conversely, MtL and HPO show similar precision, as HPO iterates further to optimize the Matthews Correlation Coefficient (MCC). Notably, the applied HPO technique is not multi-objective, leading to a sacrifice in precision to enhance MCC. On Day 3, MtL improves Recall by 8.57% and MCC by 4.19%. On the 13th day, HPO's precision at the 7th iteration is nearly zero, whereas MtL achieves a precision close to 1.

Memory usage also varies significantly. Implementing the HPO solution necessitates using the complete daily dataset from IDS2017 and IDS2018. The memory footprint ranged from a low of 18.79 MB on Day 11 to a high of 837.33 MB on Day 4. MtL, which relies solely on meta-features, consistently requires less than 36KB of memory for each day.

[20] Available at: https://www.unb.ca/cic/datasets/ids-2017.html
[21] Available at: https://www.unb.ca/cic/datasets/ids-2018.html

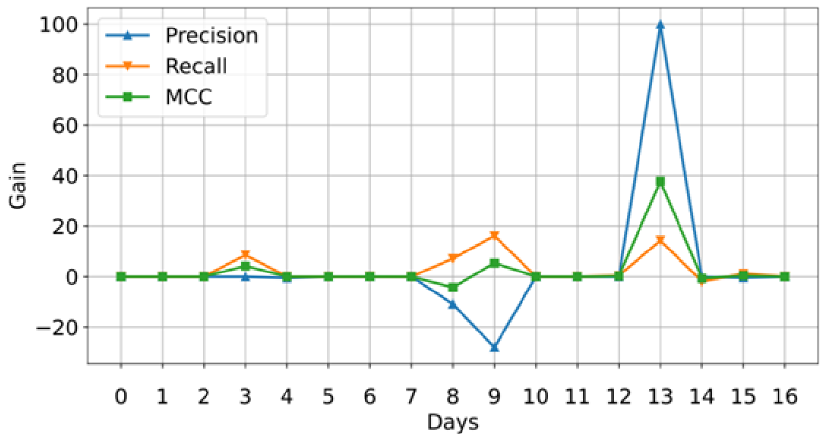


Figure 15 Relative performance of MtL calculated as a ratio between (MtL-HPO) and HPO

Meta-learning hinges on having a representative meta-dataset. Previously, our studies used 16 days of data, with each day varying in network context and attack type. In this analysis, however, we explore scenarios with fewer days to construct the meta-dataset, thus limiting prior experience. We predict configurations for days not included in each subset of days. For instance, with a 2-day meta-dataset, there are 136 combinations affecting the other 15 days, leading to 2040 unique configurations, as shown in Figure 16.

Results indicate that using more days generally improves NIDS performance, but the performance difference between closely numbered days (like 3 and 4, or 14 and 15) is negligible. Performance starts to stabilize with 8 days, achieving similar outcomes to using 12, 14, or 15 days. This finding suggests the feasibility of limiting computational resources during meta-modeling by using a smaller meta-dataset. Consequently, building the meta-target vector would require fewer iterations (2400 instead of 4800) for comparable MtL performance.

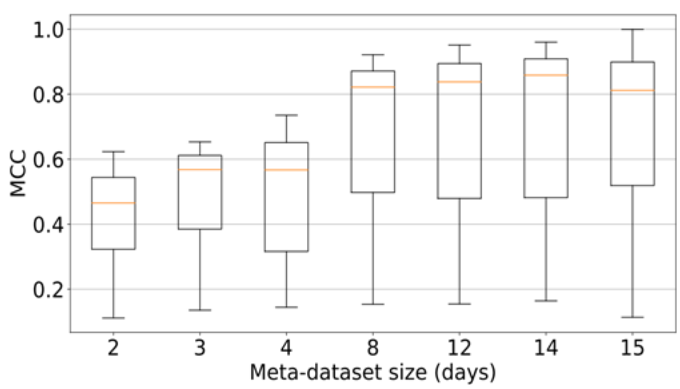


Figure 16 NDIS performance

4.4.2 Test case 2 results – Intrusion Detection for Unknown Attacks

In this test case, NetFPGAs for botnet detection were employed. The early detection of DDoS attacks is a serious concern for the reliable and secure operation of IoT networks. Since botnets can prepare such attacks in a stealthy manner, measures have to be found that enable continuous network monitoring and anomaly detection.

Due to the nature of typical botnet behaviour, an IDS optimized for the detection of stealthy botnet activity should leverage port-based analysis as a first step. Consequently, changes in network behaviour can be tracked in long-term profiles, which allows for a wide view of the network in which systemic anomalies can be distinguished from local fluctuations. This can be achieved by employing an original implementation of the SPLIT-AND-MERGE anomaly detection algorithm, which enables Collaborative IDS (CIDS).

As a reference architecture, this test case employs a software defined network (SDN), in which the data plane consists of FPGAs, which can handle network monitoring and switching decisions. The data plane communicates with the control plane, which establishes the routing rules, via a Southbound Interface (SBI). In combination, this setup allows the FPGAs to process packets on a device-level, with distributed units collaborating by communicating collected metrics to a central aggregator, which performs local anomaly detection as well as network-level monitoring and alerts. The issue of memory bottlenecks is mitigated by using data sketches via HYPERLOG.

One major advantage of this setup is that it not only allows the early detection of botnet activity, but also basic mitigation actions, such as redirecting traffic, blocking ports and allowing or disallowing specific IPs.

The solution was evaluated by testing it against the MAWI[22] Archive dataset, where it showed promising performance in detecting networks anomalies, while keeping false alarms at a manageable level. Further research should concentrate on enabling the monitoring of several ports at the same time, to maximize monitoring coverage.

[22] Please see: http://faculty.nps.edu/cabollma/MAWI_Datasets/Datasets.html

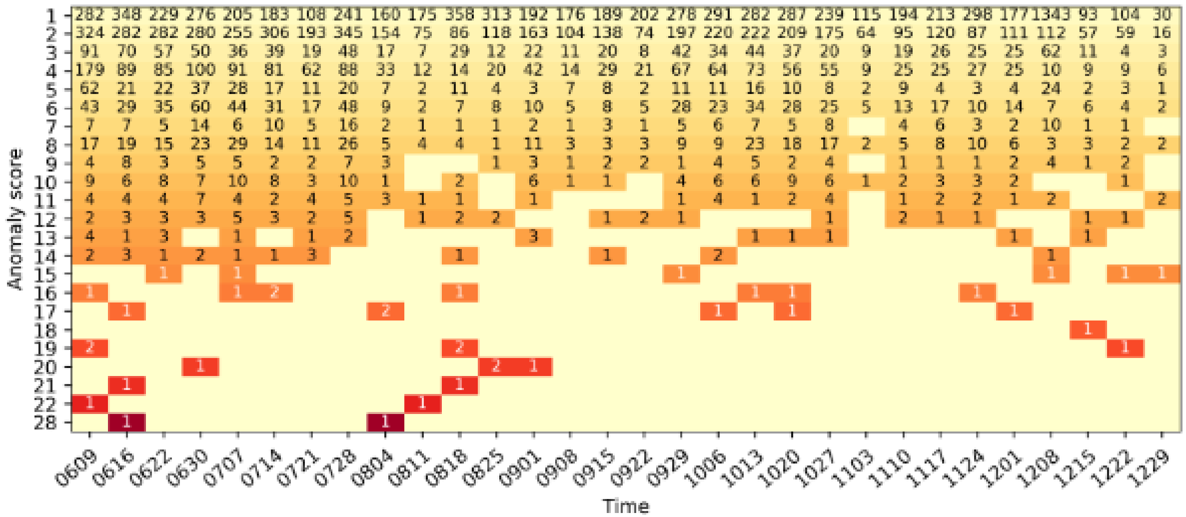


Figure 17 Anomaly Scores for the 2016 period

MAWI traces are split into nine subtraces, according to the respective subnetwork prefixes. Each subtrace is then submitted to a NetFPGA smart switch to be analysed and related per-port SPLIT-AND-MERGE metrics - namely, dst/src IP address cardinality, source port cardinality, packet count, SYN count, packet size meant/variance – are obtained. Applying SPLIT-AND-MERGE algorithm to obtained data returns a certain Anomaly Score (AS) for a given TCP port at a given time, providing a network administrator with an insight on which ports may be misbehaving at a given time.

Figure 17 shows the AS obtained for the dataset considered. Given the unlabelled nature of the archive, relevant anomalies (with AS greater than 15) have been individually investigated. As a result, the proposed CIDS provides a convenient number of alerts per day (i.e., 1 relevant alert per day on average), of which approximately 30% are false positives. Among the detected anomalies, we highlight some related to vulnerabilities that were discovered few months or even few weeks before the detection, such as the Redis scan on port 6379 (June 30 – AS 20), two potential intrusion attempts towards Squid servers on port 3128 (Aug 18 – AS 21 and Sep 1 – AS 21), a Denial-of-Service attempt on Privoxy (Aug 18 – AS 16). Also, we highlight some well-known botnets for which we detected scanning activity, such as Mirai on port 23 (Aug 4 – AS 28), ADB.Miner on port 5555 (Dec 1 – AS 17) and Cyclops Blink on port 995 (Jul 7 – AS 15).

4.4.3 Test case 3 results – Anomaly Detection

With the anomaly detection AIF based on federated learning, a framework for anomaly detection for 5G and beyond (xG) infrastructures, was developed (Figure 18). Since labelling anomalies in complex and dynamic networks is a major issue, autoencoders have been employed, as they don’t require labelling. To that end, the Long-Short-Term Memory (LSTM) approach has been leveraged for time-series analysis.

Since centralized learning is ill suited for complex networks like 5G and beyond, due to privacy and security issues as well as communication overhead, a federated learning (FL) approach has been used, in which all device data is stored and processed locally, with only the model parameters aggregated to a central node. Thus, global models can be trained locally, in a collective manner. This approach not only solves the problem of prohibitively long training times in centralized learning for complex networks, it is also uniquely suited for the monitoring of xG networks due to the fact that mobile edge computing platforms contain enough storage and processing power for machine learning.

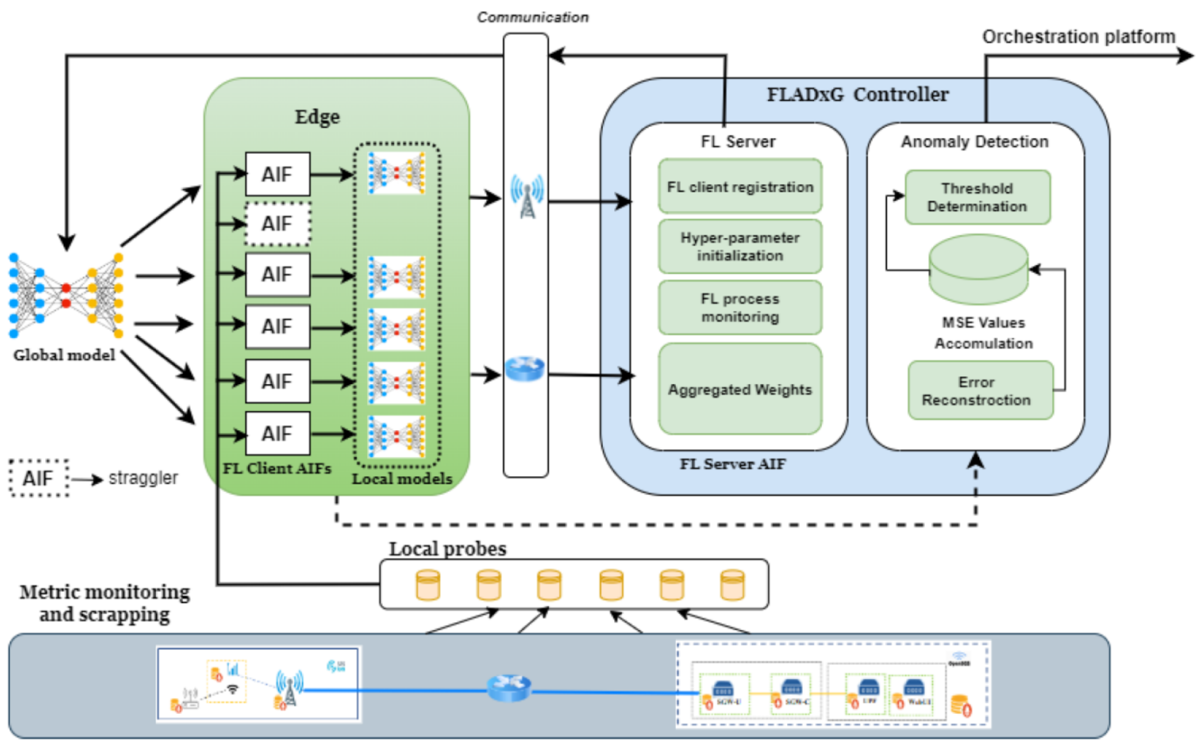


Figure 18 Framework for anomaly detection

For the evaluation of the framework, a training dataset was produced using the 5G End-to-End Emulation (5G3E[23]) platform, on normal conditions. This enables the necessary error reconstruction by identifying deviation from nominal conditions. In a second step, a threshold was defined at which events are indicated as anomalous behaviour. The test dataset was then created by injecting various anomalies:

- **CPU overload injection**: CPU overload was tested by stressing the physical CPU at 80% nominal capacity. This tests how well the autoencoders that are distributed across multiple nodes can perceive such stress. The recall which is sensitive to the anomalies that are correctly identified, varies between 0.78 to 0.9. It is approximately 8% less compared to SYRROCA[6] in the worst-case scenario, which is when the number of clients is two. For the precision, the difference is less than 0.5% which is not a severe deviation. F1-score has a similar behaviour as well.

[23] Please see: https://hal.science/hal-03698732v2

- **Link bandwidth decrease emulation**: access bandwidth bottleneck was simulated by increasing network traffic up to 15 times of its normal values. This is to preserve the effect of the change in bandwidth on the underlying infrastructure resources and how the anomaly detection framework will capture this change. The framework shows a rather large gap in recall with respect to SYRROCA, with a deviation of approximately 20%. But for the other metrics that performance is closer or in some cases better.
- **Packet loss injection**: the handling of packet loss was tested by randomly dropping up to 80% of packets. This shows similar behaviour to the scenario of CPU overload injection. The deviations from the anomaly detectin framework to SYRROCA are less than 10% at its worst case and the precision is very similar, whereas the F1-score is very close.

The results (Figure 19) show that under the various stress test conditions, the federated learning anomaly detection framework showed performance almost on par with centralized approaches, while requiring a massively shorter training time, thus enabling the fulfilment of carrier-grade requirements for post-incident service restoration as well as impairment prediction, which facilitates mitigation of reliability threats.

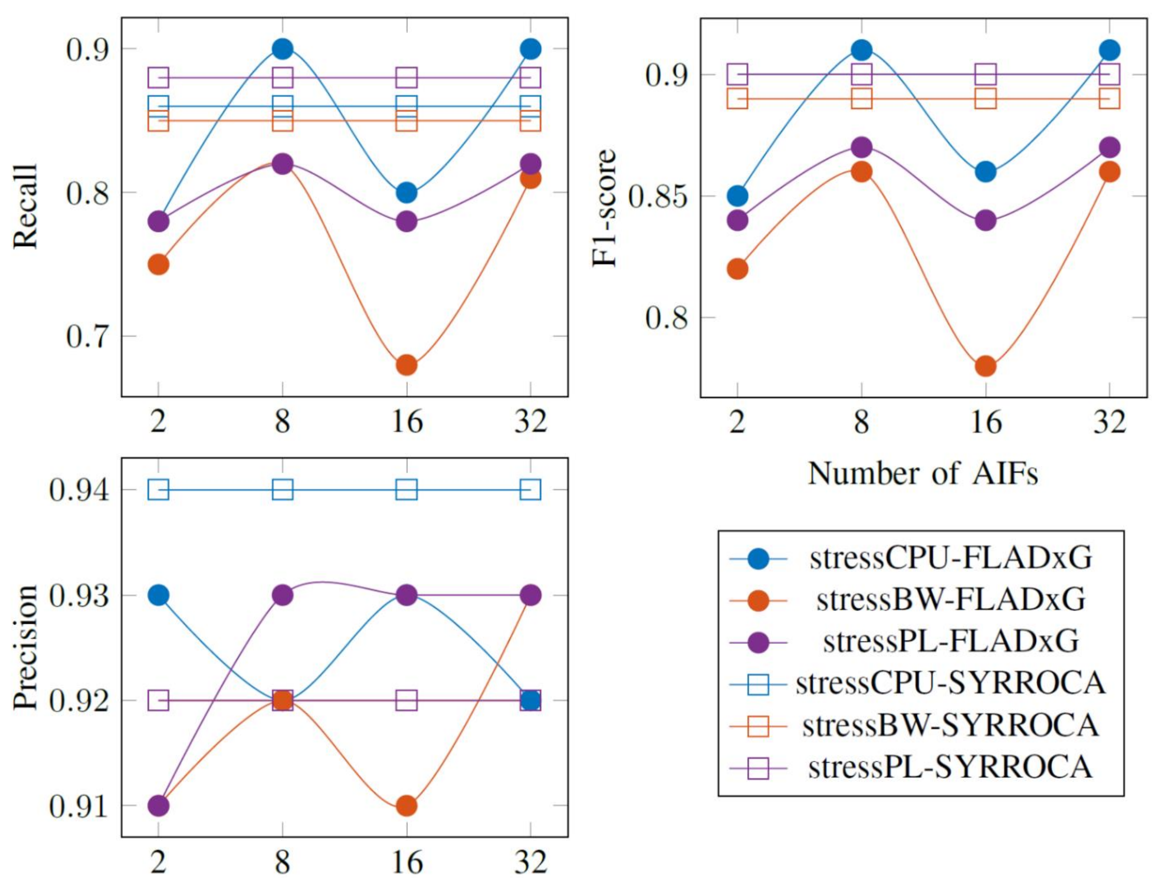


Figure 19 Anomaly Detection AIF (FLADxG) vs baseline SYRROCA comparison in terms of F1 score

4.4.4 Test case 4 results – 5G Connectivity and Local Traffic Breakout

Test Case #4.1: Connection between gNB and 5GC

No connection errors. Log messages show gNB successfully attached to the AMF. Control plane messages between the RAN and the 5GC are correctly exchanged.

Test Case #4.2: UE's attach to and detach from the 5G network

Log messages show UE successfully registered, attached, and detached to the 5GC.

Test Case #4.3: Connectivity between UE and data network (DN)

- Connectivity between UE and DN is operational.
- iPerf shows uplink/downlink traffic.
- ICMP messages are acknowledged.
- Data throughput using a Quectel RM500Q-GL (Over USB 3.0 carrier board with 30 Mhz Bandwidth):
    - Download: avg 70 Mbit/s (min 60 Mbit/s max 100 Mbit/s).
    - Upload: avg 100 Mbit/s (min 95 Mbit/s, max 110 Mbit/s).
- Data throughput using using a Samsung Galaxy S21:
    - Download: max 190 Mbit/s (30 Mhz Bandwidth), min 180 MBits/s AVG 200 Mbit/s max 240 Mbit/s (40 Mhz Bandwidth).
    - Upload: max 100 Mbit/s (30 Mhz Bandwidth), max 120 Mbit/s (40 Mhz Bandwidth).

4.5. Final remarks

In this use case the focus was on the security of interconnected devices in a smart factory environment. A 5G testbed was set up and three different security solutions were developed and validated in the scope of UC2. The validation scenario was a factory floor with AGVs controlled over the edge. The potential damage could be a production outage or even physical harm to the factory workers. The AI@EDGE platform allows for AIFs to be deployed and supports data pipelines to provide the data. The KPIs defined at the beginning of the project have been expanded by additional evaluation metrics to have more comprehensive results. Nonetheless, the presented validation methods present first positive results, there are still gaps to be filled with regards to attack and failure detection. Especially for unknown attack or for changing attack patterns it is a difficult to create sufficiently broad and labelled datasets. Although the presented auto-configuration of hyper parameters is trying to address the problem of small training sets, the problem of small test sets remains.

5. Use Case 3: Edge AI assisted drones in beyond-visual-line-of-sight operations

The primary purpose of UC3 is to monitor extensive road networks using drones in BVLOS (Beyond Visual Line of Sight) mode, enhanced by the 5G network. In this scenario, reliability and seamless data traffic are required to transmit telemetry, images, and videos with minimal delay to the operator and central office for decision-making process.

The monitoring application entails the use of advanced functionalities such as scanning, 3D modelling of infrastructures, incident identification, and notifying the drone operator. Due to the inherent constraints of drone performance, including weight, energy consumption, and other factors, it becomes imperative to minimize onboard systems. Thus, it becomes necessary to offload as many processes as possible from the drone.

By leveraging UC3's operations on the AI@EDGE platform and its MEC systems based on AI and Edge Computing supported by 5G networks, the optimal monitoring support is achieved, accelerating computational and modelling processes, improving reliability, and extending the operational range. In the context of this use case, two AIFs were developed and tested on the integrated AI@EDGE platform: the "Anomaly Detection" AIF and the "3D Reconstruction" AIF.

5.1. Validation objectives

The validation scenario of UC3 has been established based on achieving, first, an adequate integration scheme for the testbed that includes the three development areas identified and defined in D5.1, [1], as the DEVELOPMENT ENVIRONMENTS: 5G NETWORK, DRONE, and AI FUNCTIONS, as well as the integration of MEC functions of the AI@EDGE platform in the testbed for the operation of the integrated system.

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[24] 5G Network and checking that it is working successfully with the required dataflow.

And once this first stage is achieved, it is required to design specific test cases 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 KPIs defined as follows:

- For the drone operation:
  - the latency KPI, composed of two components: Control Signal latency (below 50 ms) and Video processing latency (100 ms);
  - the reliability KPI (tentative metric) in terms of control signal packet loss (below 1%);

[24] Please see: https://www.5tonic.org/

- the range of operation to be at least 20 km.

- 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

This scenario, outlined in the previous section to facilitate the achievement of established validation objectives, is explained in detail below. It is realized through the coordinated efforts of the different partners participating in the use case. On one hand, 5TONIC - Ericsson deploys the necessary devices to emulate the working environment (drone + central office + drone operator) and to establish drone control communication (C2) and video transmission from the drone, utilizing the communication network currently operational in 5TONIC (4G, 5G NSA, etc.). Meanwhile, AERO provides integrated systems onboard the drone that connect the three development nodes (AERO in Madrid, ATOS in Zaragoza, and 5TONIC), enabling continuous and efficient development to meet the objectives. Lastly, ATOS develops the AI functions and is responsible for their deployment and integration into the workflow."

The **5G NETWORK environment** is built on top of 5TONIC laboratory. 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, as depicted in Figure 20.

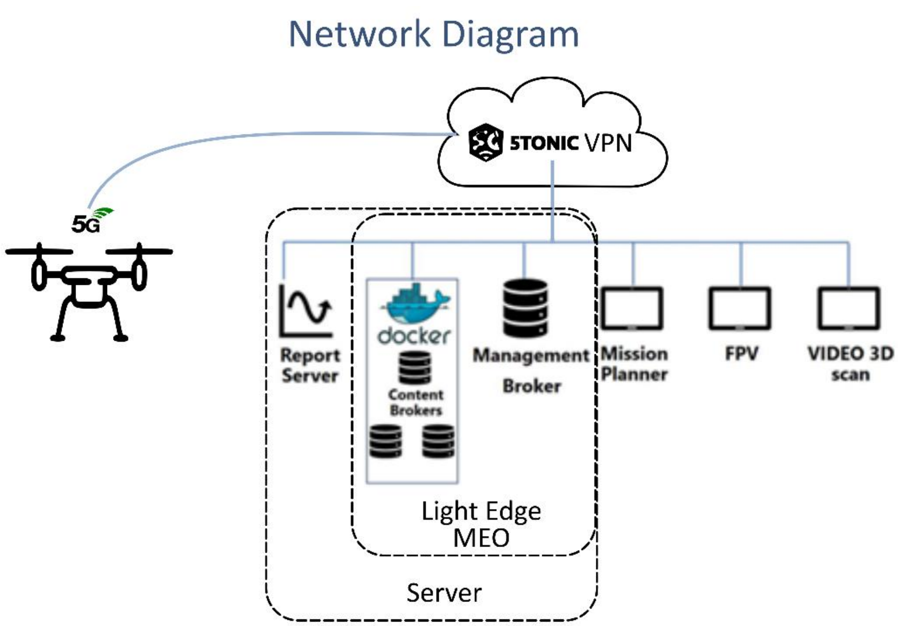


Figure 20 VPN connection for UC3 development process

The VPN deployed at 5TONIC laboratory provides the required connecting among the development environments involved in the UC3, as depicted in Figure 21.

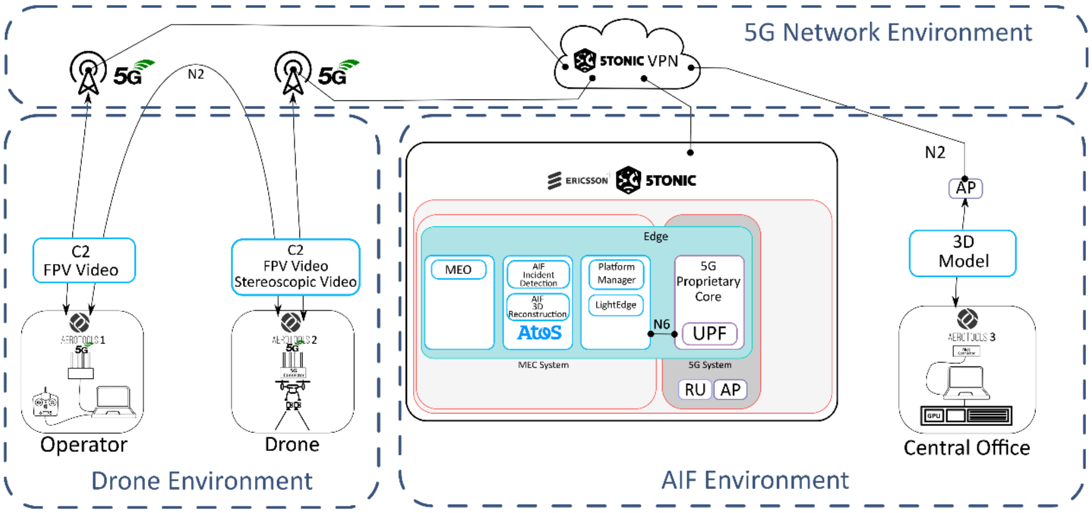


Figure 21 5TONIC VPN connecting UC3 development environments

The **DRONE environment** provides an AT6 drone, a prototype flying platform with specific equipment and integrated systems (Navigation, C2, payload, data transfer) and computing devices such as a Raspberry PI[25], as well as stereoscopic cameras in dedicated stabilized gimbal to provide high quality footage and a First Person View (FPV) camera to support drone´s operations when required by the operator.

Figure 22 UC3 Drone Environment - AT6 drone with integrated devices

Figure 23 shows the devices onboard the drone and the connections among them.

[25] Please see: https://www.raspberrypi.com/

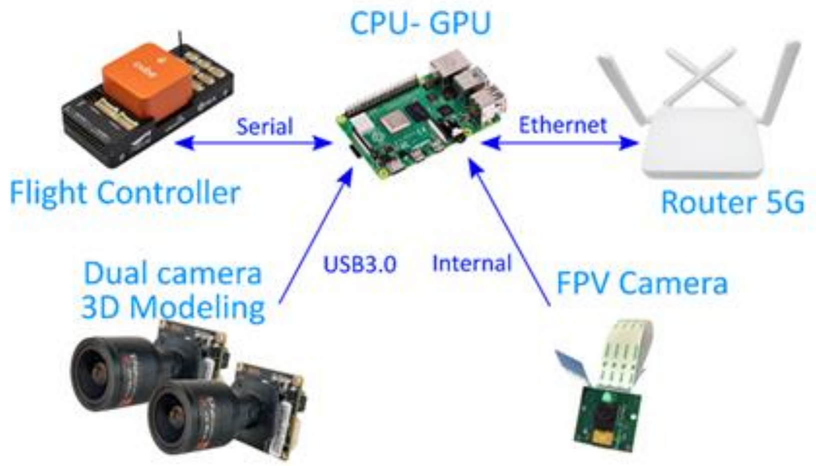


Figure 23 UC3 Drone Environment devices and connections

As shown in the figure, the Raspberry PI, labelled as CPU-GPU device, is connected to the following devices:

- Pixhawk Cube[26] (Flight Controller) using 2 serial ports:
    - Serial1: to be used by drone operator ground station (based on Mission Planner software).
    - Serial2: to be used by the Raspberry PI to read GPS and orientation data.
- CAM1: to be used by Raspberry PI to stream the FPV to drone operator ground station.
- CAM2-3: to be used by Raspberry PI to push the images to the content broker.

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

- Pixhawk Serial1:
    - Shared by Raspberry PI.
    - 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[27].
- Pixhawk Serial2:

[26] Please see: https://www.cubepilot.com/
[27] Please see: https://mavlink.io/

  - Raspberry PI 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).
  - Raspberry PI is streaming it to drone operator ground station directly.
- CAM2 – CAM3:
  - This stereo CAM generates two video streams. Raspberry PI 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 have developed and tested: Anomaly Detection AIF and 3D Reconstruction AIF.

Anomaly Detection AIF. This artificial intelligence function is designed to detect anomalies in the videos captured by the drone and pinpoint their locations. Within the context of UC3, which focuses on the inspection of critical infrastructure, this AI function analyzes data transmitted by the drone. Upon detecting an anomaly, it promptly sends notifications and supporting evidence to the pilot. Subsequently, the pilot can initiate the process of generating a 3D model of the incident scene for future evaluation using the 3D Reconstruction AIF. The implementation of this task entails configuring multiple components to preprocess the inputs and forward them appropriately to the AI function for 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.

- o Scales to 1 million messages per second.
- o Distributed.
- o JSON[28] compliance.
- o 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 (Figure 24), use DETIC[29] as main detector and CLIP[30] as visual-language model. For the scope of this project, the use of CLIP has been focused on two anomalies, persons, and cars, for the description of elements. So, the AIF is automatically fed with these descriptions and when detecting an anomaly, an event will be sent to the pilot through the message broker to trigger an order to initiate a 3D reconstruction if this is considered necessary.

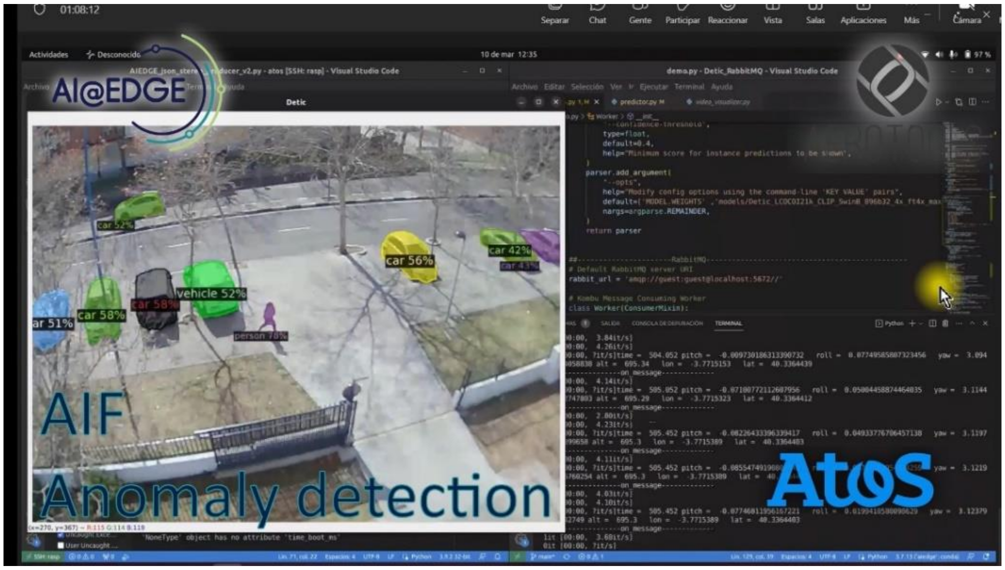


Figure 24 UC3 Anomaly Detection AIF on operation at 5G Environment in 5TONIC

<u>3D Reconstruction AIF</u>: This artificial intelligence function is responsible for creating a 3D model of the area where an anomaly has been detected (Figure 25). When the pilot receives a detection event notification from the Anomaly Detection AIF and deems it appropriate, they may initiate a circular flight (Point of Interest - POI) around the location of the detected anomaly. Simultaneously, the pilot sends a request for video and telemetry acquisition to the onboard Data Server via RabbitMQ. The Data Server then begins

[28] Details at: https://www.json.org/
[29] Please see: https://github.com/facebookresearch/Detic
[30] Please see: https://github.com/openai/CLIP

transmitting the required data to the 3D Reconstruction AIF, initiating the 3D reconstruction process. The components have the following function:

- **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.

Figure 25 UC3 3D Reconstruction AIF on operation at 5G Environment in 5TONIC

5.3. Validation procedures

For the validation process, it is necessary to specify the procedures to follow, which are structured around the Test Cases that have been developed for each of the KPIs to be verified, providing the framework to measure and record the required parameters to validate defined KPIs for this UC3.

The number of Test Cases defined are five, adapted to the goals set for the UC:

1. **Integration** of all systems in the defined environments and demonstrate visibility, connectivity among all of them and correct operation.
2. Operation of integrated systems with additional measuring elements to test **latency**.
3. Operation of integrated systems with additional measuring elements to test **reliability**.
4. Operation of integrated systems with additional measuring elements to test **range**.
5. Deployment of Anomaly Detection AIF to measure **average precision** values to compare with a reference dataset.

The detailed description and sequence of actions for each Test Case are explained in the following tables.

5.3.1 Test case 1 – Integration

Test Case	**#1**
	Integration
Slogan & Objective	Testing integration and connectivity among all systems
Test Scenario (Pre-conditions)	Designed set-up for development
Expected Results (Post-conditions)	All systems connected and sending/receiving data
General Time Plan (Validation Campaigns)	Final architecture tested as planned and achieved objective by Q3 2023. Testing throughout Q3-Q4 2022
Test Sequence	• Deployment of drone and AIFs in 5G Network Environment. • Connection of systems to VPN, checking visibility. • Start of operation with Drone take off and image shooting. • Request of AIFs

5.3.2 Test case 2 – Latency

Test Case	#2
	Latency
Slogan & Objective	Testing of communication latency for drone control and video transfer.
Test Scenario (Pre-conditions)	Integrated set-up developed plus measurement tools (iPerf, Zabbix[31], Grafana[32])
Expected Results (Post-conditions)	C2 latency ≤ 50 ms • (Drone to/from Operator PC/Ground Control Station) Video latency ≤ 100 ms • (Drone to/from Operator PC)
General Time Plan (Validation Campaigns)	Testing carried out throughout Q3-Q4 2023.
Test Sequence	• Deployment of Drone and AIFs in 5G Network Environment. • Connection of systems and preliminary communication test with iPerf. • Start of operation: Drone take off & delivery of C2&Video signals. • Monitoring, measuring & recording ping signals (Zabbix and Grafana)

5.3.3 Test case 3 – Reliability

Test Case	#3
	Reliability
Slogan & Objective	Testing of reliability on C2 signals.
Test Scenario (Pre-conditions)	Integrated set-up developed plus measurement tools (Proprietary SW tool similar to iPerf, Zabbix, Grafana)
Expected Results (Post-conditions)	C2 signal packet loss ≤ 1%. • (Drone to/from Operator PC/Ground Control Station)
General Time Plan (Validation Campaigns)	Testing carried out throughout Q3-Q4 2023.
Test Sequence	• Deployment of Drone in 5G Network Environment.

[31] Please see: https://www.zabbix.com/
[32] Please see: https://grafana.com/

	• Connection of systems and preliminary communication test with proprietary SW tool (similar to iPerf) that allows measuring E2E testing. • Start of operation: Drone take off & delivery of C2 signals. • Monitoring, measuring & recording C2 signal to/from Raspberry PI from/to Operator PC

5.3.4 Test case 4 – Range

Test Case	**#4**
	Range
Slogan & Objective	Testing of operational range in 5G Network.
Test Scenario (Pre-conditions)	Drone operating at 5TONIC spot located more than 20 km from Aerotools´ Office where the pilot is controlling the drone.
Expected Results (Post-conditions)	Drone operator will manage Drone´s C2 & Video signals from the Aerotools´ Office while requesting AIFs deployed in the 5TONIC Testbed.
General Time Plan (Validation Campaigns)	Testing carried out throughout Q3-Q4 2023.
Test Sequence	• Deployment of drone and AIFs in 5G Network Environment. • Connection of systems to VPN, checking visibility. • Start of operation with Drone take off and image shooting. • Request of AIFs

The location of test sites for this Range Test Case is shown in Figure 26.

Figure 26 Location of test sites for the Range test

5.3.5 Test case 5 – AIF Precision

Test Case	**# 5**
	Main Average AIF Precision
Slogan & Objective	Testing of Anomaly Detection AIF Precision in detecting the defined elements of an image.
Test Scenario (Pre-conditions)	Anomaly Detection Function running and tested, using images provided by developed set-up (drone at 5TONIC testbed)
Expected Results (Post-conditions)	Mean Average Precision (mAP) with Intersection over Union of 0.5, not lower than 0,6
General Time Plan (Validation Campaigns)	Generation of a reference DATASET and testing carried out throughout Q3-Q4 2023.
Test Sequence	• Deployment of drone and AIF in 5G Network Environment. • Request of Anomaly Detection Function. • Running the reference DATASET. • Measure of results.

For the validation of this KPI, the developed Anomaly Detection Function requires a reference DATASET that is used as a benchmark to measure the results obtained in detecting the characterized elements that are of interest for this Use Case. The reference DATASET is generated using images generated by the drone.

5.4. Validation Results

Once the scenario and procedures for conducting tests in each of the test cases were defined, various work sessions were carried out to verify the operation of the integrated systems under the established conditions and to measure parameters that provide objective evidence of the degree of compliance with the KPIs.

In addition to a series of partial test sessions, complete validation sessions were held on the dates 14/09/23, 22/09/23, and 05/10/23.

5.4.1 Test case 1 results – Integration

The integration of all systems, as detailed in previous sections, has been successfully completed, and the entire system has consistently operated under the specified conditions since M29 of the project.

The functional diagram that explains the architecture of the entire system is the one defined for Use Case 3 during the project, that is depicted in Figure 27.

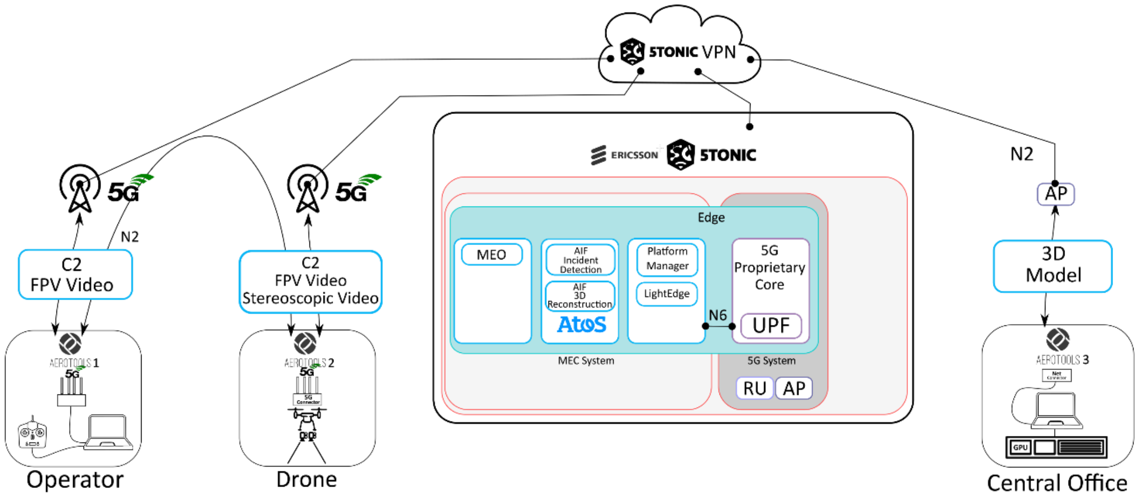


Figure 27 UC3 architecture diagram

Figure 28 shows the external 5TONIC facilities used for the drone's operations.

Figure 28 Drone operating at 5TONIC facilities

For the effective management of integrated systems, a Graphic User Interface (GUI) has been developed and fine-tuned to meet the specific requirements of this operation. A brief description of the included elements and functionalities is provided below.

GUI (Graphic user Interface)

All the communication from the user interface to the drone and to the AIFs is done through the Kubernetes cluster. The orders are sent from the interface to the MEO in the AI@EDGE Platform to open the docker, which is the AIF container, and it executes it with the parameters included in the descriptor´s file. Figure 29 shows the main page of the GUI.

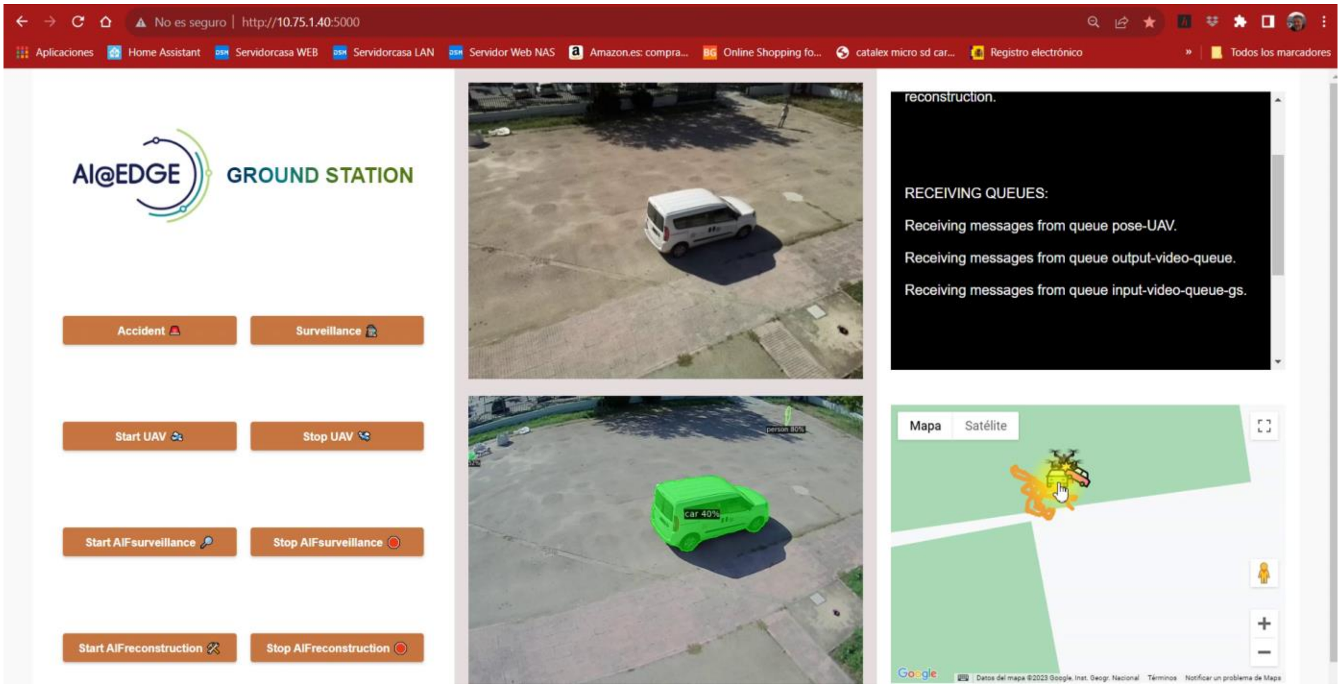
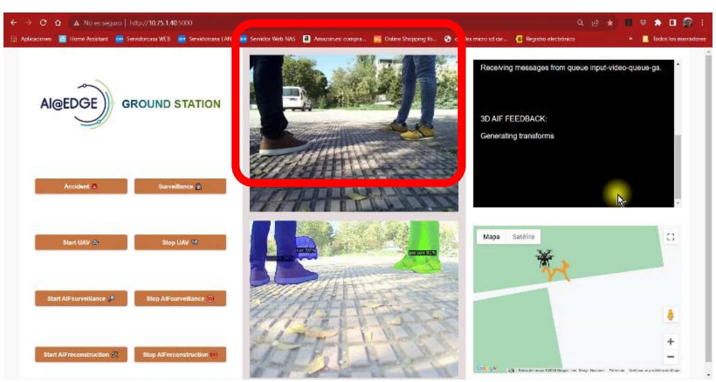


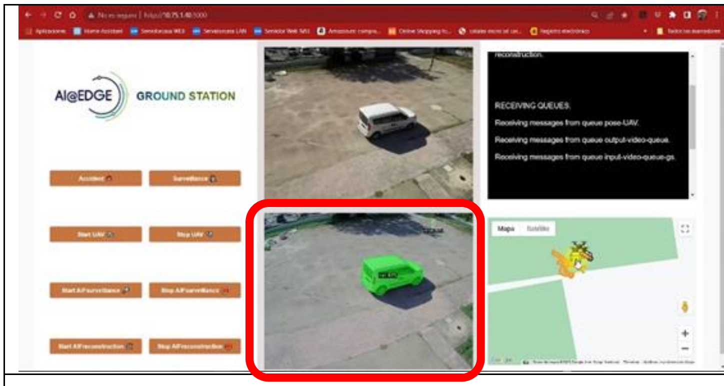
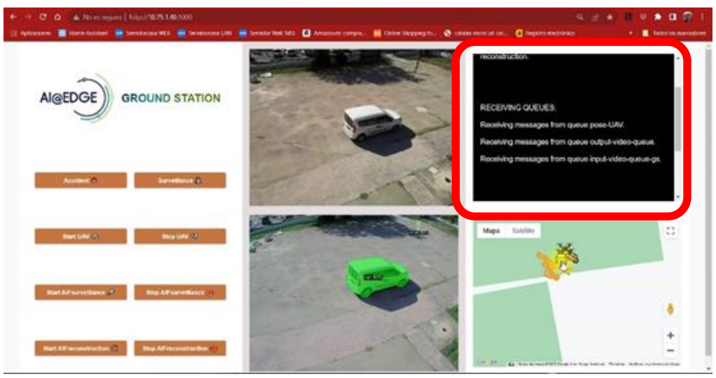
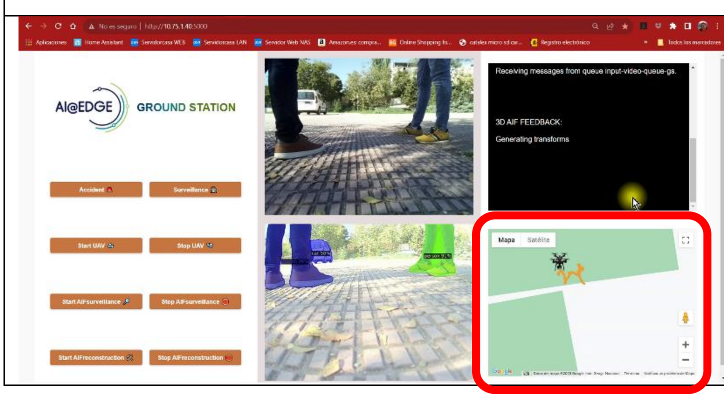
Figure 29 Main page of the AI@EDGE UC3 GUI

This web application provides the operator with a direct and simple resource to interact with all the components of the integrated system. The address is http://10.75.1.40:5000 at AI@EDGE VPN.

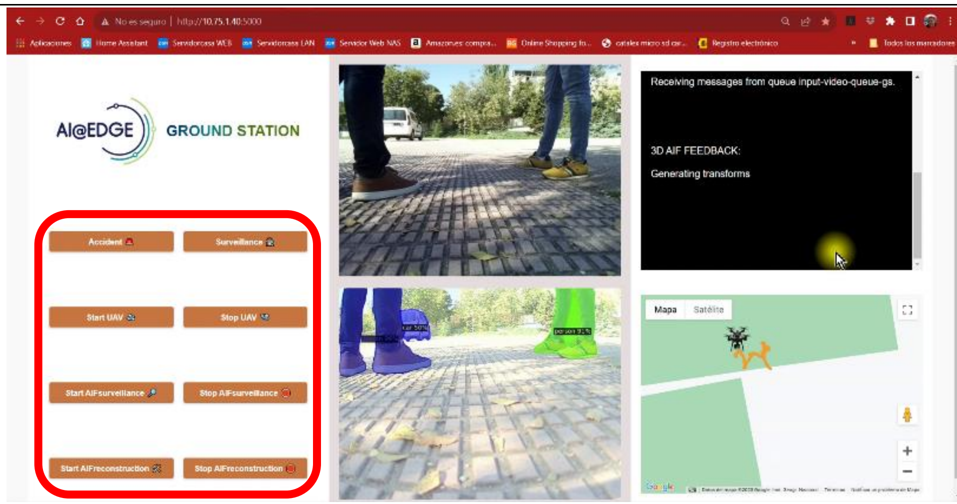
The Main page of the GUI encompasses the following elements:

Video screens

	**FPV images** These images are received in the GS through the *input-video-queue-gs* queue from the UAV. Subsequently, they are transmitted to the client browser (the Ground Station, or any supervising operator's computer) using Server-Sent Events (SSE).

 	**Annotated USB camera video from the AIF surveillance** These images are received in the GS via the *output-video-queue* from AIFreconstruction. Afterward, they are sent to the client browser through Server-Sent Events (SSE).
 	**Information Display** This section serves as a black box, emulating a terminal, and provides valuable information about the system's state. It includes details about: • **Errors** • Any errors within the system are prominently displayed in red. • **RabbitMQ queues** • This section displays both the queues that the Ground Station is awaiting information from (Waiting Queues) and the queues from which the Ground Station is receiving information (Receiving Queues). • **AIFreconstruction Feedback** • Information about the actions being performed by the AIF reconstruction is displayed here.
 	**Map** Shows the real-time position of the UAV on the map as well as the path it follows during its operation. The geoposition information is send by flight controller. Additionally, the map displays the location of accidents when they occur. The map functionality is powered by the [Google Maps API]().

Function Buttons



Start UAV	Initiates the UAV.py script on the UAV's on-board computer via an SSH GS-UAV connection. The UAV streams FPV video to the GS through the RabbitMQ queue *input-video-queue-gs* and directly to the pilot's computer via UDP sockets. In Surveillance mode, it streams USB camera video to AIFsurveillance through the *input-video-queue-AIF* queue. In Accident mode, it streams HD USB camera video to AIFreconstruction through the input-*video-queue-HD* queue. The UAV's pose is continually sent to the GS using the *pose-UAV queue*.
Stop UAV	Safely terminates the *UAV.py* script on the on-board computer.
Start AIFsurveillance	Deploys the Docker container for AIFsurveillance in the Kubernetes cluster. In Accident mode, it actively waits for images from the UAV. In Surveillance mode, it receives video sent by the UAV via the *input-video-queue-AIF queue*. The model generates panoptic segmentation masks, and annotated prediction images are streamed to the GS, along with accident notifications if detected, through the output-video-queue. Internally, the app connects to the cluster via SSH and executes the following command to deploy the Docker container: `**kubectl scale deployment uc3-aif-anomaly-detection-deployment -n surveillance --replicas=1**´.
Stop AIFsurveillance	Safely ends the execution of AIFsurveillance in the cluster by stopping the container. It executes the following command in the cluster: `**kubectl scale deployment uc3-aif-anomaly-detection-deployment -n surveillance --replicas=0**´.
Start AIFreconstruction	It deploys the Docker container of the AIFreconstruction in the Kubernetes' cluster. In **Surveillance** mode, it actively waits for the images from the UAV. In **Accident** mode, this AIF receives the HD video sent by the UAV via the *input-video-queue-HD* queue and computes the 3D reconstruction of the accident scene. During the reconstruction process, this AIF is sending feedback to the GS through the *feedback-AIF-reconstruction* queue, specifying which action it is performing then. If the reconstruction process fails, due to the appearance of too many blurry images or a bad on-flight record of the UAV, the error messages are also sent to

	the GS. The images are anyway sent to the GS to be downloaded and processed offline by the operators before the AIF starts to reconstruct. If the reconstruction succeeds, the AIF will send the 3D model to the GS, along with the generated camera path. The app is executing the following command to deploy the docker container: *kubectl scale deployment uc3-aif-3d-deployment -n 3d --replicas=1*.
Stop AIFreconstruction	It safely ends the AIFreconstruction execution in the cluster by stopping the container, executing the following command in the cluster: *kubectl scale deployment uc3-aif-anomaly-detection-deployment -n surveillance --replicas=0.*
Surveillance	This is the default working mode, so the system starts in this mode without the need of clicking the corresponding button. Each element, when active, is performing the following actions: • **UAV**: Sends 640x480 USB camera images to the AIFsurveillance RabbitMQ queue *input-video-queue-AIF*. It also streams FPV video to the pilot and the GS, along with the pose to the GS. • **AIF surveillance**: Receives UAV video, performs predictions, and sends annotated images to the GS. • **AIF reconstruction.** Actively waits for UAV video. • **Ground Station**: Displays the UAV's position on the map, FPV images, annotated predictions from AIFsurveillance, and terminal information.
Accident	Activated by the pilot when an accident occurs. It switches to Accident mode, which affects each element of the system as follows: • **UAV**: Stops sending 640x480 USB camera images to AIFsurveillance and starts sending 1280x960 images to the AIF reconstruction queue *input-video-queue-HD*. • **AIF surveillance**: Stops receiving images and, therefore, stops sending predictions to the Ground Station. • **AIF reconstruction:** Starts saving images, that will be used for the 3D reconstruction once the Accident mode stops (when the **Surveillance** button is clicked). • **Ground Station**: Displays the accident location on the map. Moreover, it stops showing the images of the AIF surveillance since they are not being generated. In the black terminal, it displays the feedback received by the AIF reconstruction in the RabbitMQ queue *feedback-AIF-reconstruction.*

Data Section

This subpage of the application (Figure 30) provides operators an efficient means to visualize, generate, and download relevant data related to the system's operation. The address is http://10.75.1.40:5000/data

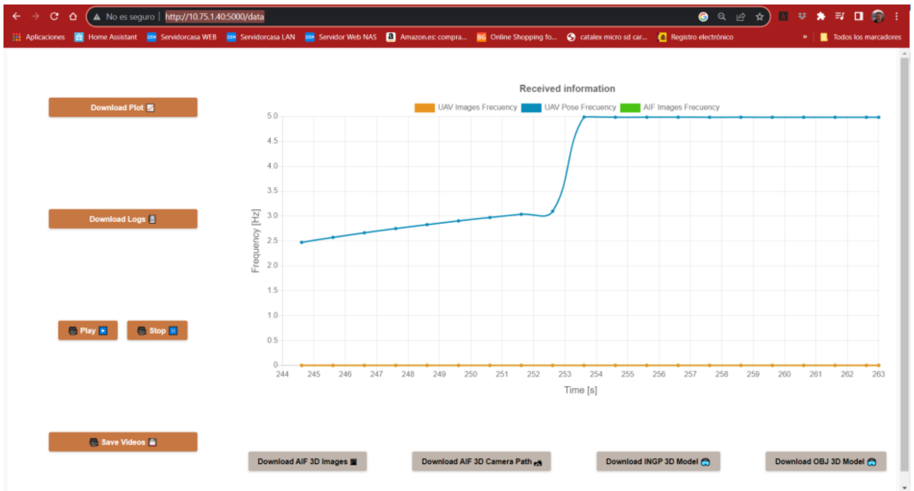


Figure 30 Data Section page

It includes the following elements:

Graphical display

Displays the real-time frequency at which streams arrive at the Ground Station server. These streams are the FPV images from the UAV, the UAV's pose, and images with annotated predictions from the AIF surveillance.

Buttons

Download Plot	Allows users to download a snapshot of the graph.
Download Logs	Downloads a text file containing stream frequency information since the launch of the Ground Station.
Play	Initiates the recording of available streams displayed in the Ground Station.
Stop	Halts the recording process.
Save Videos	Downloads any available videos in the Ground Station.
Download AIF 3D Images	Downloads the images used by the Nerf model to generate the 3D reconstruction.
Download AIF 3D Camera Path	Downloads the camera path computed from the images using Colmap.
Download INGP 3D model	Downloads the trained Nerf model checkpoint, which can be used to visualize the 3D reconstruction using the instant-ngp software.
Download OBJ 3D Model	Downloads the 3D model in OBJ format.

5.4.2 Test case 2 results – Latency

As previously detailed in the document, latency measurement is conducted in both the drone's command and control section and separately during the transfer of images generated by the drone. The subsequent paragraphs elaborate on the results obtained from these tests.

Command & Control Communication

To conduct performance measurement tests of the C2 signal between the drone and the operator, the drone is connected to the 5G network of 5TONIC, and signal control points are established between the devices indicated in Figure 31.

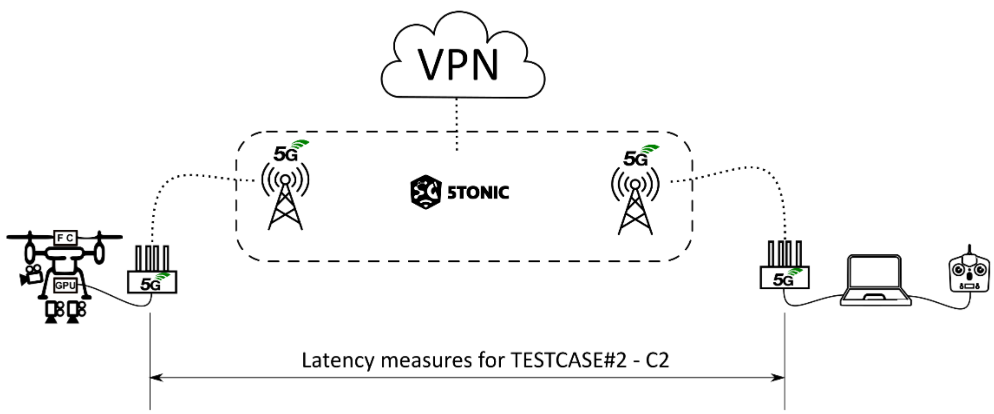


Figure 31 Latency measurements for Control and Command

At 5TONIC, Ericsson has developed a Key Performance Indicator (KPI) framework designed to efficiently gather and visualize metrics related to the utilization of the 5G System. This framework relies on a sophisticated software, named “probe”, a component adept at extracting metrics from end-user traffic with flow granularity for IP traffic. A flow is uniquely identified by a tuple consisting of the origin IP address, destination IP address, origin port, destination port, and type of protocol. This meticulous approach enables the extraction of KPIs specifically tied to application flows, providing valuable insights into the performance of individual applications within the 5G System.

The software probe is installed in the Raspberry Pi onboard the drone, specifically configured to capture application traffic generated during drone operations on the 5G system network interface. This software probe can generate key metrics, including 'TCP Round-Trip and Throughput,' for both uplink and downlink. Additionally, metrics such as Jitter can be derived from the real-time database, housed in the 5TONIC Data Center, which is populated with data exposed by the software probe. The visualization of this data is facilitated through the Grafana application, an open-source solution designed for network monitoring, providing graphical representations of the stored metrics.

The methodology for C2 Latency measurement revolves around the Round-Trip Time (RTT) of a PING, involving the sending of a PING from one end of the network and recording the complete duration for the request to traverse the network and the corresponding response to return. In this instance, a PING is initiated from the Raspberry Pi device integrated into the drone, with the opposite end situated at the PC of the drone operator.

RTT is recorded at varying intervals, ranging from 1 second to broader periods, to compile a comprehensive database of measurements, serving as evidential data. Grafana facilitates a graphical representation of the performance, illustrated in Figure 32, showcasing the RTT measurements at the Raspberry Pi. Additionally, the Jitter factor, a parameter reflective of deviation from the RTT value, is included to depict the stability of this metric.

Figure 32 RTT performance

The RTT value consistently remains below 100 ms, aligning with the definition of RTT as double the distance considered for Latency. This compliance with the 50 ms value serves as a confirmation of meeting the Key Performance Indicator (KPI) set for this parameter in the UC3.

To provide additional validation and cross-reference the results obtained through the previously explained method, several test sets were carried out with 'direct communication,' meaning that the control station and the drone were connected via Ethernet cable, while simultaneously measuring the time difference between wired connectivity and connectivity using 5G technology. This confirmation test reinforced the previous results, confirming that the latency remains below 50 ms. In the graph below (Figure 33), the latency generated when utilizing a 5G network is depicted in comparison to the scenario where the operator's PC is directly connected to the drone controller. The graph illustrates consistent values below the threshold established by the Key Performance Indicator (KPI).

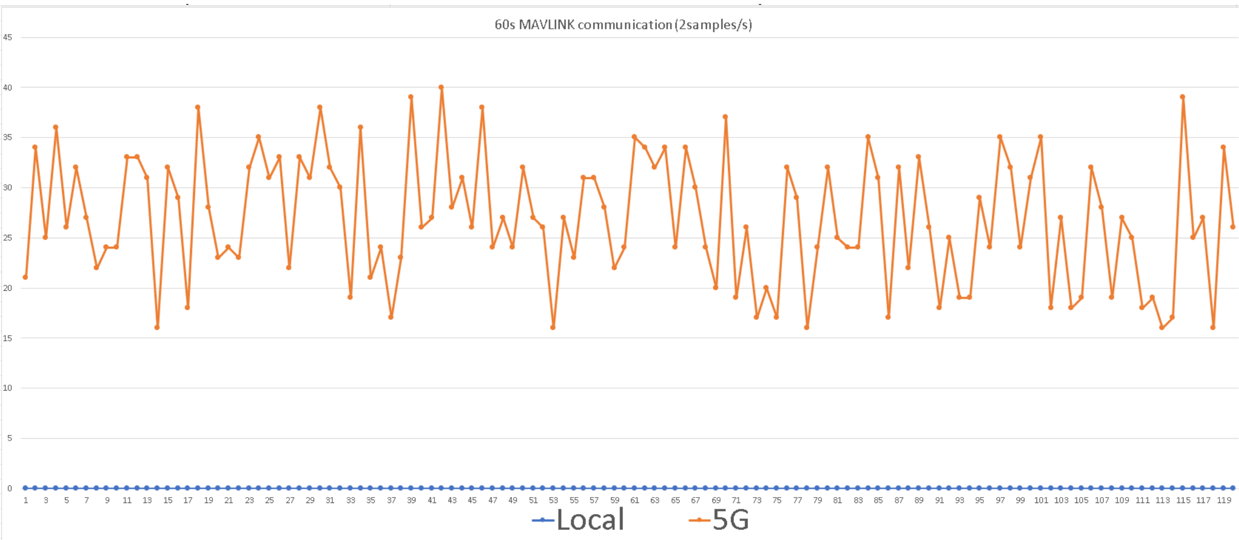


Figure 33 Latency introduced by the system in control and command communication

Finally, flight tests conducted with the drone connected to the operator through the 5G network also confirmed the feasibility of conducting drone flight control over this 5G network infrastructure, as demonstrated in the video footage.

Video Communication

To monitor the system's video performance, a similar procedure is followed as explained in the previous paragraph regarding C2 communication. The testbed configuration for this process is depicted in Figure 34.

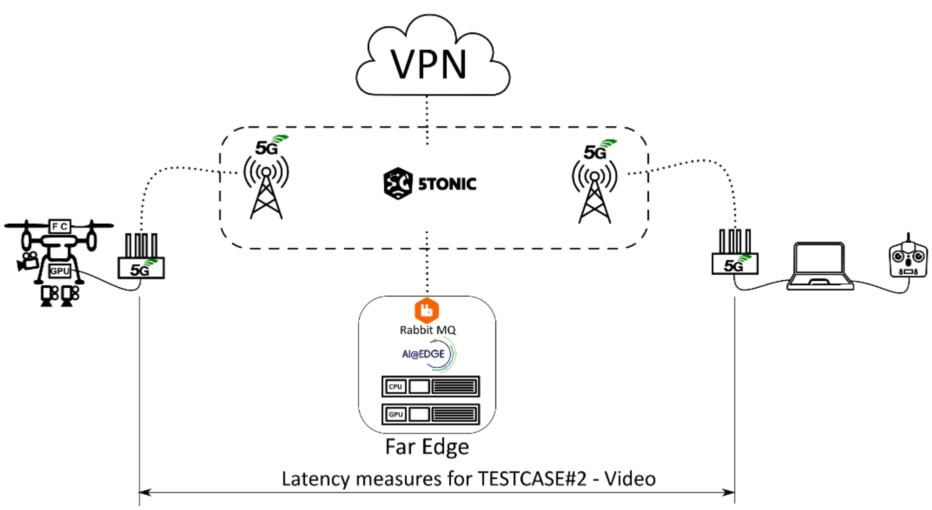


Figure 34 Configuration for measuring Latency in video communication

In this scenario, video images originate from two sources:

- Stereoscopic cameras generate video for the AIFs, which is transmitted by the Raspberry Pi to the RabbitMQ server deployed in the cloud.
- FPV streaming, sent by the Raspberry Pi to the drone operator's PC.

The latency of the video signal is monitored at the RabbitMQ server for the AIFs, and for the FPV video, the control point is established within the flying control software deployed on the operator's PC, known as Mission Planner.

The procedure for measuring this parameter relies on the Round-Trip Time (RTT) of a PING sent from one end of the network, recording the total time it takes for the request to traverse the network and for the response to return. In this case, a PING is initiated from the Raspberry Pi device onboard the drone, with the opposite end directed towards both the RabbitMQ server and the drone operator's PC.

(**Note:** It should be noted that this procedure enables measurement while other functions are concurrently running through the network. However, it holds the lowest hierarchical position, and additional latency could be introduced in specific conditions. Therefore, the measurements can be regarded as conservative.)

RTT is recorded at intervals, ranging from 1 second to broader periods, to generate a database with measurements stored as evidence. Grafana graphics offer a graphical representation of the performance, as depicted in Figure 35, where RTT is measured at the RabbitMQ (along with the Jitter factor, a parameter representing deviation from the RTT value) for the video signal used in AIFs."

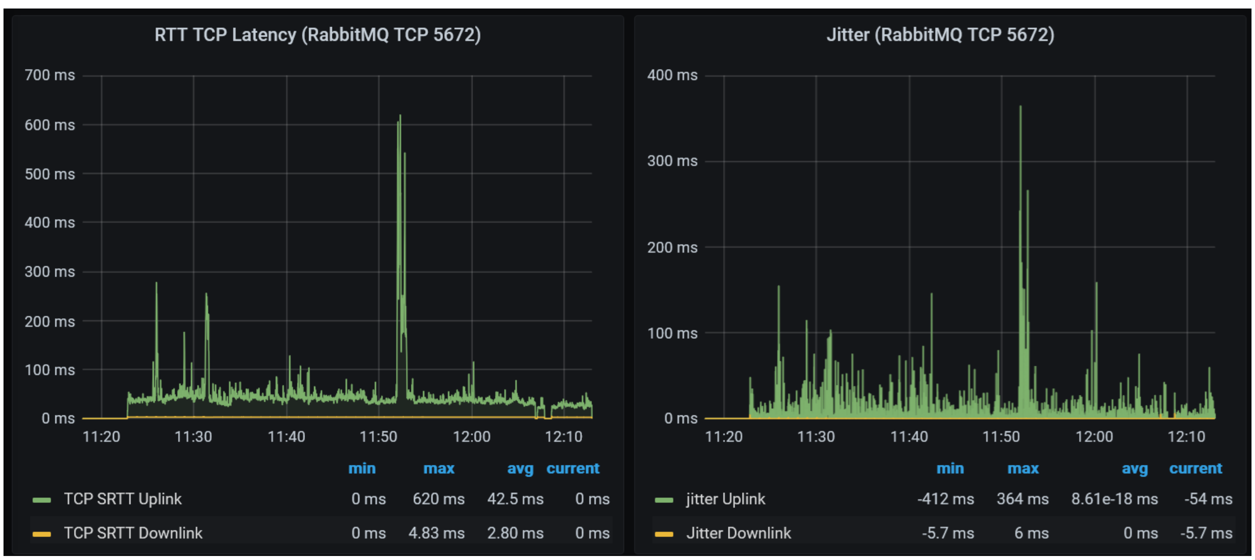


Figure 35 RTT and Jitter video signal AIF

And measured at the Mission Planner operator's PC (along with the Jitter factor) for the video signal used in FPV, as shown in Figure 36.

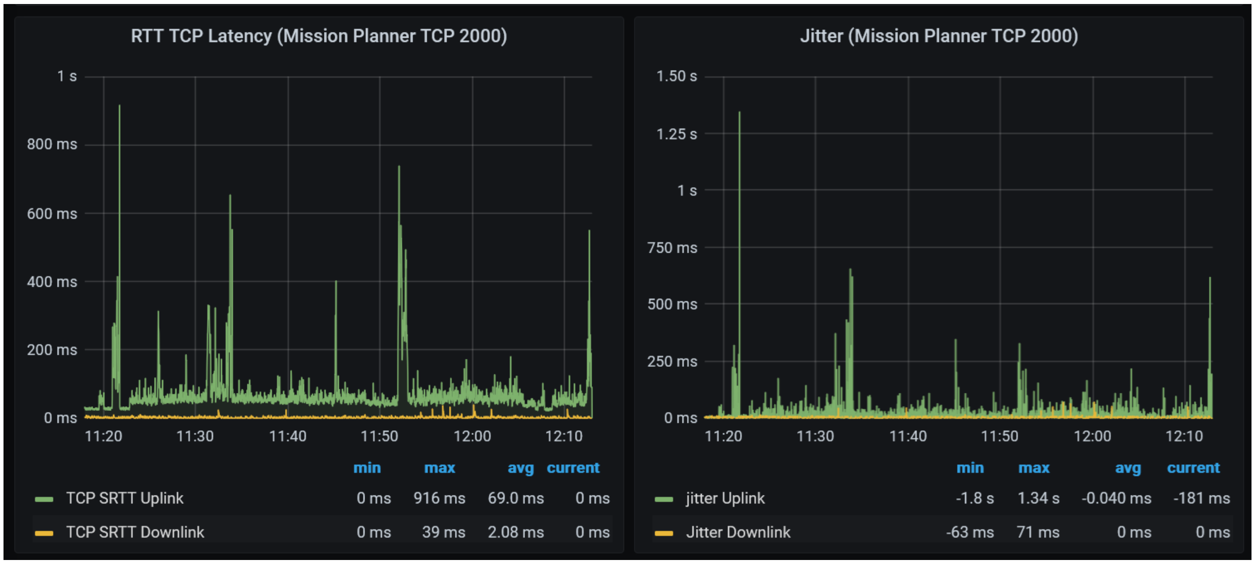


Figure 36 RTT and Jitter video signal FPV

The RTT consistently remains below 200 ms, in accordance with the RTT definition as double the distance considered for latency. This compliance with the 100 ms threshold aligns with the Key Performance Indicator (KPI) set for this parameter in the UC3.

5.4.2 Test case 3 results – Reliability

The system's performance in "packet loss" is documented throughout the trials conducted at 5TONIC, as depicted in Figure 37.

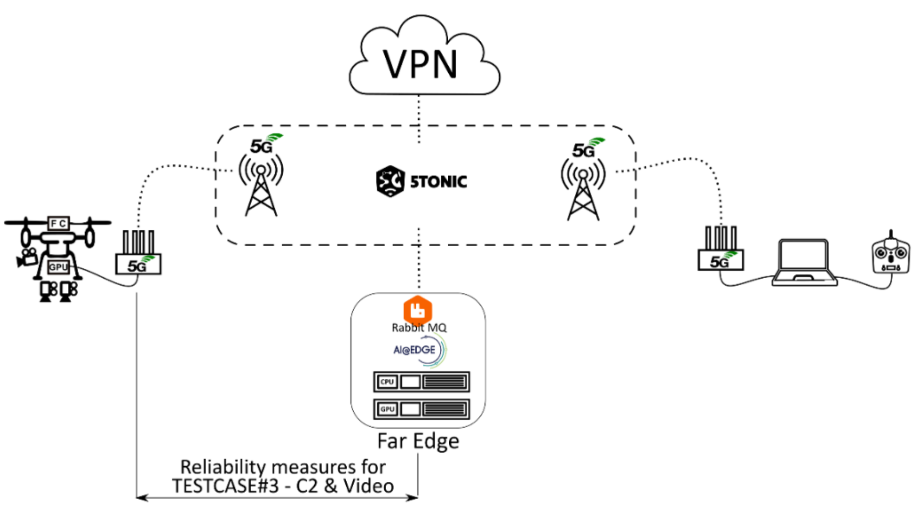


Figure 37 Packet Loss

To assess the system's performance regarding data packet loss, the software Iperf is employed to conduct tests simulating the required data flow for the AIFs Video signal (20 Mbps). The tests involve sending a data flow from the Raspberry Pi for 60 seconds and monitoring the reception of packets at the other end in the RabbitMQ server, where the video is managed. Figure 38 depicts the process of sending information and the results after monitoring, with the relevant information highlighted at each end.

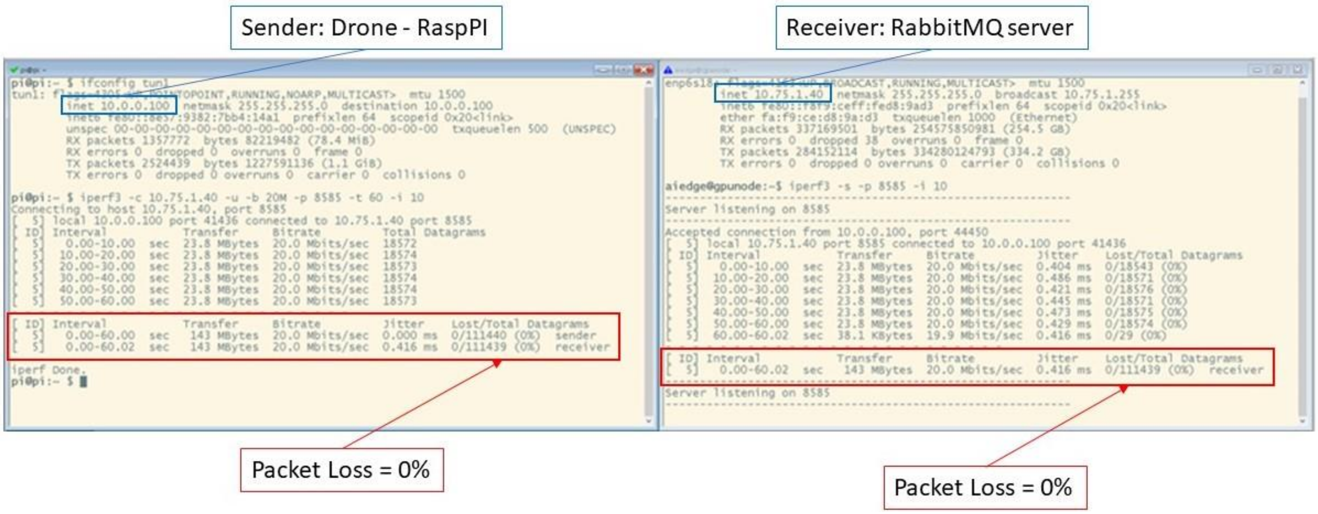


Figure 38 Sending information and monitoring results

In the conducted tests, it was observed that reliability remains uncompromised when utilizing a 5G connection. Even in simulated scenarios intentionally introducing latency, connections remain stable with no dropped connections and no lost packets. In the absence of packet loss, it becomes the responsibility of the applications to uphold open sessions for each communication. The packet loss tests performed for the project were also replicated with scenarios simulating very high latencies, consistently yielding a 0% packet loss rate.

5.4.3 Test case 4 results – Range

A critical factor for the BVLOS mode is the distance at which the drone can be controlled by the operator. In the project, a significant distance of 20 km was established as KPI, a parameter that can introduce latency or other performance-degrading factors into the system. The tests conducted to showcase the requisite performance were tailored to the deployment of the 5G network and resources at 5TONIC.

Figure 39 illustrates the adopted configuration and the locations of the two spots at each end:

- The drone is flying and operating at the 5TONIC laboratory, situated in the IMDEA building in Leganes, southwest of the central area of Madrid.
- The drone operator is located at the AEROTOOLS office in Alcobendas, north of the central area of Madrid.
- The direct line distance from one spot to the other is more than 20 km.

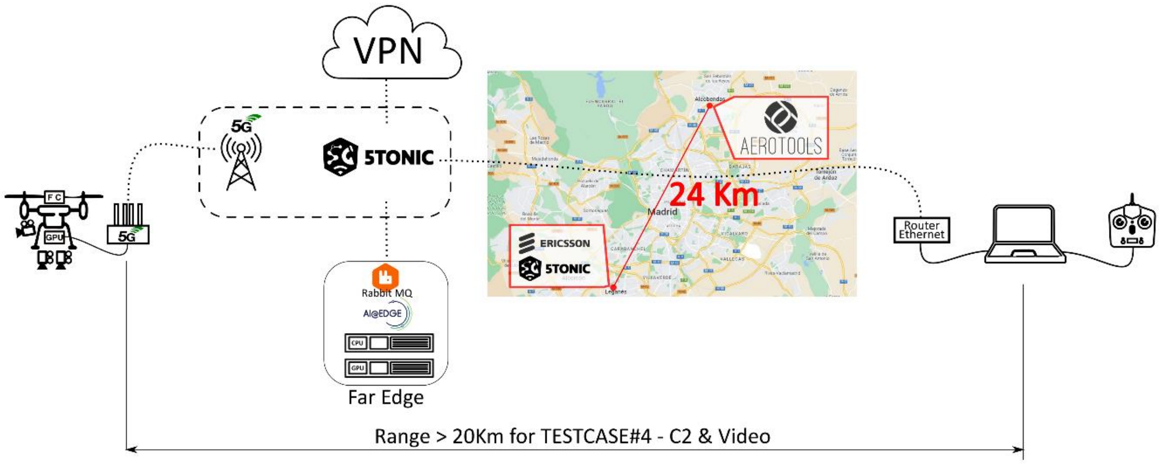


Figure 39 BLVOS configuration

The connection to the 5TONIC VPN is established through an Ethernet connection on the drone operator's side, while the drone is connected to the 5G network at 5TONIC. A BVLOS mode flight and drone operation was executed on October 5th, and the session was video recorded.

Figure 40 present views from the operator's side as well as a third person view of both the drone and the operator. In addition, Figure 41 shows the GUI while controlling the drone in BVLOS mode.

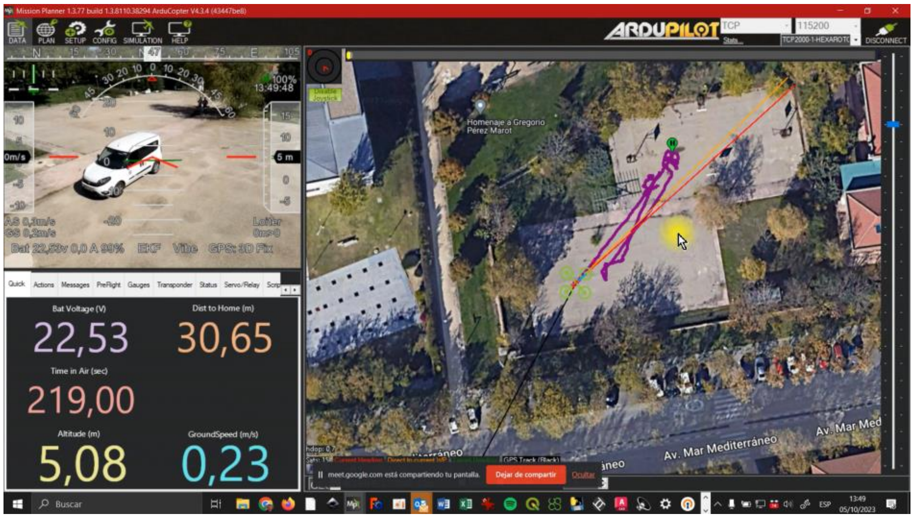


Figure 40 Different views of drone and operator

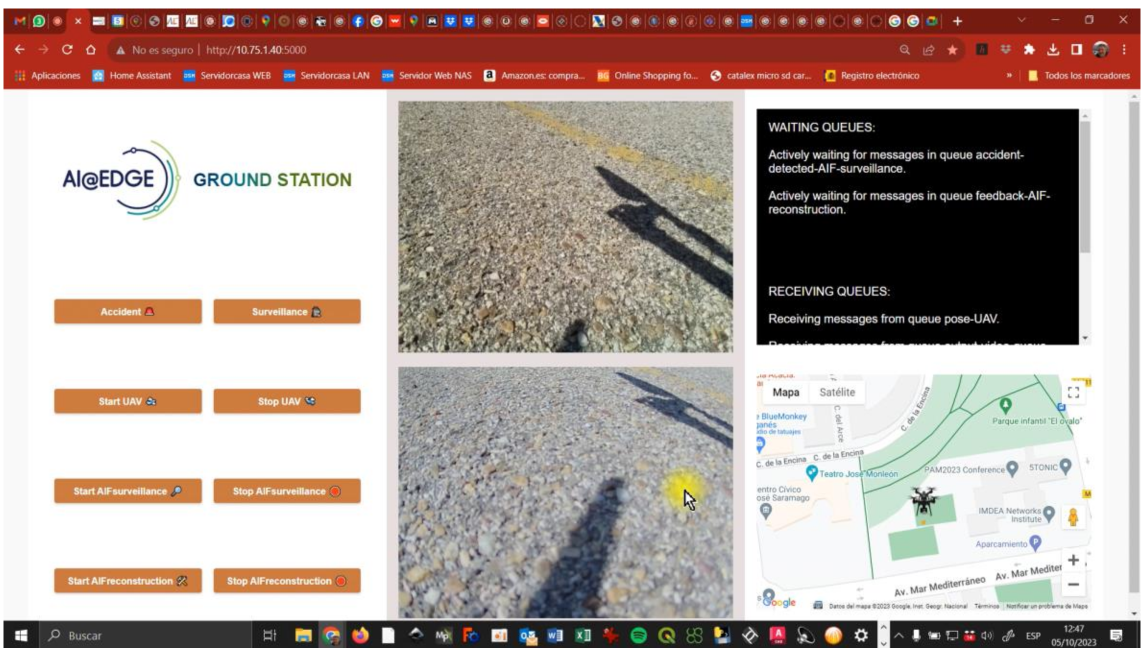


Figure 41 View of GUI controlling drone in BVLOS mode

5.4.5 Test case 5 results – AIF Precision

The performance of the Anomaly Detection AIF (AD AIF) is assessed in the test case 5 through an analysis of its capabilities in the domain of object detection. Our evaluation is based on an aerial dataset specifically generated within the testbed scenario and for relevant use-case situations.

To create this dataset, several dedicated flying sessions have been conducted for aerial surveillance above nearby roadway featuring vehicular traffic and occasional pedestrian activity, capturing on-board video footage. Subsequently, a process to select randomly up to 170 frames from the video is performed, and the frames are manually annotated according to the "car" and "person" categories.

During the project, adapting configuration of the DETIC model [3] within the AD AIF was tailored to facilitate predictions for the two labelled classes. While DETIC demonstrates proficiency in panoptic segmentation, our focus remains solely on its object detection capabilities, deemed satisfactory and pertinent for our specific use case. A summary of the model's performance standard metrics on the resulting dataset can be found in Table 10.

Table 10 DETIC model performance

	Instances	Precision	Recall	AP50	AP50-95
all	684	0.939	0.896	0.946	0.657
car	634	0.988	0.892	0.972	0.72
person	50	0.89	0.9	0.919	0.595

In this table:

- The instances column refers to the number of objects of each class present in the dataset. Precision, recall, and Average Precision (AP) are fundamental evaluation metrics in object detection tasks that gauge the performance of detection models.
- Precision refers to the ratio of correctly identified positive instances to the total instances predicted as positive by the model. It measures the accuracy of the model's positive predictions.
- Recall, on the other hand, represents the ratio of correctly identified positive instances to the total actual positive instances present in the dataset, indicating the model's ability to detect all relevant instances.
- Average Precision (AP) is a comprehensive metric that considers precision-recall pairs across various Intersection Over Unions (IoUs[33]). AP50 refers to the average precision computed for a single IoU threshold of 50%, while AP50:95 averages the AP over 10 IoUs, from 50% to 95%, with an increment of 5%. The mean AP (mAP) is a metric used to measure the accuracy of object detectors over all classes in a specific database. The mAP is simply the average AP over all classes, that is:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i$$

[33] Please see: https://giou.stanford.edu/

with AP_i being the AP in the ith class and N is the total number of classes being evaluated. Typically, higher precision and recall values signify better model performance, while AP provides a comprehensive overview by considering precision across recall levels. In our case, $N = 2$ for the 'car' and 'person' classes.

Within this evaluation framework, the AD AIF Key Performance Indicator (KPI) was established as the Mean Average Precision (mAP) with an Intersection over Union (IoU) threshold of 50%, and it was required to exceed a threshold of 0.6. Precision and recall are both computed considering an IoU of 50% and establishing the confidence threshold of the model -detections with probability above this threshold are considered, otherwise not- to 44%. This parameter was experimentally set based on the best F1 Score in numerous experiments. Note that the obtained results significantly surpass the performance benchmark, as evidenced by the highlighted first value of the AP50 column in the former table, referring to the mAP in our evaluation dataset. The performance of the model diminishes when we consider more IoUs thresholds, as presented in the AP50-95 column of Table 10. Nevertheless, within our use case, this outcome holds secondary significance, as the primary objective lies in detection rather than precise localization of individuals and vehicles. An example prediction of the model with low IoU is presented in the bottom-right image of Figure 42.

The disparity between the desired Key Performance Indicators (KPIs) and the obtained results can be attributed to various factors. These include: i) rapid advancements in state-of-the-art object detection, ii) the dataset's simplicity, which confines detection to two specific classes deemed sufficient and purposeful for the intended use case, and iii) the capture of UAV onboard images under favourable daylight conditions at a moderate altitude, facilitating object recognition within the images.

Figure 42 illustrates the performance of the model -with the confidence threshold set to 44%- over the dataset.

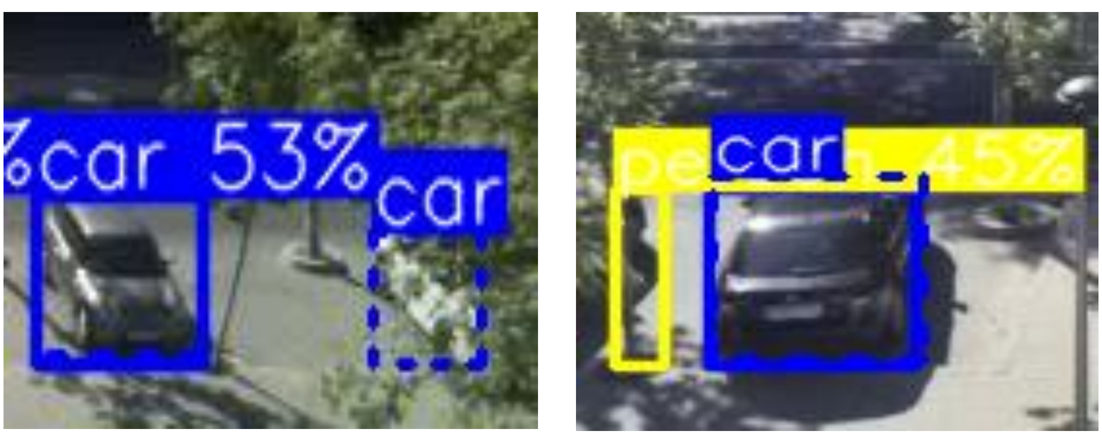


Figure 42 Illustration of the DETIC model performance over images of the dataset

The dashed boxes correspond to annotations and the solid ones to predictions. Top: Erroneous predictions. The top-left one shows a false negative occluded car (blue dashed box), and the top-right includes a false positive person (yellow box). Bottom: Correct predictions. The bottom-left image displays a full image of the dataset with all the objects properly detected, and the bottom-right one exemplifies a correct prediction with low IoU. And finally,

Figure 43 shows the DETIC model performance over images of the dataset in the way are represented in the graphic user interface (with segmentation masks).

Figure 43 Illustration of the DETIC model performance over images of the dataset

5.5. Final remarks

AI@EDGE represents a significant contribution to the advancement of BVLOS mode operations with drones in the industrial sector, spanning various applications such as infrastructure monitoring, surveillance of large areas, and inspection of substantial industrial assets. Ensuring a reliable network connection for communication is paramount for maintaining consistent operations in this mode.

Beyond BVLOS, AI@EDGE introduces powerful tools to enhance drone performance and expand their utility. The ability to access a repository of on-demand AI functions, supported by computational resources, alleviates the need to burden drones with additional devices and systems, paving the way for innovative applications.

In the realm of KPIs, achieving consistent performance of the assisted 5G network in various domains has been a primary concern. While the operational range is directly tied to network deployment, a well-defined and proven configuration to address different scenarios or available tools is essential. Moreover, ensuring low latency and high reliability is crucial, given the increasing demand for these parameters in drone operations.

More specifically, the obtained results demonstrate:

- Consolidated operation of drone systems in the 5G Network, including:
- Command and control (C2) and First Person View (FPV) video signals for drone flight.
- Payload video transfer.
- Developed working flows and interfaces for system management and decision-making.
- Access to Artificial Intelligence Functions (AIFs) repository.

Based on the obtained results, it can be affirmed that the UC3 has validated the operation of drone's operations within an AI@EDGE enhanced 5G network, incorporating AI and Edge Computing functionalities for automated monitoring of road infrastructures in BVLOS mode.

The major outcomes arising from this validation include the integration of systems to enable the flight and operation of drones in a 5G network; the development of AI functions to automate monitoring operations and the deployment of the AI@EDGE platform providing access to the Edge repository of AI functions.

These results are supported by AI@EDGE technical enablers like distributed and decentralized serverless connect-compute platform and AI-enabling application provisioning, as well as relevant technologies such as AI Functions for automating incident detection and generating 3D models; Edge Computing, by deployment of MEC system for accessing the AI Functions repository in drone operations; 5G for connection to 5G network for drone control and imagery transfer, and drones for the integration of system enabling BVLOS operation with advanced functions.

6. Use Case 4: Smart content & data curation for in-flight entertainment and connectivity services

The UC4 of AI@EDGE project focuses on the development of a test bed and presents initial experimental results aimed at providing broadband connectivity to passengers' on-board aircraft, as a step towards achieving ubiquitous access. SPI provides a detailed account of its research and experimentation conducted within the framework of the AI@EDGE research project, which encompasses a 5G network and an edge-cloud infrastructure built using both avionic-certified and off-the-shelf hardware. The edge-cloud serves as a platform for developing and testing AIFs and other MEC applications, which represent the next generation of services offered to airlines and their passengers, relying on machine learning capabilities. Moreover, the 5G network is seamlessly integrated into the SPI test-bed environment and connected to a ground-based 5G core network via a Low Earth Orbit (LEO) satellite backhaul, such as Starlink[34].

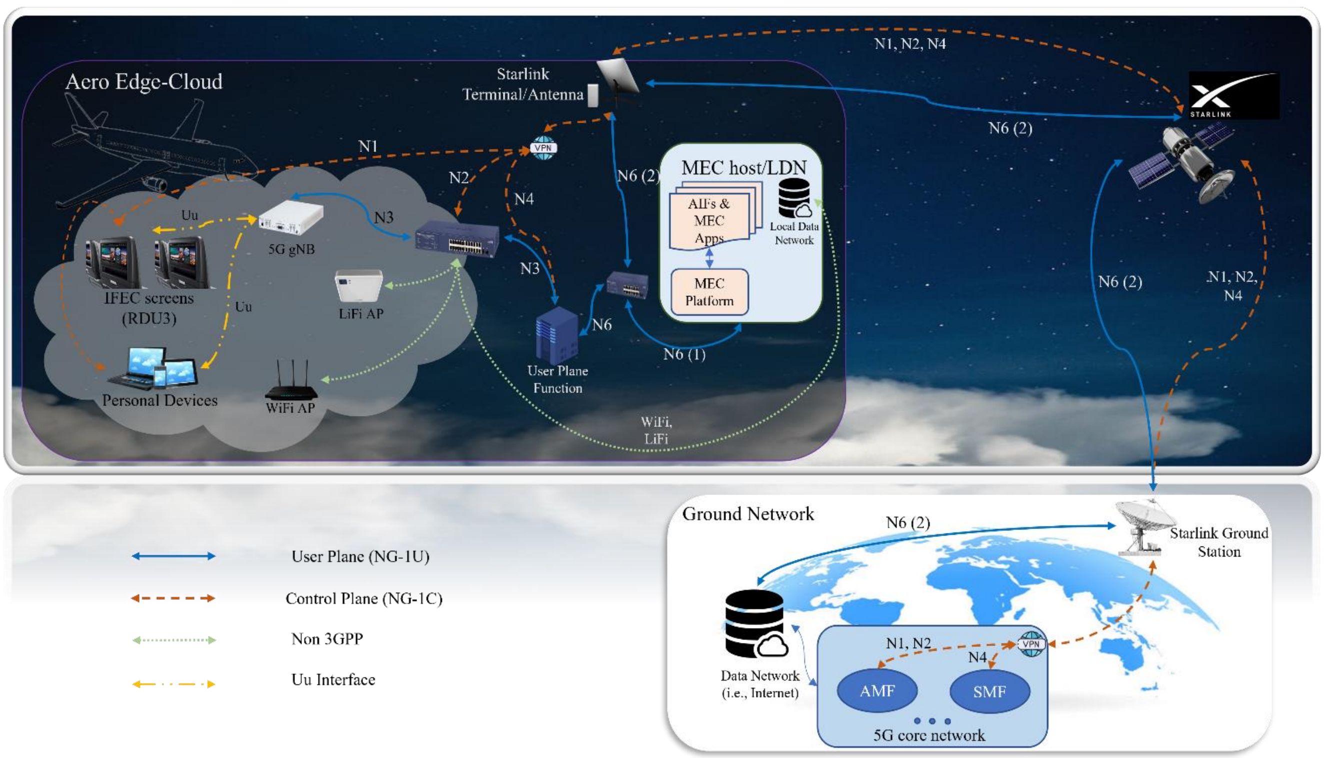


Figure 44 Overall architecture of UC4

[34] Please see: https://www.starlink.com/

Figure 44 presents the architecture of the proposed Aero Edge-Cloud and connectivity system that is implemented in the SPI test-rack designed for the UC4 of AI@EDGE. The figure offers a visual representation of the 5G system architecture that extends from the edge site of an aeronautical network to the ground infrastructure, facilitated by LEO satellite connectivity. The network is divided into three primary components:

- Aero Edge-Cloud: This component encompasses the on-board network, including the Radio Access Network (RAN), Multi-access Edge Computing (MEC) host, and the Local Data Network (LDN). It also includes the existing In-Flight Entertainment and Connectivity (IFEC) hardware, such as IFEC screens or Removable Display Units (RDUs), along with the servers.
- Satellite Backhaul: The satellite backhaul is established using the Starlink Low Earth Orbit (LEO) constellation, enabling communication between the aircraft and the ground network.
- Ground Network: The ground network comprises the 5G core network and the central data network, which refers to the internet.

By leveraging this architecture, the proposed system enables seamless connectivity and data transfer between the on-board network, the satellite backhauls, and the ground network, ensuring efficient communication and services for the aircraft and its passengers.

Within the framework of AI@EDGE project, UC4 serves as the far edge component situated at the radio access site. Within this setup, SPI hosts the user plane function (UPF) of the 5G core on its premises, while the control plane is located on the ground at FBK site. This configuration enables the aircraft to function as an edge-cloud entity, connected to the ground network through LEO satellites.

6.1. Validation objectives

Following what reported in deliverable D5.1, the focus of SPI is mostly on i) Integration, validation, test plans, and ii) Planned demonstrations and logistic, as reported in Table 11.

Table 11 SPI time plan by M36

	Integration, validation, test plans	**Planned demonstrations and logistic**
(To be completed by Q2 2023)	1. Complete integration of AI@EDGE platform components, 2. Advanced integration of the content curation AIF, 3. Advanced integration of the video transcoding, 4. Advanced integration of MP-TCP proxy AIF 5. Advanced integration of hardware acceleration.	1. Demonstration of 5GC and 5G NSA RAN, 2. Initial demonstration of 5GC and 5G SA RAN, 3. Initial demonstration of content curation AIF, 4. Initial demonstration of video transcoding, 5. 5. Initial demonstration of MP-TCP proxy AIF,

(To be completed in Q4 2023)		1. Advanced demonstration of 5GC and 5G RAN, 2. Advanced demonstration of content curation AIF, 3. Advanced demonstration of video transcoding, 4. Advanced demonstration of MP-TCP proxy AIF, 5. Demonstration of UC4 test bed integrated AIFs, performance benchmark and collection of relevant KPIs, 6. 6. Live demonstration or video recording of complete UC4 functionalities

6.2. Validation scenario

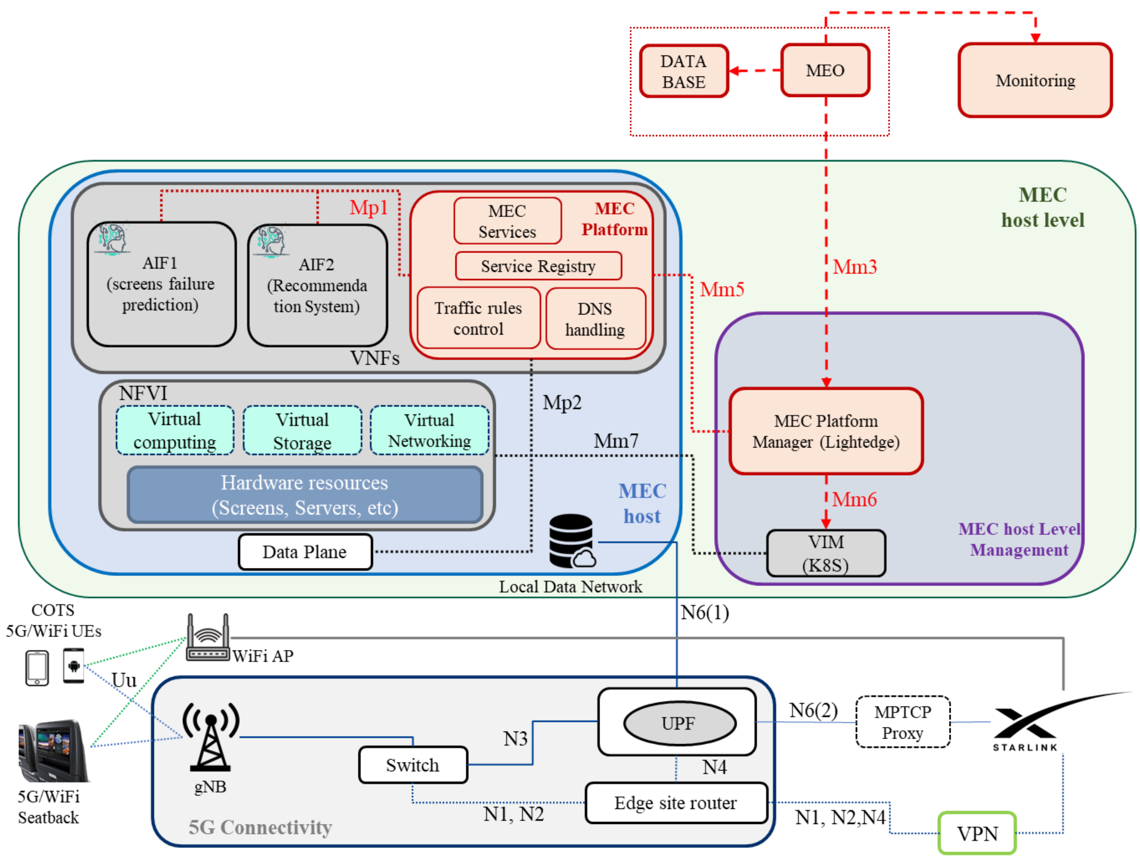


Figure 45 Detailed architecture of UC4

Figure 45 illustrates that Control Plane (CP) traffic is consistently directed to the ground-based 5G core network via the satellite backhaul, while the User Plane Function (UPF) is deployed on-board the aircraft. The Access and Mobility Management Function (AMF) and Session Management Function (SMF) are accessed by establishing the 3GPP-defined N1, N2, and N4 interfaces for user authentication and session management, as outlined in Figure 44 and Figure 45.

User Plane (UP) traffic can be routed either to a ground network or remain on-board the aircraft to access Local Data Network (LDN), such as in-flight entertainment media. Consequently, users are categorized into two groups: (i) IFEC users who access on-board content and (ii) regular users who primarily require internet access. This differentiation is achieved by employing distinct Data Network Names (DNNs) that guide users to the respective Data Network (DN). In practice, after the attach procedure, the 5G core assigns two IP address pools for users, and the UPF establishes two Protocol Data Unit (PDU) sessions for each pool of IPs, as depicted by N6(1) and N6(2) in Figure 45.

Similar to CP signals, the N6(2) traffic traverses the satellite connection, but with a major difference: while CP signals (i.e., N1, N2, and N4 interfaces) are routed to the corresponding Athonet 5G core located in FBK, (on the ground via ZiroTier Virtual Private Network (VPN)), traffic over N6(2) is directly sent to the ground to reach the internet. Within the system, the N6(1) interface establishes the PDU session to enable communication between wireless IFEC devices and on-board servers, particularly at the MEC host level and the LDN.

As depicted in Figure 45, the MEC host comprises the MEC platform integrated with the NFVI, which provides storage, compute, and network resources to execute ML-based MEC applications referred to AIFs within AI@EDGE. Placing the MEC host at the edge enhances resource accessibility in close proximity to on-board users. The test case 1 (Sec. 6.3.1) explores two AIFs specifically designed for the Aero Edge-Cloud network, running within containers on the virtual infrastructure managed by the Virtualized Infrastructure Manager (VIM).

Other components indicated by the red boxes in Figure 45, including the MEC Orchestration (MEO), MEC platform manager, monitoring, and database, are part of the AI@EDGE Network and Service Automation Platform (NSAP) and Connect-Compute Platform (CCP). These components are also integrated with the Aero Edge-Cloud, but their detailed specifics are discussed in Section 2 and here they are solely depicted in the figure for the sake of completeness and clarity as part of the AI@EDGE platform.

6.3. Validation procedures

6.3.1 Test case 1: AIFs development

6.3.1.1 Popularity- and Item-based recommendation system

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)	• Completing and advancing the application Q2 2023 • Full Integration of the application in the test-rack with AI@EDGE platform Q3 2023 • Final demonstration Q4 2023
Test Sequence	1. Integration of the application with the aero edge/cloud infrastructure. 2. Deployment of the application in the far-edge, employing AI@EDGE platform

Popularity-based

The popularity-based model is a straightforward approach used to recommend movies based on their popularity. In this model, the popularity of a movie is determined by three factors: the watching ratio obtained from PAX (passengers) logs, the IMDB[35] ratings, and the movie's release date. By combining these components, a popularity score is calculated for each movie, and the movies are sorted based on this score. The system then recommends movies with the highest popularity score to users.

As for audio recommendations, our focus has been on developing a model that allows generating recommendations without the need for retraining. This model is saved in a dump file for easy access and usage.

The following tasks are undertaken to build the movie recommendation system:

- Collect internal and external data and preprocess them to create a dataset containing the necessary features for building the popularity-based model.
- Ensure that there is no significant correlation between the features created, ensuring independence and avoiding biases in the recommendations.
- Generate recommendations tailored to different airlines and routes, taking into account the preferences and demographics of the passengers.
- Develop a configurable application that can generate diverse recommendations based on the airline and seat class selected, providing a personalized experience for users.

[35] Refer to: https://www.imdb.com/

By accomplishing these tasks, we aim to create a robust and adaptable movie recommendation system that caters to the specific needs and preferences of different airlines and their passengers.

The movie recommender system utilizes a popularity-based model, which recommends content based on its popularity among users. Popularity is determined by factors such as the number of views by passengers and the highest ratings received. This approach is beneficial as it mitigates the cold-start problem, where recommendations can be generated even for new users without historical data. To address the item cold-start problem, the release year and IMDB rating of a movie are considered. The database contains records of passengers' activities, which serve as the basis for generating recommendations.

The data processing and recommendation generation tasks are performed using Jupyter Notebooks[36] as the platform/IDE, with Python[37] as the programming language. The primary library used is Pandas[38]. The code is executed on a server, and the data is extracted from the database using the pyodbc[39] library. To establish a connection with the database, an ODBC driver must be downloaded.

The dataset is gathered and pre-processed through an SQL query. The amount of data extracted can be configured, and the extraction time varies depending on the airline and the number of records. Extracting one month of data from AFR (Air France) takes approximately 25 minutes, while extracting one day of data takes around 3 minutes. To account for new movies released weekly and added to the database, the system needs to be updated regularly to include these movies in the popularity score calculation. Choosing a suitable time interval is crucial. Using a longer interval, such as the past two months, may result in fewer viewers for movies released just a week ago, impacting their popularity score. On the other hand, using too short of a dataset may exclude highly popular movies from the calculations, leading to their exclusion from recommendations. Therefore, using one week of data provides a more up-to-date system while maintaining the reliability of the popularity score calculation.

[36] Please see: https://jupyter.org/
[37] Please see: https://www.python.org/
[38] Please see: https://pandas.pydata.org/
[39] Please see: https://pypi.org/project/pyodbc/

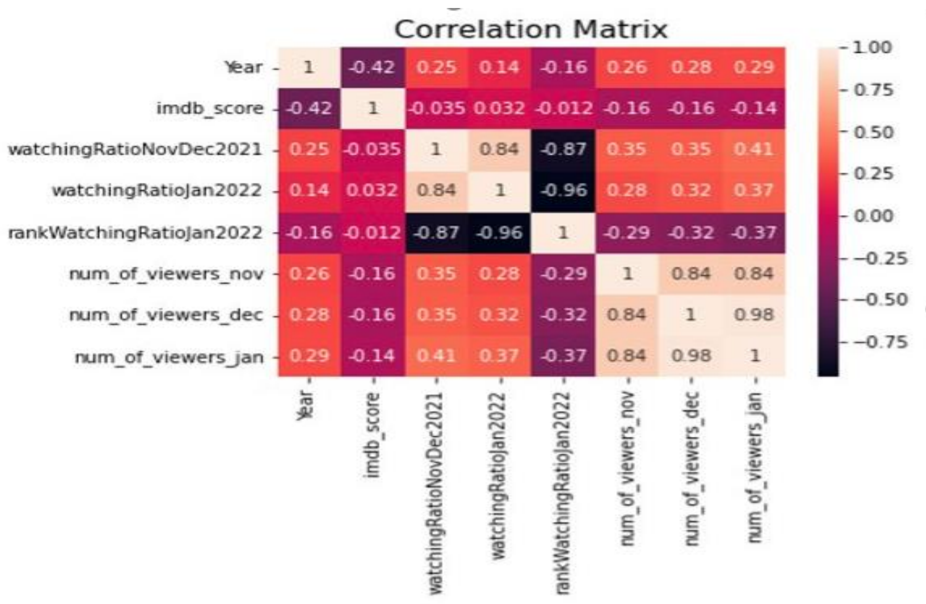


Figure 46 Correlation matrix

The correlation matrix, Figure 46, examines the relationship between different features to determine if there is any correlation among them. In this case, the correlation matrix indicates that there is no correlation between the watching ratio and the IMDB score, as well as between the watching ratio and the release year. This suggests that these three features are independent of each other. Consequently, the formula used to calculate the popularity score, which incorporates these features, is valid and appropriate for the recommender system.

Figure 47 shows the process used by the application to generate recommendations.

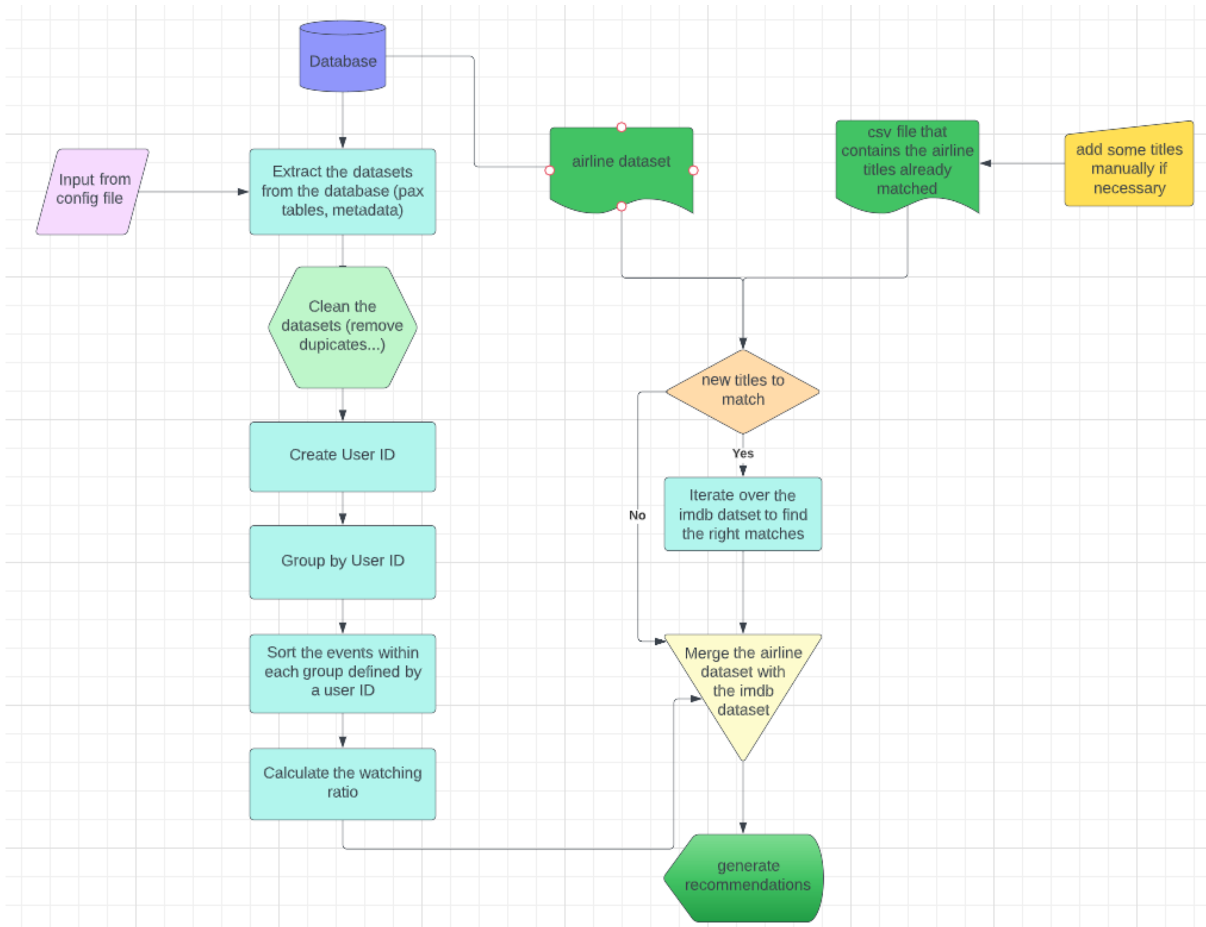


Figure 47 Flowchart illustrating the process of the application to generate recommendations

Item-based

In order to add a more precise recommendation system for a certain passenger, we have developed an AIF as the Item-based collaborative filtering approach to specialize the recommended content.

Item-based collaborative filtering is being utilized to recommend content to passengers onboard an airplane through the In-Flight Entertainment and Connectivity (IFEC) system. Here's a high-level overview of how it is being implemented in such a scenario:

- *Data Collection*: Gather data about passenger preferences and interactions with the IFEC system. This can include ratings, reviews, watched movies, preferred genres, or any other relevant information.
- *Content Analysis*: Analyse the content available in the IFEC system and extract relevant features or attributes. This could include genres, actors, directors, keywords, or any other metadata associated with the content.

- *Similarity Calculation*: Compute the similarity between content items based on their features or attributes. This can is done using k-Nearest Neighbour (k-NN) techniques.
- *User Preferences*: It collects information about the preferences and interactions of the individual passenger for whom the recommendation is being generated. This can include their historical ratings, watched movies, or any explicit feedback provided.
- *Candidate Selection*: Based on the user's preferences and interactions, identify the content items they have not yet consumed but are similar to the ones they have liked or interacted with.
- *Recommendation Generation*: The algorithm generates personalized recommendations for the passenger by suggesting the most similar and relevant content items. This can be done by considering the similarity scores between items and the user's preferences. The system can prioritize the items with higher similarity scores and present them to the passenger.
- *Real-time Updates*: Continuously update the recommendation model based on the passenger's ongoing interactions and feedback. This allows the system to adapt to the changing preferences of the passenger and provide more accurate recommendations over time.

6.3.1.2 ML-based predictive maintenance

Developments & Integration	
Test case # 1.2:	**AIF: Seatback screen predictive failure**
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.
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)	• Completing and advancing the application□ Q2 2023 • Full Integration of the application in the test-rack with AI@EDGE platform □ Q3 2023 • Final demonstration □ Q4 2023
Test Sequence	1. Integration of the application with the aero edge/cloud infrastructure. 2. Deployment of the application in the far-edge, employing AI@EDGE platform

The second AIF, named as predictive maintenance model is based on machine learning (ML) and is specifically designed to predict failures in In-Flight Entertainment and Connectivity (IFEC) screens, also known as Removable Display Units (RDUs). RDUs can experience non-functional states due to various factors such as temperature, aircraft type, software releases, software updates, and hardware types. By accurately predicting these non-functional states, the aim is to minimize maintenance time and prevent downtime of RDUs, which can have a negative impact on passengers' quality of experience.

- **Data Collection and Aggregation**: Reliable and comprehensive datasets are essential for ML models to make accurate predictions. In this study, the required dataset was obtained from a commercial SQL database that systematically stores data or logs collected from different aircraft in various airlines. Typically, such data is utilized by maintenance departments for offline analysis of aircraft system performance, including IFEC devices. The multi-label historical failure dataset for RDUs was extracted from the overall data maintained by the repair team, involving data gathering from diverse sources.
- **Overview of the Dataset**: The historical data collected comprises various features that describe the state of an IFEC screen. These features are represented by an n-tuple, including but not limited to attributes such as ID, average temperature, flight duration, and last software update. Each subset of the n-tuple in the database corresponds to different characteristics of RDUs and identifies a unique screen. The database contains comprehensive raw data representing different types of attributes. Some data fields can be directly inputted into the ML model, while others need to be discarded. The initial step involves preprocessing the raw data to obtain a meaningful subset of data. The remaining features primarily consist of categorical (textual) and numeric data types. Additionally, the dataset includes an important feature that indicates whether a screen has been previously broken or replaced, serving as the target label class. This feature is represented by a binary variable, where "0" denotes a normally functioning RDU and "1" represents an RDU predicted to be defective by the ML model and requiring replacement. Based on the historical values stored in the dataset, this use case can be categorized as a binary classification machine learning problem. Various experiments were conducted using different algorithms to address this problem.

The preliminary ML experimentation for IFEC screens (RDUs) failure prediction AIF was conducted on existing platforms before transferring to the Aero Edge-Cloud. Four different platforms, namely Azure, H2O[40], TPOT[41], and NNI[42], were evaluated based on criteria such as Graphical User Interface (GUI) support, AutoML[43] support, multi-core support, and ensemble model[44] availability. H2O was identified as

[40] Please see: https://h2o.ai/platform/ai-cloud/
[41] Please see: http://epistasislab.github.io/tpot/
[42] Please see: https://github.com/microsoft/nni
[43] Please see: https://www.automl.org/automl/
[44] Please see: http://www.scholarpedia.org/article/Ensemble_learning

the best-suited platform as it provides comprehensive functions for autoML, multi-core support, and built-in ensemble algorithms.

A pre-processed dataset, which involved removing duplicates and null values, was divided into a 20% portion for testing and an 80% portion for training. Since IFEC screens failure prediction is a binary classification problem, H2O AutoML was used with various built-in algorithms, including XGBoost[45] Gradient Boosting Machines, H2O Gradient Boosting Machines, Distributed Random Forests (DRF), Generalized Linear Models, and Stacked Ensemble models.

During each training iteration, the models were evaluated using the "Area Under The Curve" (AUC) of the Receiver Operating Characteristic (ROC) as a performance indicator. To address data set imbalance, the built-in auto-balancing method based on SMOTE (Synthetic Minority Over-sampling Technique) was employed. The top-performing model with the highest AUC was selected, and the F1 score was calculated during the testing phase, please see Table 12.

Table 12 Testing results – RDUs failure prediction

			12 Features experiments using H2O Flow		12 Features experiments using H2O R Cluster	
max model	**max secs**	**class_sampling_factors**	**running time**	**F1 Score**	**running time**	**F1 Score**
6	1800	0.1, 1	00:36:16	0,8741	00:15:61	0.880660
6	1800	0.1, 1	00:38:11	0.8757	00:22:71	0.869760
6	1800	0.1 ,1	00:39:14	0.8614	00:15:21	0.878220
6	1800	0.1 ,1	00:30:25	0.8709	00:14:36	0.880452
6	1800	0.1, 1	00:27:35	0.8654	00:15:41	0.881411
Average Result			00:34:20	0.868	00:16:68	**0.878**

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

Development & Integration	
Test case # 2:	**MPTCP**
Slogan & Objective	• To test multi-path aggregation at the transport layer in multi-connectivity multi-RAT scenarios.

[45] Please see: https://xgboost.readthedocs.io/en/stable/

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)	• Improving performance of onboard MPTCP between aero-certified nodes → Q3 2023 • Deployment of MPTCP Proxy → Q3 2023 • Deployment of Predictable MPTCP → Q3/Q4 2023
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.

The Multipath TCP[46] (MPTCP) tests were conducted between two devices, Supermicro and one of the RDUs both located in the AI@EDGE subnet. The devices were connected to the same Wi-Fi network using two Wi-Fi dongles, allowing them to have two different links for communication: Wi-Fi and Li-Fi. The Supermicro was connected to a Li-Fi Access Point via Power over Ethernet cable, while the RDU was connected to the Li-Fi Access Point using a Li-Fi dongle. The Wi-Fi network was dedicated and configured specifically for these tests, ensuring minimal interference from other devices or connections.

The Supermicro device runs Debian 11 (bullseye) server version with Linux Kernel version 5.19, while the RDU 3E runs a custom Linux-based OS from the PTXdist build system with Kernel version 5.10. Both devices have MPTCPv1 enabled by default in their kernels and are ready to use after configuration. They also have a Multipath TCP daemon built from source, which includes the "mptcpize"[47] binary. This binary enforces the creation of MPTCP sockets instead of TCP sockets on both machines.

The tests had two main objectives: measuring the achieved throughput and verifying that the end-to-end connection remains established even when one interface becomes temporarily unavailable. To achieve these objectives, an HTTP server was launched on the Supermicro side, utilizing the "mptcpize" binary to enforce the use of MPTCP sockets. On the client side, the "wget" command was wrapped with the "mptcpize" binary as well. It was used to download a large file (500 MB) from the Supermicro server.

[46] Please see: https://www.multipath-tcp.org/
[47] Details at: https://manpages.ubuntu.com/manpages/lunar/man8/mptcpize.8.html

These test configurations and procedures allowed the evaluation of MPTCP's performance in terms of throughput and connection resilience in a scenario with multiple network interfaces.

Server's MPTCP path manager configuration:

Supermicro's MPTCP path manager has been configured to establish a maximum of two MPTCP sub-flow and advertise two MPTCP endpoints: Wi-Fi and Li-Fi as can be seen in Figure 48 with the wlxec086b106ee0 (Wi-Fi) and eno1.52 (ethernet to Li-Fi AP) interfaces (Figure 49).

```
test@smicroright:/tmp/mptcp$ ip mptcp limit show
add_addr_accepted 2 subflows 2
test@smicroright:/tmp/mptcp$
test@smicroright:/tmp/mptcp$ ip mptcp endpoint show
192.168.1.133 id 1 signal subflow dev wlxec086b106ee0
10.0.52.1 id 2 signal subflow dev eno1.52
test@smicroright:/tmp/mptcp$
test@smicroright:/tmp/mptcp$ mptcpize run python3 -m http.server
Serving HTTP on 0.0.0.0 port 8000 (http://0.0.0.0:8000/) ...
192.168.1.140 - - [19/Apr/2023 18:14:27] "GET /test_file_500M HTTP/1.1" 200 -
```

Figure 48 MPTCP endpoints

Client's MPTCP path manager configuration:

RDU's MPTCP path manager has been configured to establish a maximum of two MPTCP sub-flow as shown in Figure 49. The client side shows no endpoints, as they are handled by the server.

```
root@rdu-3E:~
root@rdu-3E:~ ip mptcp limit show
add_addr_accepted 2 subflows 2
root@rdu-3E:~
```

Figure 49 MPTCP sub-flows configuration

Predictive MPTCP

Unlike TCP, where data is transmitted in only one stream, MPTCP allows multiple flows (called sub-flow) to be used concurrently for data transmission, so for each packet, the MPTCP scheduler will have to decide which sub-flow to use for transmission. Various algorithms are used for this decision-making (simple, like round-bin, or complex, which is based on the current state of the network at each point in time). Making decisions based on the state of the network allows the MPTCP-scheduler to operate more efficiently, and by default, MPTCP-scheduler relies on the TCP signal to get the network properties, and in some cases, such as packet loss, it takes time to detect.

In this test case, the information collected at the RAN is used to help the MTPCP scheduler get more information about the network health earlier, even before the data is sent. In Figure 49, through the E2 interface, an xApp at near RT-RIC will collect metrics in gNB, and the information collected will be passed to a predictor application to predict, in this test case, only the probability of a packet loss. The result will then be passed to MPTCP to update the MPTCP scheduling policy.

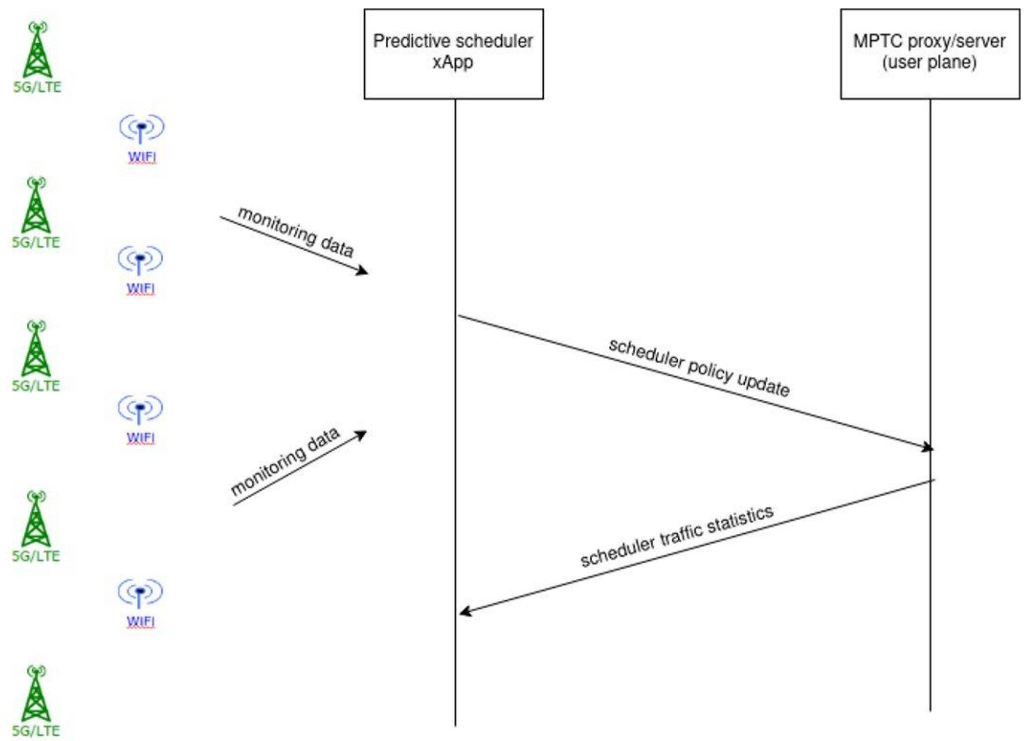


Figure 50 RAN metrics collection scheme for the MPTCP scheduler

6.3.3 Test case 3: Video Streaming for IFEC services

Video Streaming for IFEC services	
Test case # 3:	**Adaptive Video Streaming**
Slogan & Objective	Adaptive video streaming refers to the technique of dynamically adjusting the quality of video content based on the viewer's network conditions and device capabilities. It aims to deliver the best possible viewing experience by continuously adapting the video quality in real-time.
Test Scenario	• The adaptive video streaming technology allows to define a “manifest file” which describes the parameters of the video streams (e.g., bitrate)

	• We have defined two different manifest files, one for the Economy class and one for the Business class. • When selecting a video content from its RDU, the user will be automatically associated with its class via the manifest file.
Expected Results	• The selected video is streamed to the RDU according to the class (Business/Economy). • As the video plays, the video player continuously monitors the network conditions, such as available bandwidth and latency. It periodically requests and downloads small video segments, usually a few seconds in duration, from the server. • Using the information obtained from the video segments, the video player dynamically adjusts the video quality for subsequent segments. If the network conditions are good, it stays at the maximum quality level for related user class. Conversely, if the network conditions deteriorate, it may switch to a lower quality version to avoid buffering or interruptions.
General Time Plan	• Q1/2023: First trials at ITL testbed. • Q2/2023: Integration into the SPI testbed and first experiments. • Q4/2023: Final validation
Test Sequence	1. Access to the Recommendation system from the RDUs, 2. Check that the user can access the application for selecting a video among those stored locally at the MEC server. 3. Video selection by the user. 4. Verify that the quality of the video stream to the RDU complies with the user's class. 5. Force a degradation of the network condition. 6. Verify that adaptive video streaming allows user to enjoy video content without interruptions, buffering, or noticeable quality degradation.

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

The following three connectivity tests 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.

5G Connectivity and Local Traffic Breakout	
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)	• Q3 of 2023.
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).

5G Connectivity and Local Traffic Breakout	
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 2023
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.

5G Connectivity and Local Traffic Breakout	
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[48] shows uplink/downlink traffic. • ICMP messages are acknowledged.
General Time Plan (Validation Campaigns)	• Q4 of 2023.
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.

6.3.5 Test case 5: 5G performance to access the Internet over Starlink

The test case is to validate Internet connectivity through Athonet 5G core (edge UPF), SrsRAN, and LEO satellites (SpaceX Starlink).

5G connectivity	
Test case 5:	**5G performance to access the Internet over Starlink**

[48] Please see: https://iperf.fr/

Slogan & Objective	• To enable broadband connectivity onboard through LEO satellite backhaul
Test Scenario (Pre-conditions)	• 5G User plane traffic to street to Data Network through LEO satellite.
Expected Results (Post-Conditions)	• Internet Connectivity between UE and DN is operational.
General Time Plan (Validation Campaigns)	• Q4 of 2023.
Test Sequence	• Defining new IP pool for users to access internet. • Establish a new PDU session from Edge UPF to DN. • Performing speed-test to measure the throughput and latency.

6.4. Validation scenario

The UC4 testrack setup in SPI is depicted in Figure 51, featuring all AIFs and the application. The configuration comprises 16 operational RDUs, an SCU (aero-certified Server), and a SuperMicro (COTS Server), forming an edge cluster managed by K8S as VIM. Additionally, it incorporates an orchestration lever (MEO) as part of the AI@EDGE CCP.

The first row of RDUs (1A, 1B, 1C) is dedicated to the monitoring system, overseeing various metrics across the entire cluster, including CPU and memory consumption, data transmission rates, disk storage, etc for each working node. The second row serves the crew members: RDU-2A is assigned to initiate seat screen prediction failure (or RDU swapping) AIF, 2B functions as the crew panel for activating the recommendation engine for passenger movie suggestions, and 2C provides a comprehensive overview of RDUs status, running AIFs, and applications through Portainers[49] (refer to Figure 52 and Figure 53).

Rows three and four of RDUs cater to passenger interactions, offering access to the GUI and the recommendation engine AIF for content consumption. Lastly, RDUs 5A, 5B, 5C, and 6A showcase the outcomes of RDU swapping AIF on separate screens.

[49] https://www.portainer.io/

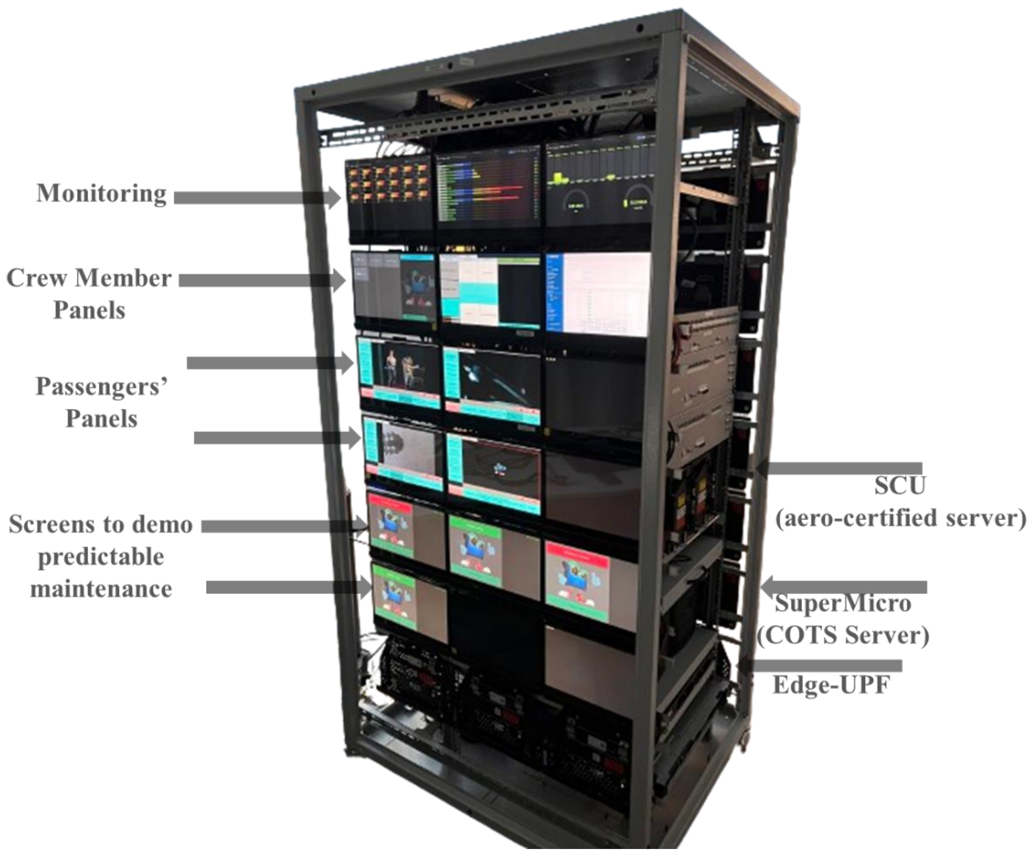


Figure 51 UC4 Test rack setup at SPI with deployed AIFs and other applications

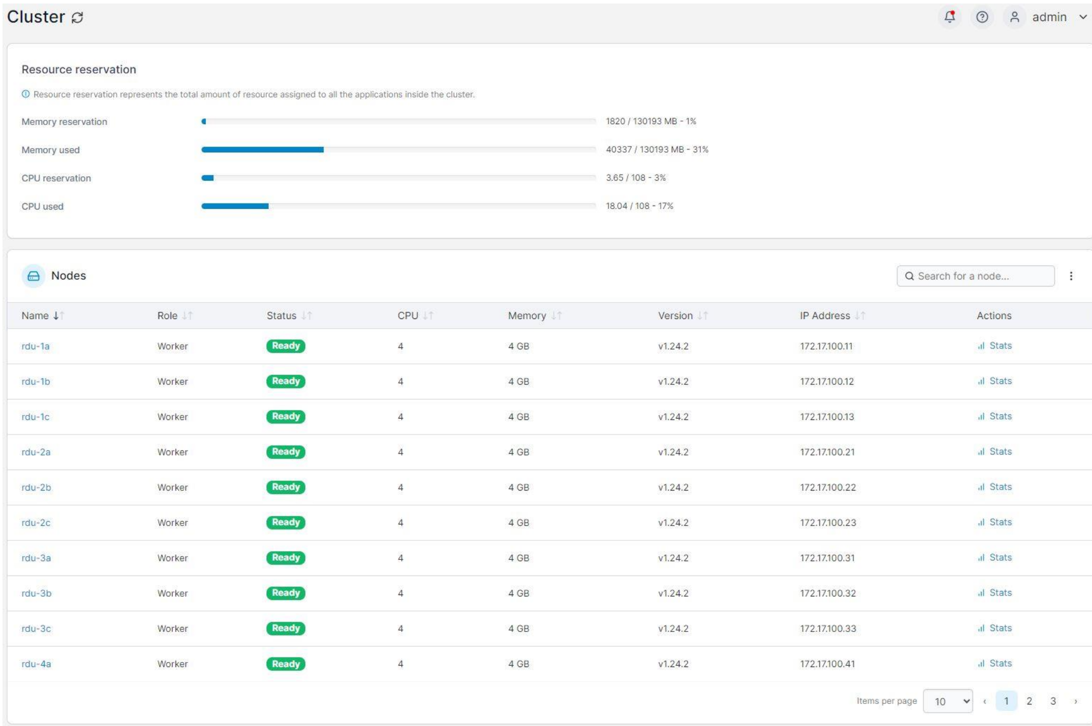


Figure 52 Visualization of cluster by Portainer

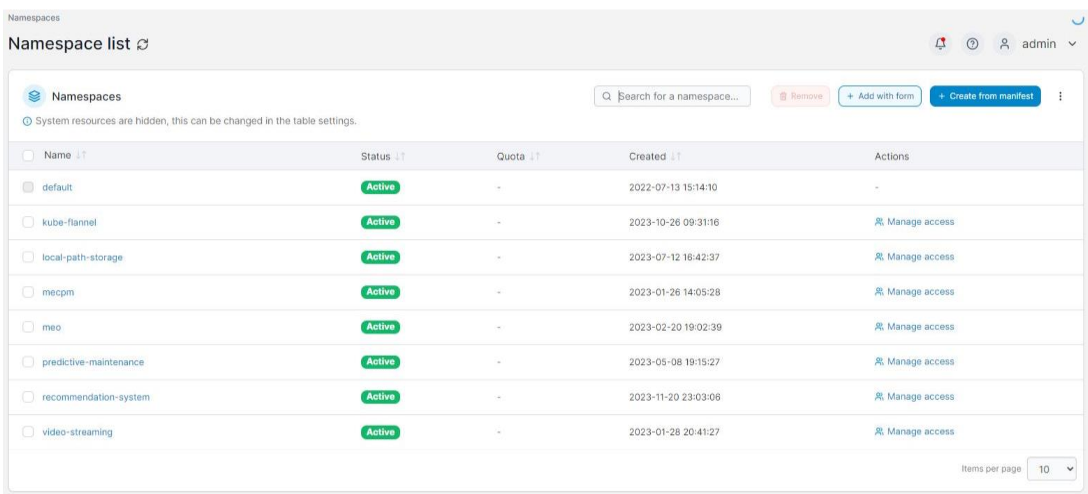


Figure 53 Visualization of namespaces by Portainer

6.4.1 Test case 1 results: AIFs development

Test Case #1.1: Recommendation system

- Service Deployment (delivery) time: < 2s (Initial recommendation)

- Service Recovery time: ~1s (when the image already pre-loaded in the node)

Test Case #2.1: Seatback screen predictive failure

- F1 Score: 0.878
- Service Deployment (prediction) time ~ 15s
- Service Recovery time: ~1s (when the image already pre-loaded in the node)

The experiments were conducted using a Dell PowerEdge server with an AMD EPYC 7402P processor, 24 CPU cores, 64 GB of RAM, and Linux OS Kernel version 4.19 with Debian 10 distribution. The training time (Figure 55) and F1 score results (Figure 56) were analysed, and it was observed that the H2O Code-based version provided lower training time. The DRF algorithm was selected as the one with the highest AUC value. The F1 score of the DRF algorithm was found to be highest in the H2O Code-based version using the top 12-feature dataset (Figure 54).

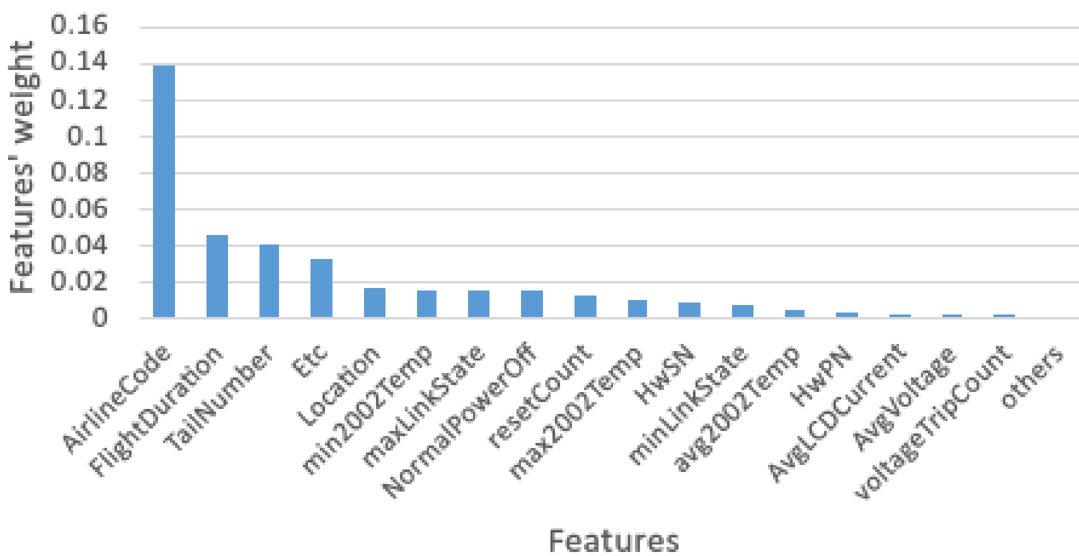


Figure 54 Considered features

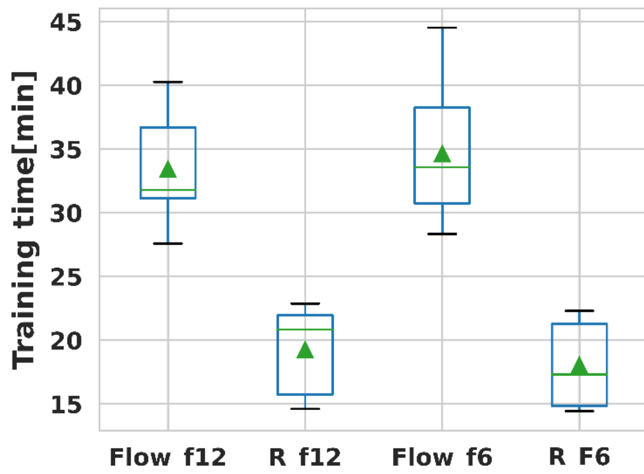


Figure 55 Training time

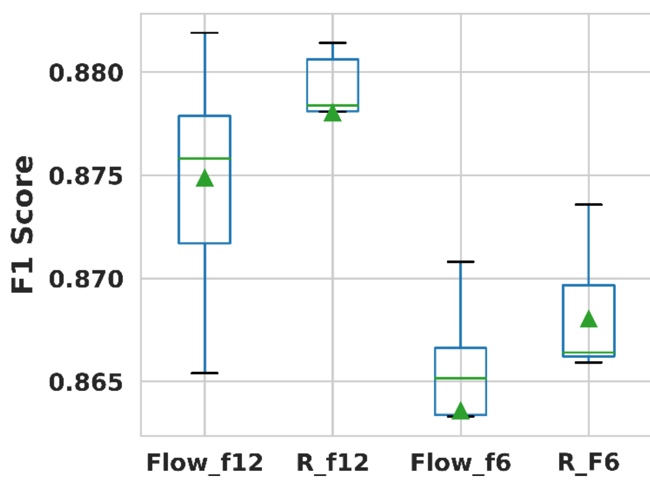


Figure 56 F1 score

Based on these results, it is concluded that the IFEC Failure Prediction App can rely on the H2O Code-based platform with the DRF algorithm for accurate predictions.

Furthermore, Figure 57 provides a report on the CPU consumption of the SCU, which hosts the RDU swapping model. This analysis focuses on the period when the crew member initiates the model through the GUI assigned to them (on RDU-2A). During the active operation of the model, the SCU experiences a peak CPU consumption, reaching up to 95% of its capacity. In contrast, when the model is not in operation, the SCU typically utilizes around 35% of its CPU resources.

In contrast, Figure 58 centres its attention on the CPU consumption of RDUs 5A and 6A during their refresh cycles to prepare for new measurements and predictions. During these refresh intervals, their CPU usage can escalate, nearly reaching half of their capacity. In standard operating conditions, outside of the refresh periods, the CPU consumption typically maintains a level of around 15-20%.

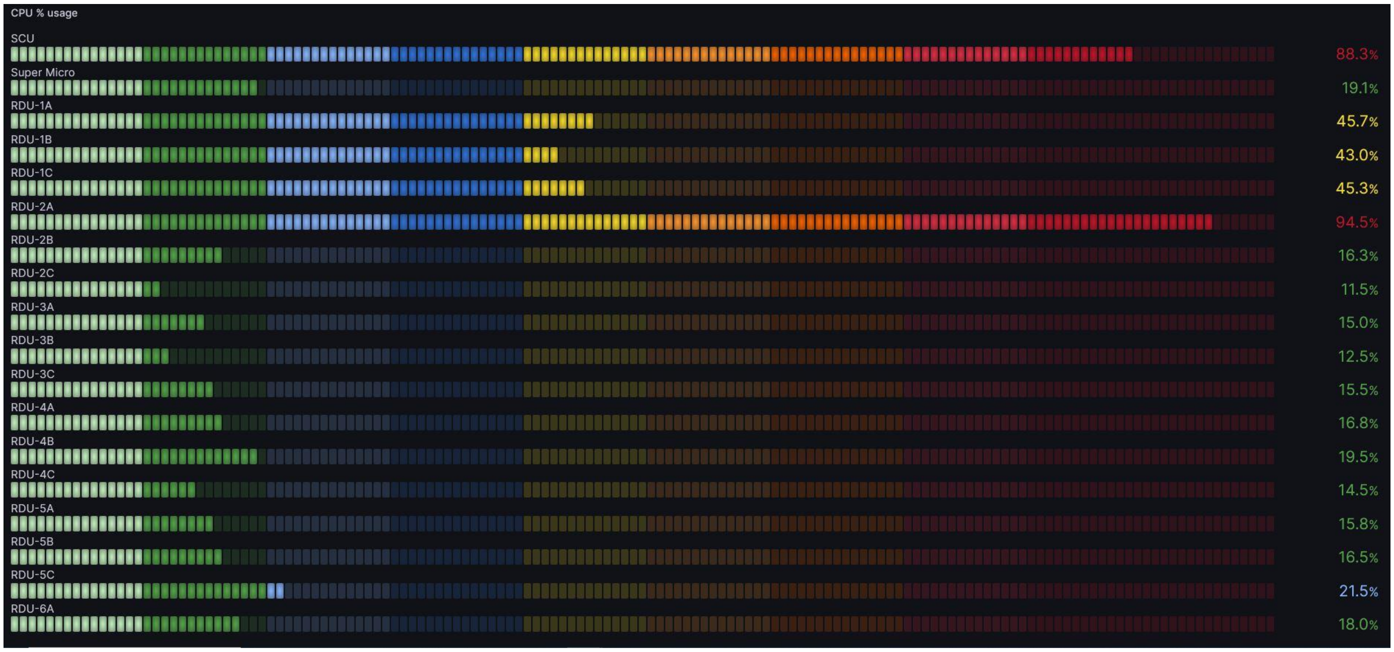


Figure 57 CPU consumption when prediction model AIF is being run

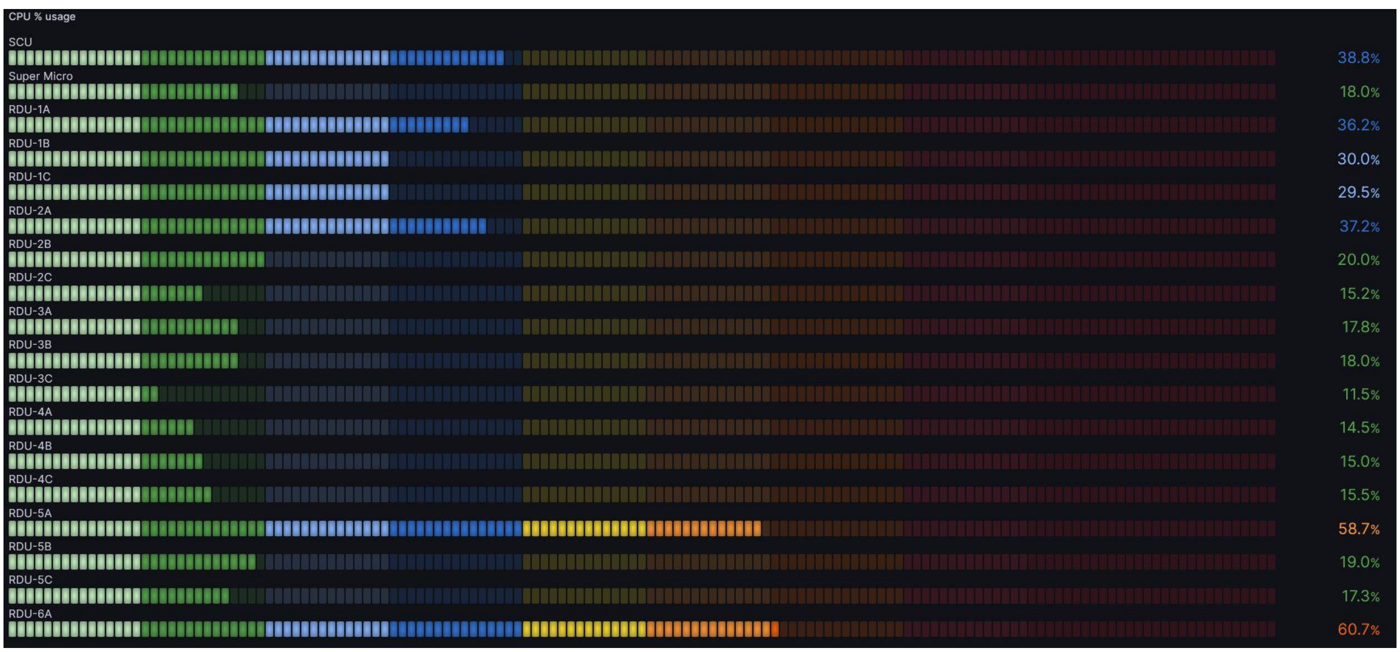


Figure 58 CPU consumption when 2 RDUs (5A, 6A) are being refreshed for a new prediction

6.4.2 Test case 2 results: Multi-path TCP and MPTCP proxy

Results:

- Connectivity between RDUs and the server is established.

- MPTCP with default schedular between RDUs and server is established.
- Max achieved on Li-Fi alone: ~180 Mbps for 1 UE
- Max achieved on Wi-Fi alone: ~80 Mbps for I UE
- Max achieved on Li-Fi/Wi-Fi aggregation using MPTCP: ~250 Mbps for 1 UE

Figure 59 demonstrates the MPTCP connection test results between the Supermicro server and the RDU device. This test was conducted using the default TCP Window Size. However, the test revealed unstable throughput performance, indicating fluctuations in the data transfer rate over time.

The figure presents a chart showing the throughput performance over time during the MPTCP connection test. The fluctuations in the graph indicate varying data transfer rates, suggesting that the default TCP Window Size may not provide optimal performance for MPTCP.

The unstable throughput performance highlights the need for further optimization to achieve consistent and reliable data transfer rates using MPTCP.

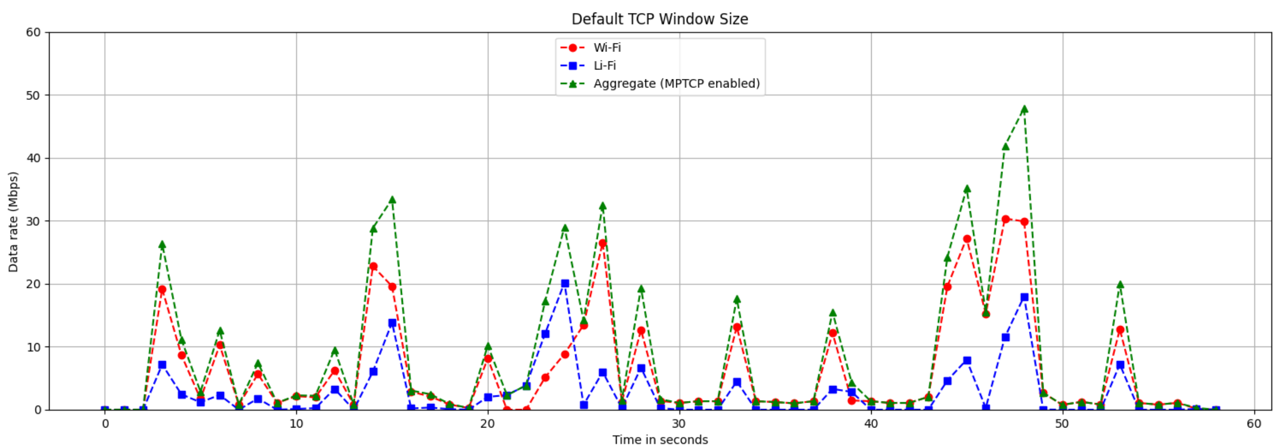


Figure 59 MPTCP connection test results - default TCP Window Size

Figure 60 displays the MPTCP connection test results between Supermicro and RDU with increased TCP Window Size on client side. We can clearly see that in this case, increasing TCP Window Size has led to slight improvement in the MPTCP performance, especially in terms of achieved throughput.

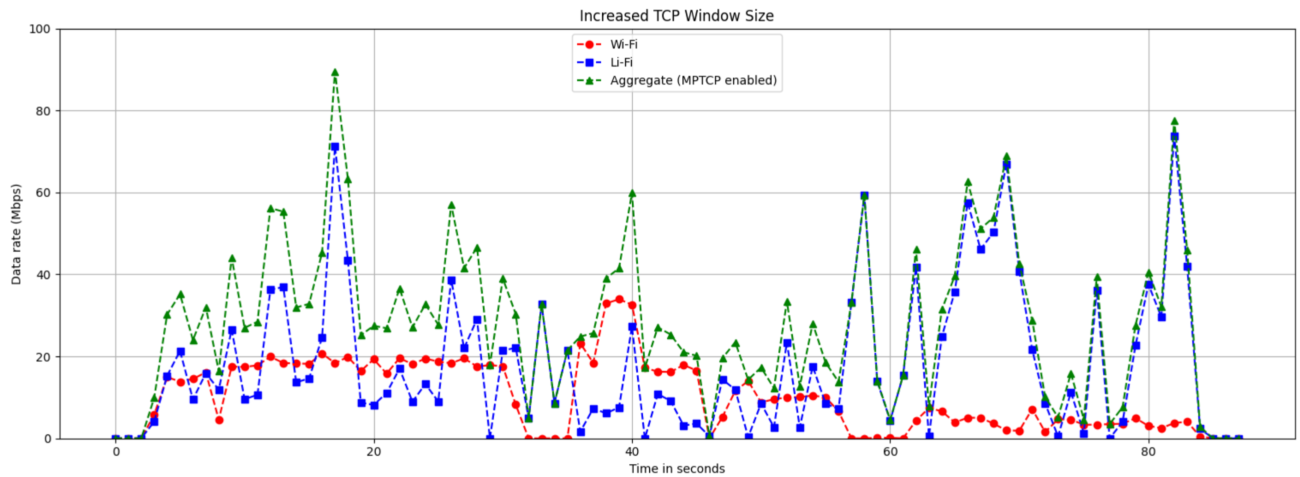


Figure 60 MPTCP test result of content loading from Supermicro to RDU 3E with increased TCP Window Size

The benefits of using MPTCP to improve network reliability are clear from the test results, as it enables the usage of both Wi-Fi and Li-Fi links for content loading. However, in these tests, the performance of MPTCP was found to be unstable. This was because the client side, RDU, was using Kernel version 5.10, which may have had some limitations in MPTCP implementation.

However, MPTCP implementation in the newer Linux Kernels provides better performance, improved network reliability, and increased data rate, as evidenced by other tests. For example, when we replaced the RDU 3E with a laptop running a Kernel version higher than 5.13 (in this case, 6.1), we observed improved performance and stability in the MPTCP connection as can be seen in Figure 61. This demonstrates the importance of using an up-to-date Kernel version when working with MPTCP to achieve the best possible results.

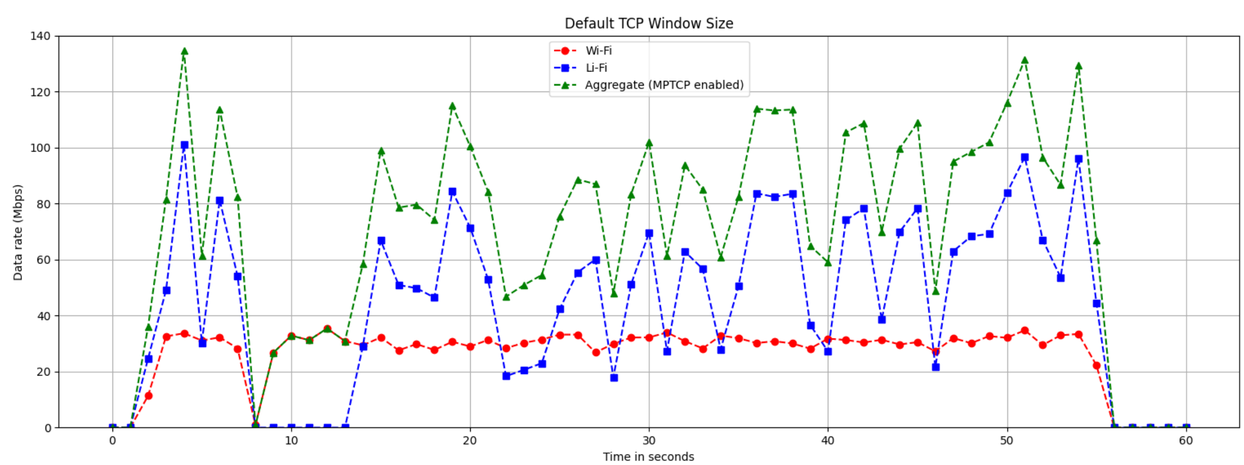


Figure 61 MPTCP test result of content loading from Supermicro to Laptop using Wi-Fi and Li-Fi with default TCP Window Size

During the MPTCP testing phase between Supermicro and a Laptop, the TCP Window Size on the client side was increased from its default value. As a result, the tests produced better results, with data transfer rates reaching up to 190 Mbps, please see Figure 62. These findings suggest that increasing the TCP

Window Size on the client side can significantly enhance MPTCP performance in the latest stable Linux Kernels.

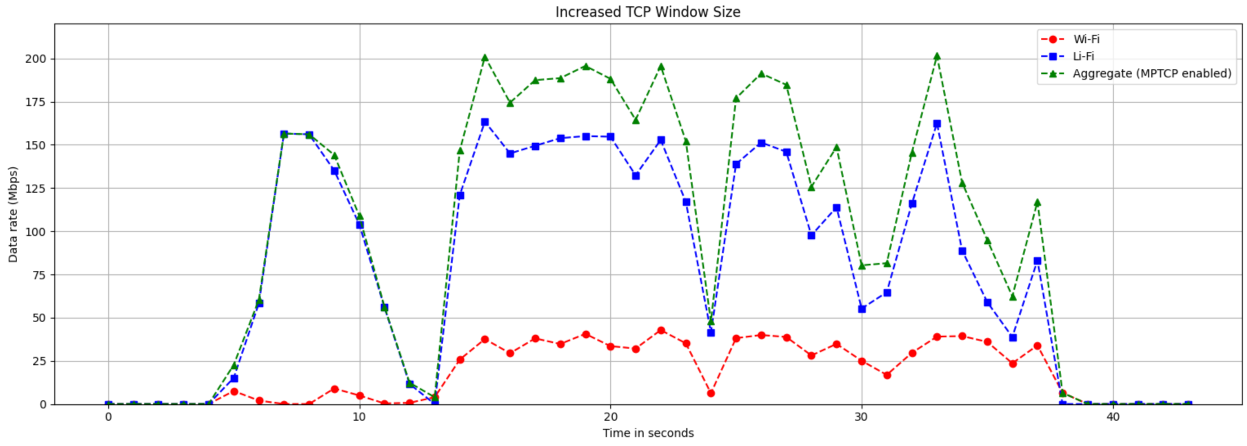


Figure 62 MPTCP test result of content loading from Supermicro to Laptop using Wi-Fi and Li-Fi with increased TCP Window Size

The results of MPTCP tests demonstrate that the server can establish a connection using multiple sub-flows with the client using different paths, allowing for increased reliability. Moreover, by increasing TCP Window Size, network performance can be enhanced, resulting in higher throughput. This configuration enables the server to fully utilize the available network resources, providing a better experience for the client.

Figure 63 shows the setup to provide both Wi-Fi and Li-Fi wireless interfaces for the nodes of UC4 test rack.

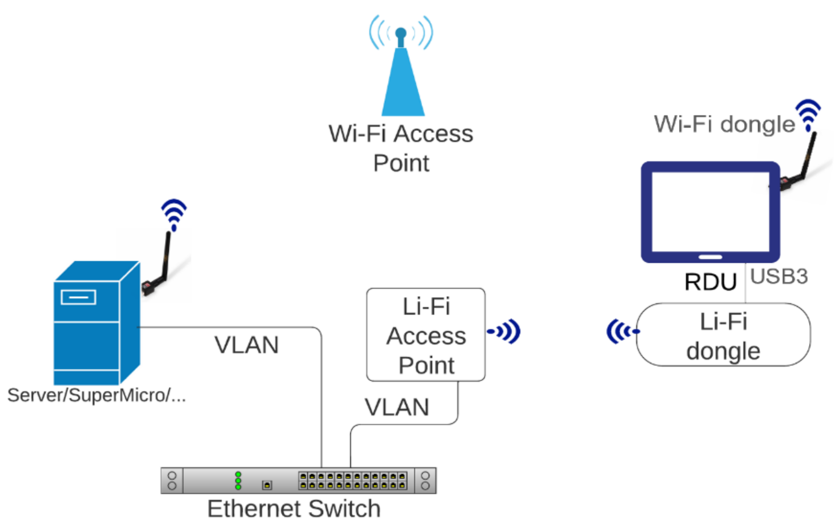


Figure 63 Wi-Fi/Li-Fi connectivity in UC4

To demonstrate the effectiveness of the default MPTCP scheduler within the UC4 testbed, we have categorized passengers into three classes: "***First class***," "***Business class***," and "***Economy class***." All nodes within each class are equipped with two interfaces of Wi-Fi and Li-Fi. However, when running an MEC application, the nodes in the ***economy class*** route their traffic exclusively through Wi-Fi, while those in the ***business class*** solely utilize Li-Fi for data transmission. In contrast, the ***first-class*** nodes leverage both interfaces for connectivity, with MPTCP enabled.

This configuration allows the first-class nodes to achieve the highest data rates compared to the other two classes, while also ensuring a more robust connection in case one of the links experiences a failure.

Furthermore, by running the MEC application, it becomes straightforward to change the class of each RDU by simply adjusting the class tag within Kubernetes (K8S), without the need for any further RDUs modifications.

Figure 64 and Figure 65 illustrate the performance of the nodes in accordance with the aforementioned scenario with three clients.

Figure 64 Integration of MPTCP in UC4 test rack

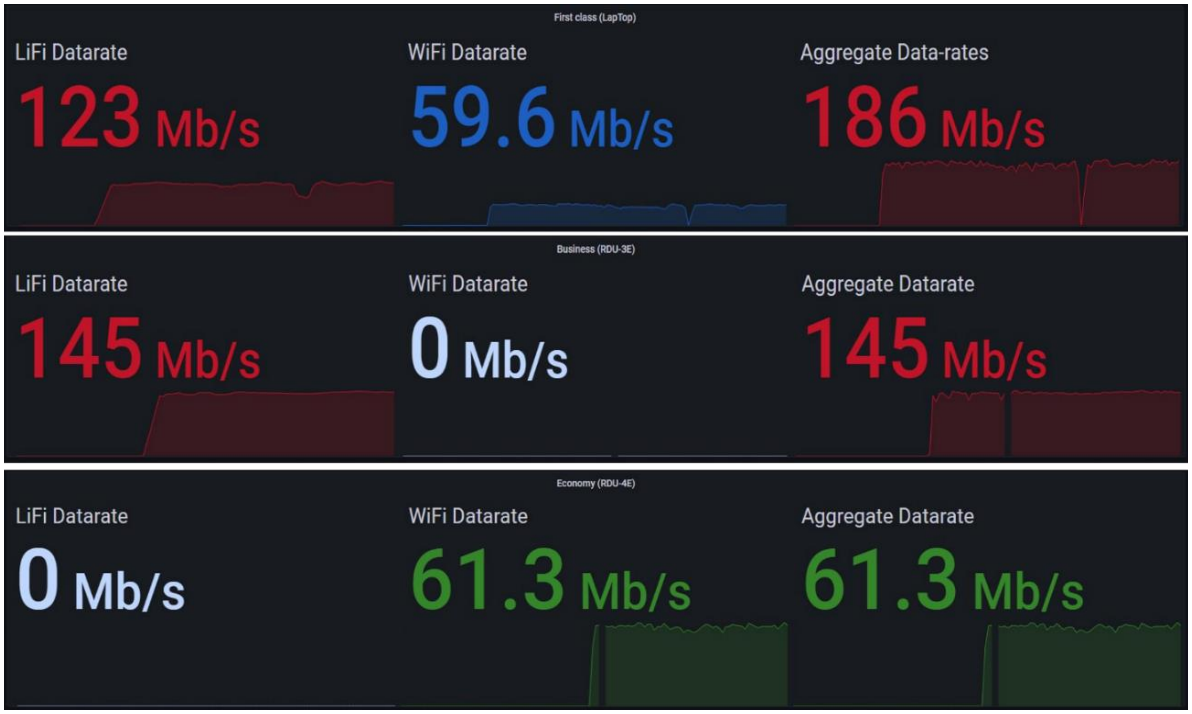


Figure 65 Data rates and performance of different nodes with different connectivity

6.4.3 Test case 3 results: Video Streaming for IFEC services

- Manifest files for adaptive video streaming technology are operational.
- Average CPU consumption of RDUs during streaming ~ 85% (Figure 66).

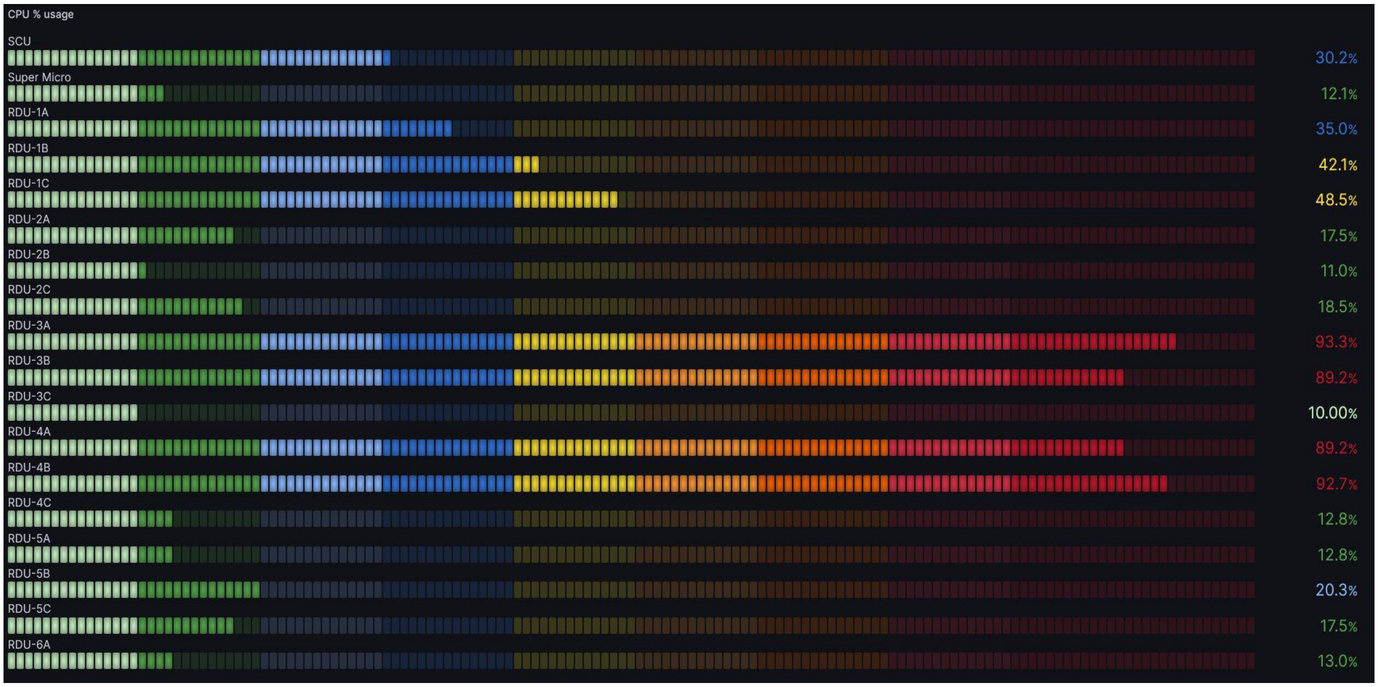


Figure 66 CPU consumption of Streaming videos on four RDUs (3A, 3B, 4A, 4B) from SuperMicro.

6.4.4 Test case 4 results: 5G Connectivity and Local Traffic Breakout

Test Case #4.1: Connection between gNB and 5GC

- No connection errors. Log messages show gNB successfully attached to the AMF. Control plane messages between the RAN and the 5GC are correctly exchanged.

Test Case #4.2: UE's attach to and detach from the 5G network

- Log messages show UE successfully registered, attached, and detached to the 5GC.

Test Case #4.3: Connectivity between UE and data network (DN)

- Connectivity between UE and DN is operational (Figure 67 shows the testbed setup).
- iPerf[50] shows uplink/downlink traffic.
- ICMP messages are acknowledged.

[50] Please see: https://iperf.fr/

- >100 Mbps in close proximity between 5G-UE and gNB with SISO and 40MHz BW.

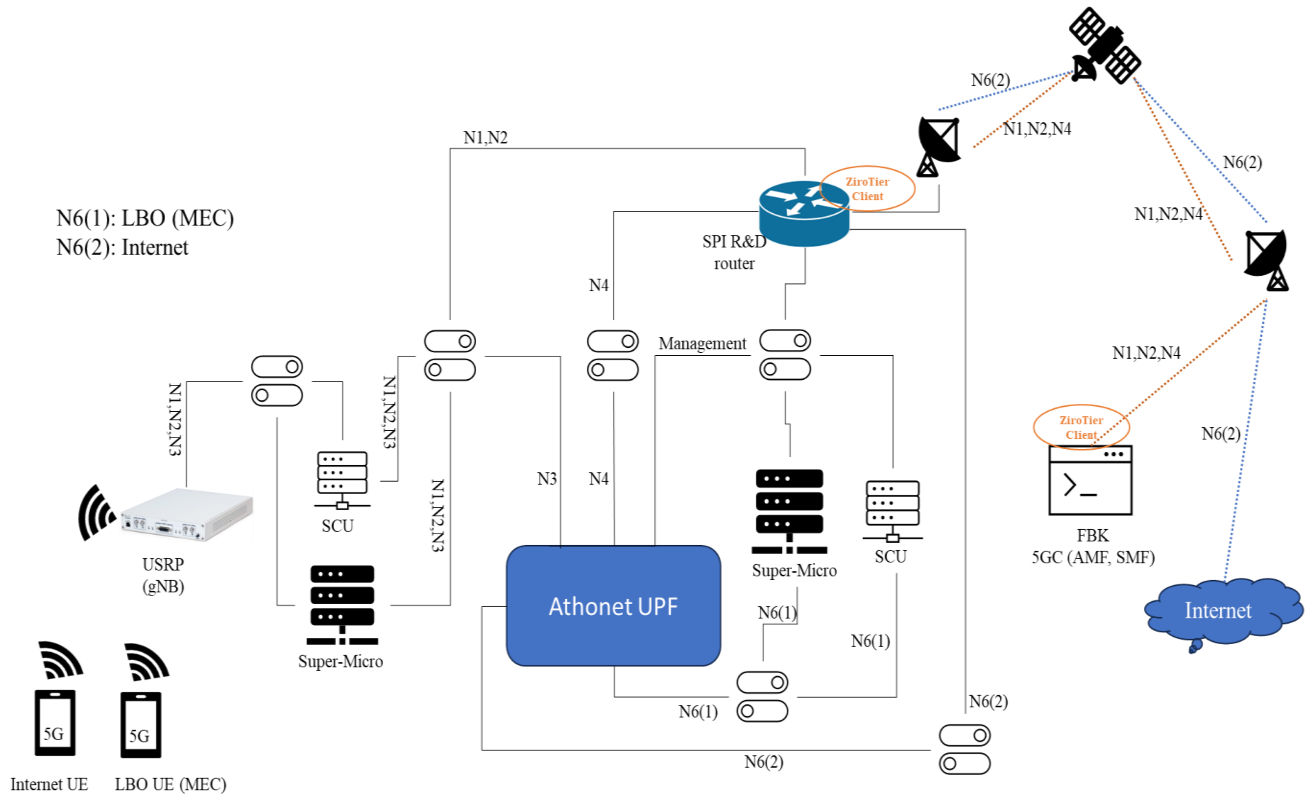


Figure 67 5G Connectivity and Local Traffic Breakout setup for UC4

6.4.5 Test case 4 & 5 results: 5G performance to access LDN and the Internet over Starlink

Following Figure 44 and Figure 45 there are two types of 5G users to access to the internet through LEO orbit satellite backhaul. The backhaul for our tests is Starlink satellites.

5G performance

Figure 68, Figure 69, and Figure 70 demonstrate analysis of the performance of the 5G User Equipment (UE) employing SrsRAN as the gNB. The experimentation utilized the SDR (Software-Defined Radio) board B210, maintaining a configuration of Single Input Single Output (SISO), and a 40MHz bandwidth consistently across all tests. The tests focused on Band n41 of NR (New Radio), providing insights into the behaviour of the 5G UE under these conditions.

A key parameter to measure throughput was the distance between the 5G UE and the gNB. By varying the proximity of the 5G UE to the gNB, we evaluated performance metrics across different scenarios. Figure 68, Figure 69, and Figure 70 serve as visual representations of the outcomes observed during these tests. The results showcased in the figures encompass performance indicators of throughput, packet loss, and the

MCS (Modulation and Coding Scheme) index. As can be seen from Figure 68, the average throughput at a close distance to the gNB is about 100 Mbps, whereas this value can go down to around 70 Mbps when the UE is about 10m from the gNB. Moreover, the results show the modulation can hardly reach to 256-QAM and it remain mostly with 64-QAM. Moving to higher modulation can cause a high packet-loss that force the gNB to switch to lower modulations immediately (e.g., 64-QAM).

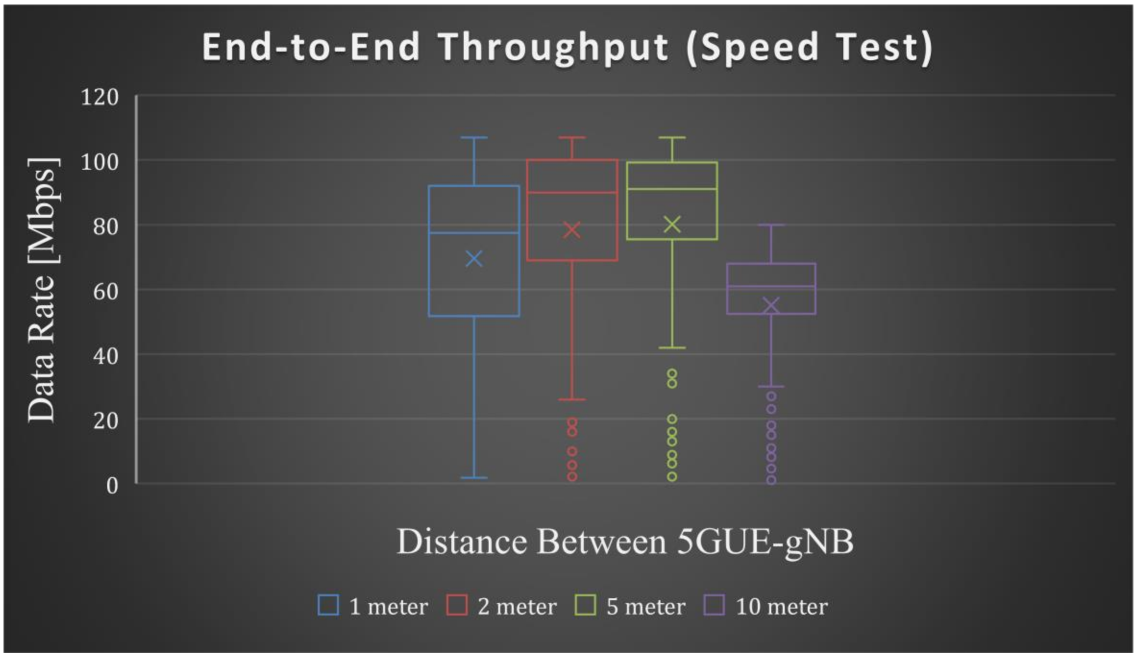


Figure 68 End to end throughput

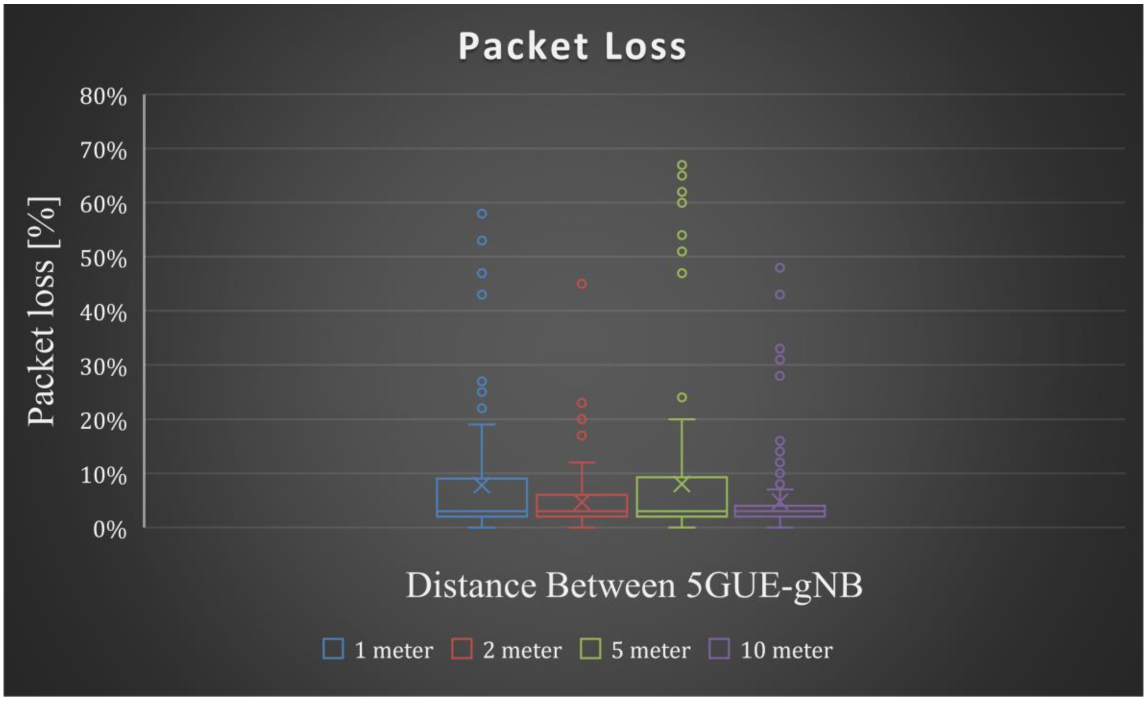


Figure 69 Packet loss

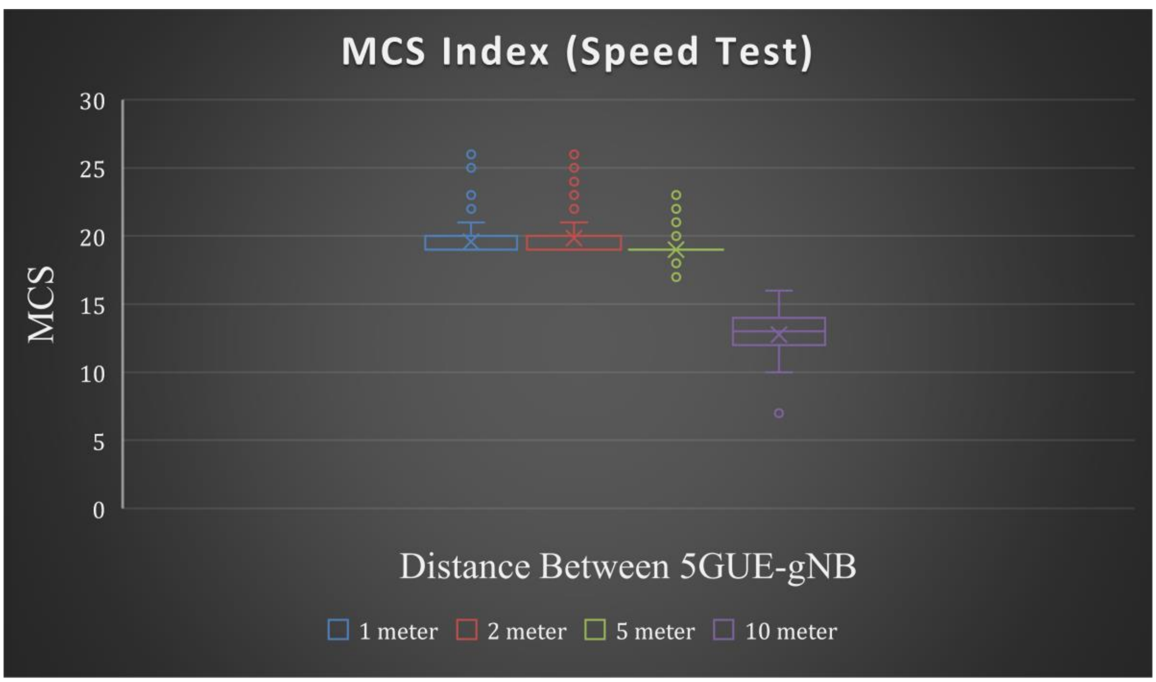


Figure 70 MSC index according to 38.214 - Table 5.1.3.1-2: MCS index table 2 for PDSCH[51]

Moreover, we observed also the round-trip time for both Control Plane (CP) and User Plane (UP). As already discussed in Deliverable D5.2 and highlighted in Figure 67, both control and data plane go through Starlink, however, for the Control Plane we are using a ZiroTier VPN to reach Athonet 5G core, deployed in FBK to establish N1/N2 and N4. (Shown in Figure 44 and Figure 67). The N6(2) interface, on the other hand, is not going through the ZiroTier VPN and reach to the Data Network (DN) (e.g., Internet) directly through the Starlink Satellite. Figure 71 report the round-trip time for both CP and UP. Starting from the left, the initial two charts illustrate the utilization of both Starlink and the ZiroTier VPN to establish point-to-point connectivity between SPI and Athonet 5G core at FBK for the control plane. The third chart, distinguished in grey, depicts the latency in the 5G data plane that traverses through Starlink without any intermediary VPN. Concluding the sequence, the yellow chart represents the latency between the 5G-UE and the MEC.

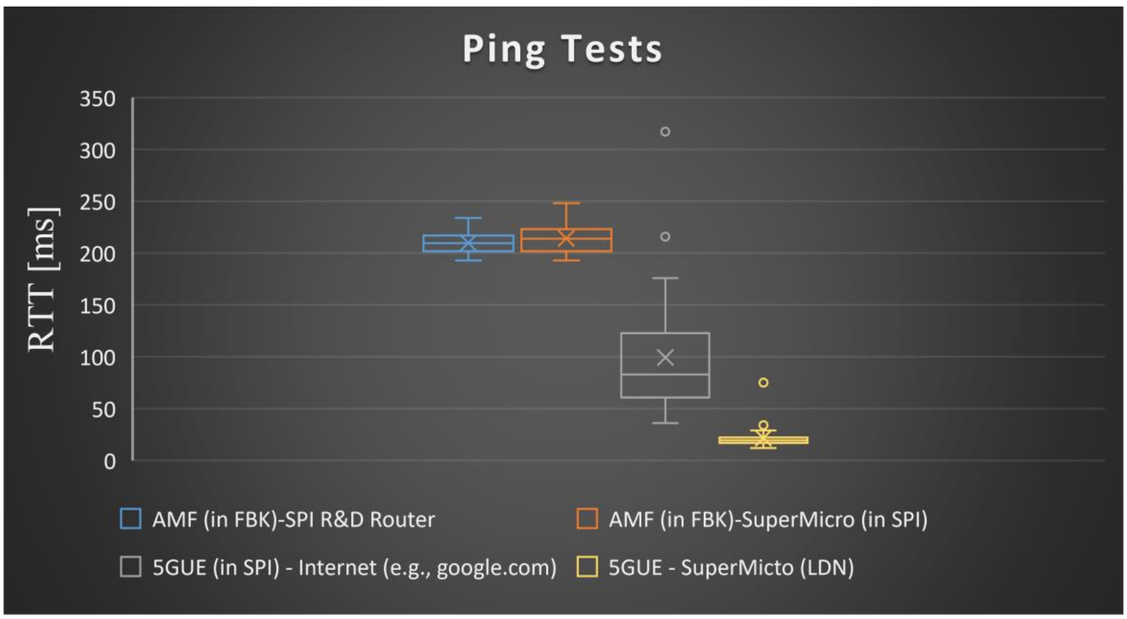


Figure 71 5G latency tests for Internet and MEC

Data Plane monitoring for MEC and Internet access

Figure 72 to Figure 75 have been included to enhance the visualization of data plane traffic. These figures depict the monitored traffic transferred and received over the N3 interface. Specifically, Figure 72 and Figure 73 illustrate the data rate for accessing LDN, representing the download of content and streaming from the edge server, respectively.

[51] Please see Specification: 3GPP TS 38.214 V18.0.0 (2023-09) - link

Moreover, Figure 74 and Figure 75 provide insights into the patterns of received and transmitted data reaching the internet. While Figure 74 illustrates the data traffic when watching a YouTube video, Figure 75 portrays the data flow during online streaming; considering that the maximum data rate for internet access was showcased in Figure 68 doing a “speed test”. It is crucial to emphasize that, for internet access, all data traffic traverses through the LEO satellite, SpaceX Starlink, utilizing the N6(2) interface.

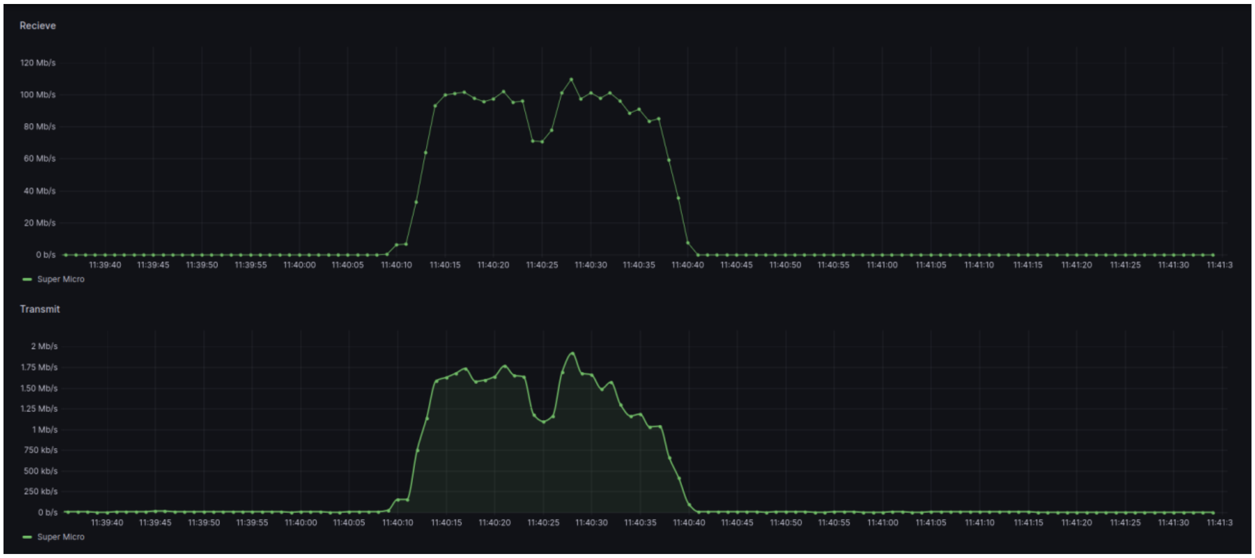


Figure 72 Downloading a content from LDN by smartphone (5G-UE)

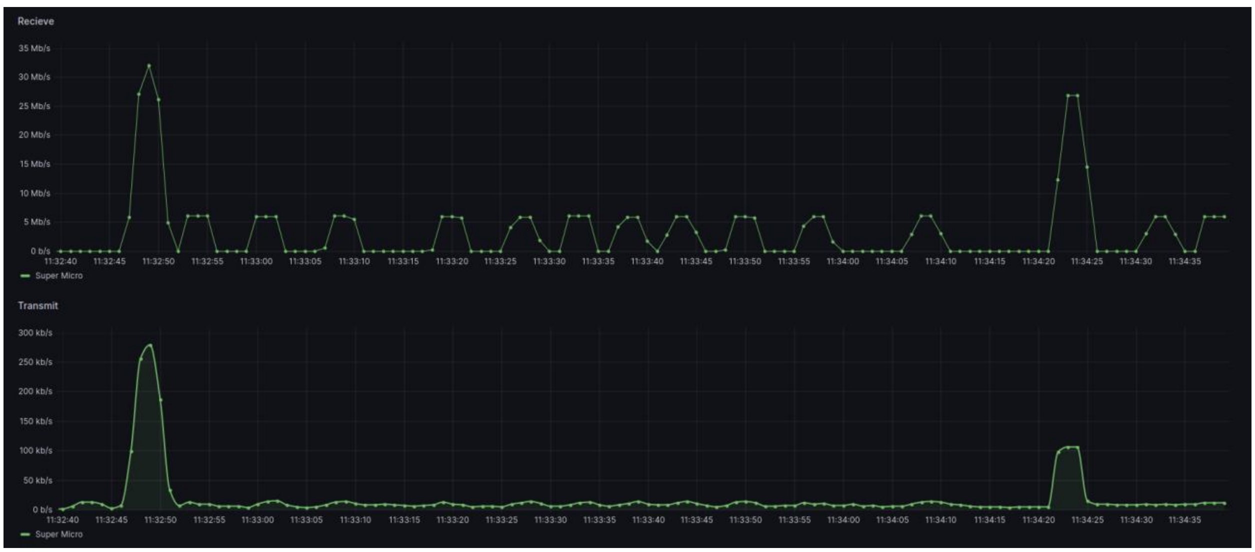


Figure 73 Streaming on Smartphone(5G-UE) from LDN

Figure 74 Watching YouTube Video by smartphone (5G-UE)

Figure 75 Watching Live Streaming online (Internet) by smartphone (5G-UE)

6.5. Final remarks

The importance of AI@EDGE platform in UC4 is multifaceted:

- Firstly, it involves integrating 5G communication with on-board infrastructure and connecting to ground-based 5Gcore, following the 3GPP 5GNTN specifications. Additionally, incorporating edge UPF on an aircraft facilitates the introduction of edge computing within the 5G and B5G network framework in the aviation industry. Utilizing real LEO satellites, such as SpaceX Starlink, enables detailed measurements on various parameters, allowing to achieve the KPIs regarding the 5G-connectivity system.

- Secondly, the implementation of a cluster and the utilization of Virtualized Infrastructure Manager (VIM) and MEC Orchestrator (MEO) through the AI@EDGE platform provide complete control over on-board AI-based applications known as Artificial Intelligent Functions (AIFs). This development sets the stage for the next generation of InFlight Entertainment (IFE) systems.
- Thirdly, the incorporation of multi-link aggregation, achieved by enabling Multi-path TCP (MPTCP), demonstrates the simultaneous use of two wireless links (specifically WiFi and LiFi) for on-board communication, resulting in more robust and higher capacity communication that contributed in achieve the KPIs regarding the onboard connectivity system.

The integration of these three main concepts of the AI@EDGE platform within UC4 contributes to the SPI main product, InFlight Entertainment and Connectivity (IFEC) systems.

The future improvement could go mainly to the direction of using the best potentials of 5G RAN (such as MIMO and higher BW), the usage of MPTCP can be also adopted for 5G link which was not feasible within the timeframe of AI@EDGE project due to limitation of proper 5G dongles.

7. Conclusions

Within WP5, the tasks 5.1, 5.2, 5.3, 5.4, and 5.5 dealt with the integration, validation and benchmarking of the project's use cases. Starting from the initial testbed's configurations, the AI@EDGE platform was deployed on all the testbeds also integrating a complete 5G network which includes, in each testbed, several HW and SW components related to the radio, Edge, and core functionalities.

Achieving this result required a very complex system integration process driven by the specific requirements, constraints, and needs of each testbed, mainly due to the heterogeneity of the four use cases.

To meet the use cases requirements, the related applications were designed and developed to leverage the project's technological enablers also integrating the AIFs and data pipeline support, for their deploying at the edge. Having the validation procedures as a reference, in the final phases of the project a test campaign was carried out to collect results and evaluate KPIs. All the outcomes have been collected and presented in this deliverable highlighting scenarios, challenges, and objectives for each of the use cases.

Sections 3.5, 4.5, 5.5, and 6.5 report the final remarks, on a use case basis, on the obtained results explaining how AI@EDGE contributed to the advancement of each of the cases, creating a basis for further improvement and optimizations that can boost the impact to the relevant markets.

Furthermore, with the submission of deliverable D5.3, the milestone MS5.6 (Final validation completed) foreseen for M36 can be declared achieved.

8. References

[1] AI@EDGE project – Deliverable D5.1: "Testing and validation methodology, planning and preparation".

[2] AI@EDGE project - Deliverable D5.2: "Preliminary validation and use case benchmarking".

[3] AI@EDGE project - Deliverable D2.3: "Consolidated system architecture, interfaces specifications, and techno-economic analysis".

[4] AI@EDGE project - Deliverable D3.2: "Final report on Systems and Methods for AI@EDGE platform automation".

[5] AI@EDGE project - Deliverable D4.2: "Results on the validation of the AI@EDGE Connect-Compute Platform".

[6] Diamanti, A.; Vilchez, J.M.S.; Secci, S. LSTM-based radiography for anomaly detection in softwarized infrastructures. In Proceedings of the 2020 32nd International Teletraffic Congress (ITC 32) IEEE, Osaka, Japan, 22–24 September 2020; pp. 28–36. [**Google Scholar**]

[7] Intelligent Transport Systems (ITS): Access layer specification for Intelligent Transport Systems operating in the 5 GHz frequency band - https://www.etsi.org/deliver/etsi_en/302600_302699/302663/01.02.00_20/en_302663v010200a.pdf