H2020 5G-TRANSFORMER Project

Grant No. 761536

5G-TRANSFORMER
Report on trials results

Abstract

This deliverable provides a final report on the five vertical proofs of concept (PoCs) associated to the different UCs from the different vertical industries that are part of the project: Automotive, Entertainment, E-Health and E-Industry and the MNO/MVNOs UCs. The PoC have been conducted to demonstrate and validate the benefits of adopting the integrated 5G-TRANSFORMER architecture components designed and developed in the context of the project.

Document properties

Document number	D5.4
Document title	5G-TRANSFORMER Report on trials results
Document responsible	Francesco D'Andria (ATOS)
Document editor	Francesco D'Andria (ATOS)
Editorial team	Francesco D'Andria (ATOS), Ignacio Domínguez (ATOS)
Target dissemination level	Public
Status of the document	Final
Version	1.0

Production properties

Reviewers	Xi Li (NECLE), Thomas Deiss (NOK-N), Carlos J. Bernardos (UC3M)

Disclaimer

This document has been produced in the context of the 5G-TRANSFORMER Project. The research leading to these results has received funding from the European Community's H2020 Programme under grant agreement № H2020-761536.

Table of Contents

List of Contributors

Partner Short Name	Contributors
UC3M	Borja Nogales, Winnie Nakimuli, Carlos J. Bernardos
TEI	Erin Seder, Fabio Ubaldi
ATOS	Jose Enrique Gonzalez, Francesco D'Andria, Ignacio Domínguez
BCOM	Olivier Choisy, Cao-Thanh Phan
NXW	Giada Landi, Juan Brenes
CRF	Aleksandra Stojanovic, Marina Giordanino
CTTC	Jordi Baranda, Luca Vettori, Josep Mangues, Ricardo Martínez, Manuel Requena, Ramon Casellas, Engin Zeydan, Javier Vílchez
POLITO	Carla Fabiana Chiasserini, Giuseppe Avino
SSSA	Luca Valcarenghi, Koteswararao Kondepu, Francesco Paolucci, Silvia Fichera
NOK-N	Thomas Deiss
NECLE	Andres Garcia Saavedra, Xi Li, Josep Xavier Salvat
IDCC	Giovanni Rigazzi, Charles Turyagyenda

List of Figures

List of Tables

List of Acronyms

Acronym	Description
5G-PPP	5G Public Private Partnership
5GT	5G-TRANSFORMER Project
5GT-MTP	5G-TRANSFORMER Mobile Transport and Computing Platform
5GT-SO	5G-TRANSFORMER Service Orchestrator
5GT-VS	5G-TRANSFORMER Vertical Slicer
AGV	Automated Guided Vehicle
AV	Audio and Visual content
A-COV	Availability related to coverage
CAM	Cooperative Awareness Messages
CIM	Cooperative Information Manager
CR	Cloud Robotics
CST	Infrastructure Cost
DC	Data Center
DEN	Device density
DENM	Decentralized Environmental Notification Message
DoA	Description of Action
E2E	End to End
EPC	Evolved Packet Core
EVS	Extended Virtual Sensing
HSS	Home Subscriber Server
HW	Hardware
KPI	Key Performance Indicator
LAT	End-to-end (E2E) latency
LTE	Long-Term Evolution
MANO	Management and Orchestration
MEC	Multi-access Edge Computing
MME	Mobility Management Entity
MNO/MVNO	Mobile Network Operator / Mobile Virtual Network Operator
MOB	Mobility
MPD	Media Presentation Description
MPEG	Moving Picture Experts Group
NFV	Network Function Virtualization
NFVI-PoP	Network Function Virtualization Point of Presence
NFV-NS	Network Service
NFVO	NFV Orchestrator
NSD	Network Service Descriptor
OPEX	Operational Expenditure
PoC	Proof-of-Concept
POS	Position accuracy
QoE	Quality of Experience
RAN	Radio Access Network
REL	Reliability
RNIS	Radio Network Information Services
RTT	Round Trip Time
RSU	Road side unit
SDI	Serial Digital Interface
SER	Service creation time
SGi	Service Gateway interface

SLA	Service Level Agreement
SW	Software
TRA	Traffic type
UC	Use Case
UDR	User data rate
UE	User Equipment
UHD	Ultra-High Definition
vCDN	virtual Content Distribution Network
vEPC	virtual EPC
VNF	Virtualized Network Function
VOD	Video On Demand
VPN	Virtual Private Network
WAN	Wide Area Network
WIM	WAN Infrastructure Manager
WP1	5GT Work Package 1
WP2	5GT Work Package 2
WP3	5GT Work Package 3
WP4	5GT Work Package 4
WP5	5GT Work Package 5

Executive Summary and Key Contributions

The main objective of the WP5 is to integrate the 5G-TRANSFORMER platform developed in WP2, WP3 and WP4 together with the different technologies available at the different 5G-TRANSFORMER test-bed sites and the Proof of Concepts (PoCs) established for the different use cases of the project [1], [2].

With this goal in mind, this deliverable provides a final report on the five vertical proofs of concept (PoCs) associated to the different UCs from the different vertical industries that are part of the project: Automotive, Entertainment, E-Health and E-Industry and the MNO/MVNOs UCs.

The PoC have been conducted to demonstrate and validate the benefits of adopting the integrated 5G-TRANSFORMER architecture components designed and developed in the context of the project.

The PoCs are used to evaluate whether the solutions developed for the 5G-TRANSFORMER framework achieve the Key Performance Indicators (KPI) expected by the considered verticals. These solutions are compared to the state of the art (benchmark) or the used ones in common practice to evaluate the performance gain achieved in terms of KPIs.

A preliminary description of the experiments' realization and evaluation has been already provided by the deliverable D5.3 [1]. This deliverable completes and improves this first work by developing and implementing refined evaluation procedures.

Consequently, those measurements have been achieved by demonstrating the 5 Vertical PoC in relevant scenarios as follows:

- The Automotive PoC has been demonstrated in Orbassano, in the CFR Test Area (Italy) and with real-cars. The PoC demonstrated the automatic deployment and scalability of the EVS/video streaming service application on edge according to priority and rules managed by arbitration and scaling functionality managed at the SO level according to monitored resources. In D5.4 the PoC assessed the Density (number of vehicles in a considered area) and Latency (from generating and sending the CAM by the vehicle, to receiving back the DENM message) KPIs.
- The Entertainment PoC has been validated in an international sport event such as The Mutuactivos Open de España". There, the consortium showed how 5G-TRANSFORMER capabilities provided an entirely new way for consumers to interact with immersive media contents in the context of a large-scale sport event. An UHD video streaming service, virtualized as a 5G-T Vertical Slicer, was orchestrated by the 5G-T Orchestrator so to be deployed at the edge with an abstraction of network and compute configuration parameters. For the Entertainment use case, there are three KPIs (Latency, User data rate and Service creation time) considered and measured. The values are compared with the current state-of-the-art showing a clear improvement by adopting the 5G-TRANSFORMER technologies.
- The E-Health PoC focused on the automation of emergency support deployment of medical services, reducing the overall reaction time, as well as providing support to AR services. The demo has been performed at 5TONIC premises, in a TRL6 event involving key business people from Telefonica,

Ericsson as well as regional government and representatives of health-related institutions. The PoC was set on two testbed sites: 5TONIC and CTTC trial sites so to show how 5G-TRANSFORMER technologies enable the management of distributed federated services and on the edge. In the context of KPIs assessment, the Service Creation Time has been evaluated compared with the benchmark and showing a clear improvement when the 5G-TRANSFORMER capabilities are adopted as well as the deployment on the edge.

- The E-Industry PoC has been demonstrated in a lab environment at 5TONIC lab in Spain. The PoC demonstrates factory service robots and production processes that are remotely monitored and controlled in the cloud, exploiting wireless connectivity (5G). The objective of the demonstrator was to verify the allocation of suitable resources based on the specific service requests to allow the interaction and coordination of multiple (fixed and mobile) robots controlled by remote distributed services, satisfying strict latency and bandwidth requirements. Three KPIs have been evaluated (Latency, Reliability and Service Creation Time).
- The MNO/MVNO PoC has been demonstrates in a lab testing environment. The PoC assessed the deployment of 3 network slices (echographer (URLLC), video (eMBB) and IoT devices (mMTC)). Two KPIs was selected for the MNO/MVNO PoC: Service Creation Time and Infrastructure Cost.

1 Introduction

This deliverable describes the validation and evaluation activities of the 5G-TRANSFORMER technology components that have been designed and developed in the 5G-TRANSFORMER Work Packages WP1, WP2, WP3 and WP4 and finally integrated in the context of the WP5, through the implementation and delivery of five vertical use cases deployed in the testbed set up in the context of activity T5.1 "Definition and set up of vertical testbeds".

The five PoCs have been implemented and showcased in dissimilar relevant industrial contexts: Automotive, Entertainment, E-Health and E-Industry and MNO/MVNOs.

The experiments have been conducted to demonstrate and validate the benefits of adopting the integrated 5G-TRANSFORMER architecture components designed and developed in the context of the project.

The evaluation has been performed, in two cycles, through POCs demonstrated in the 5G-TRANSFORMER testbed and via simulations.

The PoCs considered in the performance evaluation are: Extended Virtual Sensing (EVS) for Automotive, On-site Live Experience (OLE) and Ultra High-Definition (UHD) for Entertainment, a heart-attack emergency use case for E-Health, cloud robotics for E-Industry, and Network as a Service (NaaS) for MNO/MVNO. Therefore, the PoCs are used to evaluate whether the solutions developed for the 5G-TRANSFORMER framework achieve the Key Performance Indicators (KPI) expected by the considered verticals.

These solutions are compared to the state of the art (benchmark) or the used ones in common practice to evaluate the performance gain achieved in terms of KPIs.

A preliminary description of the experiments' realization and evaluation has been already provided by the deliverable D5.3 [1].

This deliverable D5.4 completes and improves this first work by developing and implementing refined evaluation procedures.

Following this introductory section there are three main parts to this deliverable:

- **Section 2: KPIs Overview** – provides an overview of the KPIs selected in context of the 5G-TRANSFORMER project to assess and validate the presented technology.
- **Section 3: Trials, experiments and measurements results** – describes the relevant scenarios where the PoCs have been demonstrated, the KPIs selected to assess the 5G-TRANFORMER technologies and the methodology used to make these measurements.
- **Section 4: Conclusion** - summarizes the main points of the present document, discusses the connections between the findings of each PoC and assessment and makes recommendations for future research and practice.
- **Appendix** - gives an overview of an additional use case jointly developed by 5GT-TRANSFORMER and 5G-CORAL [1]: 360º Immersive Telepresence:

[1] http://5g-coral.eu/

Remote Robotic Control PoC. 5G-CORAL[2] is an H2020 initiative featuring European and Taiwanes partners with major focus on the convergence between radio access interfaces and edge/fog computing technologies.

[2] http://5g-coral.eu/

2 KPIs Overview

Previous WP5 deliverables D5.2 [2] and D5.3 [1] provided an overview about the KPIs considered in the project and the approach used to validate them in the different PoCs. This section reports which KPIs measurement have been refined in the different PoCs and to which 5GPPP contractual KPIs they are contributing. Moreover, the relationship between the KPIs utilized in the project and the KPIs defined by 5GPPP Technical Board and by the 5GPPP-Test, Measurement and KPIs Validation (TMV) working group 5GPPP-TMV WG is also reported.

2.1 Contribution of PoCs to 5G-PPP Performance KPIs

D5.3 defined the KPIs considered in the project whose definition is reported here in Table 1 for convenience. However, for each vertical the KPI definitions might slightly differ because verticals measure them at their application layer.

TABLE 1: KPIS CONSIDERED IN 5G-TRANSFORMER

KPI	Acronym	Description
End-to-end (E2E) latency	LAT	E2E latency, or one way trip time (OTT) latency, refers to the time it takes from when a data packet is sent from the transmitting end to when it is received at the receiving entity, e.g., internet server or another device.
Reliability	REL	Refers to the continuity in the time domain of correct service and it is associated with a maximum latency requirement. More specifically, reliability accounts for the percentage of packets properly received within the given maximum E2E latency (OTT or RTT depending on the what is considered by the service).
User data rate	UDR	Minimum required bit rate for the application to function correctly.
Availability (related to coverage)	A-COV	The availability in percentage (%) is defined as the ratio between the geographical area where the Quality of Experience (QoE) level requested by the end-user is achieved and the total coverage area of a single radio cell or multi-cell area times 100.
Mobility	MOB	No: static users Low: pedestrians (0-3 km/h) Medium: slow moving vehicles (3-50 km/h) High: fast moving vehicles, e.g. cars and trains (>50 km/h)
Device density	DEN	Maximum number of devices per unit area under which the specified reliability is achieved.

Positioning accuracy	**POS**	Maximum positioning error tolerated by the application, where a high positioning accuracy means a little error.
Confidentiality	**CON**	Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information.
Integrity	**INT**	Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity
Availability (related to resilience)	**A-RES**	Ensuring timely and reliable access to and use of information
Traffic type	**TRA**	Depending on the amount of data moving across a network at a given point of time, traffic can be: • Continuous • Bursty • Event driven • Periodic • All types
Communication range	**RANG**	Maximum distance between source and destination(s) of a radio transmission within which the application should achieve the specified reliability.
Infrastructure	**INF**	• Limited: no infrastructure available or only macro cell coverage. • Medium density: Small number of small cells. • Highly available infrastructure: Big number of small cells available.
Energy reduction	**NRG**	Reduction of the energy consumption of the overall system. The most common metric that is used to characterize this KPI is the reduction in the consumed Joules per delivered bit.
Cost	**CST**	Expenditure of resources, such as time, materials or labour, for the attainment of a certain Hardware (HW) or Software (SW) module.

		Operational Expenditure (OPEX) and Capacity Expenditure (CAPEX) are important components of the overall costs.
Service creation time	**SER**	Time required to provision a service, measured since a new service deployment is requested until the overall orchestration system provides a response (a positive response implies the service has been actually provisioned).

Table 3 summarizes how the KPIs evaluated in the different PoCs are contributing to the objectives related to the 5G-PPP KPIs. With respect to a similar table (i.e., Table 4) reported in D5.3 [1] that we report here in Table 2 for convenience, here only the KPIs for which new updated evaluation has been conducted, both through simulation and experimentally, are reported.

TABLE 2: CONTRIBUTIONS TO THE 5G-PPP PERFORAMNCE KPIS BY THE CONSIDERED POCS

Use Cases						
5G-PPP KPIs		Automotive	Entertainment	E-Health	E-Industry	MNO/MVNO
	P1	TRA				
	P2	5GT-MTP Placement Algorithms				
	P3		SER	SER	SER	CST,SER
	P4	LAT, REL	LAT	LAT, REL	LAT, REL	
	P5	MOB, DEN	A-COV	A-COV, DEN, POS		

TABLE 3: CONTRIBUTIONS TO THE 5G-PPP PERFORMANCE KPIS BY THE CONSIDERED POCS[3]

		Use Cases				
5G-PPP KPIs		Automotive (Sec. 3.1)	Entertainment Sec. 3.2)	E-Health (Sec. 3.3)	E-Industry (Sec. 3.4)	MNO/MVNO (Sec. 3.5)
	P1	TRA	UDR			
	P3		SER	SER	SER	CST,SER
	P4	LAT, REL	LAT		LAT, REL	
	P5	MOB, DEN				

2.2 Relationship between considered KPIs and KPIs defined by 5GPPP working groups

The aforementioned KPI definitions slightly differ under different context of different vertical scenarios because verticals measure them at their application layer. However, the definitions are based on the 5GPPP-TMV white paper "network layer KPI definitions" [3] that we report here for reference. In the third column from the left it reported, when available, the relationship between the 5GPPP-TMV KPI name and the related 5GT KPI. However, quality of experience (QoE) and peak throughput have not been considered in 5GT.

TABLE 4: 5GPPP-TMV KPIS

Type	KPI name	Related 5GT KPI	KPI measurement points	5GPPP KPI Validated
SLA	Minimum Expected Upstream Throughput	UDR	UE transmitting IP packets to the N6 interface beyond the Next Generation Core (NGC) toward the public Internet.	P1
SLA	Minimum Expected Downstream Throughput	UDR	UE receiving IP packets from the N6 interface	P1
SLA	Maximum Expected	LAT	RTT of UE IP packets	P1, P4

[3] P1=Providing 1000 times higher wireless area capacity and more varied service capabilities compared to 2010. P3=Reducing the average service creation time cycle from 90 hours to 90 minutes; P4=Creating a secure, reliable and dependable Internet with a "zero perceived" downtime for services provision; P5=Facilitating very dense deployments of wireless communication links to connect over 7 trillion wireless devices serving over 7 billion people. They are defined in 5G-PPP, "Contractual Arrangement Setting up a Public Private Partnership in the Area of Advanced 5G Network Infrastructure for te Future Internet between the European Union and the 5G Infrastructure Association," December 2013, [Online]. Available: https://5g-ppp.eu/contract/ [Accessed 20 11 2019].

	Latency		transmitted to the N6 interface.	
	Network Reliability	REL	Transport layer packets are lost between the UE and the N6 interface	P4
	Quality of Experience	None	Measured at the UE side at application or application API level	P1, P4
	UL Peak Throughput	None	Single UE transmitting IP packets to the N6 interface.	P1
	DL Peak Throughput	None	Single UE receiving IP packets from the N6 interface	P1

Additionally, the 5GPPP Technical Board defined the service creation time by dividing it into five phases: “platform provision”, “onboarding”, “instantiate, configure, and activate”, “modify”, “terminate”. The 5GT project adopted a similar definition of service creation time by considering only the “instantiate, configure, and activate” phase because specific attention was focused on the slice activation phase.

3 Trials, experiments and measurements results

This section has a double objective. On one hand it aims at describing all technologies developed and implemented for the five PoCs integrated in the different testbeds”.

It also describes the revelant scenarios where the vertical PoC have been demonstrated. On the other hand, the section aims to evaluate the 5G-TRANSFORMER technologies developed in the different WPs (WP2, WP3, and WP4) by measuring vertical relevant KPIs through experimentations of the PoCs.

Therefore, through those measurements, the consortium assessed these technologies whether they meet the expected KPIs required by the proposed vertical PoCs by comparing them with the state of art solutions already proposed in the literature or used in the common practice to evaluate whether the 5G-TRANSFORMER platform is enhancing these KPIs.

In D5.3 [1], the consortium already performed an early evaluation analysis. This section delivers the final 5G-TRANSFORMER technologies assessment.

3.1 Automotive

3.1.1 Selected Proofs of Concept

The Automotive PoC demonstrates how the 5GT platform functionalities facilitate the correct execution of a vertical offering multiple services with different priorities running in parallel according with the available resources. In this case, the coexistent services are the Video Streaming and EVS (Extended Virtual Sensing) services, as described in D5.3 [1]. For an exhaustive provision of the KPIs, beyond the description of the field trial experiments (in section 3.1.2) that provide the latency measurement, it is reported in following sections the parallel activity of the simulations that allow the provision of density and mobility measurements.

3.1.2 Field Trial Experiments

The Automotive PoC has been developed incrementally, providing several releases, as described in D5.2 [2].

The final PoC demonstrated coexistance of the EVS and Video Streaming service applications at the mobile edge. It showcases the Vertical Service arbitration and life cycle management functionalities, introduced by the Vertical Slicer (5GT-VS), and the network service lifecycle management operations provided by the Service Orchestrator (5GT-SO) and how these services, residing in different PoPs are interconnected through the Mobile Transport and Compute Platform (5GT-MTP). Figure 1 shows the scenario setup for the Automotive final demo:

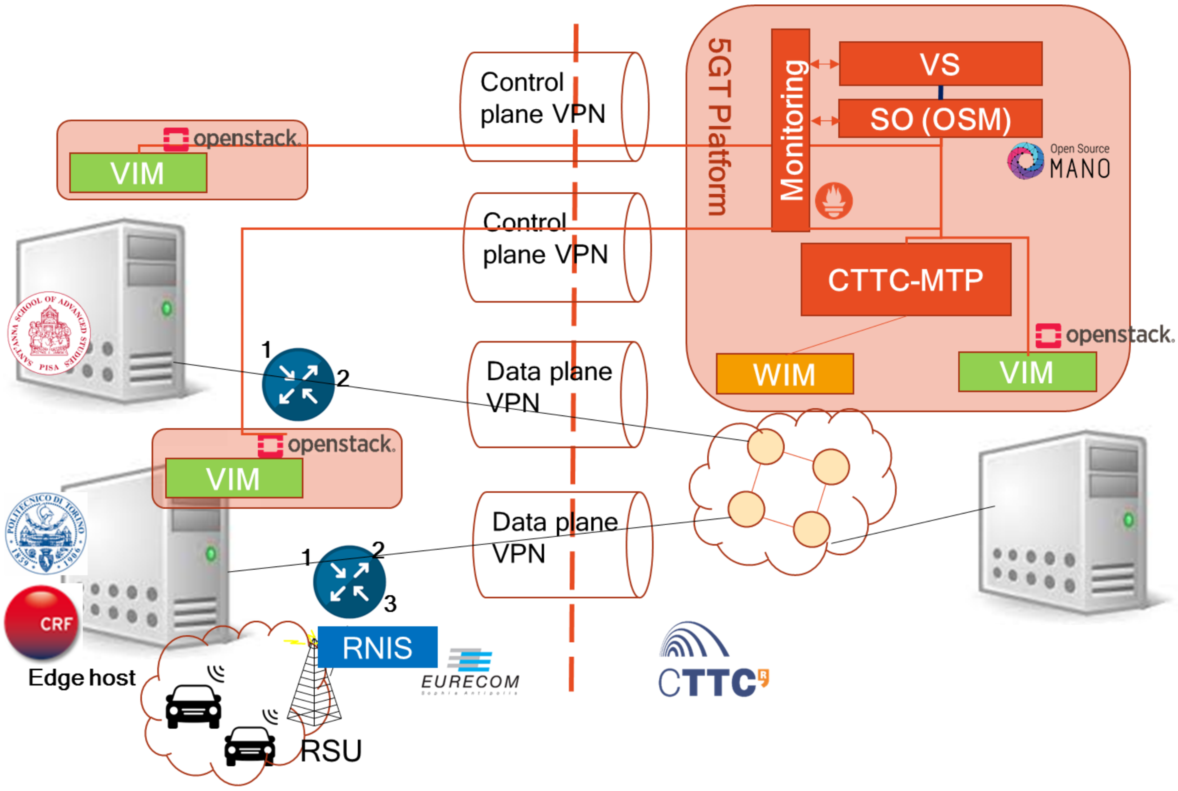

FIGURE 1: SCENARIO SETUP FOR THE AUTOMOTIVE DEMO

Two testbed sites are used for the final demo, (i) CTTC Site where the whole 5GT Stack (5GT-VS, 5GT-SO, 5GT-MTP and 5GT-MON) is deployed, (ii) and Italian site (ARNO and CRF/Polito) where the services are running. The workflow of the demo is:

1. Instantiation of two instances of video streaming service;
2. Instantiation of EVS with higher priority;
   Handle of priority by the Arbitration
   Scaling down of video streaming managed by arbitrator (only one VS instance instead of 2);

The first part of the demo shows how an automotive vertical can use the 5GT platform to instantiate two different video streaming services, just by providing high-level service parameters (and without any knowledge of the underlying infrastructure). The first service instance consists of a single VNF (VM2 in Figure 2) implementing the video server based on HTTP streaming (containing the video catalogue, a front-end, Media Presentation Destription (MPD) files and the media chunks). The second service instance consists of two VNFs: one VNF is VM2, as previously mentioned, while the other VNF (i.e. VM1) implements a video streaming controller including an optimization algorithm and a radio link manager as shown in Figure 3. The optimization algorithm uses the Radio Network Information Services (RNIS) values fetched by the radio link manager to optimize the video streaming rate. Then it notifies the allowed bit rates to the video server (VM2) Front-End that edits the MPD file with the bit rates that a user can request. For more details the reader is referred to [4].

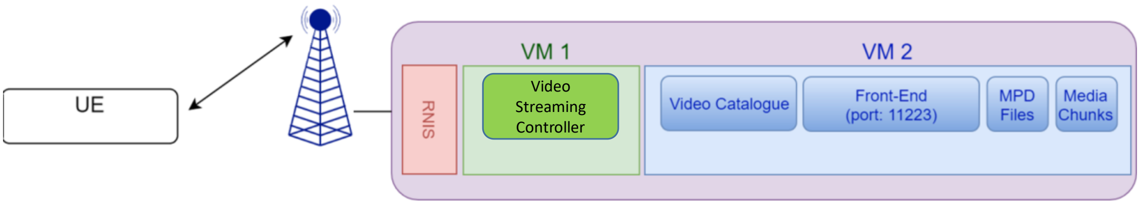

FIGURE 2: TWO INSTANCES OF VS

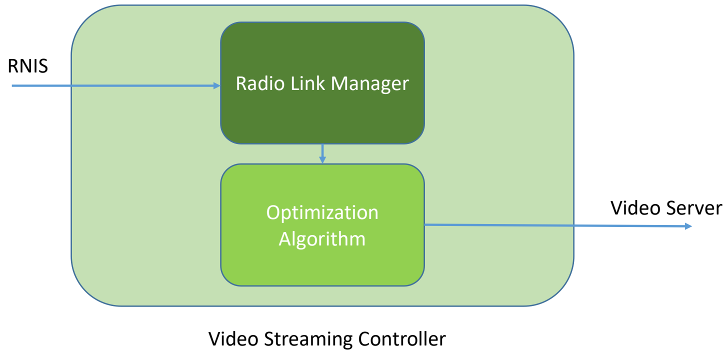

FIGURE 3: VIDEO STREAMING CONTROLLER BUILDING BLOCKS

The second part demonstrates the arbitration capabilities introduced by the 5GT-VS and the deployment of an Extended Virtual Sensing (EVS) service. For this, the Vertical requests the instantiation in the MEC (in order to meet the latency requirements) of an EVS which has a higher priority than the VS service. The SLA specified in the demo for the Vertical does not have enough resources in the MEC to accommodate the three service instances concurrenly, therefore, the arbitration module at the 5GT-VS forces the termination of the video streaming instance with the rate adaptation mechanism (i.e., VM1 and VM2) in order to release resources for the EVS. During this phase, there are two real cars which generate the information (Cooperative Awareness Message (CAM)) to feed the EVS service and, afterwards, process the incoming commands (Distributed Environment Notification Message (DENM) to demonstrate the correct operation of the service.

Hence, besides showing the correct video streaming provisioning, this second part of the demo also demonstrates that the EVS successfully meets the safety requirements. In particular, the two cars send their CAMs to the CIM; the information is delivered and then processed by the EVS alghorithm; an Alert (DENM) is generated and transmitted towards the cars. One of the cars is equipped with the Automatic Emergency Braking (AEB) system and brakes upon receiving the Alert, thus avoiding collision.

Additionally, the time at which the CAM and DENM are transmitted and received are recorded (respectively) and the obtained latency is demonstrated.

3.1.3 Considered KPI(s) and benchmark

For the Automotive use case, in D5.3 [1] are presented three measured KPIs (latency, reliability and density) and the values are compared with the current state-of-the-art.

3.1.4 Measurement Methodology

Since we have already reported the measuerement methodologies used for the different phases of the Automotive PoC (in D5.3 [1]), in the following, we only report the testings results that have not been reported yet.

3.1.4.1 Density

Simulations were performed to assess the maximum number of vehicles that would trigger the scaling of the EVS service: During the SUMO simulation (replicating the scenario with 2 intersections), the number of vehicles approaching the intersection was increasing (following Poisson distribution with different λ). In order to guarantee very low latency of the service a continuous monitoring was performed by the 5G-TRANSFORMER Platform.

3.1.4.2 Latency

Measured the average, standard deviation, maximum and minimum values of the e2e latency of the EVS service, from the transmission of a CAM that triggers a collision detection to the DENM reception. The values are computed based on the experimental trials carried out on field during the demo, using real vehicles and a MEC-edge implementation of the service.

During the field trial experiments, we measured the latency of the EVS service, as described above. It is worth stressing that the system is composed of all open-source software components.

The last measurements that will be done for the Automotive Use Case will be mobility measurements done with emulated cars.

3.1.5 Assessment Results

Vehicle density. After the initial measurements done during the initial stages of PoC development (in D5.3, where we have reported that in order to guarantee a very low latency KPI, the overall EVS processing time needs to be below a threshold set to 5 ms), we defined the scaling threshold, i.e. the number of vehicles that should trigger the scaling of the EVS service. As explained in section 3.1.4.1, we had an increasing number of vehicles approaching the intersection, see Figure 4.

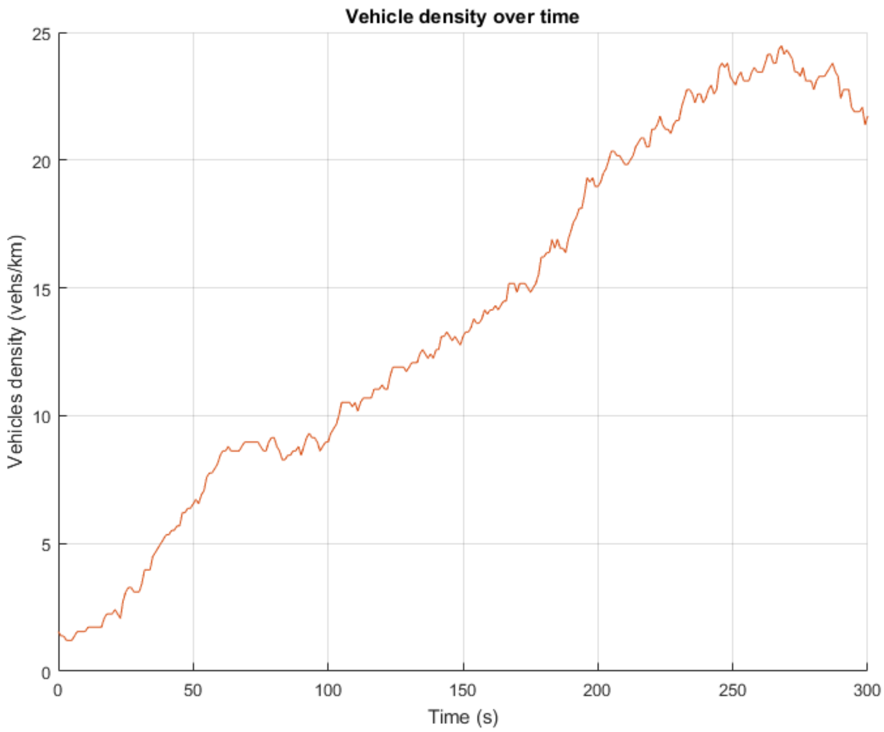

FIGURE 4: SIMULATION OF VEHICLE DENSITY[4]

After the simulation, we have evaluated the maximum vehicle density in which the processing time is below a set threshold (5 ms) for the 99,9% of time. The density is strictly related to the CPU consumption of the EVS VM, therefore, by monitoring the CPU consumption scaling can be triggered in cases where the number of vehicles is too high to maintain the processing time below the 5 ms. The 5G-TRANSFORMER monitoring platform, located at CTTC, allows monitoring of the CPU consumption and automatically triggers the scaling by instantiating a new EVS VM.

[4] Vehicles/km is used as unit when measuring the density to better reflect that vehicles are located in linear roads.

Figure 5 and Figure 6 show the results obtained during the simulations.

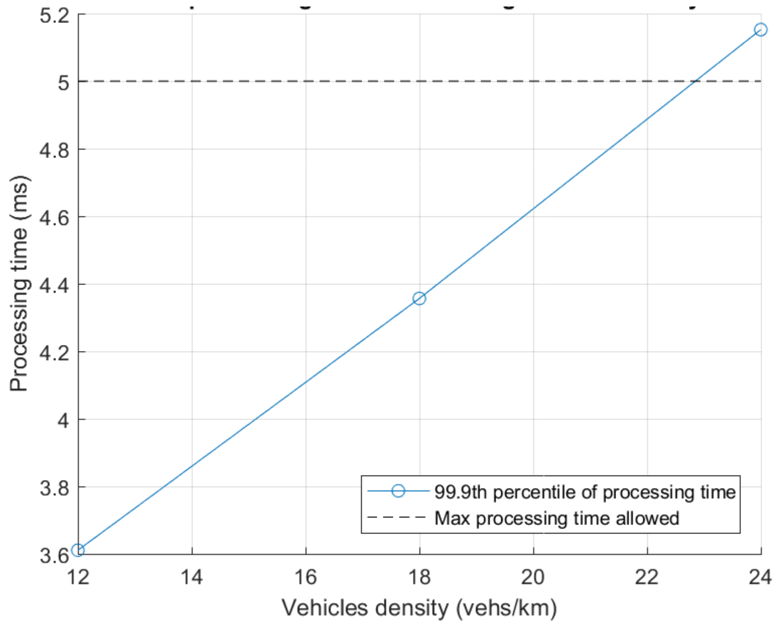

FIGURE 5: EVS PROCESSING TIME AT INCREASING VEHICLE DENSITY.

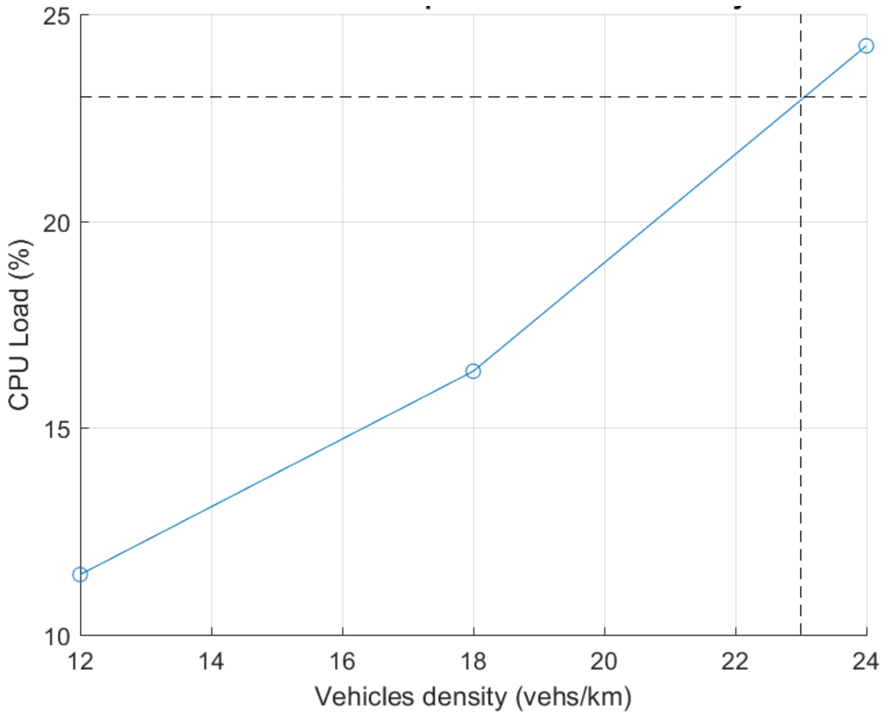

FIGURE 6: EVS CPU CONSUMPTION VS VEHICLE DENSITY

Analyzing the results, we have defined the CPU load threshold after which the performance is not guaranteed and scaling is necessary.

Latency. Figure 7 shows the e2e latency values measured for the EVS traffic, from the transmission of the CAM to the reception of the corresponding DENM. Also the following statistics were derived: average value = 8.870 ms, standard deviation dev: 1.447 ms, maximum value: 11.637 ms, and minimum value: 5.050 ms. As demonstrated by these results, the 5G-T architecture and the implementation of the

EVS in the MEC-edge fulfil the vertical requirements and actually provide a latency performance that is well below the 20 ms required to leverage the information obtained through V2I communication and merge this with other data detected through sensors aboard the vehicle .

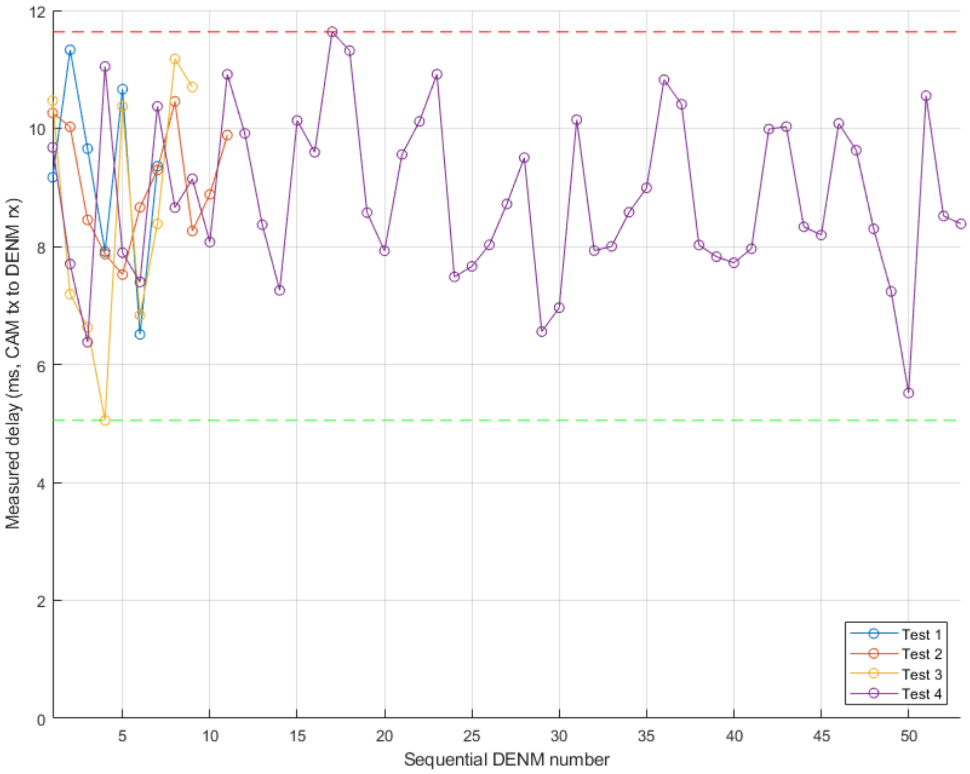

FIGURE 7: E2E LATENCY OF THE EVS SERVICE: EXPERIMENTAL RESULTS OBTAINED DURING THE DEMO WITH REAL VEHICLES

For the video streaming service in the automotive demo, the capability of the algorithm to adapt to the Channel Quality Indicator (CQI) has been evaluated. The performance evaluation parameter is the video streaming throughput defined as the bitrate of the transmitted video segments. Figure 8 shows that if the video streaming service featuring video adaptation is utilised the delivered segment size bitrate increases as a function of the CQI increase. Indeed, higher resolution segments are sent if a higher CQI (i.e., higher capacity) is available in the radio link.

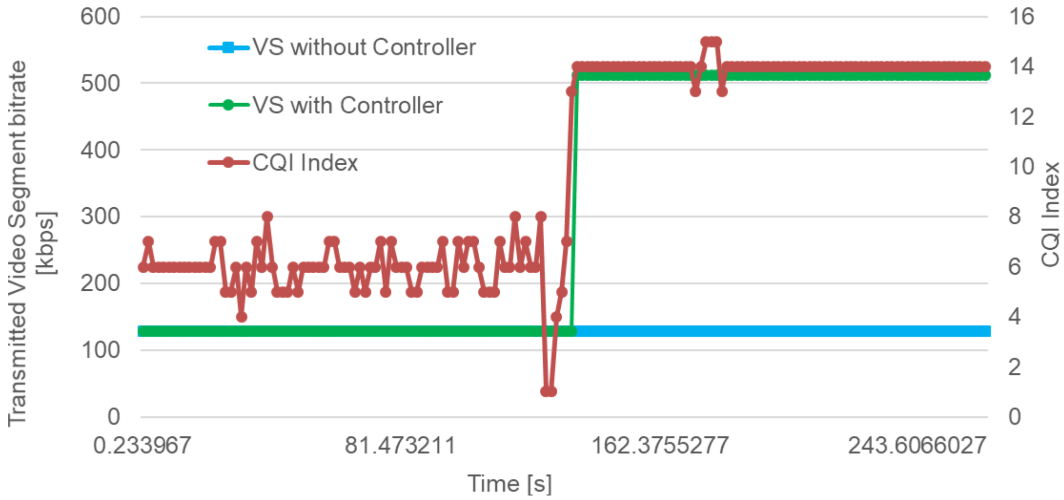

FIGURE 8: SEGMENT SIZE BITRATE AND CQI AS A FUNCTION OF TIME

3.1.6 A full-fledged MEC implementation and performance study of the EVS service

This section still focuses on the EVS service, but it presents a different implementation that has been realized using a full-fledged MEC platform. Since it aims at the evaluation of the EVS service after this has already been deployed, it does not exploit the full 5G-T architecture.

The system includes four main blocks, namely, (i) the MEC-enabled Evolved Packet-Core (EPC) Network; (ii) the procedures for service onboarding and instantiation within the MEC platform; (iii) the vehicle emulator; and (iv) the EVS and the CIM services running as MEC applications. Figure 9 provides an overview of the interactions between the building blocks. In the test-bed that has been developed, two instances of cellular User Equipments (UEs), based on Open Air Interface[5] (OAI), act as vehicles (although the same framework could be used with 802.11p V2I communications). Each UE periodically sends the information related to the position, speed, acceleration, and direction of several emulated vehicles towards a third party database, the CIM. In turn, the MEC-enabled EPC identifies the EVS traffic directed towards the CIM and applies traffic redirection to keep it at the edge. The EVS, which onboards the trajectory-based algorithm that detects approaching vehicles on a collision course, periodically retrieves the latest vehicle information received by the CIM. When needed, the EVS sends alerts towards the vehicles, exploiting again the same traffic redirection rules that the MEC-enabled EPC used for the uplink traffic.

MEC-enabled EPC. Our system builds on OAI, an open source implementation of a full LTE network, spanning the RAN and the EPC, with current developments focusing on 5G technology. On top of this, we have implemented a MEC platform, which exposes REST-based API endpoints to the MEO (Mobile Edge Orchestrator) and ME (Mobile Edge) applications, so that they can discover, register, and consume MEC services, including traffic redirection and, in our case, the EVS applications. We provide extensions to the OAI RAN and the core network elements to implement the Mp2 reference points. Core network extensions are necessary for traffic offloading to ME application instances, while specific support is needed at the RAN level for retrieving

[5] https://www.openairinterface.org

radio network information from eNBs, such as per-UE channel quality indications (CQI), and exposing them to subscribing ME applications. The Mp2 interface towards the RAN is implemented using the FlexRAN protocol, which is integrated into the standard OAIsoftware distribution. The S/P-GW has been split into two entities: S/P-GW-C and S/P-GW-U. The former is in charge of managing the signalling to establish the user data plane, while the latter is in charge of forwarding the user plane data. In our implementation, the S/P-GW-U is based on a version of OpenVSwitch (OVS), patched to support GPRS Tunnelling Protocol (GTP) packet matching. When requested over its Mp1 interface, the MEP installs traffic rules on the S/P-GW-U to offload traffic to the MEC applications by remotely executing OpenFlow commands. The MEP needs to be aware of specific UE bearer information (UE IMSI, GTP tunnel endpoint identifiers, UE and eNB IP addresses) to appropriately install these rules. This information is available at the S/P-GW-C level upon UE bearer establishment, and we have modified the OAI EPC code to communicate it to the MEP via its REST Mp2 interface. In our MEC testbed, ME applications are running on the MEC host as VMs directly on top of the kvm hypervisor. However, our MEC platform is also compatible with VIMs such as OpenStack, while it has been tested with containerized ME applications managed by lxd. As shown in Figure 6, the OAI EPC is virtualized, with the HSS, MME, and SPGW running as separate kvm VMs on a single physical machine, which also hosts the MEP. Note that the latter can also be executed as a virtual instance on the MEC host. Due to its real-time constraints, the OAI eNB software runs on a dedicated host, to which a USRP B210 RF board is attached.

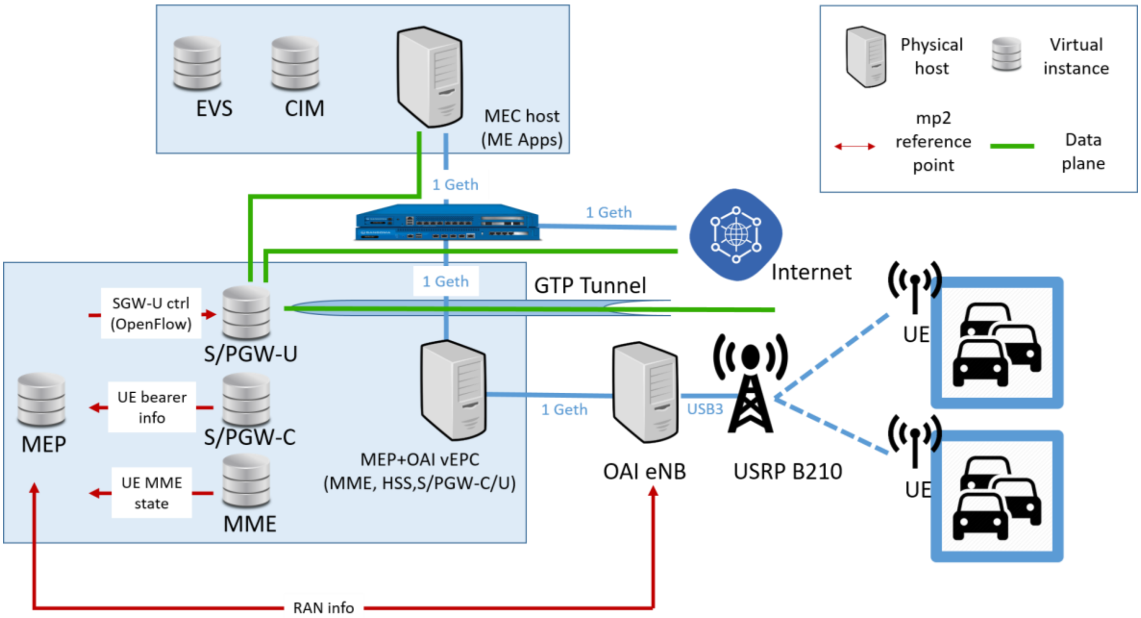

FIGURE 9: OVERVIEW OF THE FULL-FLEDGED MEC IMPLEMENTATION OF THE EVS SERVICE AND INTERACTIONS BETWEEN THE BUIDING BLOCKS

EVS as a MEC Application. The EVS MEC application is characterised by the application descriptor AppD, including a reference (URL) to the actual application image, application latency requirements, minimum requirements such as the amount of computing resources that should be allocated for an application instance, MEC services that the application exposes or consumes, and DNS rules and traffic filters. The latter ones define the characteristics of the traffic that should be offloaded to the MEC application instance (e.g., traffic flows matching a specific protocol-destination and address-port tuple). The vertical service provider submits an application package for onboarding to the OSS/BSS via the Customer Facing Service (CFS) Portal. The OSS/BSS then onboards the application package to the MEC system by communicating with the MEO over the standardized Mm1 reference point, thus making it available for instantiation. The vertical service is composed of two main components (CIM and EVS), in turn broken down into specific modules that can be run as independent services. We opted for a micro-services based implementation where each component is onboarded and instantiated separately. Standard-based service registration and discovery procedures are used so that the application components (including those aboard the vehicles) can discover and communicate with each other. Thus, the following two MEC application packages should be provided [5]:

- **CIM package**, integrating all the software related to the CIM, including both the CAM receiver and the Information Manager. The CIM features a modular design, with the CAM Receiver, Information Manager, and other components operating as separate, networked sub-services. Since one CIM instance is expected to be active per monitored area, without significant scaling requirements, we chose to package all its modules into a single application.
- **EVS package**, including the EVS manager, the collision detection algorithm, and the DENM Decider. Due to its scaling requirements, the collision detection algorithm is deployed as a separate application sub-package, whose instances can be scaled in/out independently from the others, when needed. The EVS manager and the DENM Decider, instead, are built into a single application sub-package.

Upon service instantiation, the MEO extracts appTrafficRules from the AppD and communicates with the MEPM via the Mm3 interface to apply them. The MEPM, in turn, accesses the MEP's traffic rules service (in our implementation, over a REST interface), and the latter eventually applies them to the S/PGW-U over the Mp2 reference point. This type of traffic steering builds on SDN and is transparent to the UE: CAMs are sent to the well-known IP address and port of the CIM, and the S/PGW-U offloads the traffic to the IP address/port of the MEC instance by applying packet-rewriting OpenFlow rules installed by the MEP. EVS instances, on the other hand, consume the CIM service via the Mp1 interface. When the CIM is instantiated, the MEO extracts the appServiceProduced field from the AppD. This field provides a description of the service endpoint exposed by the CIM, which the EVS Manager component needs to access to consume the input to the collision detection algorithm. Finally, the MEO adds the service to the MEP Service Registry.

Vehicle emulator. In our test-bed two UEs act as vehicles. The mobility traces describing the pattern of all emulated vehicles are obtained previously running the well known Simulation of Urban MObility (SUMO) tool[6]. We sample the mobility traces of each vehicle every 0.1 s and we record key information of vehicle movements, such as position, speed, acceleration, and direction. For each obtained sample, we create a CAM, which is transmitted towards the eNB of the OAI cellular network. The radio interface of the two UEs used in the test-bed exploits a standard OAI UE implementation. Each UE is emulated by a PC, equipped with an octa-core processor at 1.8 GHz and 16 GB RAM and connected to a USRP B210 RF board. Over-the-air

[6] https://sumo.dlr.de/index.html

communication is further improved by a pass-band filter, which reduces undesired interference at the receiver. Each UE also onboards the software for transmitting CAMs and receiving DENMs (the latter being alert messages sent by the EVS to the vehicles). This software, named VehicleSimulator, is a C++ standalone Linux application.

3.1.6.1 Experimental results

Two UEs emulate flows of vehicles traveling on the roads of a Manhattan road map. Vehicles traverse the scenario from north to south (or viceversa), and from east to west (or viceversa). Collisions happen only between vehicles crossing each other's path: no rear-end collisions are foreseen since we focus on the EVS service for collision avoidance at intersections. To simplify the DENM transmissions towards the pair of vehicles involved in a collision, we use one of the two UEs to emulate only vehicles in the north-south direction, while we use the second UE to emulate the presence of the vehicles travelling in the east-west direction. Finally, to evaluate how performance changes with the number of cars in the system, we consider three different values of vehicle density: (i) high, i.e., 20 vehicles/km, (ii) medium, i.e., 14 vehicles/km, and (iii) low, i.e., 7 vehicles/km. The inter- arrival times of vehicles into the system follows an exponential distribution, with mean set to the aforementioned values in the three different cases, respectively. For each vehicle density, we performed 5 different runs of 300 seconds each.

The end-to-end delay consists of three main components: (i) the network latency; (ii) the CIM storage and processing times; (iii) the EVS detection and DENM preparation times. The only difference between our EVS MEC implementation and an equivalent EVS cloud implementation is represented by the network latency. In order to account for the additional delay required by the traffic to reach the CIM in a cloud server, we performed two measurement campaigns. With our OAI UE, we first pinged the CIM in the MEC 10,000 times and collected the experienced network latency. Then, to evaluate the effect of traversing a real cellular EPC to reach a cloud server, we used a commercial smartphone to ping the Amazon datacenter closest to Turin (the location of our test-bed), i.e., the one in Paris.

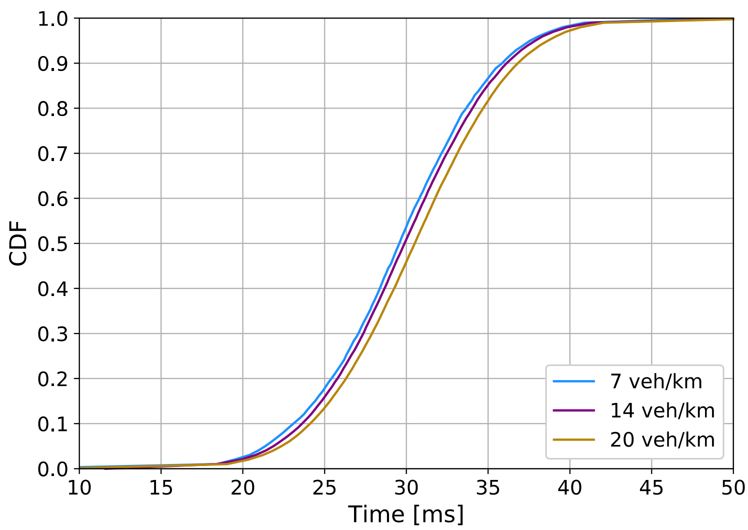

FIGURE 10: CDF OF THE END-TO-END LATENCY FOR VARYING VEHICLE DENSITY, IN THE FULL-FLEDGED MEC IMPLEMENTATION

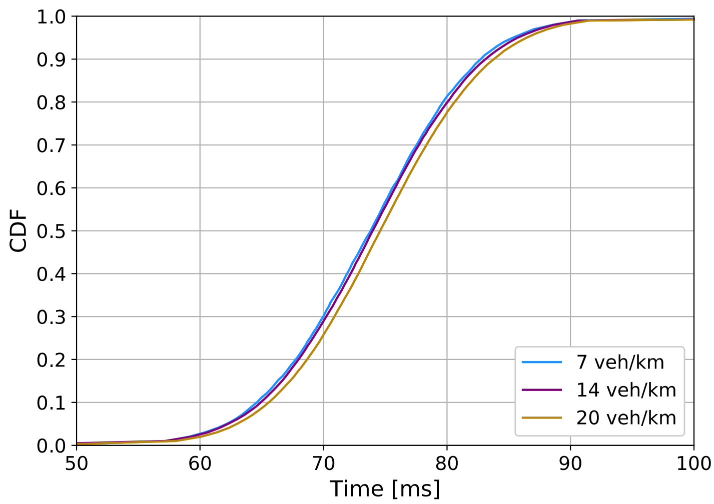

FIGURE 11: CDF OF THE END-TO-END LATENCY FOR VARYING VEHICLE DENSITY, IN THE CLOUD-BASED IMPLEMENTATION

Figure 10 and Figure 11 depict the experimental CDF of the end- to-end latency of the MEC and cloud-based implementations. Each curve is obtained considering all DENMs received by the vehicles in the 5 different runs performed for each vehicle density. In our tests, the average number of collisions is 91 for the high-density case, 44.6 for the medium density, and 11.2 for the low density. In all scenarios, in our MEC-based implementation, the 99.99% of the end-to-end latency values are below 50 ms. In particular, the average end-to-end latency is 29.55 ms for the low-density case, 29.89 ms for the medium density, and 30.5 ms for the high density. Such values are above 20 ms due to the contribution of the LTE radio; using 5G radio, the radio latency would dramatically reduce, leading to a performance that fully matches the vertical requirements. For our cloud-based implementation, instead, the end-to-end latency never drops below 50 ms: end-to-end latencies are on average 44 ms larger than the end-to-end latencies in the MEC-based implementation, which exactly corresponds to the network latency differences. In this case, even replacing the OAI interface with a 5G link, it would not be possible to fulfil the latency requirements.

To understand if the end-to-end latency achieved by our implementation is good enough, we take as a reference the cycle time of LIDAR sensors aboard vehicles. LIDAR sensors typically refresh their information every 60 ms and, for the information contained in the DENM to be coherent with on-board sensors, the maximum end-to-end latency should not exceed this value. Our MEC implementation is well within the cycle time of a LIDAR sensor, even for the worst-case end-to-end latency in the high density case. On the contrary, the cloud-based implementation of the EVS application is constantly violating the 60 ms bound, meaning that the car may act upon obsolete information. As a matter of fact, recently automotive companies are leaning towards an even more stringent end-to-end latency for the EVS applications, i.e., 20 ms. As shown above, such a latency is hardly achievable with 4G networks, even with the support of a MEC, but it can be obtained using 5G cellular networks. Given the fact that the total processing and communication times of our EVS and CIM VMs are well below 10 ms in the worst case (see Figure 12), we can conclude that our implementation is also consistent with such stringent latency requirement.

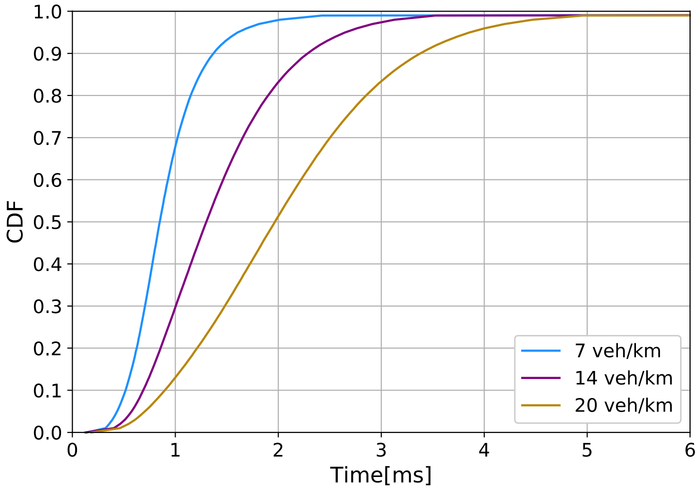

FIGURE 12: CDF OF THE EVS PROCESSING TIMES

Thanks to the SUMO error-log, we now check if our EVS application can detect all occurring collisions. The result of our evaluation is that our MEC-based application can alert in time all vehicles on a collision course, and that all crashes are avoided, under all vehicle densities. On the contrary, the cloud-based implementation cannot detect on time two of the collisions that appear in the SUMO trace under high vehicle density, for the reasons explained next. We first look at the number of alarms unnecessarily raised by the EVS service. The percentage of false positives is almost as relevant as that of correctly detected collisions, since a large number of unnecessary alerts may affect the drivers' trust in the application. When we look at false positives, results are consistent with those presented above.

Figure 13 shows the percentage of false positives obtained by the MEC and the cloud implementation of the EVS, while Figure 14 depicts the minimum distance between vehicles involved in situations that led to false positives. The additional latency suffered by the cloud version of our application causes a clear increase in the percentage of alerts directed toward vehicles that will not actually crash. Also, while in the MEC-based implementation the minimum distance between vehicles involved in false positive situations is always below 1m, in the cloud-based implementation, such a value doubles.

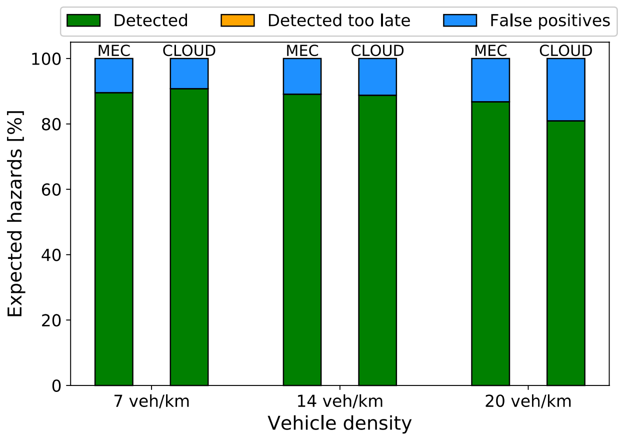

FIGURE 13: PERCENTAGE OF FALSE POSITIVES: MEC VS. CLOUD

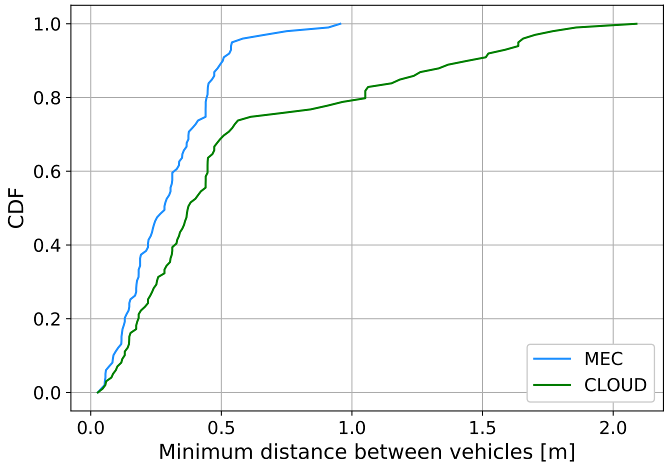

FIGURE 14: DISTANCES BETWEEN CARS INVOLVED IN FALSE POSITIVE DETECTIONS: MEC VS. CLOUD

In summary, our experiments show that the MEC is undoubtedly the key enabler for delay sensitive applications, such as our EVS, while cloud-based implementations cannot meet the automotive ultra-low latency requirements.

3.1.7 Simulation results with ns-3 LENA and SUMO

In addition to the experimental results presented above, we also wanted to evaluate the performance of the algorithm in a high variety of scenarios considering multiple speeds, vehicle densities, DENM actions taken, etc. This is why this section presents the results obtained for the defined KPIs in a joint ns-3 LENA+SUMO simulation ([8], [9], [10]).

3.1.7.1 Simulation setup

The architecture of the simulator used is presented in Figure 15. The LTE network is simulated as a single ns3 process. A collision detection client placed in UEs (i.e., cars) and a collision detection server is placed in a remote host outside the EPC, i.e., EVS traffic traverses the whole LTE RAN and EPC network to reach the server. The exit of the SUMO mobility simulator, through the TraCI server [11], is introduced in ns-3 to determine the mobility of the simulated cars, each running a TraCI client to receive this information.

SUMO offers a the TraCI client-server protocol to be able to extract simulated positions from SUMO to be used elsewhere, for instance, in the n-s3 simulator as is the case in our simulations.

CAM and DENM messages are ASN.1-encoded and sent through UDP/IP over LTE.

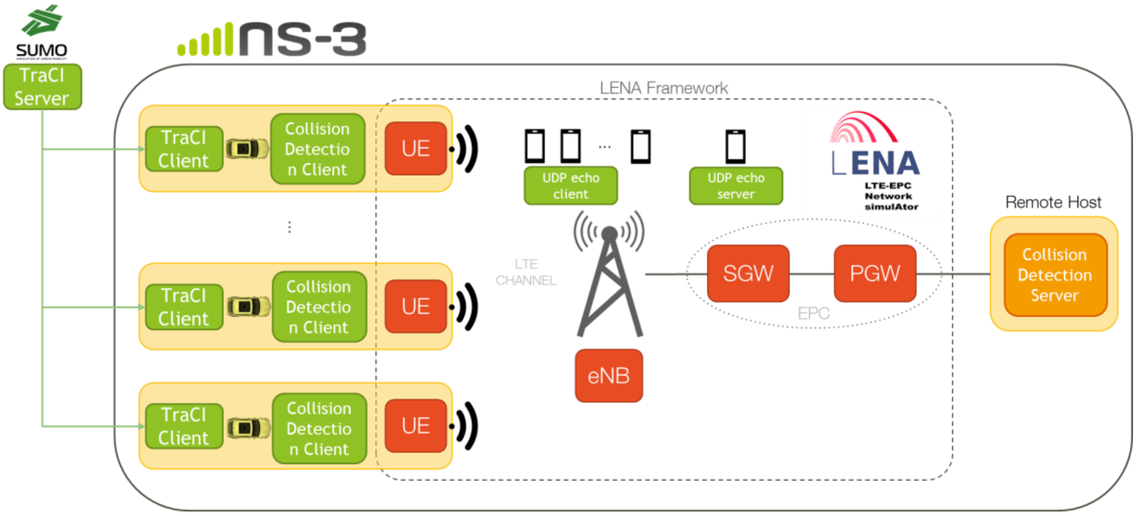

FIGURE 15: NS-3 LENA + SUMO SIMULATOR ARCHITECTURE

The following tables and figures present the relevant simulation pamaters used to obtain the results presented below as well as the simulated street layout with two crossings.

TABLE 5: SIMULATION PARAMETERS - SUMO

Parameters	Value	Description
Density	[2, 3, 4] veh/km[7]	Average number of vehicles per km
Generation Distibution	Poisson (exponentially distributed inter-arrival time)	Distribution for vehicle generation
Dimensions	W1.8 - L4.3	Vehicle dimension
Speed	[50, 75, 100] km/h	Vehicle’s max speed
Acceleration	4 m/s^2	Vehicle’s max acceleration
Deceleration	7.5 m/s^2	Vehicle’s max deceleration
speedFactor	1.5	The vehicles’ speed expected value multiplicator for lane speed limits. Used to make the mobility pattern more realistic.
speedDev	0.5	The deviation of the speedFactor.
Map size	3Km	Total length of the roads on the map
Map topology	2 intersections	See next slide
Simulation step	0.1s	Mobility updates

[7] Even if the density values may seem low, it is indeed realistic in the scenario under evaluation. Notice that these values correspond to densities for each of the streets in the scenario. Since in the simulated scenario there are three streets, the actual density per Km^2 in the global scenario is much higher. In our simulations, we have observed congestion in some crossings with these densities, which we believe makes the simulations realistic for a urban scenario. Higher densities were also simulated , but they are not presented because, in the scenario under evaluation, the high number of cars made the average speed globally decrease.

		frequency

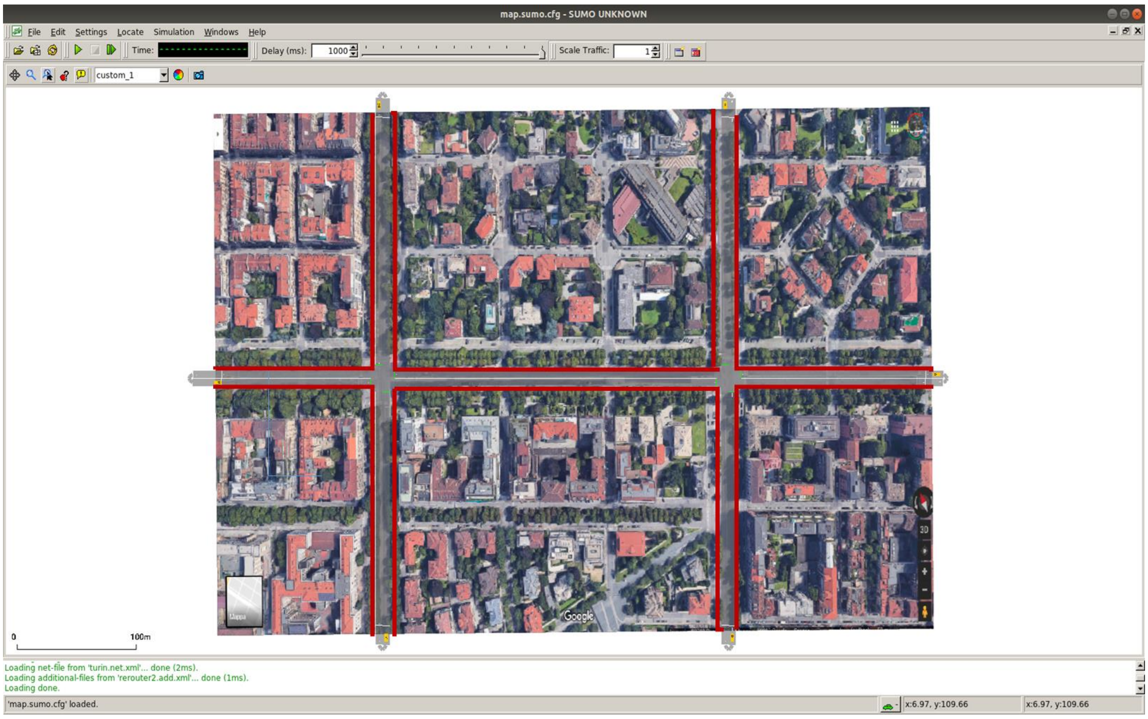

FIGURE 16: STREET LAYOUT SIMULATED IN SUMO (TWO INTERSECTIONS)

TABLE 6: NS-3 LENA PARAMETERS.

Parameters	Value	Description
Architecture	Single-process	Single ns3 process simulating all the communication layers
Communication model	LTE Channel	CAM and DENM delivered through the simulated LTE network
Transport-Network client-server	UDP-IP	UDP used over IP to create the communication sockets

TABLE 7: APPLICATION PARAMETERS.

Parameters	Value	Description
Message format	ASN.1	-
CAM frequency	10 Hz	CAM generation frequency (fixed)
CAM size	83B	Packet size at physical layer
DENM size	83B	
Human reaction time	1s	Delay between the DENM reception [at application layer] and start of driver's reaction

Reaction	[0, 1, 2, 3]*	Type of reaction after a DENM reception 0: Stop both vehicles and restart after a random period taken from U[stopping_time,stopping_time*1.5] 1: Only the leftmost vehicle stops and restarts after a random period U[stopping_time,stopping_time*1.5] 2: Only the slower vehicle stops and restarts after a random period U[stopping_time,stopping_time*1.5] 3: Only the farthest vehicle (from the crossing) stops and restarts after a random period U[stopping_time,stopping_time*1.5]

TABLE 8: SIMULATION RUN PARAMETERS.

Parameters	Value	Description
Number of simulations	20	Number of runs for each simulation parameter n-tuple
Simulation duration	300 s	Duration of each simulation run

3.1.7.2 KPIs/Metrics

The adaptation of the metrics defined in previous sections to the specificities of the automotive scenario are explained below:

LAT - End-to-end latency. Latency from the transmission of the CAM by the vehicle to reception of the DENM message

REL - Reliability. Fraction of correctly received DENM messages

TRA - Traffic. Data rate transmitted from and to the vehicles

MOB - Mobility. Performance of the EVS service (in terms of collisions avoided) as a function of mobility-related parameters (e.g., speed, density of vehicles).

DEN - Density: maximum number of vehicles in a considered area

3.1.7.3 The EVS system

The EVS system has been developed adopting a centralized architecture, as shown in Figure 6. The vehicles, acting as UEs, send their information through CAMs, specifying their position, speed, acceleration and heading.

The CAMs messages are sent unicast by each car (UE) to the remote host running the EVS service by traversing the eNB and the mobile core. The host running the EVS service gathers these messages and uses the information extracted to generate a trajectory of each vehicle.

The trajectories are then analyzed and, in case two vehicles are detected to be in risk of collision, a DENM is generated and sent to both the vehicles. As explained in Table 4, it is possible to configure the algorithm to take one of the 4 different actions when a collision is detected. Thus, the decision to stop one, the other, or both cars is taken at the server side, by looking at the information contained in the last CAMs received from the involved vehicles.

The decision of which car to stop is then encoded in a custom field included in the DENMs messages generated, called "Precedence" field, that is then read by the vehicles that will take the proper action.

3.1.7.4 Results

This section presents the results obtained in the two crossroads urban scenario for each of the relevant KPIs.

3.1.7.4.1 Traffic

Various graphs are presented for the uplink rate of CAM messages sent by the vehicles. The rate is expressed in bits per second (b/s). Since CAM messages are periodically sent every 100 ms, the traffic does not depend on the speed of the cars, but just on the number of cars in the scenario (i.e., density) as observed in the following graphs.

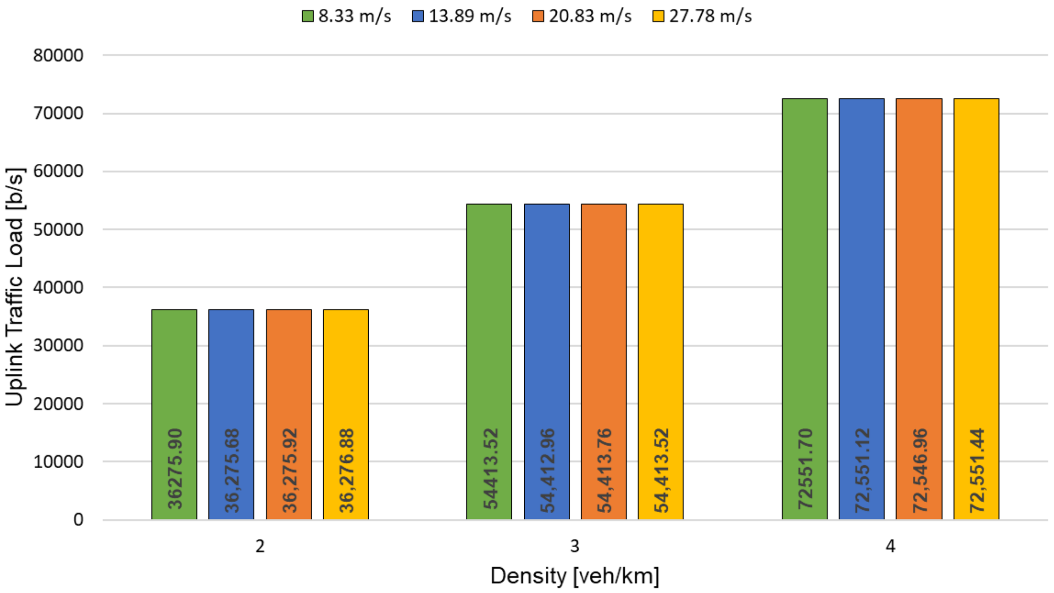

FIGURE 17: TRAFFIC IN UPLINK

As for downlink traffic, notice that downlink traffic is due to the transmission of DENM messages when the EVS algorihtm detects that there is a risk of collision based on the CAM messages received from cars.

This traffic depends on the DENM action and the speed of vehicles. The former is due to the fact that, by altering the behaviour of cars (e.g., stopping some of the vehicles at DENM reception) the number of situations with risk of potential collision also vary.

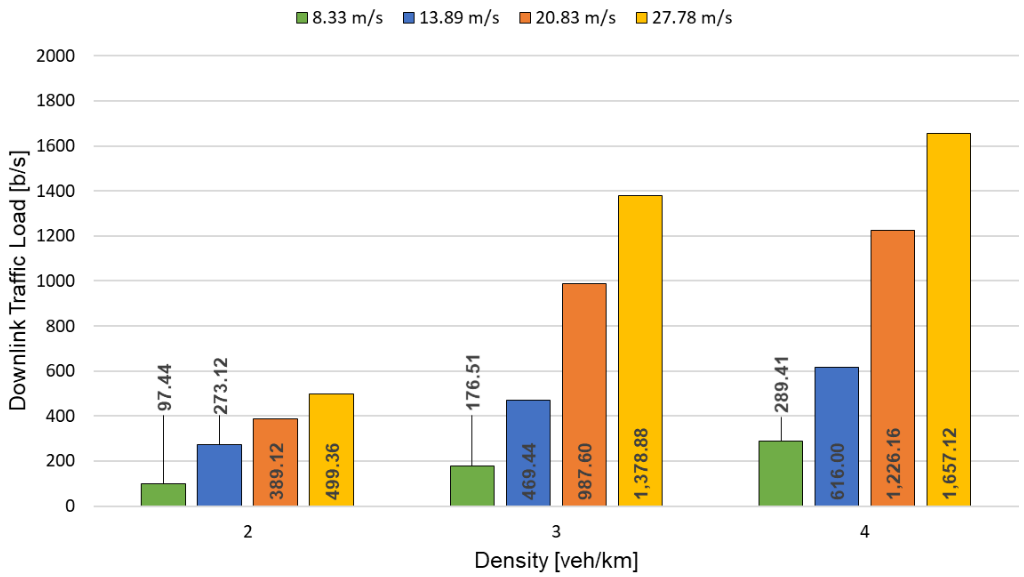

FIGURE 18: TRAFFIC IN DOWNLINK (DENM ACTION 0)

3.1.7.4.2 Latency

The tables below present the latency from the instant at which the CAM message is sent to the instant at which the consequent DENM message is received due to risk of collision. As it can be observed, its average only presents small variations and it is always below the targeted 20 ms rount-trip at the vertical collision detection application level. Furthermore, it depends neither on the speed of cars nor the car density. Multiple DENM action and density combinations are provide below.

TABLE 9: EVS LATENCY RESULTS FOR VARIOUS COMBINATIONS OF DENSITY, DENM ACTION, AND SPEED

Density [veh/km]	DENM Action	Speed [m/s]	Average Latency [s]	Std Deviation
2	0	8.33	0.015533991	4.329E-05
		13.89	0.015545345	4.926E-05
		20.83	0.015520311	4.273E-05
		27.78	0.015515757	4.214E-05
3	1	8.33	0.015527412	4.292E-05
		13.89	0.015519411	5.475E-05
		20.83	0.015523116	4.527E-05
		27.78	0.015523851	4.301E-05
4	2	8.33	0.015495844	3.611E-05
		13.89	0.01552572	3.839E-05
		20.83	0.015545963	4.276E-05
		27.78	0.015549914	3.167E-05
2	3	8.33	0.01550654	1.087E-04
		13.89	0.015493513	1.360E-04
		20.83	0.015520912	9.390E-05
		27.78	0.015529376	6.388E-05

3.1.7.4.3 Reliability

Reliability of DENM messages is always above 99% on average for all combinations of speeds, densities and DENM actions. Multiple DENM action and density combinations are provide below.

TABLE 10: EVS DENM MESSAGE RELIABILITY RESULTS FOR VARIOUS COMBINATIONS OF DENSITY, DENM ACTION, AND SPEED

Density [veh/km]	DENM Action	Speed [m/s]	REL average [%]	Std Deviation
2	0	8.33	100	0
		13.89	100	0
		20.83	99.902	0.438
		27.78	100	0
3	1	8.33	100	0
		13.89	100	0
		20.83	99.626	1.274
		27.78	99.840	0.352
4	2	8.33	100	0
		13.89	100	0
		20.83	99.262	1.976
		27.78	99.714	0.570
2	3	8.33	100	0
		13.89	100	0
		20.83	100	0
		27.78	99.960	0.180

3.1.7.4.4 Mobility

The following graphs present the performance of the service under various combinations of (DENM action, density, speed). In each figure, the histogram represents the percentage of collisions avoided, with respect to the case in which the EVS service is not active and vehicles do not take any action in case of risk of collision.

Four different DENM actions are presented below. By analysing the complete set of results one can observe that there is a dependency of the results on DENM action and density (as well as speed). In general, the observed behaviour shows a reasonable pattern, i.e., slight decrease of performance as speed increases. However, there are a few cases in which increasing the speed causes the algorithm to detect a higher percentage of collisions. This happens because the mobility traces change as the speed changes, and this causes the final number of collisions to be different, and at the same time, the values of collision detected oscillates.

The reader should also notice that the algorithm was tuned for 50Km/h (i.e., 13.89 m/s), since it is the more realistic speed in an urban scenario like the simulated one. However, we also wanted to evaluate the algorithm under more demanding conditions for which it was not tuned, so we adjusted the algorithm to bear with higher speed. The main conclusion is that, as expected, the algorithm performs really well for the scenarios for which it was tuned (speeds up to 50Km/h). There would be various options to reach 100% collisions at higher speeds. In fact, and as mentioned before, the algorithm should be tuned to the speeds at which it is expected to work. Furthermore, it could also be complemented with other car security measures.

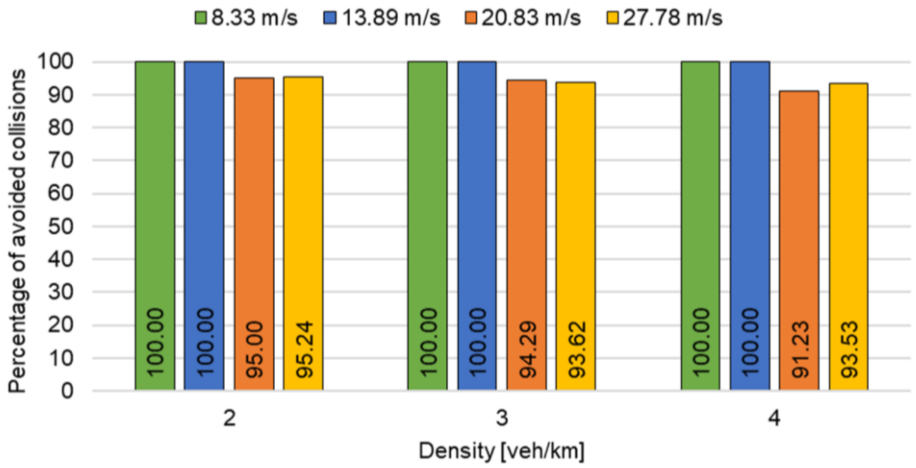

FIGURE 19: MOBILITY KPI RESULTS FOR DENM ACTION 0

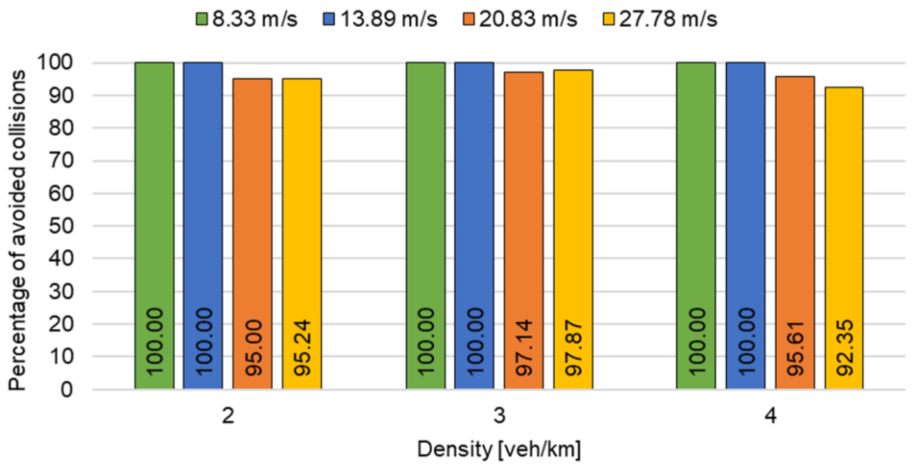

FIGURE 20: MOBILITY KPI RESULTS FOR DENM ACTION 1

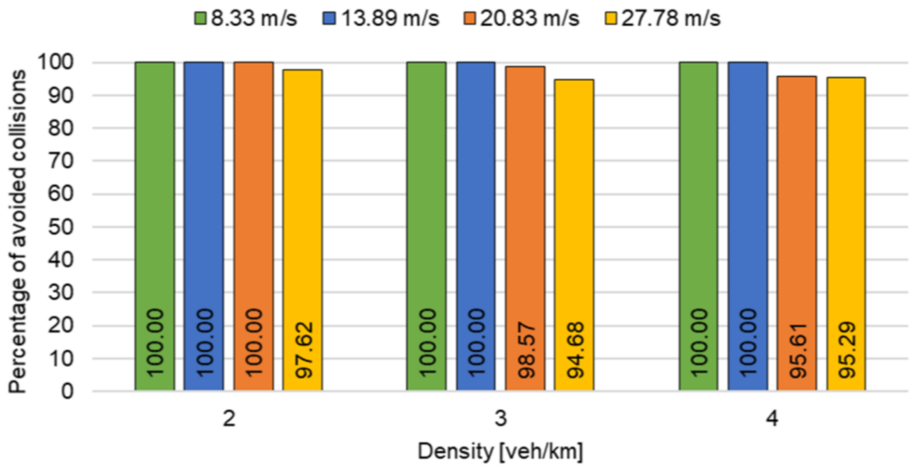

FIGURE 21: MOBILITY KPI RESULTS FOR DENM ACTION 2

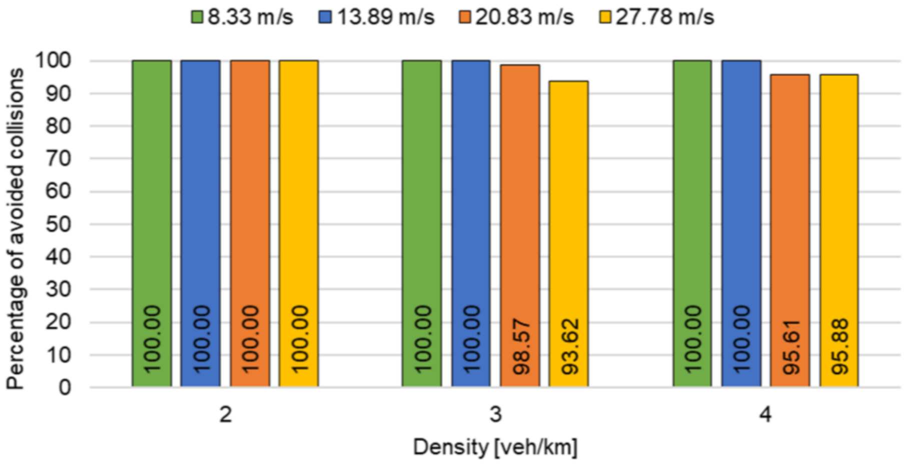

FIGURE 22: MOBILITY KPI RESULTS FOR DENM ACTION 3

3.1.7.4.5 Density

The traffic, latency, and reliability KPIs certainly have a direct influence on the performance of the EVS service, which is reflected in the mobility KPI of the previous section. This KPI is evaluated for densities equal to 2, 3, and 4 vehicles per Km. Higher densities were also simulated, but they are not presented because, in the scenario under evaluation, the high number of cars made the average speed globally decrease.

This was due to the long queues that formed at each crossing when some of the cars take an action after DENM reception to avoid the collision, which in turn results in the following cars reducing speed. Therefore, the evaluated densities for which performance results presented in the previous sections result in high average speeds of cars with a high number of cars, which are used to stress the algorithm.

It can be observed that for all DENM actions, the algorithm is capable of avoiding all collisions for the speeds for which it was designed (i.e., <50Km/h), which are the speeds of a typical urban scenario like the simulated one. Furthermore, higher speeds (unrealistic in a urban scenario) were also evaluated to stress the algorithm.

Even in this case, the percentage of collision avoided was above 92% in all cases that were simulated. As mentioned above, this could be improved by appropriately tuning the algorithm for those speeds and by possibly combining it with other security measures introduced in cars.

In any case, the main conclusion is that for the simulated urban scenario and for a variety of realistic urban densities and speeds, the algorithm is capable to avoid all collisions.

3.2 Entertainment

3.2.1 Selected Proofs of Concept

The Entertainment use case aims to provide a video streaming service to deliver an immersive and interactive experience to users attending a sports event. The entertainment vertical consists of two PoCs regarding On-site live experience (OLE) and Ultra High-Definition (UHD) foreseeing the streaming of UHD live feeds that can be consumed on-demand by the users.

The objective of the demonstration is to prove that 5G-TRANSFORMER platform can deploy a video service in the same or in different locations and which can be used by a group of users simultaneously. The 5G-TRANSFORMER platform can place the resources near the users, ensuring the availability of the network and reducing significantly the end-to-end latency of the network, thereby allowing a better experience to the fans in a sports venue. These features are essential since the source feed of the video can be local to a venue and the service must be able to provide the users an immersive experience by means of an optimal use of the network infrastructure.

The 5G-TRANSFORMER platform allows the Entertainment vertical to instantiate the streaming service dynamically in seconds, providing a transparent abstraction of the network infrastructure and auto-scaling functionalities to manage different load conditions.

Figure 23 describes the different applications involved in the virtual Content Delivery Network (vCDN) use case that delivers a video streaming service. A Content Delivery Network is used to serve a group of servers placed in different parts of the network that have local copies of some media content originally stored in other geographically remote servers, being able to deliever such content efficiently to end users.

The video encoder uses Serial Digital Interface (SDI) to receive the video signal from the video source and then sends the audio and visual (AV) data, using Moving Picture Experts Group (MPEG-4) for compression, to the video recorder for streaming. The video encoder and recording applications can be deployed on a Cloud server or in a Multi-access Edge Computing (MEC) server encoding, and recording the video feeds to serve them after through the cache server.

The local video distributor application is deployed on an edge cloud or in the MEC to be close to the users in order to validate user access and serve the video feed from the video recorder to the users.

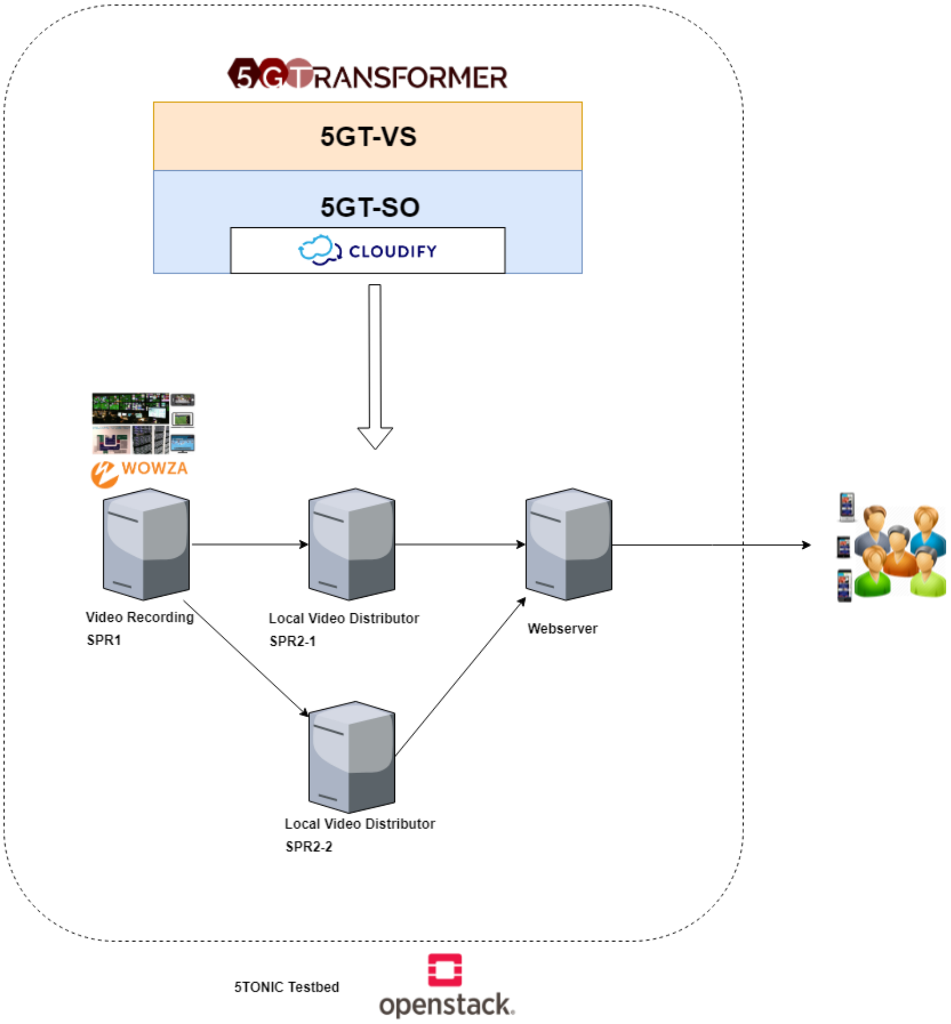

FIGURE 23: DESIGN OF OLE AND UHD USE CASES

3.2.2 Trial organization (Experiment Scenario, deployment view/map)

The best way to demonstrate the capabilities of the 5G-TRANSFORMER platform was to show an immersive demo in a real-world environment, defined as the first PoC, On-site Live Experience. After weighing different sport events, the chosen one was the *Mutuactivos Open España*.

This golf tournament was celebrated in the first week of October in 2019 in Madrid, it is part of *The Race to Dubai*, spanning 47 tournaments in 31 countries across four continents, a season-long competition to crown the European Tour's Number One player.

The demo showed a 360 live video in a mobile player with the capability to switch between two different streams provided by two 360-degree cameras strategically placed in two different places of the training area. The player also was allowed to show the 360 videos in VR mode using a Google™ Daydream Head Mounted Device. The training location was selected by suggestion of the client (the organizer of the event, the European Tour) due to the media rights.

After study the venue, the location of the cameras and the demo booth was selected considering different logistic aspects and the customer needs. We had two different cameras, one with wired connection used in the driving range, placed in the shooting line of the players and a second one with Wi-Fi connection placed in the putting green practice area.

As a first approach the proposed demo scenario aimed to use a real 5G connection to serve the video to the mobile phones, but the antenna was not ready at the date of the event. As a workaround, the video access to the demo mobile phones was provided by a dedicated Wi-Fi 802.11a connection. With all the considerations, the final deployment was the next one:

- Two 360 degree cameras. One wired and another one wireless.
- One Wi-Fi access point for the wireless camera.
- One Wi-Fi access point to serve the video to the mobile phones.
- Two Intel™ NUC servers, with the 5GT-VS and the 5GT-SO deployed on site.
- A network switch to interconnect all the components of the demo.

The final location of all the components in the venue is detailed in the next figure:

FIGURE 24: LOCATION OF THE COMPONENTS OF THE ENTRATEINMENT DEMO IN THE VENUE

The demo was shown to the tournament public in the booth through a mobile application with the capability to play the 360 video both in standard way and VR mode.

3.2.3 Considered KPI(s) and benchmark

For the Entertainment use case, there are three KPIs considered and measured. The values are compared with the current state-of-the-art using another .

TABLE 11: ENTERTAINMENT USE CASE CONSIDERED KPIS.

KPIs	Acronym	Before	Future performance
Latency	LAT	>20ms	<20ms (ITU-R), <5ms (5G PPP)
User data rate	UDR	≥10 Mb/s	≥1 Gb/s (5G-PPP)
Service creation time	SER	>10 hours	≤ 90 min (5G-PPP)

3.2.4 Measurement Methodology

In the previous deliverable D5.3 [1], the measurements were realized in the 5TONIC site in Madrid deploying the components of the vCDN using the 5GT-VS and 5GT-SO using Openstack as the edge cloud infrastructure. In this deliverable the measurements was also obtained in the 5TONIC servers but with the fully integrated 5G-TRANSFORMER stack composed of the 5GT-VS, 5GT-SO and the 5GT-MTP.

3.2.4.1 Latency

The latency was calculated for the UHD PoC using VOD content. The process consist on measure the RTT, using traffic dumps, between the video recording component (SPR1) and the Edge Cache server (SPR2). This means that this measured latency does not take into account the latency introduced in other links or the radio access network latency.

3.2.4.2 User data rate

The User Data rate values were extracted from the monitoring information present in the Web Server video player. The specific metric used was the downloading rate of the different video segments, in other words, the rate at the UE receives the video from the Edge Cache server.

3.2.4.3 Service creation time

The service creation time is the time between the instantion request is made from the 5GT-VS until the complete deployment and configuration of all the VNFs in the VIM. This implies the whole 5G-TRANSFORMER stack, 5GT-VS, 5GT-SO and 5GT-MTP.

3.2.5 Assessment Results

3.2.5.1 Latency

To obtain a latency measure, the Round Trip time (RTT) was calculated analyzing the traffic captured between the video respository (SPR1) and the cache server (SPR2) with Wireshark. The obtained results are depicted in Figure 25.

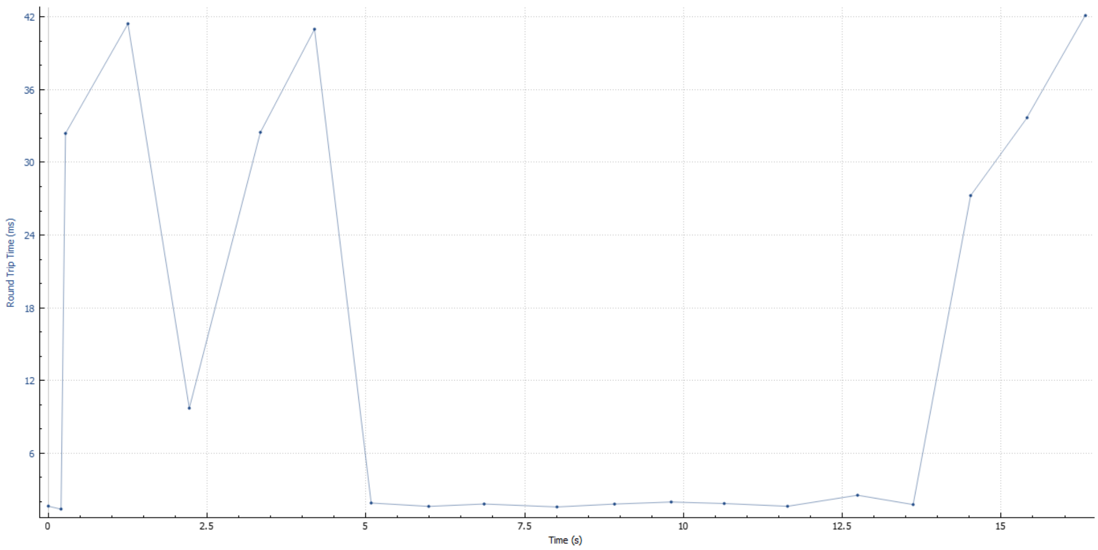

FIGURE 25: ROUND TRIP TIME BETWEEN THE ORIGIN SERVER AND THE CACHE SERVER

The peaks of the graph represents the moments where the client requests the different video segments, obtaining an average of 40 ms.

3.2.5.2 User Data Rate

For this particular experiment, the metrics of the vCDN service have been collected using a real sport video that had been recorded originally in 1080i format. It was transcoded in ABR, using the H264 codec for the video and AAC codec for the audio, the content was also encapsulated in 8 seconds HLS fragments.

This formatting gives us a maximum quality with a target bit rate of 2.7 Mbps.

The measurements detailed in the Figure 26, depicts that the minimum bitrate was 19.27 Mbps and the maximum 35.23 Mbps, this means that the video was downloaded much more quickly than the video was played, having a good amount of seconds cached.

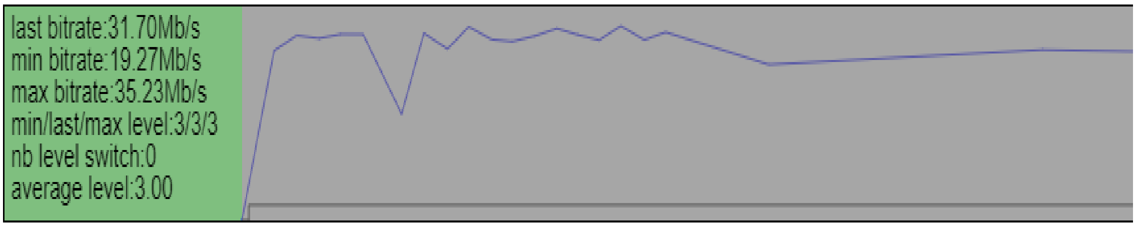

FIGURE 26: USER DATA RATE OBTAINED FROM THE METRICS OF THE VIDEO PLAYER

3.2.5.3 Service creation time

Regarding the service creation time, the time was measured inlcudes the 5GT-VS service definition (starting from sending a pre-created VS descriptor to the 5G-TRANSFORMER platform), the processing of the NFV Network Service in the 5GT-SO, the 5GT-MTP and also the creation and the configuration of all the vCDN VNFs.

To reach a stable result, the test was made ten times and the results were averaged and shown in the Figure 27.

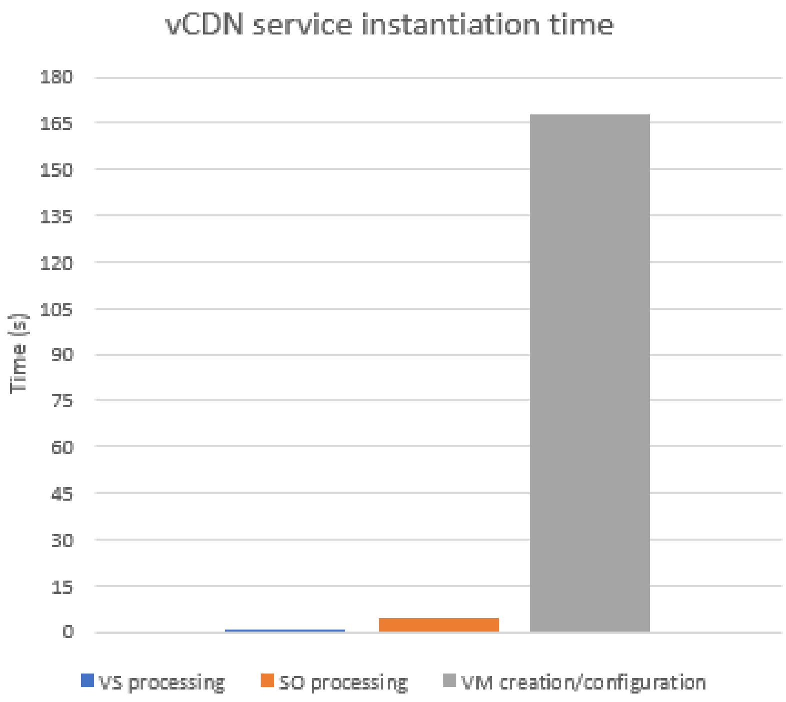

FIGURE 27: COMPLETE VCDN CREATION TIME

As the Figure 27 depicts, the part that takes most time is the VM creation and configuration while 5GT-VS and 5GT-SO processing time is only few seconds. Besides, the totoal service creation time is below 200 seconds, which demonstrates the feasibility of the 5GT platform for instantiating a complete vCDN service in minutes, meeting the KPI on desired service creation time. .

Taking use of the 5GT-SO serivce auto-scaling feature corresponding to the available resources, another test was performed. The CPU alert and threshold was created for the Cache server VNF using the 5G-TRANSFORMER Monitoring platform. If the CPU usage increases, The 5G-TRANSFORMER triggers the auto-scaling action defined in the NFV-NS descriptor.

In this case, the VNF scale-out time was measured since the alert is triggered to the complete deployment and configuration of a new Cache server at 5TONIC server. The test was also repeated ten times to have an averaged result. The VNF scale-out time is shown in the Figure 28 The time to process the scaling action at the 5GT-SO is much smaller, in the range of 10 seconds.

Due to this scaling time, it is necessary to set the correct threshold to avoid a service failure.

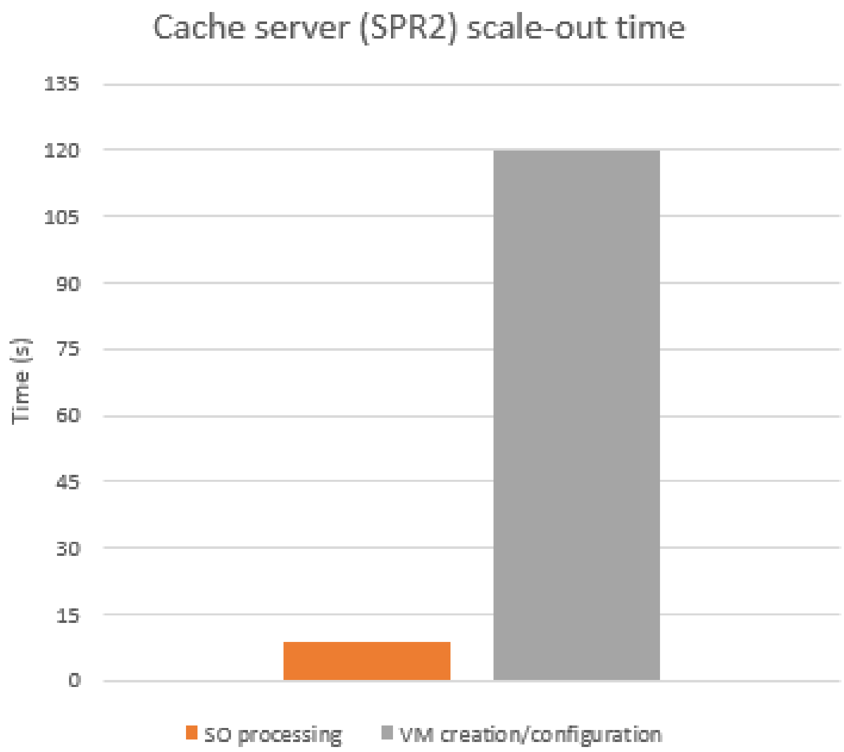

FIGURE 28: CACHE SERVER VNF SCALING TIME

3.3 E-Health

3.3.1 Selected Proofs of Concept

For the E-Health PoC we are demonstrating the Monitoring & Emergency use case that is described in D5.3 [1].

3.3.2 Trial organization (Experiment Scenario, deployment view/map)

The demonstration of the E-Health PoC use case is set on two testbed sites: 5TONIC and CTTC trial sites. The physical location where the demonstration takes place is at the 5TONIC premises.

Figure 29 presents the demonstration site of the E-Health use case. A basketball player simulates a patient with heart-attack, wearing a smart-wristband or a watch that triggers an emergency and deployment of the emergency service. The emergency service is deployed until the ambulance crew reach the emergency site. The ambulance crew then would use the AR/VR googles to locate the patient and obtain important information about the paitent from the newly deployed local eServer (in 5TONIC).

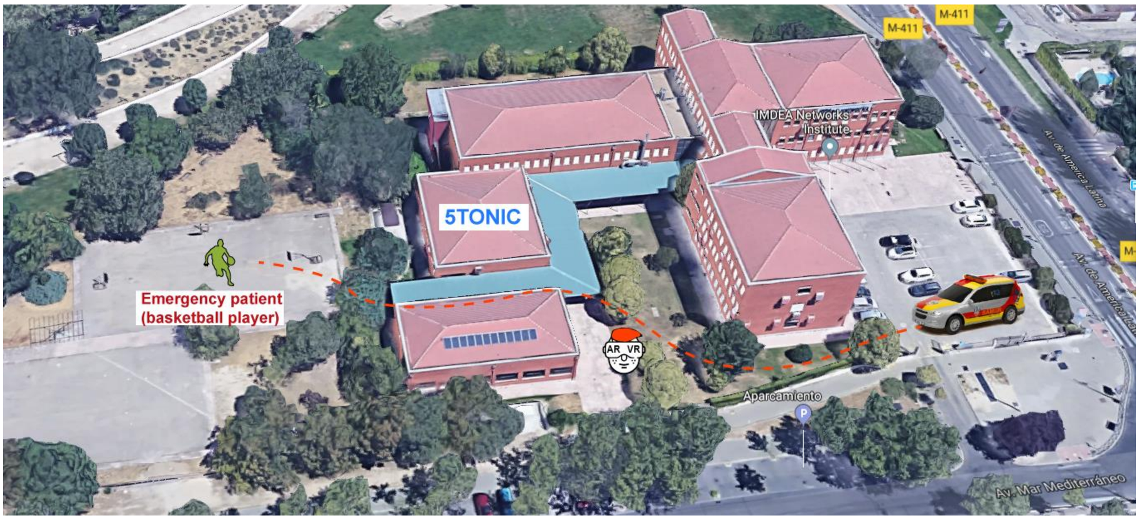

FIGURE 29: E-HEALTH DEMONSTRATION SITE

The E-Health use case is realized through instantiation of two NFV-NS: Monitoring NFV-NS and Emergency NFV-NS. The Monitoring NFV-NS is instantiated at the beginning as a single-domain NFV-NS. Upon emergency trigger (explained in D5.3 [1]), the Emergency NFV-NS is deployed.

The Emergency NFV-NS contains two nested services, the Monitoring (that is already running) and the Edge NFV-NS. The Edge NFV-NS is instantiated through a service federation procedure. Figure 30 presents the visualization of the scenario setup.

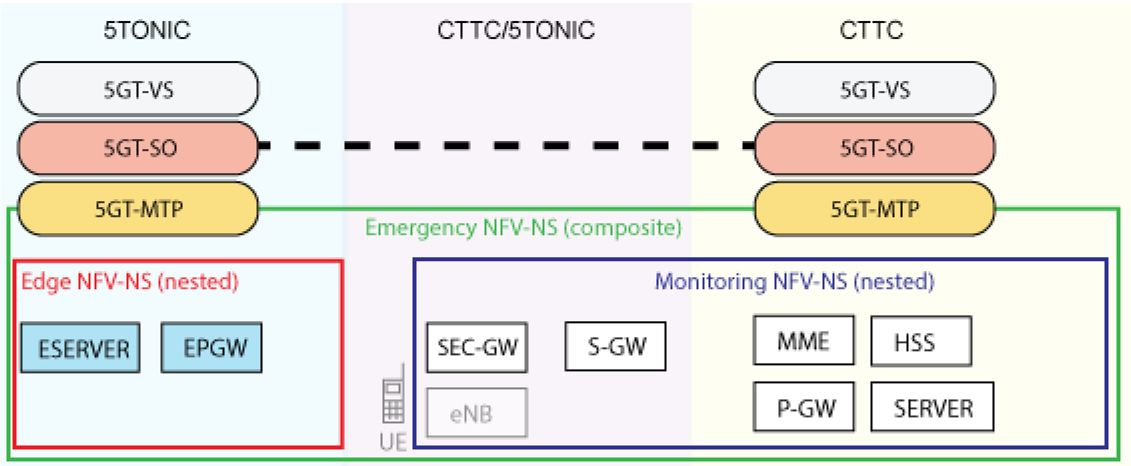

FIGURE 30: E-HEALTH COMPOSITION OF NFV-NSS AND VNFS AT DIFFERENT TEST-SITES

The CTTC testbed site is referred as an entry-point and represents the administrative domain where the Monitoring NFV-NS is deployed. The Monitoring NFV-NS contains 6 VNFs. The SERVER, HSS, MME and P-GW VNFs (see Figure 30) are deployed on the CTTC premises. The S-GW and the SEC-GW are deployed on a CTTC-controlled infrastructure at the 5Tonic premises. In other words, the SEC-GW and the S-GW are physically present in the 5TONIC infrastructure, but belong to (or are being controlled by) the CTTC 5GT platform. The main reasons for the setup are (i) to have realistic demonstration of deployed mobile network; (ii) the need of direct connection to the PHY eNB to connect directly to users via the UEs.

Upon an emergency trigger received from the SERVER, the 5GT-VS initiates the instantiation of the composite Emergency NFV-NS. The CTTC 5GT-SO has the Monitoring NFV-NS already instantiated.

For the Edge NFV-NS, the CTTC 5GT-SO sends service federation request to the 5TONIC 5GT-SO. The requested (Edge NFV-NS) is instantiated and federated at the 5TONIC 5G-TRANSFORMER platform containing two VNFs ESERVER and EPGW.

Once the federation is successfully executed, the Server issues a redirection rules, where all the UE traffic is redirected from the S-GW (of the Monitoring NFV-NS) towards the EPGW. This way the UE of the ambulance unit is connected via the eNB PHY, SEC-GW, S-GW, EPGW to the ESERVER. The AR/VR glasses at the UE receive data regarding the patient (location, health records, etc..) directly from the ESERVER.

3.3.3 Considered KPI(s) and benchmark

For the E-Health use case, the considered KPIs are presented in D5.3 [1] shown in Table 12. Most of the KPIs are already measured, except the Service Creation Time (SER). The results are comparable with the state-of-the-art and with the emergency response time data.

TABLE 12: CONSIDERED KPIS FOR THE E-HEALTH USE CASE

KPIs	KPI	Before	Future performance
Latency	**LAT**	<120 ms	<35 ms (using the local eServer)
Service availability	**REL**	98%	99.999%
Positioning	**POS**	<12 m	<12 m
Total connected devices	**TCD**	Single device connected to local eServer	More devices per local eServer (depending on the provided features)
Service creation time	**SER**	≤ 90 min (5G PPP)	270 seconds

3.3.4 Measurement Methodology

The measurements performed in D5.3 [1] are described in detail. In this document, we are updating the measurements for the Service Creation Time (SER) KPI.

The measurements are done on the CTTC premises, using the 5G-TRANSFORMER platform. The measurements are done with a goal to test the successful deployment of the composite Emergency NFV-NS using the federation feature of the 5G-TRASNFORMER platform for instantiating part of the service in an external domain. The results are shown in the next section.

3.3.5 Assessment Results

The results shown on Figure 31 demonstrate that the deployment of the Emergency NFV-NS over two different domains takes around 270 seconds. Compared to the 5G PPP expected service creation time this is in the order of minutes (significantly less than 90 min). The results (in Figure 31) are derived from measuring the instantiation time (from left to right): 1) & 2) for each single nested NFV-NSs of the eHealth composite NFV-NS (e.g., the Nested-MB and the nested-vEPC); 3) the eHealth composite NFV-NS instantiated in a single NFVI-PoP; 4) the eHealth composite NFV-NS instantiated over multiple NFVI-PoPs; 5) the eHealth composite NFV-NS instantiated using service federation. In the worst case (when service federation is used) the instantiation time is maximum 270 seconds.

TABLE 13: SERVICE CREATION TIME FOR E-HEALTH EMERGENCY NFV-NS

KPIs	KPI	Before	Future performance
Service creation time	SER	≤ 90 min (5G PPP)	270 seconds

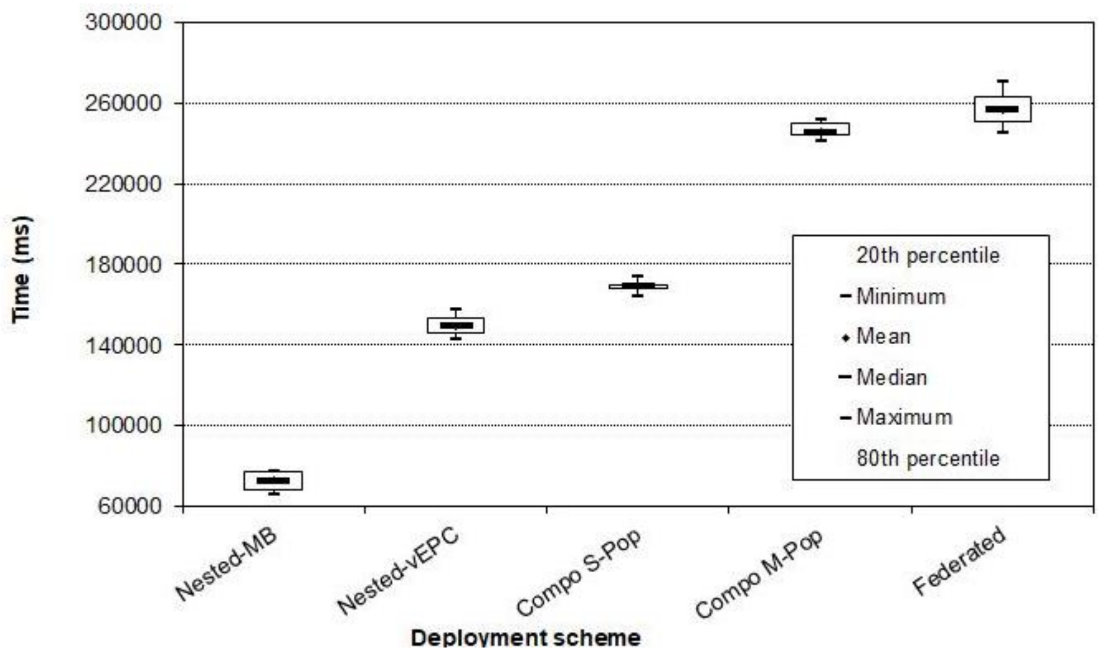

FIGURE 31: E-HEALTH COMPOSITE SERVICE CREATION TIME USING FEDERATION

3.4 E-Industry

3.4.1 Selected Proofs of Concept

The E-Industry Cloud Robotics (CR) PoC simulates factory service robots and production processes that are remotely monitored and controlled in the cloud, exploiting wireless connectivity (5G) to minimize infrastructure cost, optimize processes, and implement lean manufacturing. The objective of this PoC is to verify the capability of 5G-Transformer platform to perform the allocation of suitable resources based on the specific service requests to allow the interaction and coordination of multiple (fixed and mobile) robots controlled by remote distributed services, satisfying strict latency and bandwidth requirements.

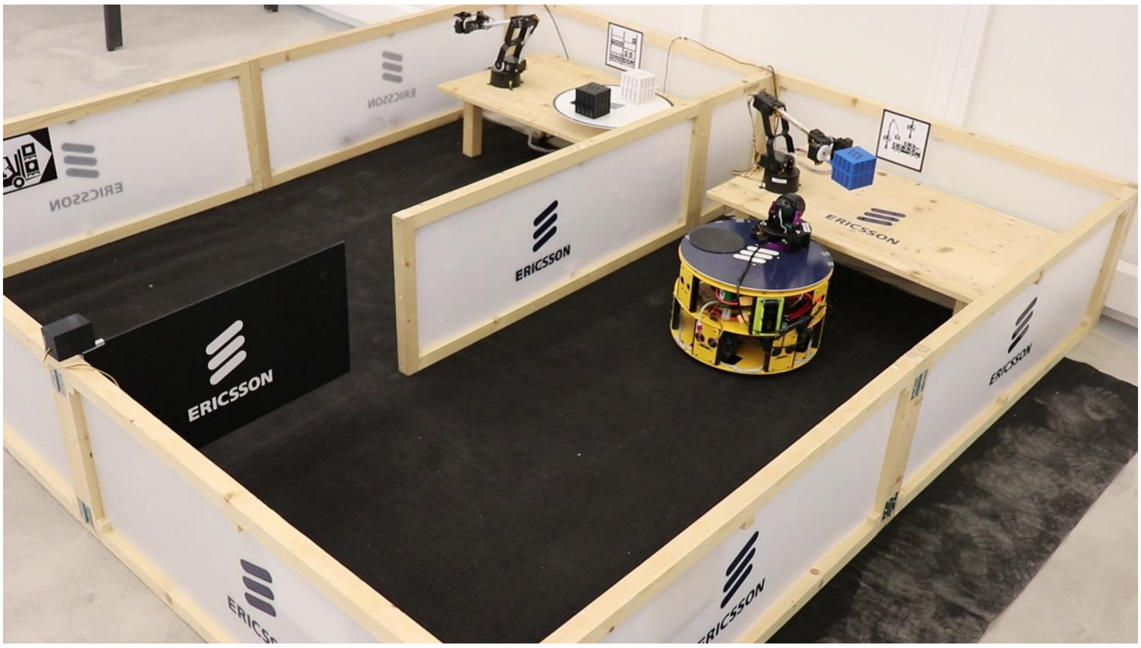

FIGURE 32: E-INDUSTRY CLOUD ROBOTIC FACTORY POC

The Cloud Robotics PoC, as depicted in Figure 32, includes an autonomous mobile robot shuttling materials between work cells in a factory by means of image processing navigation algorithms. A factory control tablet is used to select a customized set of factory tasks, e.g. a pallet transfer from one cell of the factory to another. The request is handled on the Cloud by a main control server which orchestrates the multiple factory robots' tasks as well as executes other control functions including image processing from the autonomous mobile robot. In addition to the mobile robot, the factory includes two robotic arms which are used to load and unload goods from the mobile robot, as shown in Figure 33.

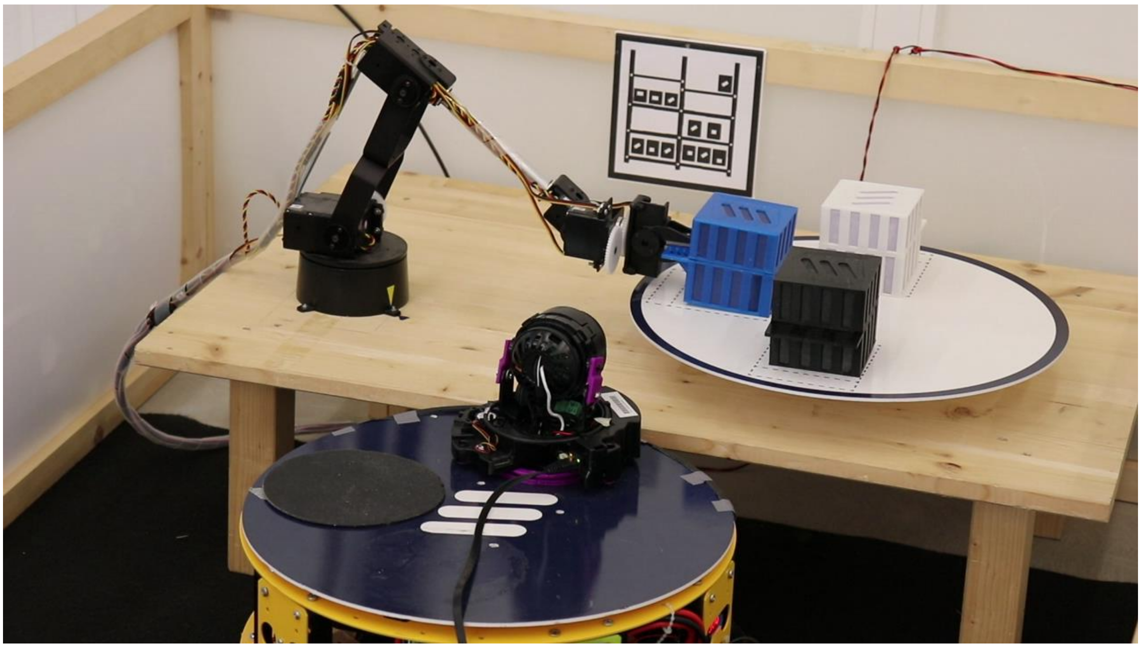

FIGURE 33: E-INDUSTRY CLOUD ROBOTIC ARM LOADING A PALLET ONTO A MOBILE ROBOT

An automated warehouse is simulated by a rotating platform, and an automated door is placed along the navigation tracks to show a flexible and optimized shuttling of materials between work cells. The entire sequence is monitored and controlled by the remote server through radio communication using the EXhaul optical network infrastructure.

EXhaul serves as both backhaul and fronthaul to convey radio traffic on an optical infrastructure. The cornerstones include a novel photonic technology used to provide optical connectivity complemented by a dedicated agnostic framing, a deterministic switching module, and a flexible control paradigm based on a layered and slicing concept to facilitate optimal interactions of transport and radio resources while preserving a well demarcated mutual independence. A detailed description of EXhaul can be found in [7].

3.4.2 Trial organization (Experiment Scenario, deployment view/map)

The physical demo is comprised of 3 areas located at the 5TONIC testbed site: a Server room containing the cloud (XenServer running a VM) and v-EPC, interfaced via a router; Table area containing the 5G-TRANSFORMER Software stack, EXHAUL DWDM ring, remote radio site, and the user interface for the VM (XenCenter) where the user interface and 5G-TRANSFORMER Software stack connect to the radio via network router and Wi-Fi switch; and Demo area containing the factory (2 work cells and an automated guided vehicle (AGV) and tablet), as shown in Figure 34.

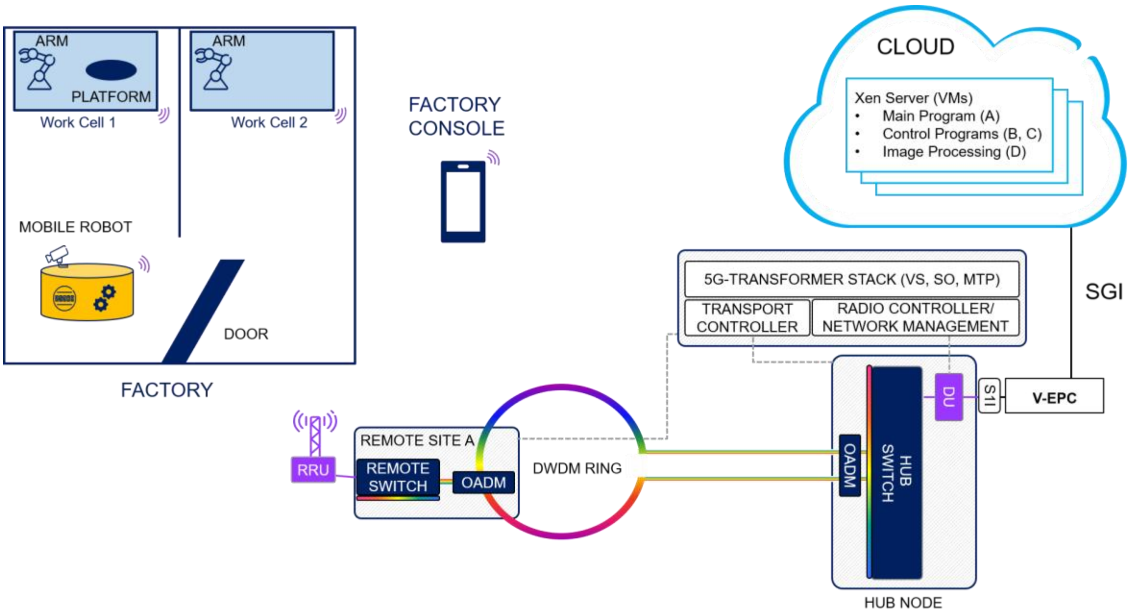

FIGURE 34: SCHEMATIC OF THE E-INDUSTRY CLOUD ROBOTICS POC

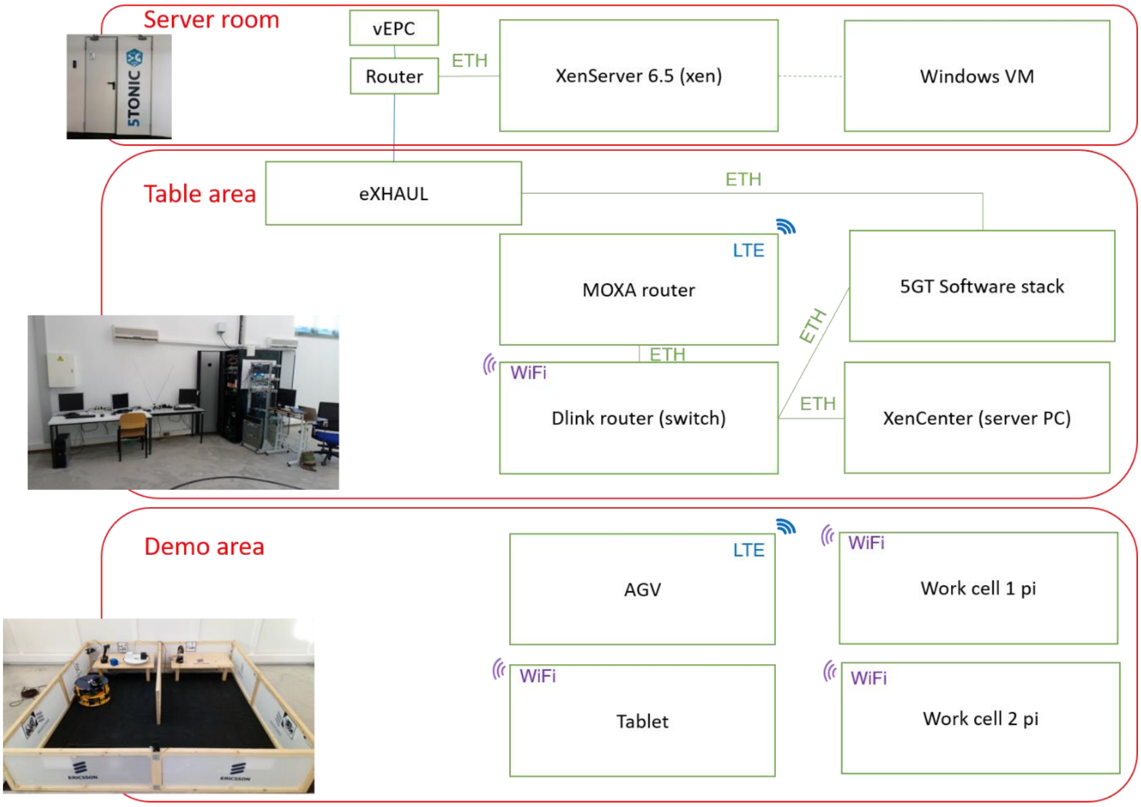

FIGURE 35: E-INDUSTRY CLOUD ROBOTICS NETWORK SCHEME

The demo uses the complete 5G-TRANSFORMER stack, as illustrated in Figure 36 that consists of:

- The 5GT-VS to load the blueprint and VNFD and transfer the NSD to the 5GT-SO.
- The 5GT-SO, which using the descriptors, orchestrates the E-Industry service, placing the VNFs and selecting the logical link using the abstract view provided by the 5GT-MTP. The 5GT-SO also triggers the set up of both the network connectivity and compute resources for VNFs.
- The 5GT-MTP provides the network and cloud resources, interacting with the dataplane to complete the setup.

The demo interacts with a real radio system and controls the radio-realted configurations via the developped radio plug-in to communicate with a Radio controller and exploits the Radio Abstraction feature developed in the project.

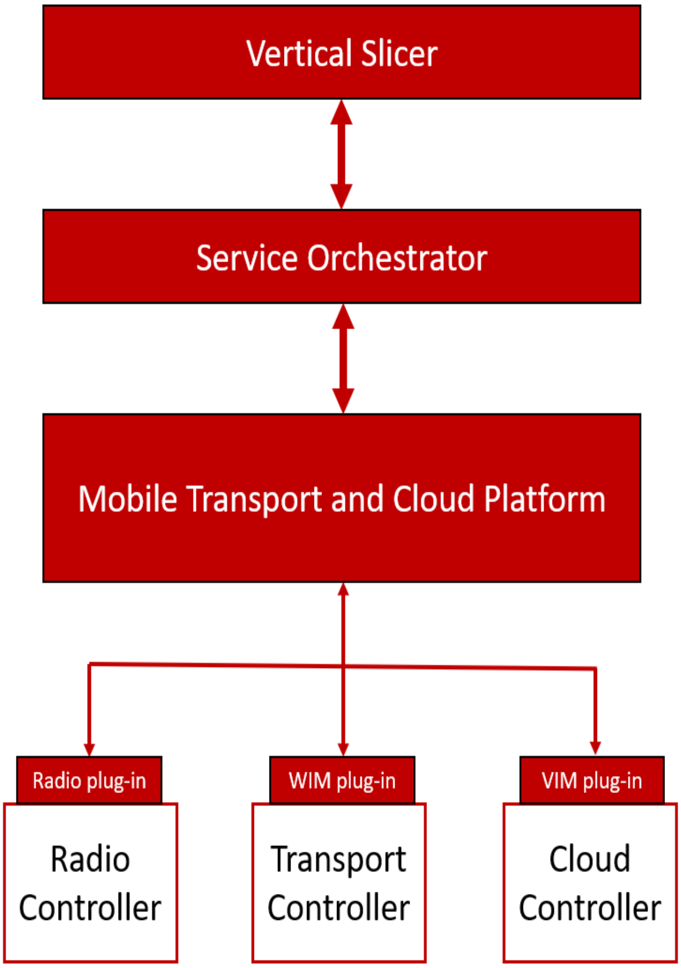

FIGURE 36: 5G-TRANSFORMER SOFTWARE STACK USED IN THE E-INDUSTRY POC

3.4.3 Considered KPI(s) and benchmark

The E-Industry use case has 3 associated KPIs: Latency, Reliability, and Service creation time. Latency (LAT) is the time it takes from the time when a data packet is sent from the transmitting end, e.g. factory service robot, to the time when it is received at the receiving entity, e.g. core network. In the CR conext, RTT Latency is considered, i.e. the round-trip time of communication between the factory and the cloud. The KPI Reliability (REL) is the percentage of messages that have been sent and received correctly. In CR, it involves measuring the availability of the service for duration of a factory task(s) (e.g. pallet transfer, navigation, etc.).

Finally, the KPI Service creation time (SER) is the time required for the network and compute setup and teardown of a service. Table 14 maps these KPIs to the current performance specifications and future targets set by the ITU-R and 5G PPP projects, where applicable.

TABLE 14: KPI MAPPING TO CURRENT PERFORMANCE SPECIFICATIONS AND FUTURE TARGETS

KPIs	Acronym	Before	Future performance
Latency	LAT	>20ms	<20ms (ITU-R), <5ms (5G PPP)
Reliability	REL	<99%	1 - 10^{-5} success probability (ITU-R), 99.999% success probability (5G PPP)
Service creation time	SER	Not available	≤ 90 min (5G PPP)

3.4.4 Measurement Methodology

The E-Industry use case is comprised of 3 PoCs, E-Industry phase 1, E-Industry phase 2, and E-Industry phase 3. In phase 3, the full 5G-TRANSFORMER software stack (5GT-VS, 5GT-SO, and 5GT-MTP) is used, running the Cloud Robotics Factory in the test bed of 5TONIC. In phase 2, only the 5GT-MTP component is implemented, again running the Cloud Robotics Factory in the test bed of 5TONIC, and in phase 1 only the 5GT-MTP component is implemented running the Cloud Robotics Factory using the Ericsson radio network in Pisa.

3.4.4.1 Latency

The E-Industry use case contains two 5G-TRANSFORMER PoCs (E-Industry phase 1 and E-Industry phase 2) for which the KPI Latency is measured.

Measurement scenarios differ between the two PoCs in that, due to physical location, E-Industry phase 1 relies on the Ericsson Stockholm EPC interface to the cloud while E-Industry phase 2 makes use of the vEPC located at the 5TONIC testbed.

Table 15, maps these Proofs of Concept to the measurement methodology used for the Latency KPI measurement.

TABLE 15: LATENCY MEASUREMENT METHODOLOGIES FOR E-INDUSTRY POC RELEASES

Proof of Concept (PoC)	Measurement Methodology
E-Industry phase 1	Measuring preliminary RTT latency (sample size 10,000 ping packets) from the cloud controller to the mobile robot located in Ericsson Pisa using the Ericsson EPC located in Stockholm
E-Industry phase 2	Final measurement of RTT latency (sample size 10,000 ping packets) from the cloud controller to the mobile robot using the 5TONIC testbed and vEPC

3.4.4.2 Reliability

The E-Industry use case contains one Proof of Concept for which the KPI Reliability is measured (E-Industry phase 2). The CR is reliability critical as all factory requests are handled on the Cloud by a main control server which orchestrates the multiple factory robots' tasks as well as executes other control functions including image processing from the autonomous mobile robot.

A reliability of less than 99.999% would result in asynchronous robotic control sequences. Table 16 maps this Proof of Concept to the measurement methodology used for the Reliability KPI measurement of the CR.

TABLE 16: RELIABILITY MEASUREMENT METHODOLOGIES FOR E-INDUSTRY POC RELEASES.

Proof of Concept (PoC)	Measurement Methodology
E-Industry phase 2	Measuring the availability of the service (%) for duration of a factory task(s) (e.g. pallet transfer, navigation, etc.).

3.4.4.3 Service creation time

The E-Industry use case contains two Proofs of Concept (E-Industry phase 2 and E-Industry phase 3) for which the KPI Service creation time is measured. The difference in the measurements stems from the timeline of the software integration as described in D5.2 [2]. Table 17 maps these Proofs of Concept to the measurement methodology used for the Service creation time KPI measurement.

TABLE 17: SERVICE CREATION TIME MEASUREMENT METHODOLOGIES FOR E-INDUSTRY POC RELEASES

Proof of Concept (PoC)	Measurement Methodology
E-Industry phase 2	Measuring the time of the network and compute setup and teardown for the CR service from the MTP.
E-Industry phase 3	Measuring the time of the network and compute setup and teardown for the CR service from VS/SO/MTP.

3.4.5 Assessment Results

3.4.5.1 Latency

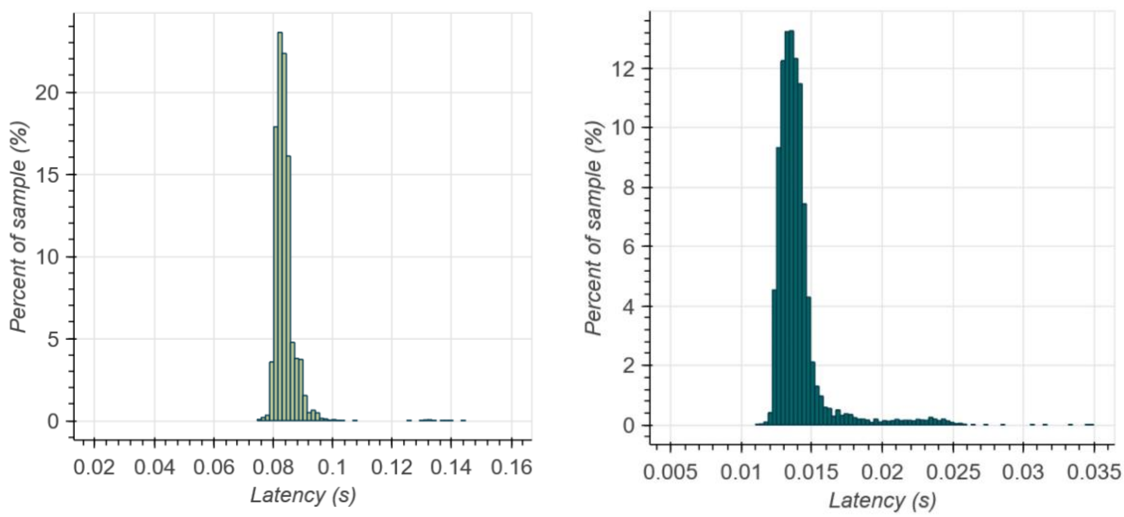

FIGURE 37: ROUND TRIP-TIME (RTT) LATENCY, POC E-INDUSTRY PHASE 1 (LEFT) AND POC E-INDUSTRY PHASE 2 (RIGHT)

The KPI RTT latency measurement was performed using the ping utility and wireshark packet analyzer to measure network latency between the AGV and the virtual machine running on the cloud. RTT latency in this context is defined as the time in seconds of the path from the core network to the service robotics and back. The sample size for each measurement is 10,000 packet pairs. The measurement for E-Industry phase 1

(Preparatory experiment for CR service activation) is shown in Figure 37 left. The associated demonstration can be viewed at https://youtu.be/-Ox14nzRHu0.

The preparatory experiment for CR service activation was based at Ericsson Pisa and the network setup for E-Industry phase 1 did not utilize the vEPC located at the 5TONIC test lab, as described in Figure 34. Instead, an Ericsson EPC, located in Stockholm was used for the initial experiment. As a result, the value of latency measured includes the time to transport from the Digital Unit of EXhaul in Pisa to the EPC in Stockholm. The mean value of the distribution is 83.88 ± 0.05(statistical error) ±2 (systematic error) ms. The large systematic error is attributed to fluctuations in the core network latency as a function of time.

Figure 37 (right) shows the latency KPI measurement for E-Industry phase 2, using the 5TONIC testbed and vEPC, unlike the measurement for E-Industry phase 1. Again, KPI RTT latency measurements were performed using the ping utility and wireshark packet analyzer to measure network latency between the AGV and the virtual machine running on the cloud.

The mean value of the distribution is 14.05 ± 0.02(statistical error) ms.

The large difference between the E-Industry phase 1 and E-Industry phase 2 latencies is due to the network setup difference (EPC vs. vEPC) as described in the previous paragraph. The result, obtained using the 5TONIC 5G testbed, is inline with the expected 5G performance outlined by ITU-R of below 20ms (Table 14).

3.4.5.2 Reliability

The reliability of the service (%) for the duration of the complete pallet transfer factory task was verified to meet the 99.999% expected performance (Table 15) using 10 executed trials. Each trial task took a time of approximately 3.5 minutes.

3.4.5.3 Service creation time

The KPI SER for PoC ID E-Industry phase 2 has been taken at the MTP level. The measurement has been repetead using the same measure methodology using the full 5GT stack (VS,SO,MTP) Specifically, the SER was measured triggering the service setup/teardown command from VS and measure the time (reported as “date” command from the softwares). The measurementare reported in Figure 38 has been done for 10 trials to minimize the contribution from fluctuations from database access of different softwares. The mean time for network and compute setup is 1563 ms and 3432 ms, respectively. Similarly, the mean time for network and compute teardown is 1530 ms and 2961 ms, respectively. The measurement shows that the 5GT stack introduces a service creation time of 100-200 ms that is inline with KPI. The main critical operation is the VM deployment that is strictly dependent from the applications (it takes 2.5-3 seconds in the PoC). To note that we have shorter setup time respect to other PoC because we use a different DC software (XenCenter) that manages the VM setup in asynchronous way

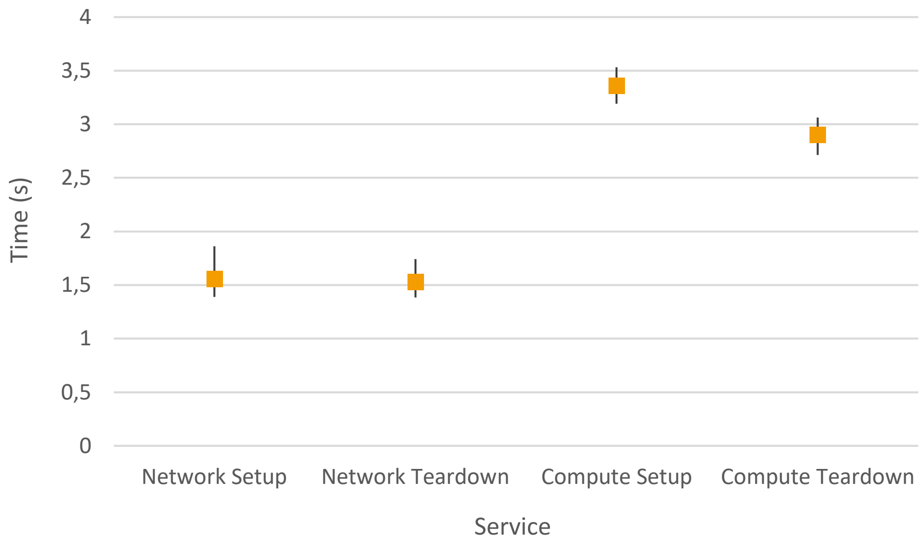

FIGURE 38: KPI SER FOR POC 4.3 TAKEN AT THE MTP LEVEL

3.5 MNO/MVNO

3.5.1 Selected Proofs of Concept

The MVNO use case consists of the provision of a network slice as a service (NSaaS) by a mobile network operator (MNO) hosting a virtual mobile network operator (MVNO). This leads to the allocation to the MVNO of a network slice instance that contains a wireless core network service providing communication services to the MVNO's users.

The network slice instance can offer several types of access combining 3GPP access technologies such as LTE and non 3GPP like WiFi and LoRa. Indeed, the MNO composes network service using the VNF from the software solution named 'Wireless Edge Factory' (WEF) [4] to realize a network slice.

The WEF is a modular convergent software offering "à la carte" (customizable) connectivity services such as WiFi (IEEE 802.11), cellular (Long Term Evolution) and IoT (LoRa) access. The WEF offers operators the ability to build a network slice that meets connectivity needs in a multi-access environment from the ingredients that are the virtualized network functions.

The VNF composition in a network service allows the sharing of features between access types such as the authentication and key agreement (AKA) method for the user authenticaton, the host configuration and IP allocation based on DHCP, the function chaining mechanism and the customer care. Besides the pooling of functions, the WEF architecture is based on control and user plane separation (CUPS) leveraging SDN paradigm to manage efficiently session connections and data flows by a SDN controller and OpenFlow switches. As a result, it enhances the flexibility of the slice design in term of enabled features, performance sizing and capacity distribution.

This proof of concept of MNO/MVNO is related to the description provided in the deliverable D5.3 [1], with a focus on the deployment of a network slice instance comprising a Evolved Packet Core (EPC) network providing wireless data communications.

The below figure illustrates the structure of the EPC network service.

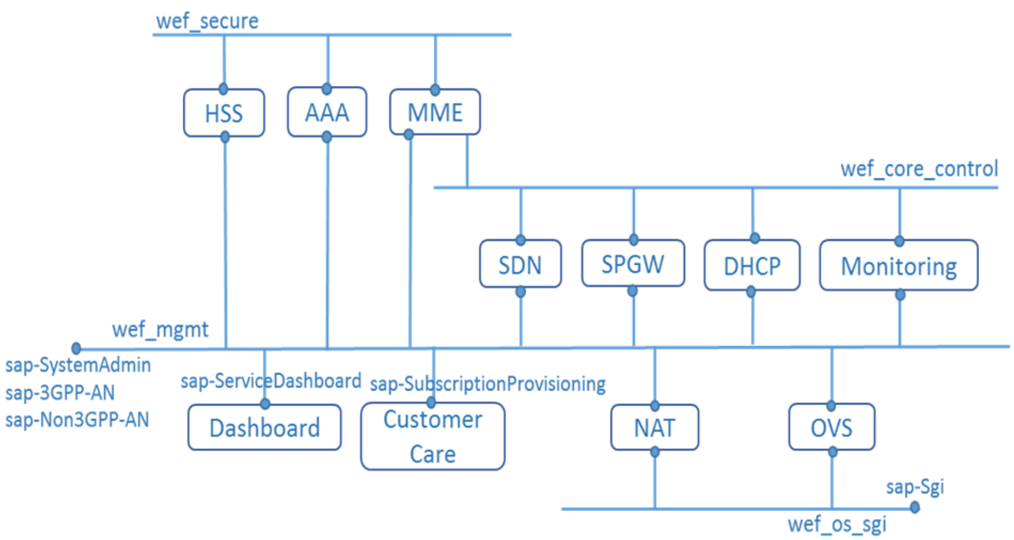

FIGURE 39: ARCHITECTURE OF THE EPC NETWORK SERVICE DEPLOYED FOR THE MVNO USE CASE

3.5.2 Trial organization (Experiment Scenario, deployment view/map)

For the trial, we opted to deploy the data plane along with the control plane when creating a MVNO network slice instance, it has the effect of inserting in the EPC network service descriptor (NSD) two VNFs which are the OpenVswitch (OVS) and the network address translation (NAT) function. These functions deal respectively with the processing (switching, de/encapsulation, binding) of data flows and the mapping of IP addresses. Related to EPC architecture, the OVS plays the role of servicing and packet gateway user plane (S/PGW-U) and it is operated by a SDN controller and the application S/PGW-C.

We also included two functional blocks called element manager (EM) in the ETSI NFV architecture framework [3] with the main responsibility for network management functionalities fault, configuration, accounting, performance, security (FCAPS) for a VNF. The first is the customer care enabling the MVNO managing its user suscriptions, and the second is the dashboard displaying the active user session information. The addition of these 2 functions makes it possible to realize the defined concept network slice as a service (NSaaS) whose specifications in [6] express the exposure of:

- network slice management data (user session dashboard), and
- network slice management capability (customer care).

To deliver the NSaaS to a Communication Service Consumer, we implemented a network slice manager (NSM) which manages the lifecycle of network slice instances. The description of the network slice manager features is provided in details in the delivrable D3.4. The NSM uses the 5GT-VS's South Bound Interface (SBI) to operate with the NFVO orchestrator. The 5GT-VS's SBI is compliant to the specifications of ETSI NFV IFA013, its implementation can be based on OSM Release 5 interface or on the 5GT-SO driver. This first option was chosen for this trial.

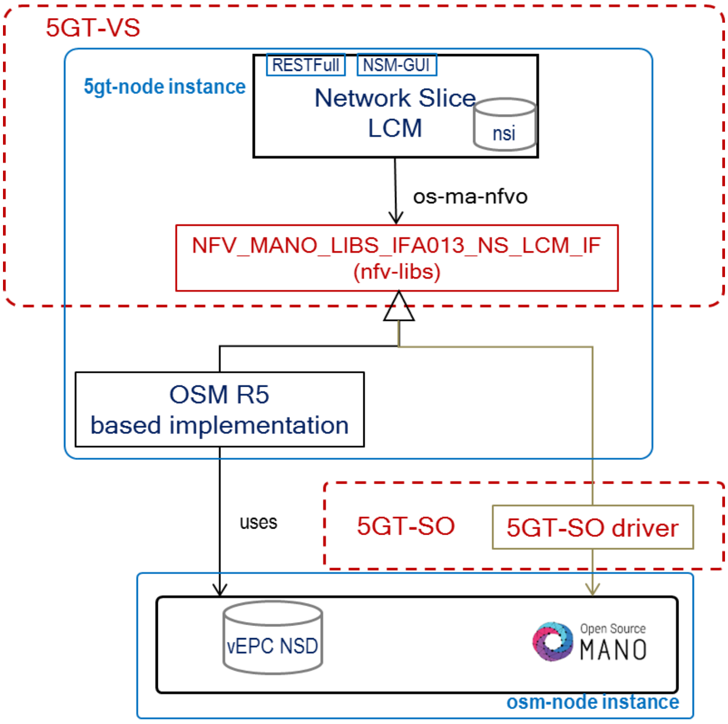

FIGURE 40: THE NETWORK SLICE INSTANCE LIFECYCLE MANAGEMENT FUNCTION ARCHITECTURE.

The VNF images and the VNFD and NSD of the virtualized EPC are on-boarded in the orchestrator OSM beforehand. The network operator uses the NSM GUI to retrieve the available NSD list and perform the creation, the activation and the termination of network slice instance.

When a virtualized EPC network slice instance has been activated, its exposed Service Access Points (SAP) can be connected to the RAN and to SGi LAN. For wireless access, an software solution based on OAI-SIM [5]is used to emulate an User Equipement carrying out an LTE attach with the core network. This creates a user bearer enabling the IP connection between the UE and the SGi LAN.

3.5.3 Considered KPI(s) and benchmark

We selected two KPIs for the MNO/MVNO use case:

1. SER: measuring the creation and the activation of network slice instances while varying the size (quantity of CPUs and RAMs) of VNFs
2. CST: Computing the infrastructure cost of network slice instances of various sizes

The table lists the KPIs that we selected for the use case MVNO.

TABLE 18: MVNO CONSIDERED KPIS

KPIs	Acronym	Before	Future performance
Service Creation Time	SER	Not provided	< 90min (3GPP)
Infrastructure Cost	CST	Not provided	Not Provided

3.5.4 Measurement Methodology

In the EPC architecture, the MME is the main network function communicating signaling data with the Radio Access Network (RAN) at the interface S1-C. It performs several procedures such as the user authentication, session and bearer lifecycle management. As a consequence, it appears as the first element to be congested when the signaling load increases significantly. In the MNO/MVNO use case, the CSP may require a NSI able to cope with a certain user session rate which direcly impacts on the size of the MME to be deployed. This was brought to light by the observation of the session setup time of several UEs during the gradual increase in the rate of attachment procedures to a given MMS flavor. The result of the tests shows the lower the MME flavour, the sooner the attachment time starts to differ and it clearly identifies the MME as the primary element impacting the performance of the EPC network service. Thus, we relate the KPIs SER and CST fluctuations to the virtualized resources CPU and RAM allocated to the MME. To better compute the considered KPI, we deploy the NSaaS several times (up to 50) at different network service instantiation levels (IL), which varies with the MME flavours by increasing the number of CPU and/or the amount of RAM in steps of {1, 2, 4, 8, 16}.

3.5.5 Assessment Results

For results to be consistent between measurements and undisturbed by external elements, we isolate our measurement environment on a platform in our lab. The platform consists of dedicated hardware, two physical machines to run the VIM Openstack, while another machine hosts the NSM and the NFV orchestrator OSM. We present below the results of our assessment of the service creation time and the cost of the infrastructure needed to deploy our use case at each level of instantiation.

Figure 41 depicts the Service Creation Time (SER) Cumulative Function Distribution (CDF) for our EPC network service, measured by increasing the flavours of the MME. We used four flavours, namely c1r1, c2r2, c4r4 and c8r16, where c*x*r*y* corresponds to the flavour that counts *x* CPU and *y* gigabytes of RAM.

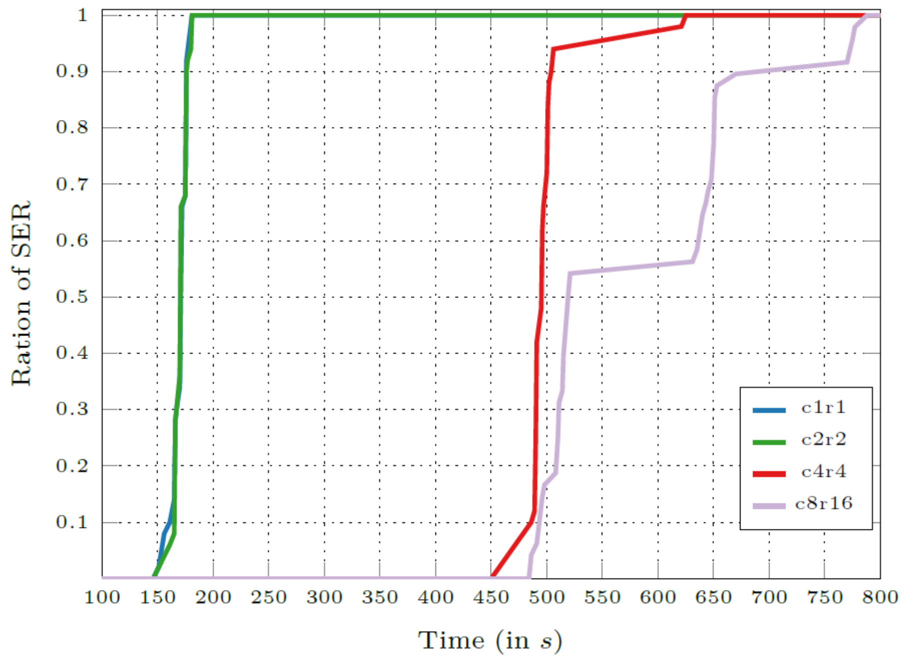

FIGURE 41: CDF OF SERVICE CREATION TIME FOR THE MVNO USE CASE WHEN INCREASING THE FLAVOUR OF THE MME

In Figure 42, we used box-plots to present the time needed to instantiate the service (SER) and to quantify the its variability. These two figures are showing:

- Time needed to create MVNO service using small flavours such as c1r1 and c2r2 for the MME is not differentiating and it had low fluctuation between measures. The hypervisor on the compute node can easily allocate small amount of virtualized resources.
- When the MME is using a flavour larger than the flavour c2r2, the value and variability of the SER increase significantly with the size of the flavour (c4r4 and c8r16). We believe that requests for large amounts of resources are more demanding for the hypervisor which has to find room to create and boot up eleven VNF instances of the MVNO service. As the datacentre we are using has limited physical resources (4 cores x 2 threads), the service creation leads to overload available physical resources.

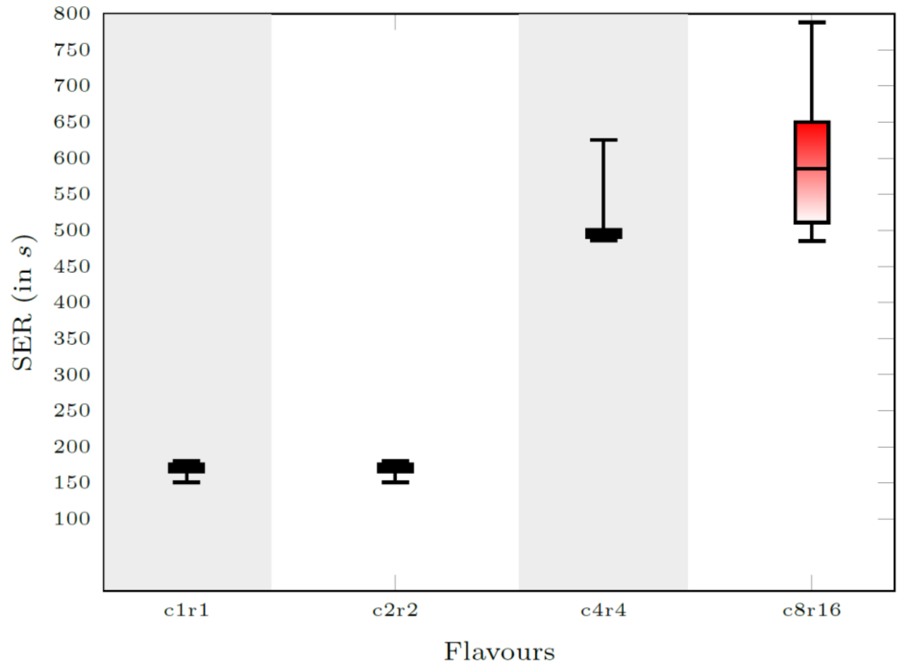

FIGURE 42: VARIABILITY OF THE SER WHEN INCREASING THE FLAVOUR OF THE MME

Regarding the considered KPI CST, we evaluate the cost of the service according to the size of the hosting cloud, by increasing the size of the flavours of the MME and by choosing either the scaling mode in horizontal or vertical. To increase the capacity of the MME, if we increase the amount of virtualized resources (CPU and/or RAM) in same virtualized deployement unit (VDU) then it is vertical scaling. While horizontal scaling means adding more instances of VDU to the pool of MME where the load can be distributed.

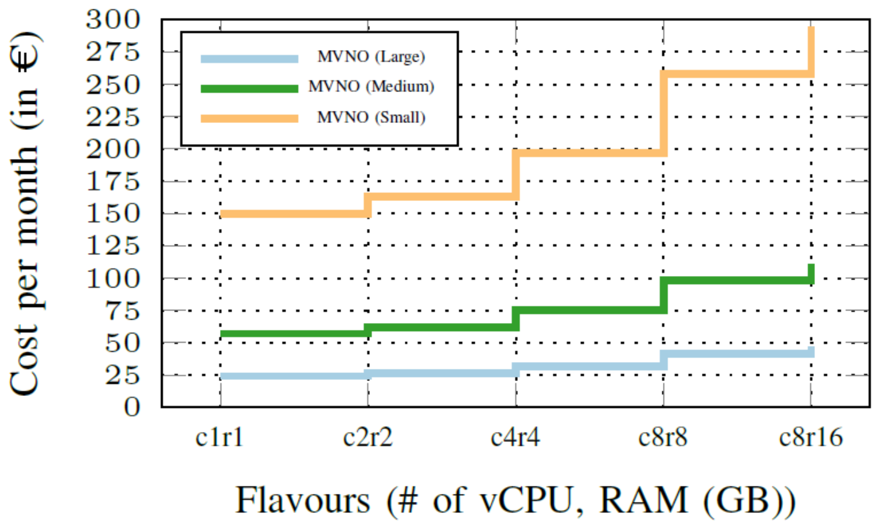

FIGURE 43: INFRASTRUCTURE COST PER MONTH (CST) CALCULATED FOR THE MVNO SERVICE DEPLOYED ON DATACENTRE OF THREE SIZES (SMALL, MEDIUM, LARGE), CONSIDERING VERTICAL SCALING OF THE MME

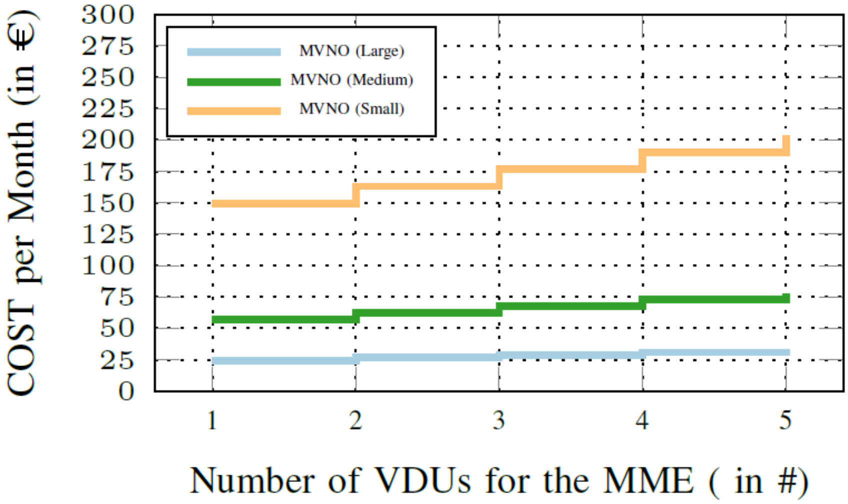

FIGURE 44: INFRASTRUCTURE COST PER MONTH (CST) CALCULATED FOR THE MVNO SERVICE DEPLOYED ON DATACENTRE OF THREE SIZES (SMALL, MEDIUM, LARGE), CONSIDERING HORIZONTAL SCALING OF THE MME

From Figure 43 and Figure 44, we may note the following remarks:

- As a DU is composed of resources, CST is proportional to the quantity of resources and the unit cost of the resources
- As stated in D1.4 [6], the unit cost of a virtualized resource such as a vCPU or a GB of RAM decreases as the size of the datacentre grows by volume effect on the hardware. This implies the CST for small datacenter is higher than the CST for medium datacenter, which in turn higher than the one for large datacentre.
- Horizontal scaling is more economical than vertical scaling because it allows to choose the granularity of the scaling step via the VDU. That is to say, the allocation of 3 VDUs of the flavour c1r1 for the MME (so, we obtain 3 vCPUs, and 3GB RAM) in the horizontal scaling consists of three instantiation levels while vertical scaling requires an allocation of one VDU of the flavour c4r4 in one instantiation level.

The bottom line is the 5G-Transformer project is contributing to offer more flexibility to a MNO to provision a network slice as a service that hosts a MVNO. By judiciously choosing VNF flavours by level of instantiation, the network operator can make the overall network slice deployment operation faster and speeds time-to-market. In addition, smaller horizontal scaling steps better adjust the amount of allocated virtualized resources to the demand thus they lead to an effective infrastructure cost reduction.

4 Conclusion

The deliverable provided a final report on the five vertical PoCs associated to the different UCs from the different vertical industries that are part of the project: Automotive, Entertainment, E-Health and E-Industry and the MNO/MVNOs UCs.

The document demonstrated how the WP5 has been able to develop a fully integrated 5GT platform composed of the final release of the platform building blocks developed in WP2, WP3 and WP4, implementing and building the five vertical PoCs to demonstrate the innovative features of the developed 5GT system from the project.

The report described, furthermore, the planning of the second and final evaluation phase which, completes the evaluation reported in the deliverable D5.3 delivered on M18. The evaluation focused on assessing whether the technologies developed for the 5G-TRANSFORMER framework achieved the Key Performance Indicators (KPI) expected by the considered verticals. These measurements results were compared to the state of the art (benchmark) or the used ones in common practice to evaluate the performance gain achieved in terms of KPIs.

Overall, we can conclude that there were some positive indications from KPIs measures in all five evaluation experiments. In some cases, these indications were based on demonstrating the POCs in the 5G-TRANSFORMER testbed and via simulations.

The Automotive PoC demonstrated with a real field trial how the 5GT platform functionalities facilitated the vertical service arbitration of a vertical offering in an automotive context involved multiple services (EVS service and video streaming service) with different priorities running in parallel according with the available resources. The adoption of the 5GT technologies demonstrated that the platform can ensure the required Low Latency KPIs of below 20ms by maintaining the Reliability (above 99.9%) also when vehicles density increases by scaling up the requested services at the edge with the application of edge computing technology. A different set of evaluation of the automotive EVS service with the aid of simulation has allowed to assess the EVS algorithm performance, from the results we can conclude that for the simulated urban scenario and for a variety of realistic urban densities and speeds, the designed EVS algorithm running over a simulated LTE network infrastructure (across LTE RAN and EPC) is able to avoid all collisions.

The Entrainment PoC demonstrated how the 5GT platform performed the automated service provisioning and orchestration of an UHD video streaming service, at the edge based an abstraction of network and compute topology and resources. The benefits achieved by using 5G-TRANSFORMER technologies in this use case were demonstrated by measuring KPIs such as E2E Latency, User data rate and Service Creation Time. For this specific use case, reducing the Service Creation time to a few minutes is the aspect to be highlighted since, until now, deploying a service of this nature could take several hours or even days.

The E-Health PoC mainly focused on demonstrating the multi-domain service orchestration (i.e., service federation) along with service life-cycle management across one or multiple administrative domains. Thus, the PoC was set on two testbed sites: 5TONIC and CTTC test sites. Two independent 5G-TRANSFORMER platforms, acting as administrative domains on each site, show how 5G-TRANSFORMER technologies enable the management of distributed federated services and edge functionalities. KPIs measurements were mainly performed in D5.3 [1] where the service availability, latency and positioning are shown that perform better then or within the nominal KPI values. In D5.4 the E-Health PoC demonstrated the benefits achieved by using 5G-TRANSFORMER by significantly improving the Service Creation Time, in the order of minutes (other than in hours) utilizing the service federation.

The E-Industry PoC demonstrated the management of automatic allocation of network and cloud resources across 5G RAN, optical-based EXhaul (fronthaul and backhaul) and the core network functionalities (vEPC) in a factory environment for Cloud Robotic service.

The service allows the interaction and coordination of multiple (fixed and mobile) robots controlled by remote distributed services, satisfying strict latency and bandwidth requirements. Three KPIs were evaluated such as Latency, Reliability and Service Creation Time. The use case shows that the platform can reduce the service creation time in order of minutes (from several hours or even days), maintaining the reliability (99,999% of service lifecycle as recommended by ITU). In addition, thanks to 5G the latency is reduced; this allows the “migration” of several functionalities from the robots to the cloud reducing its complexity and power consumption.

Finally, the MNO/MVNO PoC assessed the deployment of 3 network slices (echographer (URLLC), video (eMBB) and IoT devices (mMTC)) considering two KPIs such as the Service Creation Time and Infrastructure Cost.

References

[1] 5G-TRANSFORMER, "D5.3 Experimentation results and evaluation of achievements in terms of KPIs," June 2019.
[2] 5G-TRANSFORMER, "D5.2 Integration and proofs of concept plan," 2019.
[3] ETSI, "ETSI GS NFV-MAN 001," December 2014.
[4] F. Giannone, P. Frangoudis, A. Ksentini, L. Valcarenghi, "Experimental evaluation of 5G MEC-based Adaptive Video Streaming Service", submitted for publication to Elsevier Compute Networks.
[5] G. Avino, P. Bande, P.A. Frangoudis, C. Vitale, C. Casetti, C.F. Chiasserini, K. Gebru, A. Ksentini, G. Zennaro, "A MEC-based Extended Virtual Sensing for Automotive Services," IEEE Transactions on Network and Service Management, in press.
[6] 3GPP, "3GPPP TS 28.530," December 2018.
[7] P. Iovanna, F. Cavaliere, F. Testa, S. Stracca, G. Bottari, F. Ponzini, A. Bianchi, and R. Sabella, "Future Proof Optical Network Infrastructure for 5G Transport," IEEE/OSA Journal of Optical Commununications and Networking, vol. 8, no. 12, pp. B80-B92, 2016.
[8] N. Baldo, M. Miozzo, M. Requena-Esteso, and J. Nin-Guerrero, "An Open Source Product-oriented LTE Network Simulator Based on Ns-3,"in Proceedings of the 14th ACM International Conference on Modeling,Analysis and Simulation of Wireless and Mobile Systems, ser. MSWiM'11, New York, NY, USA: ACM, 2011, pp. 293-298. [Online]. Available:http://doi.acm.org/10.1145/2068897.2068948.
[9] ns-3 network simulator. Available at: nsnam.org.
[10] Simulation of urban mobility (SUMO). Available at: https://sumo.dlr.de/index.html
[11] TraCI Protocol Documentation. http://sumo.dlr.de/wiki/TraCI

5 Appendix

5.1 360º Immersive Telepresence: Remote Robotic Control

The “360º Immersive Telepresence: Remote Robotic Control” PoC was designed to showcase synergies between 5G-TRANSFORMER and 5G-CORAL, a H2020 initiative featuring European and Taiwanes partners with major focus on the convergence between radio access interfaces and edge/fog computing technologies. The PoC was initially demonstrated at the European Conference on Networks and Communications (EuCNC) 2019 in Valencia, Spain, nevertheless its development and validation are still ongoing. Future demonstrations and trials are expected to be conducted within ongoing research projects, such as 5G-VINNI and 5Growth, with the goal of assessing the benefits of such technology in an industrial environment.

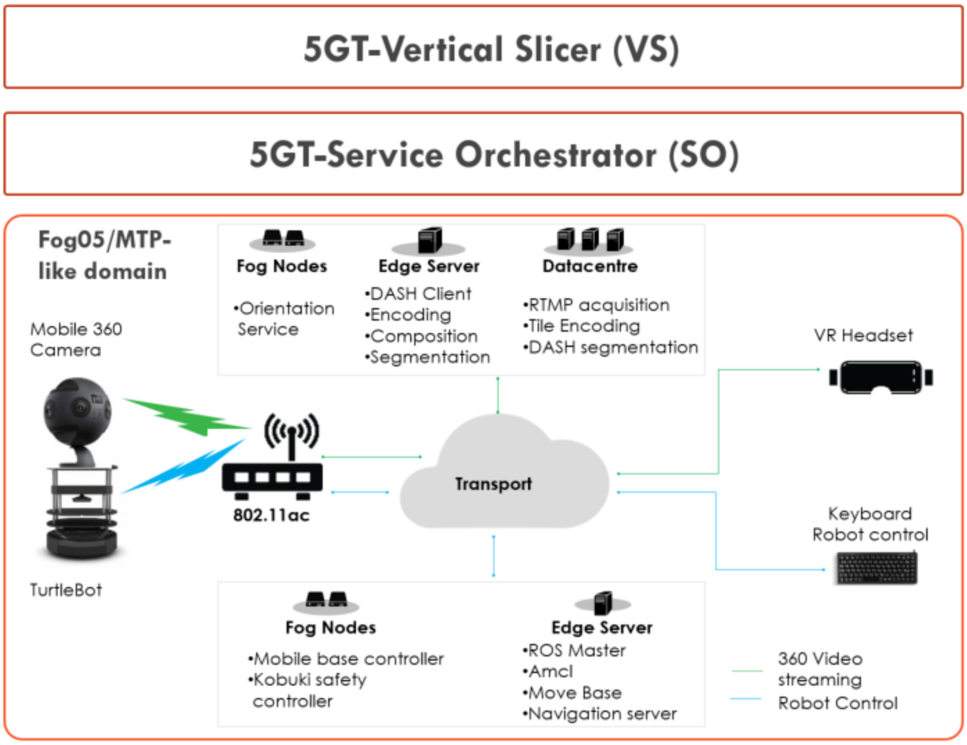

FIGURE 45: OVERVIEW OF THE POC ARCHITECTURE

Figure 45 illustrates the high-level architecture of the PoC. Two key components are the 5GT-VS and 5GT-SO, while the MTP is provided by 5G-CORAL, which features a fog computing operating system called Fog05, capable of managing, maintaining and provisioning computing, storage and connectivity resources distributed across multiple tiers, i.e., cloud, edge and fog. In this context, the use case intends to validate the benefits of having network slicing together with a distributed and dynamic resource allocation spanning across multiple layers. More specifically, a moving robot equipped with a 360° video camera is remotely controlled by an operator through a keyboard and a VR headset. The vertical customer creates such service by instantiating two slices: a high bandwidth, deployed to accommodate the 360° video streaming service, and a low latency slice, deployed to convey commands sent by the operator to the robot.

As shown in the figure, the two services rely on a number of computing tasks distributed among the different layers and coordinated by fog05, which receives requests from the service orchestrator, that processes the slice instantiation requests sent by the VS. Such approach delivers more flexibility and agility during the service establishment and takes advantage of the fog/edge resources to decompose the video and robotic service in multiple tasks that can be executed by heterogeneous devices. For instance, as illustrated above, some of the modules responsible for the robot actuation are assigned to fog nodes, whereas the navigation modules are executed by the edge server. Furthermore, the PoC features an adaptive tile-encoding streaming mechanism able to reduce the bandwidth required to deliver the 360 video content to the client by taking into account the user orientation. Figure 46Figure 46: shows a 33% bandwidth reduction by using the adaptive tile-encoding technique (21 Mbps) when compared against a non-optimized encoding approach (31 Mbps), where the 360-degree frames are all delivered with the highest quality. For the sake of clarity, the figure also reports the case (blue line) where the video frames are encoded at the lowest quality.

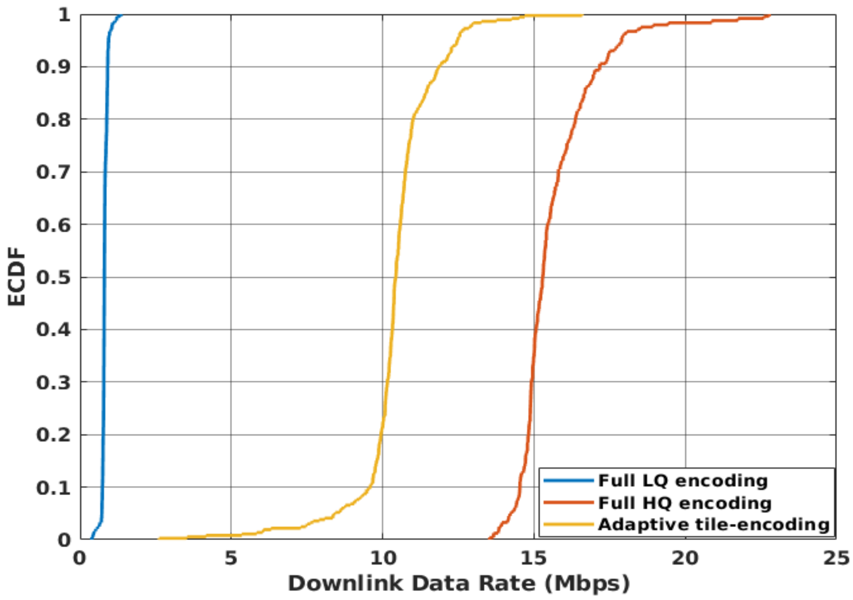

FIGURE 46: EMPIRICAL CDF OF THE DOWNLINK RATE MEASURED AT THE USER SIDE