H2020 5G-Coral Project

Grant No. 761586

D1.3 – 5G-CORAL refined system design and future directions

Abstract

This is the third and last deliverable from Work Package 1 (WP1), targeting the final 5G-CORAL system design, the definition of both procedures and techniques for incremental deployment of the 5G-CORAL platform into existing networks as well as the identification of the future directions of the 5G-CORAL solution. The refined 5G-CORAL architecture takes the outcome of validation process within the Project into account and details at system level the interaction between the EFS (Edge and Fog computing System) and the OCS (Orchestration and Control System) needed for establishing an end-to-end service making use of Edge and Fog resources. Furthermore, the capability of the OCS to manage heterogeneous virtualization platforms has been assessed on a per use case basis. Aiming at completing the analysis of the federation mechanism in 5G-CORAL (already provided in previous WP1 deliverables) the business requirements for each of the players in the 5G-CORAL ecosystem have been identified. Moreover, this deliverable also reports the State-of-the-Art for both the performance metrics and technical features taken into account by the 5G-CORAL use cases. Future directions of the 5G-CORAL platform have been elaborated in order to improve the applicability of the 5G-CORAL solution to the Project's use cases as well as new potential ones, such as the on-device Artificial Intelligence (AI) and Machine Learning (ML) as well as private wireless networks.

Document properties

Document number	**D1.3**
Document title	5G-CORAL refined system design and future directions
Document responsible	TI
Document editor	Damiano Rapone
Target dissemination level	Public
Status of the document	Final
Version	1.0

Table of contributors

Partner	Contributors
TI	Damiano Rapone
ITRI	Samer Talat, Chen-Hao Chiu, Ibrahiem Osamah, Gary Huang
UC3M	Sergio González, Luca Cominardi, Kiril Antevski, Carlos J. Bernardos
IDCC	Giovanni Rigazzi
AZCOM	Riccardo Ferrari, Giacomo Parmeggiani
SICS	Bengt Ahlgren
EAB	Chenguang Lu, Daniel Cederholm
NCTU	Hsu-Tung Chien
ADLINK	Gabriele Baldoni
TELCA	Aitor Zabala, Pedro Bermúdez

Production properties

Reviewers	**Antonio De La Oliva Delgado, Damiano Rapone, Bengt Ahlgren, Aitor Zabala, Hsu-Tung Chien**

Document history

Revision	Date	Issued by	Description
0.1	9 January 2019	TI	Draft
0.2	6 March 2019	TI	ToC reviewed after Feb.19 F2F meeting
1.0	31 May 2019	TI	Public Release

Disclaimer

This document has been produced in the context of the 5G-Coral Project. The research leading to these results has received funding from the European Community's H2020 Programme under grant agreement N° H2020-761586.

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Definition
(e)CPRI	(evolved) Common Public Radio Interface
5G	Fifth Generation
5GC	5G Core network
5GS	5G System
AF	Application Function
AI	Artificial Intelligence
AP	Access Point
API	Application Programming Interface
APN	Access Point Name
AR	Augmented Reality
AWS	Amazon Web Services
B2C	Business-to-Client
BBU	Baseband Unit
BSS	Business Support System
CAGR	Compound Annual Growth Rate
CAPEX	CAPital EXpenditures
CBRS	Citizen Broadband Radio Service
CD	Computing Device
CDMA	Code Division Multiple Access
CORD	Central Office Re-architected as a Datacenter
CPE	Customer Premises Equipment
CUPS	Control/User Plane Separation
D2D	Device to Device
DASH	Dynamic Adaptive Streaming over HTTP
DC	Data Center
DECOR	DEdicated CORe network
DL	Deep Learning
DN	Data Network
DSRC	Dedicated Short Range Communications
EFS	Edge and Fog computing System
EFS-VI	EFS Virtualization Infrastructure
EPC	Evolved Packet Core
EPS	Evolved Packet System
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
GAA	General Authorized Access
GPS	Global Positioning System
GPU	Graphic Processing Unit
HSS	Home Subscriber Server
IaaS	Infrastructure-as-a-Service
IF	Inference
IoT	Internet-of-Things
IP	Internet Protocol
IT	Information Technology
KPI	Key Performance Indicator
LBO	Local BreakOut
LLS	Lower Layer Split
MCPTT	Mission Critical Push To Talk
MEC	Mobile/Multi-access Edge Computing
ML	Machine Learning

MOCN	Multi-Operator Core Network
MQTT	Message Queuing Telemetry Transport
MTC	Machine Type Communications
MVNO	Mobile Virtual Network Operator
N3IWF	Non-3GPP InterWorking Function
NEF	Network Exposure Function
NF	Network Function
NFV	Network Functions Virtualization
NFVI	Network Functions Virtualization Infrastructure
OAM	Operation And Maintenance
OBD	On-Board Diagnostics
OBU	On-Board Unit
OCS	Orchestration and Control System
O-RAN	Open-RAN
OS	Operating System
OSS	Operations and Support System
OTT	Over-the-Top
OvS	Open virtual Switch
PAL	Priority Access Licensee
PDN	Packet Data Network
PGW	PDN GateWay
PHYD	Pay How You Drive
PoC	Proof-of-Concept
PoS	Point of Sale
QoS	Quality-of-Service
R&D	Research & Development
RAN	Radio Access Network
RAT	Radio Access Technology
RB	Resource Block
RF	Radio Frequency
RH	Radio Head
ROI	Return Of Investment
RRU	Radio Remote Unit
RSU	Road Side Unit
RTT	Round Trip Time
SAI	Switch Abstraction Interface
SBA	Service Based Architecture
SDK	Software Development Kit
SDN	Software Defined Network
SD-WAN	Software Defined Wide Area Network
SGW	Serving GateWay
SON	Self Organizing Network
UPF	User Plane Function
V2X	Vehicle-to-Everything
VIM	Virtualization Infrastructure Manager
VM	Virtual Machine
vMME	virtualized Mobility Management Entity
VNF	Virtual Network Function
VR	Virtual Reality
WAN	Wide Area Network
WHO	World Health Organization
WiMAX	Worldwide Interoperability for Microwave Access
WP	Work Package

Executive Summary

This is the third and last deliverable (Deliverable 1.3 – D1.3) in Work Package 1 (WP1) of the 5G-CORAL Project. The first deliverable D1.1 [1] focused on the initial system architecture along with the identification and prioritization of the most promising use cases from both the technical and economical point of view. The deliverable D1.2 [2], instead, was mainly focused on the analysis of the business perspectives of the 5G-CORAL technical solution, also reporting an update of the initial architecture according to the work conducted in Project's technical Work Packages and the activities related to demonstrations and Proof-of-Concepts (PoCs).

Departing from those findings, this third deliverable focuses on the final 5G-CORAL system design, the definition of both procedures and techniques for incremental deployment of the 5G-CORAL platform into existing networks as well as the identification of the future directions of the 5G-CORAL solution.

The key achievements of this deliverable are reported below:

- Identification of the business requirements from each of the player in the 5G-CORAL ecosystem – i.e. the telecom operator, software vendors, Over-The-Top (OTT) service providers, Cloud providers, hardware vendors and Vertical companies – which can be used to motivate the need to have federation mechanisms in 5G-CORAL
- Analysis of the State-of-the-Art for both the performance metrics and technical features taken into account in each of the 5G-CORAL use cases, to be used as a baseline for the assessment of performance and functional gains achieved when adopting the 5G-CORAL platform (this activity is in charge of WP4, see deliverable D4.2 [7])
- System level refinement of the 5G-CORAL architecture, summarizing how each use case makes use of multiple Radio Access Technologies (RATs) and how the EFS (Edge and Fog computing System) and the OCS (Orchestration and Control System) interact with each other in order for the OCS to deploy a certain Application/Function in the EFS so that another Application/Function could make use of its own services
- Definition of both procedures and techniques for the incremental deployment of the 5G-CORAL solution into existing networks, taking the viewpoints of both the telecom operator – highlighting the main differences between current 4G networks and incoming 5G systems – and the Vertical company into account
- Identification of the future directions of the 5G-CORAL solution aiming at improving the applicability of the solution itself to the already selected use cases as well as new potential ones, such as on-device Artificial Intelligence (AI) and Machine Learning (ML)
- Assessment (on a use case basis) of the OCS capability to manage heterogeneous virtualization platforms (i.e. Virtual Machines (VMs), containers and native applications) as well as the technical and business-related evaluation of the 5G-CORAL solution when applied in the context of private wireless networks. This is according to the outcome of the 5G-CORAL first technical review related to the Year 1 activities of the Project.
- A discussion on the perspectives of the 5G-CORAL solution from the system-level and business-oriented viewpoint of WP1, where the compliance of the 5G-CORAL platform is evaluated with respect to the key trends of the incoming years, that is, position with respect to Cloud, bandwidth cost savings, proper handling of Internet-of-Things (IoT) applications, Service Based Architecture (SBA) design approach, development of AI-based services.

1 Introduction

The overall aim of 5G-CORAL is to investigate and design an integrated Edge and Fog platform which is based on existing industrial frameworks, mainly ETSI Network Functions Virtualization (NFV) and ETSI Multi-access Edge Computing (MEC), complementing them with new features allowing to integrate also the Fog infrastructure. The initial system architecture was drafted in the first deliverable D1.1 [1], along with the most promising use cases – with related functional and non-functional requirements – from both the technical and economical point of view. The architecture, composed by two main building blocks termed 1) Edge and Fog computing System (EFS) and 2) Orchestration and Control System (OCS), was then revised and updated in deliverable D1.2 [2], based on the first outcome of the work conducted in Project's technical Work Packages (WPs), i.e. WP2 (for the EFS) and WP3 (for the OCS), also considering the activities related to demonstrations and Proof-of-Concepts (PoCs) in the scope of WP4. Together with the revised architecture, the deliverable D1.2 mainly focused on the analysis of the business perspectives of the 5G-CORAL technical solution.

After the identification of the business requirements for each of the players in the 5G-CORAL ecosystem and the investigation of the State-of-the-Art for both the performance metrics and technical features taken into account in 5G-CORAL use cases, D1.3 focuses on the final 5G-CORAL system design, by detailing the interaction between the EFS and the OCS – deemed as essential for establishing an end-to-end service making use of Edge and Fog resources – as well as the procedures and techniques for incremental deployment of the 5G-CORAL solution into existing networks. Furthermore, future directions of the 5G-CORAL solution are reported, in order to either improve the applicability of the solution itself – not only to the already addressed use cases but also considering new potential use cases, such as the on-device Artificial Intelligence (AI) and Machine Learning (ML) – or enhance its performance via exploitation of all the features characterizing the 5G-CORAL platform.

Finally, according to the outcome of the 5G-CORAL first technical review analyzing the Year 1 activities of the Project, D1.3 also focuses on 1) the capability of the OCS to manage heterogeneous virtualization platforms (this has been done on a per use case basis) and 2) the applicability of the 5G-CORAL solution in the context of private networks, along with related business considerations.

This deliverable is structured as follows:

Section 2 basically takes up and completes the analysis performed in the last WP1 deliverable, i.e. D1.2, where the business perspectives of the 5G-CORAL Edge and Fog solution applied to the Project's use cases have been analysed. In that context, the 5G-CORAL ecosystem has been defined, along with key players and roles. In D1.2 it was also highlighted that 5G-CORAL use cases typically comprise several components that might belong to different business actors or different administrative domains. As each of these components needs to interoperate with each other perfectly to deliver a plethora of end-to-end services to the end-user, the concept of *federation* among different administrative domains has been introduced. This allows the players of the ecosystem to create business relationships aiming at extending or creating new end-to-end services by importing external features or resources (being used for other services) from already deployed Edge and Fog systems. Federation is beneficial for all the business players targeting to expand their footprint, introduce new service spectrum, increase their competitiveness and keep track of increasing customer expectations. However, the business requirements motivating the need of resource federation for each of the player acting in the 5G-CORAL ecosystem – i.e. the telecom operator, software vendors, Over-the-Top (OTT) service providers, Cloud providers, hardware vendors and Vertical companies – have not been defined so far, hence they are listed for each of these players.

Section 3 analyzes and reports the State-of-the-Art for both the performance metrics and technical features taken into account in each of the use cases belonging to the scenarios foreseen in 5G-CORAL, i.e. the Shopping Mall (low mobility scenario), the Connected Cars (medium mobility scenario) and the High-Speed Train (high mobility scenario). This analysis has been conducted in order to establish a "bridge" towards the activities in WP4 related to the use cases' demonstrations and Proof-of-Concepts (PoCs), in order to have a baseline with respect to which the performance improvements and/or system enhancements achieved when adopting the 5G-CORAL platform can be assessed.

Section 4 initially focuses on the refinement of the 5G-CORAL architecture at system level, hence leaving the details of the internal characterization of both EFS and OCS to the corresponding final deliverables. It can be anticipated that the validation process within the Project has not identified the need for further modifications to the reference architecture, which basically is still the same as described in D1.2. Furthermore, since the 5G-CORAL platform has been designed having the convergence of multiple Radio Access Technologies (RATs) in mind, a summary of how each use case makes use of multiple RATs is also reported. Another aspect considered in the section is the system-level EFS-OCS interaction, which illustrates (at a high level) the process according to which the OCS deploys a certain Application/Function in the EFS in order to allow for another Application/Function to make use of the services provided by the previously deployed EFS Application/Function. In this analysis, also the interfaces involved in the EFS-OCS interaction are highlighted, along with the corresponding functionalities.

In addition, following the indications from the first technical review of the Project, the capability of the OCS to manage heterogeneous virtualization platforms, i.e. Virtual Machines (VMs), containers and native applications, has been assessed for each of the 5G-CORAL use cases. The analysis allows to demonstrate that the 5G-CORAL architecture is able to support multiple virtualization platforms properly, chosen according to the characteristics of the resources being involved.

Finally, procedures and techniques for the incremental deployment of the 5G-CORAL solution into existing networks are presented in this section, taking the viewpoints of both the telecom operator and the Vertical company into account. For the former, the analysis initially considers current 4G networks, where initial deployments of Edge and Fog solutions have to be necessarily designed as add-ons on top of the existing network. For the incoming 5G systems, instead, networks will be designed to efficiently support Edge and Fog since the beginning, so to flexibly achieve both high performance and quality of experience. When it comes to the Vertical company, the benefits in terms of security, user access control and service provisioning which could derive when deploying the 5G-CORAL solution within its own premises are reported.

Future directions of the 5G-CORAL solution are identified in **Section 5** aiming at improving the applicability of the solution itself to the already selected use cases as well as new potential ones, such as the on-device Artificial Intelligence (AI) and Machine Learning (ML) and, according to the outcome of the 5G-CORAL first technical review, also private wireless networks. Also, enhancements to the integration capabilities of the 5G-CORAL platform for full exploitation of all the features/external network entities needed to achieve higher performance are reported. It is noteworthy to highlight that, for the private wireless networks, also business considerations have been drafted, along with the technical feasibility of applying the 5G-CORAL solution in this scenario.

Finally, in **Section 6**, a conclusion is drawn to summarize the findings in this deliverable as well as setting the prospects for future work.

2 5G-CORAL business requirements

This section reports the business requirements for the different players acting within the 5G-CORAL ecosystem, i.e. the ones identified in deliverable D1.2 [2]: the telecom operator, software vendors, OTT service providers, Cloud providers, hardware vendors and Vertical companies. Departing from the findings of deliverable D1.2, the aim of this section is to complete and finalize the previous analysis, where it was highlighted that 5G-CORAL use cases typically comprise several components that might belong to different business actors or different administrative domains. As each of these components needs to interoperate with each other perfectly in order to deliver a plethora of end-to-end services to the end-user via full exploitation of the 5G-CORAL Edge and Fog platform – also striving to dynamically re-use and combine as many 5G-CORAL system components as possible – the concept of *federation* among different administrative domains has been introduced. This allows the players of the ecosystem, from different or same field of interest, to create business relationships aiming at extending or creating new end-to-end services by importing external features or resources (being used for other services) from already deployed Edge and Fog systems, without the need of deploying dedicated infrastructure or building new complex architectures. The benefits of resource federation have been demonstrated in deliverable D1.2 by applying it to Fog-assisted Robotics (also known as Cloud Robotics), Virtual Reality (VR) and High-Speed Train use cases, highlighting that federation is beneficial for all the business players targeting to expand their footprint, introduce new service spectrum, increase their competitiveness and keep track of increasing customer expectations. In order to properly implement those use cases requiring federation, the business requirements need to be investigated for each of the player acting in the 5G-CORAL ecosystem.

In the following subsections, the business requirements for each of the players within the 5G-CORAL ecosystem are reported: these requirements have been taken into account for the 5G-CORAL architecture design and validation, allowing use cases such as the Software Defined Wide Area Network (SD-WAN) to work properly, as it fully relies on federation. Each business requirement is labelled with an identifier (ID) in the form of "*B-Px-yz*", where "*B*" stands for *Business*, "*Px*" for *Player number x* and "*yz*" is the number assigned to each requirement: for instance, a requirement identified as B-P2-01 is the first business requirement specific for the 2nd player in the ecosystem, i.e. the software vendors.

2.1 Telecom operators

In [8] it is stated that the telecom operator needs to evolve its business model(s) by innovating across a combination of market, channel and network initiatives. In this sense, at the beginning, operators started to reconsider their core value proposition in such a way to expand their footprint into completely new sectors (e.g. utilities markets in the context of smart grid solutions, an example of which is the Deutsche Telekom and ABB partnership [9]) or to modify their offerings to address evolving business and consumer realities (e.g. for the mobile broadband pricing, operators such as AT&T and T-Mobile US are moving from the flat rate model into the tiered-based pricing model to stem monetization problems due to the adoption of flat rates [10][11]). Furthermore, telecom operators are making radical changes to traditional revenue models, by also considering the possibility to allow their own customers to be active parts in the infrastructure ownership, an example of which is the Norwegian fiber-optics operator Lyse whose customers have been involved in building fiber-based networks [12]. Moreover, other operators are working together with companies offering Business-to-Consumer Mobile Virtual Network Operator (B2C MVNO) services to support their particular businesses, such as Blyk which rents networks from larger operators to offer customers free of charge calls in exchange for advertisements [13].

Finally, operators are also trying to act directly on the areas impacting their cost base, e.g. by working towards simplifying complicated Operation Support Systems / Business Support Systems (OSS/BSS) such done by KPN [14] or by outsourcing significant parts of network operations to equipment vendors, e.g. Sprint signed a network outsourcing contract with Ericsson [15]. Other means also encompass the possibility to reduce the operators' environmental impact via an intelligent reduction in the power consumption of mobile networks, e.g. AT&T effort for achieving the 2020 goal of reducing the electricity consumption in the network by 60% as compared to 2013 baseline [16].
Table 2-I lists the business requirements related to the telecom operator which shall be considered in the design of the 5G-CORAL technical solution.

TABLE 2-I – 5G-CORAL BUSINESS REQUIREMENTS FOR THE TELECOM OPERATOR.

ID	Requirement	Description
B-P1-01	Customer-centric service/application development and delivery	In an increasingly competitive environment with shortened lifecycles, more complex distribution channels and ever-better informed customers, it is requested that the entire value chain is flexible enough to react quickly to cope with varying customer requirements, while keeping development and delivery of new services/applications faster, more flexible, and profitable. Customized services/applications will allow telecom operators to distinguish themselves and market unique solutions.
B-P1-02	"Fragment and Mash" approach	Companies can build skill sets across a range of sub-activities that are required to deliver a larger activity: by standardizing the best practices and by creating enablers for letting such secondary skills flourish, companies can create platforms for new services based exclusively on these skills. The result is the break-down of the classical sequential value chain into a value network, where company's competencies – which have been built up during the delivery of core services/applications – are used to help create new services/applications.
B-P1-03	"Long Tail" approach	Based on technological enablers allowing telecom operators to profitably serve niche segments that were not attractive to serve previously. The basis of this approach is that minimal/next-to-zero incremental cost makes targeting niche groups a potentially profitable business model, hence creating a strong value proposition for customers.
B-P1-04	Strategic partnering with other players in the business ecosystem, realizing dynamic and open business models	Business models need to be flexible enough to work in highly competitive markets but also open to allow any stakeholder to enter the ecosystem and enrich the value network, enabling an exchange of technology and revenue among the different players. The operator expects its service-layer technology to host applications supplied not only by its own development processes but also by 3rd-party developers, by equipment and software vendors and by OTT players. This can also consider the multi-operator partnering, in order to benefit from each other's existing solutions, establish joint applications/services' Research & Development (R&D) activities, reduce costs, exploit joint new and innovative business models, potentially enter new regional markets and extend the footprint of each other's services/applications.
B-P1-05	Disintermediation in making the "service/application provider" ←→ "end-user" linkage	To improve the delivery of new services/applications there is the need to reduce the number of stages between the service/application provider and the end-user. In extreme cases, the end-user even takes over tasks originally belonging to producers, indicating the importance of involving users of a service/application in creating the core that can then be monetized by the service provider.
B-P1-06	Outsourcing of activities	Outsourcing to external players (which might also be either

	traditionally done in-house	current or prospective customers) of activities such as service/application design as well as R&D in order to reduce costs and time-to-market.
B-P1-07	Multi-vendor environment	The telecom operator needs to run a network encompassing equipment of multiple vendors in order to avoid a lock-in with a single vendor. Multi-vendor network equipment need to be easily integrated within the operator's OSS to allow uniform operability of the network.
B-P1-08	Presence of a strongly regulated and standardized environment	This requirement plays an important role, as it is the basis for a highly competitive environment for both vendors and operators but also for a reduced flexibility in the service/application offering (with the risk of discouraging innovation in service/application development and delivery). Hence, a trade-off between these two conflicting aspects is needed, also considering that competitiveness might determine cost pressure on operators for timely deploying and operating the infrastructure.
B-P1-09	Support of legacy systems across operator's digital transformation process	Typically, legacy systems are seen as obstacles to the digital transformation process but replacing them is not always the best choice: means to adapt them to meet the demands of today's business world without "starting from scratch" are needed. This because an all-new system represents a serious effort not only in terms of CAPital EXpenditures (CAPEX) but also in time and resources required to train staff and to redefine working policies.
B-P1-10	Security	Security, integrity, data protection and privacy of end-users' data via rules which apply to all providers offering equivalent services.
B-P1-11	High efficiency and flexibility in transformation of operational structures	It is well known that operators acting within the traditional telecommunications market have to deal with two contrary conditions: stagnating (or even decreasing) prices and extensive investments in network infrastructure. Therefore, in order to realize the financial flexibility for investments in innovative applications/services, operators have to reduce costs and, at the same time, increase operational efficiency properly.

2.2 Software vendors

Software vendors are evolving their business models since Cloud environments are opening new business opportunities. In deliverable D1.2 [2] three possible roles in the 5G-CORAL ecosystem (and hence three business models) have been envisioned for software vendors: i) Edge and Fog application/service provider, ii) Edge and Fog application/service software developer and iii) Edge and Fog system software vendor. The dynamic nature of the 5G-CORAL platform, whose components might belong to different players within the ecosystem, makes the concept of federation a key mechanism allowing the unification of all the resources as a cohesive entity. From the point of view of a software vendor in the 5G-CORAL ecosystem, the federation of resources and services can be applied to the first two models, to offer services/applications to the end customer or even to 3rd-party developers. The federation of resources is more critical in the third business model where a software vendor develops the EFS system software, since it integrates resources from different administrative domains at different levels (Cloud, Edge and Fog). In the three business models both static and open federation can be applied. Due to the dynamicity of the 5G-CORAL platform, the open federation is more suitable. In terms of administrative domains involved, these business models can fit into horizontal, vertical and hybrid federation, since the 5G-CORAL components might belong to different tiers (Cloud, Edge and Fog). The flexibility and dynamicity provided by the 5G-CORAL ecosystem introduce new business opportunities but, at the same time, pose new business requirements, the resources are more volatile and dynamic and require federation and integration of different platform layers, better resource utilization, etc. These business requirements are detailed in Table 2-II.

TABLE 2-II – 5G-CORAL BUSINESS REQUIREMENTS FOR THE SOFTWARE VENDOR.

ID	Requirement	Description
B-P2-01	Scalability	The software vendor shall ensure the provided software can handle a growing number of end-users by making a proper use of the EFS resources. The software vendor shall monitor the status and capacity of the resources to determine if it is enough to handle the current workload.
B-P2-02	Security and Privacy	The software vendor shall properly secure and isolate the end-user's data making a correct use of the information.
B-P2-03	Programmability	The software vendor shall provide tools, libraries and frameworks to 3rd-party developers to develop services/applications or to interact with other services/applications. The software vendor shall develop the software facilitating future upgrades and improvements.
B-P2-04	Integration	The software vendor shall be able to integrate resources from different federated administrative domains and from different platform layers (Cloud, Edge and Fog).
B-P2-05	Quality of experience	The software vendor shall provide a good quality of experience to end-users and 3rd-party developers. The software vendor shall monitor the status and the resources to place/migrate the software to the most suitable platform.
B-P2-06	Dynamicity	The software vendor shall make use the different EFS layers (Edge and Fog) dynamically depending on the software requirements.
B-P2-07	Maintenance & Support	The software vendor shall maintain and update the developed software and provide support to 3rd-party developers and end-users.

2.3 OTT service providers

OTT service providers play a major role as stakeholders in 5G-CORAL ecosystem, providing OTT services, over the 5G-CORAL platform [17]. Requirements are defined so OTT providers can comply with strict SLAs with their customers. In Table 2-III, 5G-CORAL platform gathered the most important foreseen requirements [18]. Requirements range from resource utilization to the quality of experience of the end-user. It should be noted that resource utilization is a key requirement to consider, as it defines the ease of interaction of the OTT provider and 5G-CORAL platform. Hereunder, isolation, security and privacy requirements are key requirements to rise the interest of providers handling critical and/or sensitive information, guaranteeing trust in the platform. Furthermore, subscription management requirement is necessary to ensure OTTs are being billed correctly. Finally, quality of experience requirement ensures OTT providers can provide their services correctly.

TABLE 2-III – 5G-CORAL BUSINESS REQUIREMENTS FOR OTT SERVICE PROVIDERS.

ID	Requirement	Description
B-P3-01	Resource utilization	OTT provider shall be able to utilize computing and network resources in order to satisfy Quality-of-Service (QoS). An OTT provider shall be able to control computing and network resources by leveraging configuration parameters. This will allow the control of the QoS level, multi-RAT selection policy, etc. Additionally, OTT providers should be able to control computing and network resources in real-time such as the bandwidth reservation, real-time RAT selection, etc.
B-P3-02	Isolation	OTT providers shall be convinced that a resource (video streaming for example) in the Edge and Fog system are fully

		utilized to support OTT services as like isolated virtual resources, which enable OTT providers to supply faithful OTT services to other players.
B-P3-03	Security and Privacy	OTT providers shall be guaranteed that an OTT service is secure and isolated inside a private network. Additionally, OTT services shall be guaranteed that their end-users information will not be compromised.
B-P3-04	Subscription Management	OTT providers shall be able to support differentiated OTT services according to subscriptions, which could be managed by some functions such as access control, QoS handling, etc. This is important as an OTT provider will be able to provide their own services targeting just their subscribers. In addition, OTT providers shall be able to account the service utilization on the Edge and Fog system for billing.
B-P3-05	Quality of experience	OTT providers shall provide a good quality of experience to end-users. OTT providers shall monitor the status and the resources to place/migrate the container content (media for example) to the most suitable platform.

2.4 Cloud providers

Cloud providers are expected to be major stakeholders in the 5G-CORAL ecosystem, being one of the players responsible for supplying computing, networking and storage resources, and even entire platforms or development environments. Ultimately, the Cloud layer is meant to provide most of the resources required by 5G-CORAL end-users, since Cloud providers can typically lease them at competitive rates. Table 2-IV presents the business requirement of the 5G-CORAL platform to be desired by Cloud providers. It can be noted that capability assessment and resource utilization are key requirements, as a Cloud provider needs to have a clear visibility of the resources utilized in the platform in order to report available and suitable Cloud resources to 5G-CORAL. Also, security and privacy must be ensured, along with subscription management, which allows the provider to account the service utilization and bill accordingly. Finally, it is pointed out that data service accessibility must be guaranteed, since a provider may benefit from information provided by the Edge and Fog system, and programmability must be possible, as Cloud providers may need to offer software development kit, including Application Programming Interfaces (APIs) and testing tools.

TABLE 2-IV – 5G-CORAL BUSINESS REQUIREMENTS FOR THE CLOUD PROVIDERS.

ID	Requirement	Description
B-P4-01	Capability and Capacity assessment	Cloud provider shall be able to assess the capability of Edge and Fog system such as computing resources, network resources, Operating System (OS) availability, etc. Also, Cloud provider shall be able to monitor the status of capacity in real-time. Correspondingly, Cloud provider could determine which Cloud services are available.
B-P4-02	Resource utilization	Cloud provider shall be able to utilize computing and network resources in order to satisfy QoS. Cloud provider could control computing and network resources by configuration value in average manner such as QoS level, multi-RAT selection policy, etc. Either, Cloud provider could control computing and network resources in real time such as bandwidth suggestion, real-time RAT selection, etc.
B-P4-03	Isolation	Cloud provider shall be convinced that a computing resource and a networking resource in the Edge and Fog system are fully utilized to support Cloud services as like isolated virtual computer and network, which enable Cloud provider to supply faithful Cloud services to other players.
B-P4-04	Security and Privacy	Cloud provider shall support a Cloud service to be secured

		and isolated to private network.
B-P4-05	Subscription Management	Cloud provider shall be able to support differentiated Cloud services according to subscriptions, which could be managed by some functions such as access control, QoS handling, etc. In addition, Cloud provider shall be able to account the service utilization on the Edge and Fog system for billing.
B-P4-06	Data service accessibility	Cloud provider shall be able to access useful data services provided by Edge and Fog system, which could fertilize an efficient Cloud service. For an instance, location service could provide footprint of a service over the geographical service area map.
B-P4-07	Programmability	Cloud provider shall provide easy implementation tools and interface be interacted commonly for software developer and system developer.

2.5 Hardware vendors

As depicted in deliverable D1.2 [2] the business environment is changing drastically for the hardware vendors as more and more functions are virtualized and a white-box approach, where the hardware equipment is very much standardized, becomes more common. This leads to some vendors targeting the white-box market with very large volumes (but typically with quite low margins), while other vendors are trying to climb up the value chain by also taking on the Edge and Fog system vendor role and/or selling managed services.
Table 2-V lists the Edge and Fog business requirements for the hardware vendors acting in the role of traditional equipment producers. For hardware vendors also acting as system vendors the requirements will overlap with the software vendors and can thus be found in Table 2-II.

TABLE 2-V – 5G-CORAL BUSINESS REQUIREMENTS FOR THE HARDWARE VENDORS.

ID	Requirement	Description
B-P5-01	Standardized interfaces and APIs	The hardware vendors shall support relevant standardized interfaces and APIs to make the hardware more open and accessible for different players. For example, Switch Abstraction Interface (SAI) [19] is a standardized set of APIs used in white-box switches. Another example is Common Public Radio Interface (CPRI)/eCPRI/Open-RAN Lower Layer Split (O-RAN LLS) standards for the fronthaul interface between the Remote Radio Unit (RRU) and the Baseband Unit (BBU).
B-P5-02	Modularity	The hardware vendors shall be able to deliver equipment in different form factors, suitable for different deployments. This is especially important for Fog environments where the equipment often must be small with low power consumption and hardened to tolerate vibrations and high temperatures. The equipment should also be easy to replace in case of failures.
B-P5-03	Security	The hardware vendor shall apply adequate tampering protection for equipment indented to be located in open locations.
B-P5-04	Operations and Maintenance	The hardware vendor shall provide an abstraction of Operation And Maintenance (OAM) features so that the equipment can be managed, and the health and resource utilization can be monitored by the EFS.

2.6 Vertical companies

In deliverable D1.2 [2] it was stated that, in the context of the 5G-CORAL ecosystem, the Vertical company could act the role of the Edge and Fog site owner which manages and owns Edge and Fog nodes as facilities. These facilities might be lent by the Edge and Fog system provider, e.g. the telecom operator, since it may be difficult for it to host the facilities, especially for Fog nodes, or it needs to extend both the network coverage and capacity. For instance, Edge and Fog site owners, e.g. Shopping Mall owner and road-side/on-board unit owners in case of the Connected Cars scenario, could be connectivity providers by sharing/lending their facilities. In other words, the telecom operator (Edge and Fog system provider) is able to integrate 3rd-party resources – provided by the Vertical companies (Edge and Fog site owners) – as part of its network in order to provide an end-to-end service by leveraging and integrating the already deployed corporate Data Centers (DCs) and IT environment. Therefore, there is a new business model for Vertical companies, according to which several new business requirements will be posed for them.

Table 2-VI presents the Vertical companies' business requirements, with the Vertical company acting as the Edge and Fog site owner interfacing with the Edge and Fog system provider which provides the 5G-CORAL platform.

TABLE 2-VI – 5G-CORAL BUSINESS REQUIREMENTS FOR THE VERTICAL COMPANIES.

ID	Requirement	Description
B-P6-01	Scalability	The Vertical company shall ensure that the provided Edge and Fog nodes can fulfil the EFS requirements of Edge and Fog system providers.
B-P6-02	Security	The Vertical company shall make the locations safe to properly secure Edge and Fog nodes.
B-P6-03	Stability	The Vertical company shall monitor the status of Edge and Fog nodes to ensure nodes are alive.
B-P6-04	Maintenance & Support	The Vertical company shall maintain and upgrade their Edge and Fog nodes to support an EFS.

3 5G-CORAL use cases' KPIs: State-of-the-Art values for the performance metrics

This section analyzes and reports the State-of-the-Art for both performance metrics and technical features taken into account in each of the 5G-CORAL use cases. This analysis has been conducted to establish a link to the activities in WP4 related to the use cases' demonstrations and PoCs (see deliverable D4.2 [7]), in order to have a baseline with respect to which the performance improvements and/or system enhancements achieved when adopting the 5G-CORAL platform can be assessed.

3.1 Shopping Mall scenario

The following subsections report the State-of-the-Art for performance metrics and technical features characterizing each use case in the context of the Shopping Mall, i.e. the low mobility scenario.

3.1.1 Augmented Reality (AR) Navigation

Augmented Reality (AR) Navigation provides end-users within the Shopping Mall an intuitive way to reach target shop(s) or to identify a shop on the screen by adding digital content on top of the video captured from the smartphone camera. The user will see a guiding line on top of the real-world image displayed on the smartphone screen so to remind a real object, i.e. a pointer, to the desired destination. Moreover, end-users will be able to see shop promotions on their screen whenever they pass by the store, enhancing the shopping experience for the customers. Current State-of-the-Art for the AR technology such as ARcore released by Google [20] provides AR service on the end-terminal without any external computation support, hence relying on up to 1000 reference images which can be set in a single image database [21]. This, however, may not be feasible in a big Shopping Mall because many more than thousands of images may be needed to support AR Navigation and shop identification simultaneously. By leveraging the concept and design of Fog Computing Devices (Fog CDs) and Edge computing architecture – both providing more computation power at the Edge and, at the same time, reducing the overall communication delay – the AR Navigation can be technically feasible. Therefore, it is expected that more reference images can be stored in a physically distributed database. Table 3-I shows the State-of-the-Art performance metrics related to the AR Navigation use case. The main performance metrics in this use case are latency, delay, processing power by segmented database and localization precision.

TABLE 3-I – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO AR NAVIGATION.

Performance metrics	Description	Values
Latency	Round Trip Time (RTT) between issuing a command from an end-terminal and execution from a Fog CD	500 ms between the end-terminal and the Fog CD and 850 ms between the end-terminal and the Cloud [22].
Delay	Communication delay between the Fog CDs (for distributed computing)	11.8 s for a 521 MB file and 25.6 us for a small file [23].
Processing Power by Segmented Database	Time required by a Fog node to complete a job	~30 ms at the Cloud and ~200 ms at the end-terminal [24].
Localization Precision	Error between true location points and computed location points	500 ms between the end-terminal and the Fog CD and 850 ms between the end-terminal and the Cloud [22].

3.1.2 Fog-assisted Robotics

Fog-assisted Robotics benefits from the low-delay communication provided by the Fog to move the intelligence of the robots to the network. This approach enables a better coordination, synchronization and multi-robot collaboration while reducing the computation requirements at the robot side.

The Fog-assisted Robotics use case comprises two possible situations: the first envisions the robots cleaning the common areas of the Shopping Mall. This cleaning task is automatically triggered when deemed necessary, based on the video recorded by the Shopping Mall cameras, processed and analyzed by the Fog-assisted Robotics application. The second scenario, instead, envisions the delivery of goods by a group of robots working synchronously. Robots are coordinated and synchronized to deliver goods from the Shopping Mall warehouse to the shops.

Table 3-II shows the state-of-the-art for the performance metrics related to Fog-assisted Robotics use case. The main performance metrics in this use case are delay, latency, deployment time and migration time.

TABLE 3-II – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO FOG-ASSISTED ROBOTICS.

Performance metrics	Description	Values
Delay	Communication delay between the Fog CDs and the robots	~50 ms RTT between robots and Cloud [25].
Low-latency	The latency between issuing a command remotely (from the robot intelligence application) and execution of the command by the robot	~3 ms to set speed to the robot from Edge server [26] ~6 ms to read Robot's sensors from Edge server [26] ~35 ms to take a photo from robot's camera from Edge server [26].
Deployment time (of Access Points, APs)	The time it takes to create the new APs in the Fog CDs using the hostapd	6~8 s to deploy Docker containers of different sizes [27].
Migration time	The time it takes to migrate the robotic intelligence from the central point, closer to the edge. As well as the time it takes to migrate from one to another Fog CD, due to mobility of the robots	8~10 s to migrate Linux Containers (LXC) [28] 370 ms for initial 802.11r handover (including the full 802.1X authentication), ~30 ms for all consecutive handovers [29] 6.3~70 s to migrate LXC containers, (the migration time depends on the container size) [30].

3.1.3 Virtual Reality (VR)

The Virtual Reality (VR) use case showcases an end-to-end 360 immersive video streaming service, where computing tasks are distributed over three tiers, i.e. Cloud, Edge and Fog, with the goal of enhancing deployment flexibility, offloading computational complexity from the terminal device and improving network efficiency. The video service relies on the viewport adaptive streaming, which enables tile-based encoding and significantly reduces the bandwidth required to accommodate the streaming service. To this end, Fog nodes are deployed in the proximity of the end-users and collect the viewing orientation, which is then forwarded to the Edge server in order to optimize the Dynamic Adaptive Streaming over HTTP (DASH) operations. Table 3-III lists related performance metrics and reports corresponding reference values.

TABLE 3-III – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO VR.

Performance metrics	Description	Values
Bandwidth	Viewport adaptive streaming approach reduces the bandwidth required to stream 360-degree video while maintaining the same viewing experience	180 Mbps [31] (Video service with 60 fps and 4K resolution).
Service Setup Time/Takedown Time	Service setup time refers to the time it takes for the OCS to instantiate and provision all the VR components, distributed across the EFS, to produce an end-to-end VR service	90 ms [32].
End-to-end Delay	End-to-end application delay among: Fog CDs, Edge server, Cloud Data Center (Cloud DC) and terminal clients.	20 ms [31] (Motion-to-Photon latency).
Power consumption	Power consumed by the Graphic Processing Unit (GPU) on the Cloud DC when all the tasks are executed on the same machine or when offloading is employed.	No reference (reasonable distribution ratio of power consumption among multiple substrates).

3.1.4 Multi-RAT IoT

The idea of this use case is to centralize and cloudify the baseband processing (together with the higher radio protocol layers) of massive Internet-of-Things (IoT) RATs to Edge servers (or Edge DC), instead of running locally in the single-RAT Access Points (APs) which implement the full stack of the considered RAT. Radio Heads (RHs) with lower complexity instead of full-stack APs are deployed for service coverage. They are mainly handling the Radio Frequency (RF) related operations/processing and the RH-Edge interface. In this way, multiple RATs can be supported at the Edge. The system provides much more scalability and flexibility than traditional deployments, since the Cloud infrastructure and technologies well-developed for Web services can potentially provide "unlimited" scalability. It is also more future proof as the RAT-specific processing is minimized in RHs. The software upgrades are on the Edge. This increases the life-time of the radio infrastructure. Further, it also facilitates the possibility for radio coordination features as the communication stacks are centralized. Table 3-IV lists both performance metrics and features to be considered for the Multi-RAT IoT use case and reports corresponding reference values.

TABLE 3-IV – STATE-OF-THE-ART VALUES FOR BOTH THE PERFORMANCE METRICS AND FEATURES RELATED TO MULTI-RAT IOT.

Performance metrics	Description	Values
Latency	Application layer end-to-end latency	<10 s for cellular IoT (5G requirement) [33], < a few 100s ms for non-cellular IoT.
Bit rate	Air interface bit rate	a few 10s bps – a few100s kbps for cellular IoT [33], a few 100s kbps to 2 Mbps for non-cellular IoT.
Connection density	The number of devices can be supported per km^2	1 million devices/km^2 for cellular IoT (5G requirement) [33].
Range	The distance the transmitter and the receiver can communicate given a reliability requirement.	164 dB maximum coupling loss at a rate of 160 bps for cellular IoT (5G requirement) [33], 10-100 meters for non-cellular IoT.
Multi-RAT support	Support multi-RAT or not	Generally, no support for multiple massive IoT RATs.
Future proofness	System is future proof or not	For cellular IoT, it is future proof with BBU support. For non-cellular IoT, not future proof.

3.1.5 Software Defined Wide Area Network (SD-WAN)

The SD-WAN use case aims at validating resource federation between domains in 5G-CORAL. In the scenario of a Shopping Mall, it is expected that a shop, for instance, deploys a database and Point of Sale (PoS) web application in the same domain. As a result of static federation, its current domain negotiates with another domain also located in the Shopping Mall in order for the source domain to federate a resource and deploy a second instance of the web application in the federated node, closer to the end-user. Table 3-V lists SD-WAN performance metrics and reports corresponding reference values.

TABLE 3-V – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO SD-WAN.

Performance metrics	Description	Values
Latency	RTT between a PoS terminal and a typical Cloud provider where the PoS Cloud service could possibly be instantiated.	18 ms [34].
Bandwidth	Bandwidth Uplink and Downlink KPI for a typical Cloud provider (Amazon Web Services).	9 Mbps uplink and 70 Mbps downlink [35].
Number of connected users	Bottleneck in the SD-WAN use case is in the RAT AP. So, the State-of-the-Art performance metric is the maximum 802.11 supported concurrent associations.	2007 concurrent associations.
Service Deployment Time	Instantiation of the federation components including the SD-WAN deployment.	150 s [Initial experimental result based in fog05 local set-up].

3.2 Connected Cars scenario

This subsection reports the State-of-the-Art for performance metrics and technical features characterizing the safety use case in the context of the Connected Cars, i.e. the medium mobility scenario. The safety use case for drivers and passengers requires low and stable latency communications. The 5G-CORAL safety use case aims at showing how this is possible by leveraging the deployment of Fog CDs nearby the vehicles, e.g. Road Side Unit (RSU) or On-Board Unit (OBU). This enables a distributed and quicker information exchange among the vehicles that reduces the end-to-end network latency and, at the same time, keeping it stable. There are several metrics related to this use case, in Table 3-VI the State-of-the-Art of the main metrics, i.e. latency and reliability, are shown.

TABLE 3-VI – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO SAFETY.

Performance metrics	Description	Values
Latency	Maximum tolerable elapsed time from the instant a data packet is generated at the source application to the instant it is received by the destination application.	100 ms* [36] 20 ms** [36].
Reliability	Maximum tolerable packet loss rate at the application layer, a packet is considered lost if it is not received by the destination application within the maximum tolerable end-to-end latency for that application.	90 %* [37] 95 %** [37].
* Mutual vehicle awareness use case ** Pre-crash sensing		

3.3 High-Speed Train scenario

This subsection reports the State-of-the-Art for performance metrics and technical features characterizing the High-Speed Train use case, i.e. the high mobility scenario.

When travelling on the train, various services utilized by passengers on board require a seamless connection, a requirement which is very challenging to fulfil due to frequent handovers. Moreover, it is expected that a massive signalling will originate from these handovers due to a large number of passengers on the train. Therefore, the 5G-CORAL's Edge and Fog infrastructure aims at achieving seamless connection via breakout of the mobility functions on the Fog/Edge that mitigates the burden of passenger's mobility signalling on the backhaul links, also reducing the overall signalling amount on these links. The High-Speed Train use case leverages the deployment of Fog CDs on-board with virtual Mobility Management Entity (vMME) functionality and service migration capabilities which allows for both content and context information to be migrated from the on-board to the on-land Fog CDs in advance when passengers are approaching the train station. This will reduce signalling storm during handover process. Also, this will minimize the service downtime for passengers.

There are several metrics related to this use case, Table 3-VII shows the State-of-the-Art of main ones: latency, service downtime and migration time.

TABLE 3-VII – STATE-OF-THE-ART VALUES FOR THE PERFORMANCE METRICS RELATED TO HIGH-SPEED TRAIN.

Performance metrics	Description	Values
Signalling reduction	The amount of control signalling between vMME on-board and MME located in the Core Network	Forward relocation request (response): 700 (300) bytes/user [38].
Handover Latency	RTT to complete inter-MME handover	0.9s for a 2-hop architecture implementation in lab [39][40][41].
Service Downtime	The amount of time from the moment the container is frozen at the source until it is restored at the destination. The migration process impacts on the QoS perceived by the end-user	2s [42][43][44][45].
Total Migration Time	The amount of time from the beginning of the migration process to occur until the container restoration at the destination node is finished. This includes the in-memory state pre-copying while the container continues to run at the source node.	3s [42][43][44][45].

4 Refined 5G-CORAL architecture and system design

The initial 5G-CORAL architecture was detailed in deliverable D1.1 [1] and updated in deliverable D1.2 [2]. The architecture is based on ETSI MEC and ETSI NFV frameworks and is composed of two major building blocks:

- *Edge and Fog Computing System (EFS)*: an EFS is a logical system subsuming Edge and Fog resources that belong to a single administrative domain. An administrative domain is a collection of resources operated by a single organization. An EFS provides a Service Platform, Functions and Applications on top of the available resources and may interact with other EFS domains. The EFS design leverages virtualization technologies that decouple the EFS entities from the underlying EFS resources, i.e. compute, storage and network. Some of the benefits derived from virtualization include: resource consolidation, isolation, faster provisioning, disaster discovery, dynamic load balancing, faster development and test environment, reduced hardware vendor lock-in, improved system reliability and security. While the EFS architectural design is compliant with ETSI MEC and ETSI NFV frameworks, the EFS provides two notable extensions: first, the EFS Virtualization Infrastructure (EFS-VI) extends the ETSI NFV reference architecture to incorporate mobile and volatile resources that have different levels of availability, mobility, storage, computing, networking and power capabilities. Furthermore, the EFS entities extend the ETSI NFV network functions to include EFS Applications and an EFS Service Platform [3].
- *Orchestration and Control System (OCS)*: an OCS is a logical system in charge of composing, controlling, managing, orchestrating and federating one or more EFS(s). An OCS comprises Virtualization Infrastructure Managers (VIMs), EFS managers and EFS orchestrators. An OCS may interact with other OCS domains. OCS design choices have been based on the requirements and architecture addressing topics such as heterogeneous resources support, dynamic resources support, dynamic migration support, monitoring support and 3rd-parties support [5].

Figure 4-I depicts the 5G-CORAL architecture, including the aforementioned EFS and OCS with all the interfaces between them and towards external entities. The details of the EFS and OCS interfaces can be found in relevant, corresponding 5G-CORAL deliverables [3] and [5], respectively.
In the next subsection, only interfaces relevant for the EFS-OCS interactions will be taken into account. In 5G-CORAL technical Work Packages, i.e. WP2 and WP3, the work conducted at this time of the project is basically on the internal characterization of EFS and OCS, respectively. The final 5G-CORAL architecture is reported in Figure 4-I.
During the validation process no further modifications to the reference architecture as described in deliverable D1.2 have been identified (or deemed as necessary). The main features of the final releases of both the 5G-CORAL EFS and OCS – which are out of the scope of this deliverable – are extensively described and reported in [4] and [6], respectively.
Since the main objective of 5G-CORAL is to develop a system framework being able to support convergence of multiple RATs, it is noteworthy to report how each of the 5G-CORAL use cases takes this objective into account. To this end,
Table 4-I reports the RATs used in each use case.

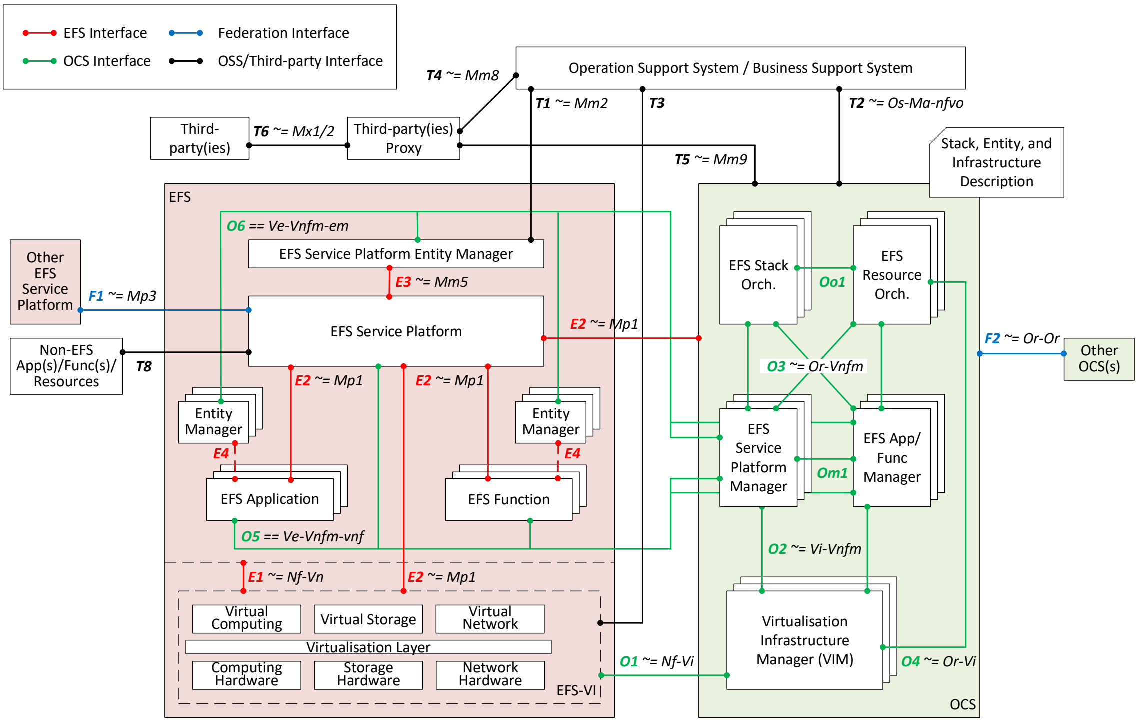

FIGURE 4-I – 5G-CORAL SYSTEM ARCHITECTURE.

TABLE 4-I – USE OF MULTIPLE RATS IN THE 5G-CORAL USE CASES.

Scenario	Use case	RATs
Shopping Mall	Augmented Reality (AR) Navigation	Wi-Fi Bluetooth (iBeacon)
	Fog-assisted Robotics	Wi-Fi LTE Bluetooth (D2D connectivity)
	Virtual Reality (VR)	Wi-Fi LTE
	IoT multi-RAT	NB-IoT IEEE 802.15.4 LoRa
	Software Defined Wide Area Network (SD-WAN)	Wi-Fi LTE
Connected Cars	Safety	LTE Wi-Fi
High-Speed Train	High-Speed Train	LTE Wi-Fi

4.1 EFS – OCS interaction

In this section, an insight of how EFS and OCS interact with each other is reported. By taking Figure 4-I into account, the interfaces involved in the EFS-OCS interactions are listed below, each briefly described in order to assess which is (are) the functionality(-ies) enabled:

- *E2*: it allows the EFS Service Platform to interact with OCS beyond to EFS Applications, EFS Functions and EFS-VI. By exploiting this interface, the EFS Service Platform would leverage efficient orchestration and management via provisioning of various monitoring data for the OCS. Compared to ETSI NFV framework, such interface has been added to support the publication and consumption of data from and to the EFS Service Platform.

Furthermore, as the EFS is also compliant with ETSI MEC [46], the corresponding ETSI MEC *Mp1* interface [47][48] has been adopted as the basis for the E2 interface to support functionalities such as forwarding of configuration information, failure events, measurement results and usage records regarding Edge and Fog resources for monitoring purposes. This is similar to what is done in ETSI MEC for service registration, service discovery and communication support for services. As explained in the next subsection, the E2 interface plays a key role in the EFS service workflow by connecting the involved entities via Message Queuing Telemetry Transport (MQTT) and RESTful protocols depending on the different scenarios. Further details on the functionalities enabled by the E2 interface – especially the ones involving internal EFS entities – can be found in [4].

- *O1*: this interface allows for exchanges between the VIM and the Edge and Fog resources encompassing the EFS, i.e. the EFS-VI. The *ETSI-NFV-Nf-Vi* interface has been adopted as the basis for the O1 interface in order to support functionalities such as 1) allocation of Functions and Applications with indication of compute/storage resource, 2) update, migration and termination of Functions and Applications including their resource allocation, 3) creation, configuration and termination of connection between Functions, Applications and the Service Platform.
  Compared to ETSI NFV environment, such interface also needs to support intermittent connectivity on volatile low-end devices with heterogeneous virtualization support.
- *O5*: together with O6, this interface is for OCS orchestration and allows for EFS Functions/Applications to connect with the EFS Service Platform Manager within the OCS. It also allows for EFS Functions/Applications to connect with the corresponding EFS App/Func Managers within the OCS. The *ETSI-NFV-Ve-Vnfm-Vnfm* interface has been adopted as the basis for the O5 interface in order to support functionalities such as 1) notification from the Application/Function/Service Platform to the corresponding EFS Managers of events, measurements and usage records regarding the Application/Function/Service Platform itself and 2) verification that the Application or Function is still alive/functional [5].
- *O6*: together with O5, this interface is for OCS orchestration and allows for the EFS Service Platform Entity Manager and the entity managers of EFS Functions/Applications to connect with the EFS Service Platform Manager and the EFS App/Func Managers within the OCS. The *ETSI-NFV-Ve-Vnfm-em* interface has been adopted as the basis for the O6 interface in order to support functionalities such as 1) Application/Function/Service Platform instantiation, update, termination, scaling in/out and up/down, 2) Application/Function/Service Platform instance query to retrieve any run-time information and 3) configuration of events and measurements regarding the Application/Function/Service Platform itself [5].

All the remaining interfaces foreseen in the 5G-CORAL architecture depicted in Figure 4-I which are relevant for both the EFS and the OCS have been extensively described in the corresponding 5G-CORAL deliverables, e.g. Table 2-2 in [4] for the EFS and Table 2-3 in [5] for the OCS.
An example of how the OCS deploys an Application/Function in the EFS, together with the procedural steps performed by the involved Application/Function entities and by the EFS Service Platform, are shown in Figure 4-II, where it is assumed that a service of the deployed Application/Function is then used by another Application/Function.

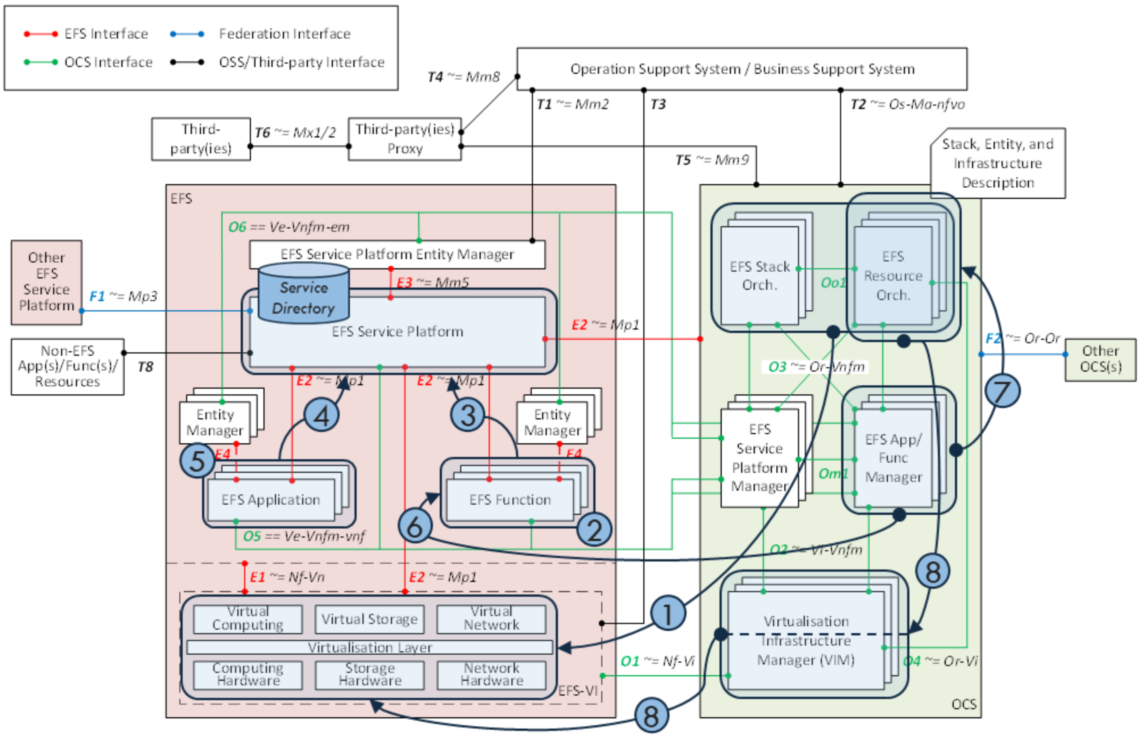

FIGURE 4-II – 5G-CORAL SERVICE WORKFLOW.

Referring to Figure 4-II, the following steps have been identified (details can be found in [4]):

1. OCS performs the deployment of an Application/Function in the EFS.
2. The deployed EFS Application/Function identifies the EFS Service Platform interface (E2 interface) by exploiting one of the mechanisms available at the OCS.
3. The EFS Application/Function registers its services with the EFS Service Platform by using the E2 interface.
4. Another EFS Application/Function identifies the services (provided by the EFS Application/Function as in step 3) by using the E2 interface.
5. The EFS Application/Function of step 4 carries out its operation by accessing the needed EFS service based on the information received by the EFS Service Platform.
6. The EFS App/Func Manager discovers via O5 interface that the EFS Application/Function from step 2 cannot deliver a reasonable service quality.
7. The EFS App/Func Manager invokes scaling using the O3 interface to the EFS Resource Orchestrator.
8. The EFS Resource Orchestrator allocates a larger number of CPU cores for the Application/Function via the VIM (O4 and O1 interfaces).

4.2 OCS management of heterogeneous virtualization platforms

The 5G-CORAL architecture integrates resources from Cloud to Fog, considering a wide variety of computing devices, including volatile and non-stationary. This volatility raises the need of dynamic deployment and migration of elements in the EFS, a suitable scenario for virtual environments. 5G-CORAL ecosystem supports virtualization platforms to cope with this dynamicity and heterogeneity, focusing mainly on containers and Virtual Machines (VMs).

A VM encapsulates a guest operating system (including the guest kernel) and applications in an instance running on top of the host kernel by using a hypervisor (e.g. KVM [49]). This approach

offers a great level of isolation but incurs in inefficient use of resources, requiring a larger amount of computing and storage capabilities, as well as considerable deployment time.
As an alternative to VMs, containers are software units that package code, libraries and dependencies to run on top of a host kernel. There are two levels of containerization: i) operating system level (e.g. LXD [50]), where the container includes all the operating system libraries, sharing only the kernel with the host machine and ii) application level (e.g. Docker [51]), where the container shares the operating system libraries and the kernel with the host, offering a more lightweight package. Containers also provide good isolation level using *namespaces* and allows to limit computational resources using *cgroups*. The resources required by a containers are lower than those required by a VM since the kernel and, in some cases, the operating system libraries are shared with the host machine, for this reason containers are more suitable for less powerful computing platforms (e.g. Fog). In some cases, e.g. when the host has a limited amount of resources or applications need direct access to very specific hardware with strong driver or kernel dependencies, VMs and containers could not be supported. For this reason deploying native applications in the Fog environment is also possible, to allow an entirely new set of devices to become part of the infrastructure. The following subsections illustrate how the 5G-CORAL use cases are deployed and how virtualization platforms are used depending on the necessities and requirements.

4.2.1 Shopping Mall scenario

4.2.1.1 Augmented Reality (AR) Navigation

AR Navigation aims at providing end-users an augmented reality navigation assistance within a Shopping Mall. AR Navigation requires real-time interaction with end-users, hence a massive amount of compute resources is expected in order to provide timely image processing for AR functionality. To this end, Fog computing has been introduced to offer a localized, distributed and compute-rich infrastructure which natively provides compute capabilities at the Edge, meaning that real-time object recognition/detection becomes possible. In addition, Fog computing provides fast-retrievable local storage for AR on environmental feature detection information, shop advertisement information, etc. which further enables timely response and content provision back to AR Navigation end-users. Combining these two characteristics, an AR Navigation end-user, for example, can get a timely guiding arrow to a destination shop at a multi-exit intersection. Yet another example, once an AR Navigation end-user reaches the destination shop, the end-user itself is able to obtain an e-menu, e-coupons or even video advertisement of the specific shop of interest by searching for representative logos of that shop via the AR Navigation application installed on end-users' mobile phones. This use case is composed of five software components: Image Recognition, Localization, Statistical Distributor, Job Dispatcher and AR Application. Requirements of each software component and the feasibility of the different deployment options have been analyzed and the results are summarized in Table 4-II.

TABLE 4-II – AR NAVIGATION DEPLOYMENT ANALYSIS.

Software component	Requirements	Native	Container	VM	Choice
Image Recognition	IP connectivity	✓	✓	✗	Container
Localization	IP connectivity	✓	✓	✗	Container
Statistical Distributor	IP connectivity	✓	✗	✗	Native
Job Dispatcher	IP connectivity	✓	✓	✗	Container
AR Application	Camera access, Wi-Fi access, Bluetooth access on smartphone	✓	✗	✗	Native

The Image Recognition provides object detection functionality to identify store brands or scenes in order to add virtual objects on end-user's screen for navigation. This software component can be either deployed in a container or as a native application. The container-based solution provides an isolated environment to update the image recognition algorithm and libraries. Also, choosing container-based deployment for image recognition is useful to extend the Fog node capabilities in the future. For these reasons, the container-based approach has been chosen to deploy the Image Recognition software component.

The Localization component provides end-user's location information based on sensed beacon information. The component has been containerized in order to allow for quick deployment onto multiple Fog CDs.

The Statistical Distributor software component maintains the status of the system, synchronizing every Fog node about the current loading of the whole system. This allows for the Job Dispatcher to work properly. The Statistical Distributor has been deployed as a native application in order to maintain short latency needed for fulfilling the database I/O requirements.

The functionality of the Job Dispatcher component is very similar to the Image Recognition component. It is in charge of handling navigation requests from end-users. If the serving Fog CD is found itself overloaded, by referencing load statistics, the Job Dispatcher is able to allocate some requests to low-loaded Fog CDs in near proximity. It can be deployed either as a container or as a native application but, in order to ensure isolation and quick deployment whenever needed, containers have been chosen.

Finally, the AR Application software component runs as a native application in the end-user's mobile phone. It captures images using the camera in the phone, sends the images to the Image Recognition component and gets the recognition result.

Based on the deployment options mentioned above, Table 4-III shows the hardware requirements for the execution of each software component, as well as the virtualization technology support requirements.

TABLE 4-III – AR NAVIGATION REQUIREMENTS FROM REFERENCE IMPLEMENTATION.

Software component	Hardware requirements	Virtualization technology support requirements
Image Recognition	2 cores, 1 GB of RAM and 500 MB disk space	Containers (LXD)
Localization	16 MB of RAM, 4 MB of disk space	Containers (Docker)
Statistical Distributor	16 MB of RAM, 4 MB of disk space	-
Job Dispatcher	16 MB of RAM, 4 MB of disk space	Containers (LXD)
AR Application	200MB of RAM, 100MB of disk space	-

The whole demonstration scenario is depicted in Figure 4-III: it is composed of 3 Fog CDs (Fog01, Fog02 and Fog03) and 2 Wi-Fi APs. In this use case two different deployment options have been mixed: containers for Image Recognition, Localization and Job Dispatcher and native applications for Statistical Distributor and AR Application. In Figure 4-III, Fog01 is running a Docker-enabled operating system with Image Recognition, Localization and Job Dispatcher software components and with a native Statistical Distributor application installed for AR Navigation service provisioning. Both Fog02 and Fog03 are running in the same operating system environment and with the same containers installed. The application Statistical Distributor is absent in Fog02 and Fog03 as only the Statistical Distributor at Fog01 associates and synchronizes compute resource usage status among all the Fog CDs. AR Navigation application is running on the mobile phone's operating system, e.g. Android, interacting with the Job Dispatcher at Fog01.

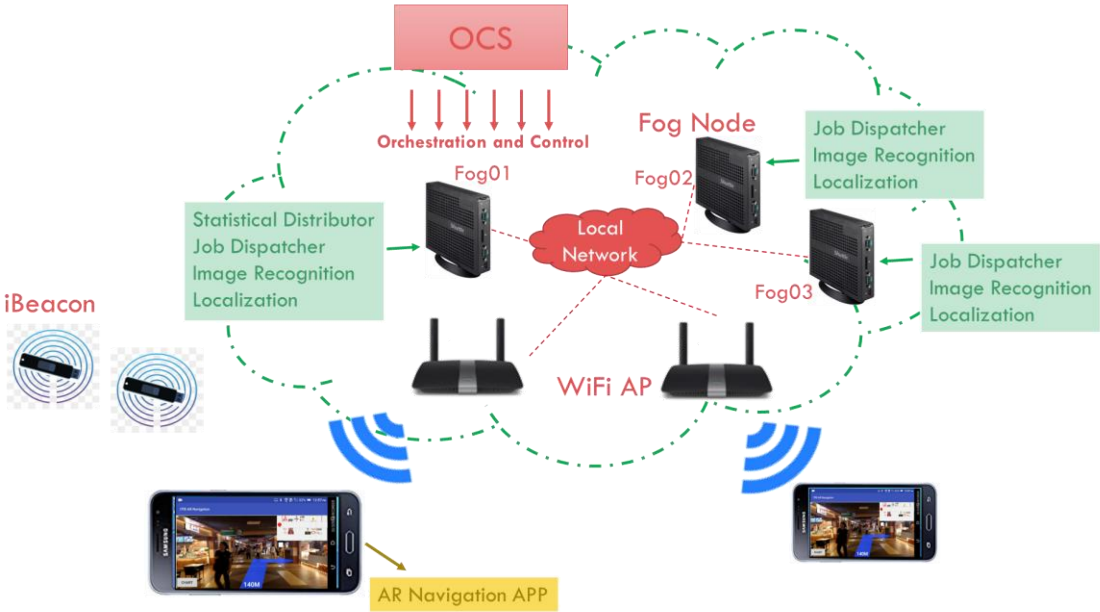

FIGURE 4-III – AUGMENTED REALITY NAVIGATION USE CASE.

An extension of virtualization technologies in this use case allows the deployment of a set of applications that cooperatively enable load balancing in the Fog environment. This flexibility implies that, in order to improve service performance, OCS has to make use of the native load-balancing mechanism of an application, not merely through compute-scaling operations.

4.2.1.2 Fog-assisted Robotics

Fog-assisted Robotics use case leverages on Fog computing capabilities to place the robot's intelligence making the robots lightweight and low cost. Fog computing offers a converged infrastructure that allows to share the state and information of the robots, enabling better synchronization and coordination. Fog-assisted Robotics can perform the following two applications in a Shopping Mall:

- *cleaning robots*: robots clean common areas of the Shopping Mall automatically based on videos recorded by Shopping Mall cameras
- *cooperative delivery of goods*: robots are coordinated and synchronized to deliver goods from the Shopping Mall warehouse to the shops.

TABLE 4-IV – FOG-ASSISTED ROBOTICS DEPLOYMENT ANALYSIS.

Software component	Requirements	Native	Container	VM	Choice
Robot intelligence	Kernel driver installation	✓	✗	✓	VM
Wi-Fi virtual AP	Access to the hardware device (Wi-Fi card)	✓	✓	✗	Container
Wi-Fi monitoring	Access to the hardware device (Wi-Fi card)	✓	✓	✗	Container
Gateway	IP connectivity	✓	✓	✓	Container
Robot agent	Kernel driver installation Direct access to hardware sensors	✓	✗	✗	Native
Network-assisted Device-to-Device (D2D)	Access to the hardware device (Wi-Fi card)	✓	✓	✗	Native

This use case is composed by six software components: the Robot intelligence, Wi-Fi virtual AP, Wi-Fi monitoring, Gateway, Robot agent and Network-assisted Device-to-Device (D2D). Requirements of each software component and the feasibility of the different deployment options have been analyzed and reported in Table 4-IV.

The Robot intelligence controls the robots, processing all the data from the sensors. This software component cannot be deployed in a container since it requires the installation and configuration of several kernel modules and it has several kernel dependencies, therefore it can be deployed in a VM or as a native application. VMs introduce a huge overhead compared to a native execution but also offer a more isolated environment and software migration. For these reasons VM have been chosen to deploy the Robot intelligence.

The Wi-Fi virtual AP provides connectivity to the robots using the hostapd Linux daemon. This software component cannot be deployed in a VM since it requires direct access to the Wi-Fi card, it can be deployed in a LXD container or as a native application. It has been decided to deploy this component in a container-based environment because this provides fast migration and isolation, introducing only a little overhead compared to the native application.

The Wi-Fi monitoring software component analyzes the Wi-Fi signal strength; this information is jointly used with the localization of the robots in the Shopping Mall by the *brain*, hence it can trigger the Wi-Fi virtual AP migration. It has the same requirements than the Wi-Fi virtual AP, hence the same deployment option has been chosen.

The Gateway software component is in charge of giving IP connectivity to the robots and can be deployed in each of the three available options. Containers have been chosen as the deployment option because they provide isolation and migration introducing a minimum overhead compared to the native execution and VMs. VMs also provide isolation and migration but with a much higher overhead.

The next software component is the Robot agent, which runs at the robot and sends the information from the sensors to the brain. The Robot agent requires a direct access to the sensor driver and it has a lot of kernel dependencies, hence the only feasible deployment option is the native execution.

Finally, the Network-assisted D2D is in charge of connecting two robots to provide a direct connection in case the connection to the brain fails. It needs direct access to the Wi-Fi card, so the VM deployment is not an option. The native execution has been chosen over containers because the computational power of the robots is limited and there is also the need to reduce the overhead to the minimum to increase the battery life, reduce latency, etc.

TABLE 4-V – FOG-ASSISTED ROBOTICS REQUIREMENTS FROM REFERENCE IMPLEMENTATION.

Software component	Hardware requirements	Virtualization technology support requirements
Robot intelligence	4 cores, 2 GB of RAM and 12 GB disk space	KVM based VMs
Wi-Fi virtual AP	16 MB of RAM, 200 KB of disk space and an 802.11 interface	LXD containers
Wi-Fi monitoring	256 MB of RAM, 3 MB of disk space and an 802.11 interface	LXD containers
Gateway	16 MB of RAM and 200 KB of disk space	LXD containers
Robot agent	128 MB of RAM and 1.5 GB of disk space	-
Network-assisted D2D	16 MB of RAM, 4 MB of disk space and an 802.11 interface	-

Based on the deployment options selected in Table 4-IV, Table 4-V shows the hardware requirements for the execution of each software component, as well as the virtualization technology support requirements.
The whole demonstration scenario is depicted in Figure 4-IV and is composed by 3 Fog CDs and 2 robots. In this use case three different deployment options showed in Table 4-V have been mixed: KVM VMs for the Robot intelligence, LXD containers for the Wi-Fi virtual AP, the Wi-Fi service and the Gateway, while native applications running in the robots have been considered for both the D2D communications and the Robot application.
Fog01 is running a KVM VM with Robot Operating System (ROS) v1 installed and the required ROS packages for the Robot intelligence. Fog02 and Fog03 are running a Wi-Fi virtual AP and the Wi-Fi monitoring service; in addition Fog02 also runs the Gateway and, as previously explained, all these three software components are running in LXD containers. Robot01and Robot02 natively run ROS with all the required drivers for the robot sensors to communicate with the Robot intelligence in Fog01 in addition to the D2D communication.

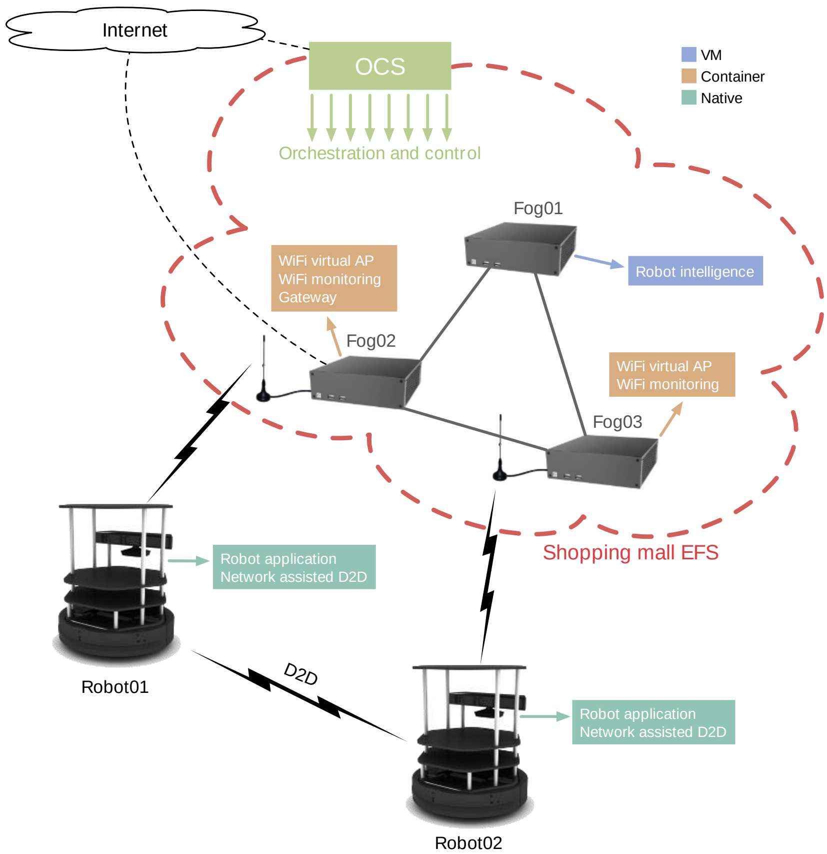

FIGURE 4-IV – FOG-ASSISTED ROBOTICS USE CASE.

As shown, the usage of different virtualization technologies allows an entire new set of applications to be developed in the Fog environment and this flexibility reflects also in the possibility for the developers of these application to choose the best technology for their applications. All this heterogeneity is vital for the Fog environment as it allows to exploit and efficiently use the whole infrastructure. It is worth to say that the management overhead of having different virtualization technologies is hidden by the usage of the OCS that itself is able to discover and choose among the today's most used virtualization technologies, hence hiding the complexity and allowing for seamless integration in the Fog environment.

4.2.1.3 Virtual Reality (VR)

The Virtual Reality (VR) use case is based on a multi-tier system architecture enabling distributed computing across Cloud, Edge and Fog layers. Specifically, the end-to-end service consists of multiple software functions consuming heterogeneous resources hosted on industrial-grade DCs, high-capable Edge servers and low-end Fog nodes. A list of the software components necessary to deliver the VR multimedia content is presented in Table 4-VI.

TABLE 4-VI – VR DEPLOYMENT ANALYSIS.

Software component	Requirements	Native	Container	VM	Choice
Insta360 dashboard	Connectivity to the camera	✓	×	×	Native
Decoding and DASH reassemble	Hardware acceleration (GPU)	✓	×	×	Native
Wowza streaming engine	Internet connectivity for license validation	✓	×	×	Native
Encoding and DASH segmentation	HW acceleration (GPU)	✓	×	×	Native
Orientation service	Wi-Fi connectivity	✓	✓	✓	Native
Video player	Enabled Android developer options and debugging	✓	×	×	Native

As reported in the Table 4-VI, most of the software components are deployed as native applications and might not be easily virtualized or containerized, as they heavily rely on hardware acceleration supplied by high-capable graphic cards (note that graphic cards such as NVIDIA GeForce P5000 and GTX 1080 are currently not supported by the NVIDIA Virtual GPU technology [52]). As a consequence, this dictates various constraints in the deployment process and limit its flexibility, as resources cannot be easily replaced or moved to another machine. More specifically, operations involving DASH, transcoding, encoding and decoding must be hardware-accelerated and be hosted by a Cloud DC or Edge servers. It should also be noted that the Orientation service may be executed either on a VM or a container, provided that the Wi-Fi connectivity is available. This can ensure more flexibility and allows the Orientation service to migrate from a host to another. Finally, applications interacting with the video source and the terminal, i.e. Insta360 dashboard and Video player, must run on the same device or must be connected to it, therefore virtualization cannot be employed.

4.2.1.4 Multi-RAT IoT

In this use case, the baseband processing (together with the higher radio protocol layers) is mostly centralized to Edge servers (or Edge DCs) while the RHs are mainly handling the RF-related operations and processing. In this way, different RAT protocol stacks can be implemented at the Edge DC. The RHs are deployed for service coverage for different RATs, e.g. using different bands. RHs can be either single-RAT capable or multi-RAT capable equipment. They are connected to the Edge servers via Ethernet and/or IP networks. Further, according to the 5G-CORAL system design, an EFS Service Platform providing data services is also deployed in Edge servers, via which the EFS Functions and Applications can exchange context data. As an implementation example of this use case, the RAT communication stacks deployed can publish IQ data via the EFS Service Platform, which is currently implemented using MQTT. There is an interference analysis application which subscribes the IQ data service and performs interference analysis using the IQ data received. In the following, Table 4-VII gives an overview regarding the software components developed in the implementation example. Docker containers have been chosen due to the considerations of lower deployment time and the integration with Kubernetes as the orchestrator.

TABLE 4-VII – MULTI-RAT IOT DEPLOYMENT ANALYSIS BASED ON POC DESIGN.

Software component	Requirements	Native	Container	VM	Choice
RH software	Ethernet/IP connectivity providing access to the Edge communication functions, running in RHs	✓	✓	✗	Native
15.4 communication stack function	Ethernet/IP connectivity providing access to RH and Edge Service Platform, running at Edge servers	✓	✓	✓	Docker Container
NB-IoT communication stack function	Ethernet/IP connectivity providing access to RH and Edge Service Platform, running at Edge servers	✓	✓	✓	Docker Container
LoRa communication stack function	Ethernet/IP connectivity providing access to RH and Edge Service Platform, running at Edge servers	✓	✓	✓	Docker Container
Interference analysis application	Ethernet/IP connectivity providing access to Edge Service Platform, running at Edge servers	✓	✓	✓	Docker Container

4.2.1.5 Software Defined Wide Area Network (SD-WAN)

The SD-WAN use case leverages SD-WAN technology to federate resources belonging to different domains, providing a virtual network for management, control and data planes.
The software components needed in the use case are listed below and comprise the overall picture of the use case, from infrastructure to user applications.

TABLE 4-VIII – SD-WAN DEPLOYMENT ANALYSIS BASED ON POC DESIGN.

Software component	Requirements	Native	Container	VM	Choice
SD-WAN Function	Kernel version and Open virtual Switch (OvS) modules	✓	✓	✓	Container
Wi-Fi virtual AP	Access to the hardware device (Wi-Fi card)	✓	✓	✗	Container
SD-WAN Controller	Public/routable IP address	✓	✓	✓	Container
SD-WAN Manager	Public/routable IP address	✓	✓	✓	Container
Point of Sale (PoS) Web Application	LAMP and connectivity to the Wi-Fi AP. Connection to SD-WAN function	✓	✓	✓	Container
PoS database	Connectivity to the web applications. Connection to SD-WAN function	✓	✓	✓	Container

The SD-WAN Function behaves as a virtual switch. Other containers can be connected to the ports of the SD-WAN Function, gaining access to the virtual network. It needs an specific Linux kernel version and also load the Open virtual Switch (OvS) modules to work properly. When working with containers, the parent device must also have the kernel version and modules. If it is deployed in a VM or container, it should export ports that behaves as WAN port and LAN ports. As a native application, the device must have physical ports.

The Wi-Fi virtual AP manages a Wi-Fi card and redirects the traffic to the SD-WAN Function. The virtualization technique must control the hardware device, be able to run the hostapd Linux software and control the interface. It is easier to be implemented either as a native application or container.
The SD-WAN Controller controls the data plane of the SD-WAN Function, being the control plane of a Software Defined Network (SDN). It runs as a Java application and can be executed in any environment while it exposes a routable IP to both the SD-WAN Function and the SD-WAN Manager.
Also, the SD-WAN Manager runs as a Java application and can be executed in any environment. It has to use different ports than the SD-WAN Controller. The SD-WAN Manager receives the descriptors that characterize the desired virtual network and manages the SD-WAN Controller and SD-WAN Function to deploy the network. Again, it needs a routable IP address.
Finally, the two application components, i.e. the Point of Sale (PoS) Web Application and the PoS database, can be deployed on any environment and the only requirement to be fulfilled is the connection with the SD-WAN Function, through a physical network or via elements such as virtual Ethernets to connect using virtual network.

4.2.2 Connected Cars scenario

In the context of the Connected Cars scenario, the Safety use case aims at increasing safety for drivers and passengers exploiting the deployment of Fog CDs located nearby the vehicles via, e.g., Road Side Units (RSUs) or on the vehicles themselves via On-Board Units (OBUs). Both RSUs and OBUs collect information regarding speed, direction and position of the vehicles, process them to generate and send warnings when needed, i.e. collision avoidance, in order to increase the safety during the driving.
The software components used in this use case, i.e. the Applications and the Wi-Fi virtual AP, are described in Table 4-IX.

TABLE 4-IX – SAFETY DEPLOYMENT ANALYSIS.

Software component	Requirements	Native	Container	VM	Choice
Application	Kernel driver installation	✓	×	×	Native
Wi-Fi virtual AP	Access to the hardware device (Wi-Fi card)	✓	✓	×	Container

All the software needed to process the data received and then to generate the warning messages, i.e. the ones used for collision avoidance, emergency vehicle approaching, etc., is indicated as “Applications”. These Applications are implemented both on the OBU and RSU as a native execution since the OBU has a low computational power, hence containers and VMs cannot be deployed.
The Wi-Fi virtual AP provides the secondary RAT (for multi-RAT exploitation) and it is implemented in an LXD container.
Table 4-X shows the hardware requirements to be fulfilled for the execution of each software component.

TABLE 4-X – SAFETY REQUIREMENTS FROM REFERENCE IMPLEMENTATION.

Software component	Hardware requirements	Virtualization technology support requirements
Application	1 core, 256 MB of RAM and 512 MB disk space	-
Wi-Fi virtual AP	16 MB of RAM, 200 KB of disk space and an 802.11 interface	LXD containers

The whole demonstration scenario is depicted in Figure 4-V and it is composed by 2 Fog CDs – one on the OBU (the vehicular Fog CD), the other one on the RSU – and an Edge DC at the eNB side.
The vehicular Fog CD integrates a commercial LTE modem and a limited computation element being able to run the Applications. The Fog CD on the RSU works as a Wi-Fi virtual AP. In addition, an instance of the MQTT broker and the Applications are deployed. Vehicles can transmit and receive the information from both the broker on the eNB and the RSU, and then elaborate this information and exchange the generated warnings among them.

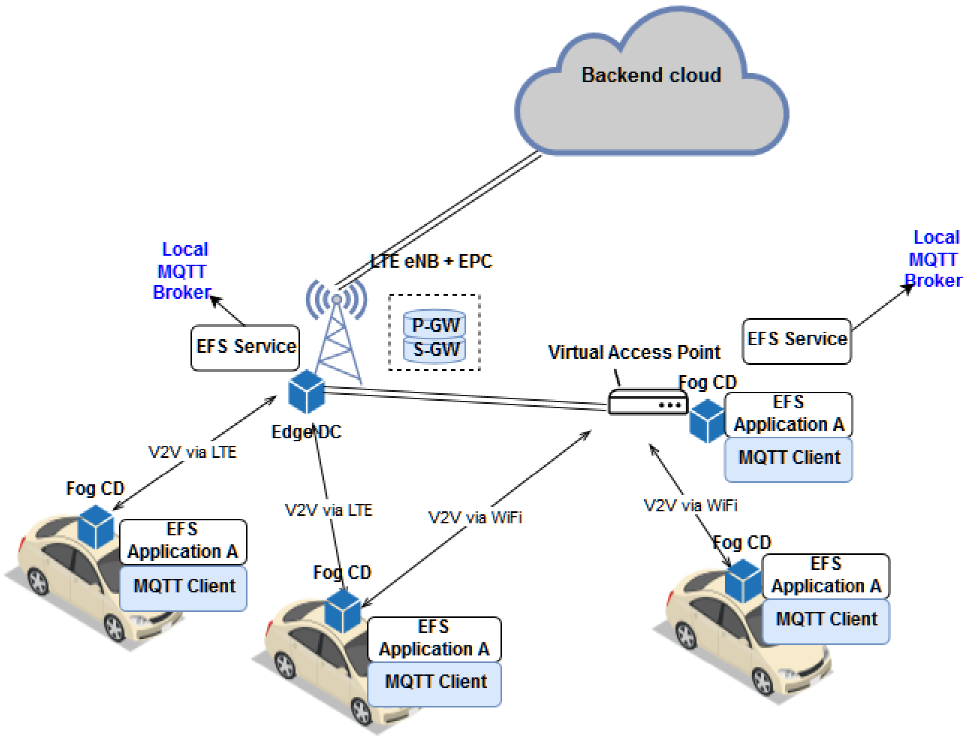

FIGURE 4-V – CONNECTED CARS SCENARIO.

4.2.3 High-Speed Train scenario

As explained in subsection 3.3, the High-Speed Train use case, i.e. the high mobility scenario, aims at reducing the service downtime for passengers exploiting the deployment of Fog CDs on-board and vMME functionalities. Based on the information provided by the user classifier function within the on-board Fog CD and the train approximate location, the vMME will execute handovers for groups of end-users as the train approaches the train station. Each group of end-users is determined based on end-users' application types and related QoS requirements. This classification is very important to reduce the signalling overhead during the switching from on-board to on-land and to maintain the service interruption at the minimum.
The software components used in this use case, described in the Table 4-XI, are the EPC emulator (NextEPC [53]), the vMME and a Multiple Users emulator.

TABLE 4-XI – HIGH-SPEED TRAIN DEPLOYMENT ANALYSIS.

Software component	Requirements	Native	Container	VM	Choice
NextEPC	Kernel driver installation	✓	✗	✗	Native
vMME	Ethernet	✓	✓	✓	Container
Multiple Users emulator	Ethernet	✓	✓	✓	Container

In the following Table 4-XII Table 4-XII – High-Speed Train requirements from reference implementationthe hardware requirements needed to be fulfilled for the execution of each software component are shown.

TABLE 4-XII – HIGH-SPEED TRAIN REQUIREMENTS FROM REFERENCE IMPLEMENTATION.

Software component	Hardware requirements	Virtualization technology support requirements
NextEPC	Intel i7 4 cores, DDR4 16GB, 256GB disk space, Small Cells connect to Next EPC	-
vMME	Raspberry pi 3, on-board Customer Premises Equipment (CPE)	LXD containers, Docker containers
Multiple Users emulator	100 MB of RAM, 200 KB of disk space and an Ethernet interface	LXD containers, Docker containers

The whole demonstration scenario is depicted in Figure 4-VI and is composed by two Fog CDs – one on-board and the other one on-land and located in the train station – and two LTE small cells.
The on-board Fog CD connects to a commercial Customer Premises Equipment (CPE), which contains three EFS functions: vMME, user classifier and service migration handler. The on-land Fog CD contains service migration handler and a local gateway. The vMME hosted in the on-board Fog node close to the passengers copes with the signalling generated by hundreds of passengers using their terminals in the train. Also, this vMME reduces the service interruption time during handover process by carrying multiple passenger devices into a single control message towards the LTE core network.

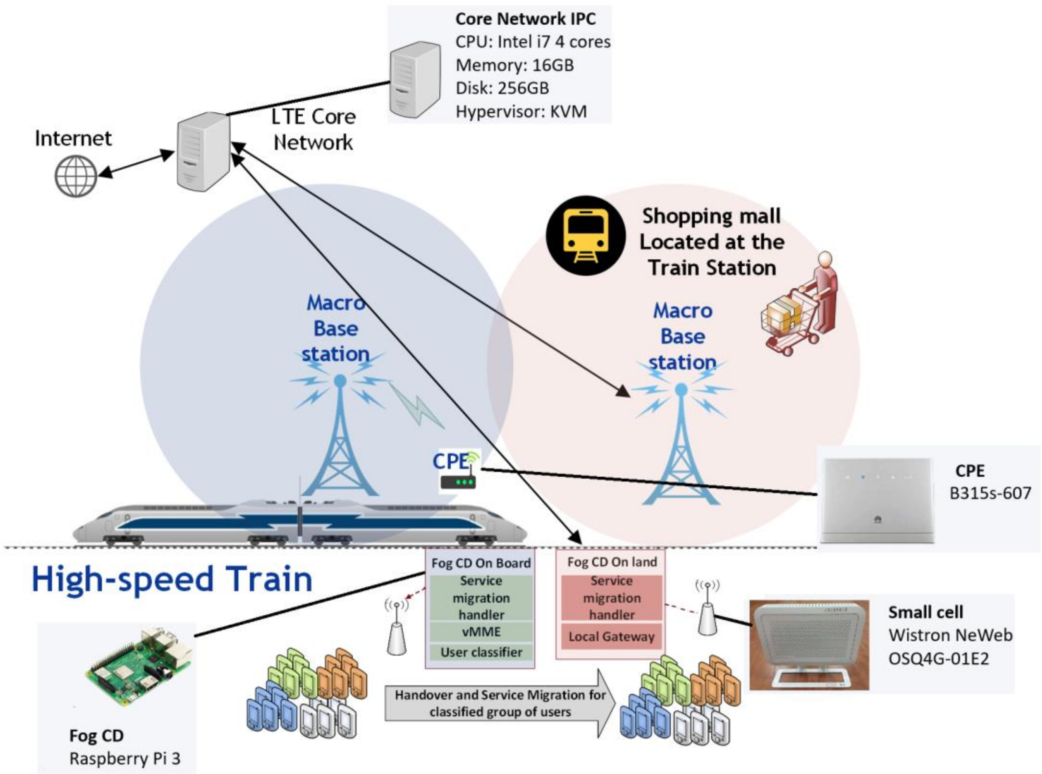

FIGURE 4-VI – HIGH-SPEED TRAIN SCENARIO.

4.3 Procedures and techniques for incremental deployment of 5G-CORAL solution into existing networks

This subsection takes the viewpoints of both the telecom operator and the Vertical company into account regarding deployment considerations of the 5G-CORAL technical solution within existing networks. From the telecom operator perspective, the analysis initially considers current 4G networks (EPC-based, with LTE RAT), where initial deployments of Edge and Fog solutions have to be necessarily designed as add-ons on top of the existing network. For the incoming 5G systems, instead, networks will be designed to efficiently support Edge and Fog since the beginning, so to flexibly achieve both high performance and quality of experience. When it comes to the Vertical company, the benefits in terms of security, user access control and service provisioning which could derive when deploying the 5G-CORAL solution within the Company's premises are also discussed.

4.3.1 Deployment of the 5G-CORAL solution: the telecom operator perspective

Edge and Fog computing paradigms are becoming of great importance for operators not only as enablers for fulfilling the requirements of the diverse 5G use cases (hence representing new business opportunities for them) but also as means to realize cost synergies due to, e.g., investments on the cloudification/virtualization of different network segments (RAN, Core Network, Central Offices, etc.). However, it is a common understanding that, when it comes to its deployment, Edge/Fog computing is often considered as a 5G-only feature. However, current 4G networks (i.e. based on pre Rel-15 EPC and LTE) are expected to still be largely exploited in the coming years, hence there is the need to identify opportunities and challenges in deploying Edge and Fog technical solutions in existing 4G networks. This of course is also valid for the 5G-CORAL solution, whose architectural pillars are inherited from existing industrial frameworks, i.e. ETSI NFV, ETSI MEC and OpenFog, in order to create a physically-distributed computing fabric where Fog includes Edge and goes through the access infrastructure up to the end-users' terminals (see Figure 4-VII).

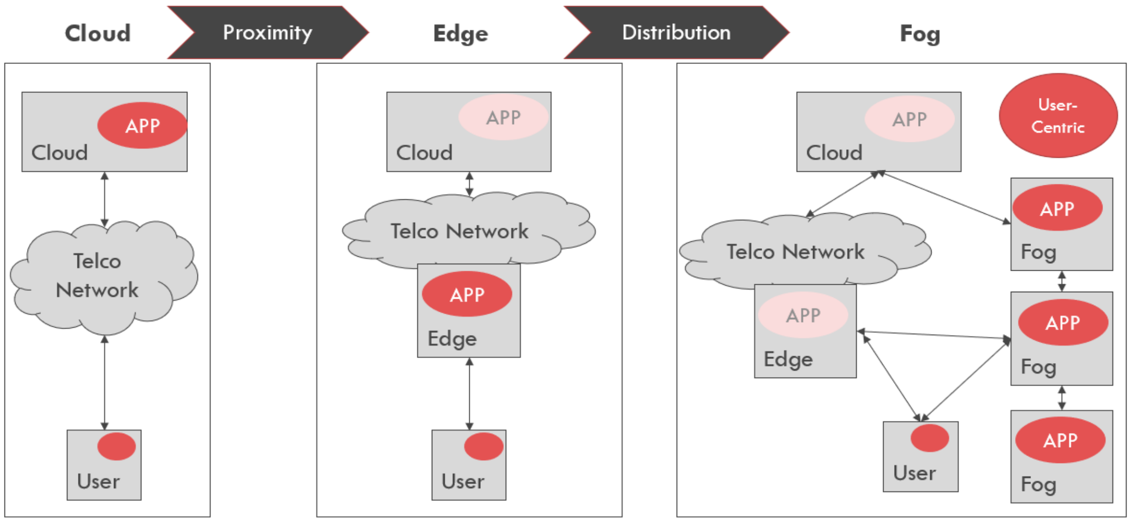

FIGURE 4-VII – 5G-CORAL: FROM THE CLOUD TO THE "EXTREME" EDGE.

In a recent survey report among several telecom operators [54], it was stated that Edge is not considered only, or even mainly, a 5G capability: just the 14% of respondents, in fact, thinks at Edge as *a set of 5G features* whilst the two much larger responses, with 35% and 38%, respectively, think Edge is either *an important new aspect of Cloud* or an *opportunity to create a new set of low latency applications*. From an operator perspective, this basically means opportunities for their business, but they are aware that Edge is also a wider movement they can take advantage of since the beginning, not limited to 5G only.

Potential applicability of Edge and Fog solutions within current networks is also motivated by concrete initiative such as MobiledgeX, a new U.S. company founded by Deutsche Telekom in 2018 [55] as the result of an extensive two-year study on Edge computing that Deutsche Telekom performed internally. MobiledgeX is deploying Edge computing today within 4G infrastructures using an expanding, global federation of mobile operator partners, and emphasizing the importance of building an Edge now, hence not waiting for *full* 5G, in order to gain experience with the many nascent opportunities earlier. MobiledgeX, which is already live in production networks in Germany and Poland, deploys its own services by using Cloudlets running on existing mobile operator's physical or virtual infrastructure and developed an Software Development Kit (SDK) to adapt mobile applications to integrate with these services. In this way new business opportunities can be created not only for operators but also for companies outside the mobile network industry that want to leverage the unique assets enabled by mobility.

While the demand drivers for Edge and Fog computing are clear (low latency, location-based services and so on) they do not necessarily define an Edge/Fog architecture either: instead, architectures are heavily influenced by the players involved, with the Edge and Fog infrastructure residing anywhere outside of DCs all the way down to end-point devices. Operators generally favor the collocation of Edge and Fog equipment within their network infrastructure, whether in access network nodes as well as pre-aggregation and aggregation locations in transport networks (see Figure 4-VIII).

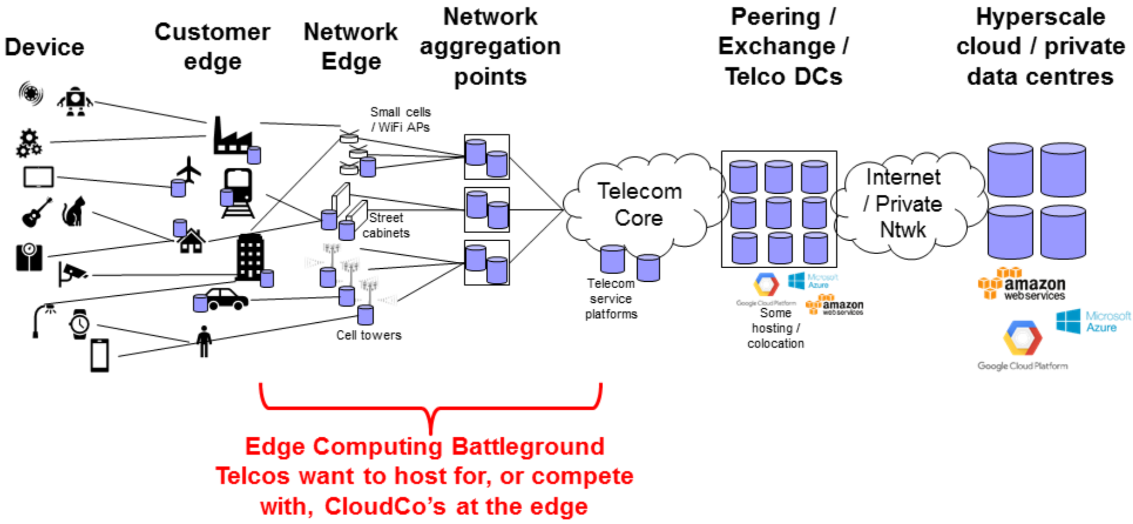

FIGURE 4-VIII – EDGE LOCATION WITHIN THE OPERATOR'S NETWORK. RE-ELABORATION OF [56].

Due to the proliferation of Cloud providers eroding the value of their networks, operators started to seek opportunities for Edge and Fog computing to bring value back to their networks, examples of which include MEC and Control/User Plane Separation (CUPS) – both network-centric initiatives – as well as the Central Office Re-architected as a Datacenter (CORD) initiative, which is applying DC technologies in telecom's central office environments [57]. Strong market drivers, however, do not imply Edge and Fog solutions to come without challenges: incumbent services that make use of typical Cloud platforms create strong incentives for status quo to be maintained, therefore, rather than disrupting conventional Cloud services, Edge and Fog computing should be complementary in making future business opportunities thriving and profitable.

An important aspect which might reduce the barriers in adopting Edge and Fog solutions, i.e. the 5G-CORAL platform – within current networks is the possibility to deploy it as part of the operator's NFV environment, where 5G-CORAL Functions and Applications will be deployed as Virtual Network Functions (VNFs) and will cooperate with Evolved Packet System (EPS) functions in composing an end-to-end network service. In this sense, also considering that the 5G-CORAL architecture is based on the ETSI MEC framework, the deployment options identified in [58] for a MEC system can be exploited for assessing the technical feasibility of deploying the 5G-CORAL solution into current 4G networks, also allowing the telecom operator to properly plan the network investments deemed as essential for the mobile network evolution towards 5G. Examples of deployment, with the MEC system being connected to the user plane, are listed below:

- *Bump in the Wire*: deployment scenario in which the Edge and Fog platform is located between the eNB and the EPC
  - Low latency is supported by installing the platform very close to the eNB location, or in other locations that ensure minimal latency
- *Distributed EPC*: deployment scenario in which the Edge and Fog platform logically includes all or part of EPC components
  - In this case the communication with the operator's core site is optional/infrequent hence it can serve Mission Critical Push to Talk (MCPTT) and Machine Type Communications (MTC) services. Moreover, it is also possible to deliver exactly the QoS and configurability features requested by, e.g., an enterprise customer
- *Distributed S/PGW*: similar to the Distributed EPC with exception of Serving GateWay (SGW) and PDN GateWay (PGW) being deployed at Edge whilst control plane functions – MME and Home Subscriber Server (HSS) – are at operator's network core
  - SGW and PGW run as VNFs together with the Edge and Fog applications on the same NFV platform, allowing for traffic offloading on a Access Point Name (APN) basis
- *Distributed SGW with Local Breakout (SGW-LBO)*: here both the SGW and the Edge and Fog applications are in the form of VNFs being hosted in the same platform at the network Edge
  - SGW-LBO offers the possibility to steer traffic based on any operator-chosen combination of policy sets such as APN and other IP-level parameters (e.g. IP version) and allows users to reach both Edge and Fog applications and operator's core supported ones over the same APN
- *Control/User Plane Separation (CUPS)*: all the above options which distribute the EPC gateways at the network Edge can be realized by using the CUPS paradigm standardized in 3GPP Rel-14, with the user plane being hosted in the Edge and Fog platform.

With 4G networks already having been deployed for a number of years, initial deployments of Edge and Fog solutions have necessarily been designed as add-ons on top of a 4G network, therefore these systems are to a large extent self-consistent, covering everything from management and orchestration down to interactions with the data plane for managing specific traffic flows. 5G design and architecture principles differ from previous mobile generations and allow building open networks with open interfaces where, in principle, any level of network control can be offered to the outside, with the network designed to natively support NFV and SDN paradigms. In 5G it is expected that Edge and Fog will be one of the key technologies required to support low latency along with mission critical and future IoT services, therefore the overall 5G System (5GS) was designed since the beginning to efficiently support Edge and Fog for achieving both high performance and quality of experience.
From the architectural viewpoint, one of the most significant change with respect to the EPS is in how core network functions communicate to each other; in 5GS there are two available options:

one is the traditional reference point and interface approach whilst the other exploits the novel concept of the *Service Based Architecture* (SBA) reported in Figure 4-IX.
With SBA, Network Functions (NFs) can be classified as functions producing services and functions that consume services, with the framework supporting flexible procedures to efficiently expose and consume services. Different messaging models can be used: for simple services or information requests a *request-response* model can be used while, for more advanced and complex processes, the framework supports a *publish-subscribe* (and *notify*) model. The latter model was also adopted in 5G-CORAL as the messaging system within the EFS.

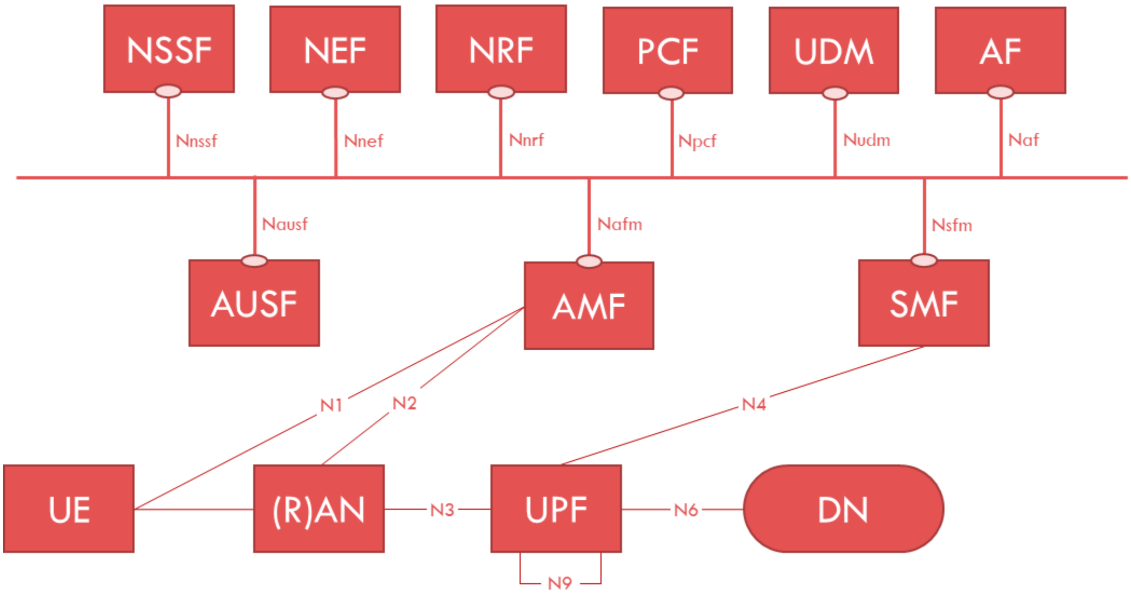

FIGURE 4-IX – 5G SYSTEM (5GS): SERVICE BASED ARCHITECTURE (SBA) REPRESENTATION [89].

Following the outcome of the work in [59] and focusing on the SBA representation, also for the 5G-CORAL platform it could be possible to consider a mapping of the platform itself to the Application Function (AF in the above Figure) which can make use of the services and information offered by the 5GS NFs based on the necessity and policies/permissions. Some of the services provided by NFs are accessible only via the Network Exposure Function (NEF), which is also available to untrusted entities (external to the operator's domain) to access these services.
The User Plane Function (UPF) has a key role in integrating the 5G-CORAL platform in a 5G network as it handles the traffic towards specific 5G-CORAL Applications/Functions and towards the network. Given the distributed nature of the 5G-CORAL architecture and the possibility to flexibly locate the UPF based on the requirements of the service to be provided, several deployment options can be considered, with the 5G-CORAL platform located in a place which could be close to the RAN node (LTE ng-eNB and/or NR gNB) up to the central Data Network (DN). 3GPP Rel-15 technical specifications also support non-3GPP access networks, connected to the 5G Core Network (5GC) via the Non-3GPP InterWorking Function (N3IWF). Non-3GPP access includes access from, e.g., Wi-Fi, Worldwide Interoperability for Microwave Access (WiMAX), fixed and Code Division Multiple Access (CDMA) networks and can be either trusted or untrusted. Trusted access is often assumed to be an operator-built access with encryption in the RAN and a secure authentication method; conversely, untrusted access encompasses any type of access over which the telecom operator has no control, e.g. public hotspots, subscribers' home Wi-Fi and Corporate Wi-Fi. It also includes Wi-Fi access that does not provide sufficient security mechanisms such as authentication and radio link encryption. In 3GPP Rel-15 only untrusted non-3GPP accesses are supported.

An exemplary deployment of the 5G-CORAL solution within the 5GS – encompassing both 3GPP and untrusted non-3GPP radio access networks – is represented in Figure 4-X where the OCS, acting as an AF, can interact either directly with the needed NFs within the 5GC or via the NEF, depending on whether the OCS is owned by the operator or by a 3rd-party. At EFS level it is the EFS Service Platform that can interact with these NFs, again in the role of an AF. The remaining EFS entities could be deployed in a DN in the 5GS.

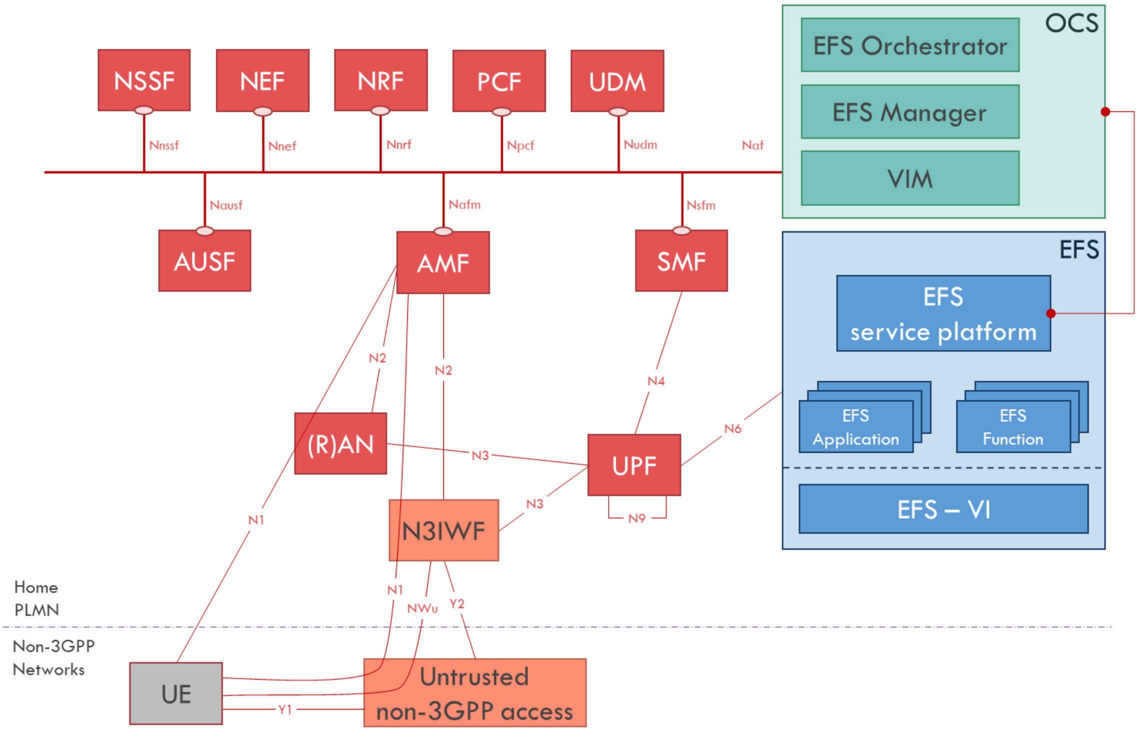

FIGURE 4-X – EXEMPLARY DEPLOYMENT OF THE 5G-CORAL SOLUTION WITHIN THE 5GS.

So far, the discussion was more focused on the technical feasibility of deploying the 5G-CORAL solution but, for the sake of completeness, also business-related aspects should be analyzed. According to the previously mentioned survey report in [54], the majority of operators (44%) tends to consider that Edge-enabled applications represent a new category of applications that can be offered across operators, but a not negligible percentage of the survey respondents (31%) thinks Edge is a *mix of something unique to their networks and business plans, enhancements to today's Cloud applications and a new category of cross-operator applications*. This implies that the market should think about Edge and Fog as the enabler for a new category of services that incorporates input from a broad ecosystem of participants, hence requiring the presence of a system integrator. On this aspect, a large majority of respondents (69%) declares that they will primarily exploit existing vendors and integrator partners for their Edge deployment, even if new integrator partners might become important (47% of respondents), not involving internal company resources (option chosen by only the 27% of respondents): likely, a combination of these options will prevail. And this could be exploited for the 5G-CORAL adoption in existing networks due to the possibility for the operator to act not only as the Edge and Fog system provider – which, as analyzed in deliverable D1.2 [2], is the central role in creating the ecosystem and making it successful – but also as the system integrator "orchestrating" the whole ecosystem of different players in order to provide tailored and reliable Edge-enabled solutions' implementations.

4.3.2 Deployment of the 5G-CORAL solution: the Vertical company perspective

Edge and Fog systems are foreseen to trigger new business opportunities and models for Vertical companies acting as Edge and Fog site owners (e.g., the Shopping Mall owner) to enable solutions aiming at improving both their advertisement reachability and effectiveness. As of now, though with legacy infrastructures available at sites of middle or large scale such as Shopping Malls, train stations, underground shopping streets, etc., Edge and Fog computing has not yet been introduced into such sites. This brings challenges to Vertical companies to deploy new application solutions into those sites. For example, if an infrastructure of a site does not provide Wi-Fi access, Fog CDs, Edge DCs, etc. the network will not be capable of provisioning new applications which require wireless access and massive compute power. The methodology allowing to cope with these challenges is the incremental deployment of new equipment within the already existing network infrastructure. To successfully incrementally introduce the 5G-CORAL platform into a legacy infrastructure, the following considerations are worth to address:

- *Security*: as the Edge and Fog application/service provider may not be the same as the site infrastructure owner, the access control of Edge DCs and Fog CDs should be managed. That is, the site infrastructure owner should provide a way for the Edge and Fog application/service provider (and/or its OCS) to access both Edge DCs and Fog CDs, with limited access right.
- *User access control*: the infrastructure should be able to control end-users' access to Edge DCs and Fog CDs: only end-users registering services have the right to access the resources available at both Edge DCs and Fog CDs.
- *Service provisioning*: the Edge and Fog application/service provider should provide a way to let the site owner to configure parameters and prepare local information for the applications/services deployed in the Edge and Fog system. For example, the Shopping Mall owner uploads shop advertisement information in Fog CDs for the end-users, which are then used via a service application installed in end-users' smartphones (as in the AR Navigation use case).
- *Contribution to 5G requirements* such as 1 ms objective, 1000-fold mobile traffic increase, connection density increase, etc.

5 5G-CORAL future directions

This section reports some indications on how the 5G-CORAL solution should evolve in order to improve the applicability of the solution itself to the already selected use cases (subsection 5.1) – also considering enhancements to the integration capabilities of the 5G-CORAL platform for full exploitation of all the features/external network entities needed to achieve higher performance – as well as new potential use cases, such as the on-device Artificial Intelligence (AI) and Machine Learning (ML) (subsection 5.2) as well as private wireless networks (subsection 0). It is noteworthy to highlight that, for the private wireless networks, also business considerations have been drafted, along with the feasibility assessment of applying the 5G-CORAL solution in this scenario.

5.1 5G-CORAL use cases' future directions

The following subsections are related to the future directions of the use cases already considered in the 5G-CORAL Project, highlighting potential enhancements which could be taken into account for extending the applicability of the platform for a specific use case, improving its own achievable performance, including other device types in the ecosystem, etc.

5.1.1 Shopping Mall scenario

5.1.1.1 Augmented Reality (AR) Navigation

Real-time AR Navigation application in the Shopping Mall scenario requires both low-latency communications for timely responsiveness and massive compute power for real-time image processing. As a result, Fog computing has been adopted to be part of the network architecture to enable AR Navigation. However, considering a massive number of customers in a Shopping Mall, the Fog environment should be responsive enough so that a high percentage of customers will be satisfied when using the application. To this end, a possible future direction for this use case is to enable cooperation among several Fog CDs in proximity to each other, in order to provide an AR Navigation service that is capable of serving a massive number of customers simultaneously with a satisfactory quality of experience. Also, AR Navigation can be adopted in contexts other than the Shopping Mall, such as educational institutes, historical museums, libraries, train stations, airports and so on.

5.1.1.2 Fog-assisted Robotics

The Fog-assisted Robotics use case is composed of two possible applications: the automatic cleaning and the coordinated delivery of goods. Those scenarios are characterized by being sensitive to the communication latency and by requiring a frequent interaction among the robots themselves as well as between the robots and 3rd-party entities. While the 5G-CORAL Project has thoroughly investigated low-level resource interaction (e.g., resource discovery and integration), the interaction between the resources and high-level entities has not been fully defined. For instance, the integration of 3rd-party OSSs has been defined in the architecture but no mechanisms have been defined. For what concerns Fog-assisted Robotics, the capability to properly integrate the robotics company's OSS into the 5G-Coral system is crucial for the future evolution of the use case and any potential business aspects.
The Fog-assisted Robotics use case envisages a constant monitoring and lifecycle management in order to operate correctly. Specifically, robots' synchronization and localization as well as the AP migration require a constant flow of up-to-date information to the OCS. In the current 5G-CORAL architecture the EFS lifecycle is managed at different levels by two OCS entities, i.e. the EFS Stack Orchestrator and the EFS Manager. While the integration of 3rd-party lifecycle managers within the OCS is possible at architectural level, the actual procedure for integrating them has not been investigated yet. This aspect is particularly important since some services or

applications may need the instantiation of different interfaces across the data and control planes. For the Rog-assisted Robotics use case, the integration is required, in the control plane, to properly isolate the manager necessary for monitoring the robots' signal quality and localization so as to trigger the AP migration. Without proper isolation, the manager may be required to run on the data plane making difficult ensuring the control plane performance.

The robots' owner is expected to provide a management interface for remotely accessing and controlling the robots. However, the current architecture focuses on the federation of the data plane and the control plane leaving aside the federation of the management plane. The Fog-assisted Robotics use case, in fact, may require the federation of the management plane to control such remote access as well as federation with 3rd-parties.

Finally, the Fog-assisted Robotics use case exploits the live access to information from the sensors and video cameras to locate the robots and to trigger the cleaning procedures. Such information, which is transmitted via the EFS Service Platform, is considered to be sensitive. As a result, privacy and security aspects should be considered in such a way it can be ensured that sensitive data are not shared with non-authorized applications, services and external entities. If an application or service is consuming or publishing data, there is the need to ensure that the data and the channel are properly secured and enciphered. While the current definition of the EFS Service Platform focused on the messaging protocol, aspects like authorization, accounting, and violation-detection for EFS services has not been investigated in depth. In order to make the Fog-assisted Robotics feasible in a commercial environment those aspects should be considered.

5.1.1.3 Virtual Reality (VR)

Scalability is a major concern in entertainment applications, such as video streaming and gaming. In particular, VR applications dictate fairly strict QoS requirements, ranging from high and constant data rate to very low latency. As future work, how to address scenarios where a very high number of VR end-users wish to consume the VR service in a limited area, e.g. a concert venue or a conference center, might be considered. In such conditions, not only does the downlink communication need to be reliable and scalable, but also the uplink channel must be able to support the tile reporting periodically performed by each end-user. For this reason, the adoption of custom-tailored multicast communication protocols capable of aggregating video streams, associated with a certain tile, to multiple end-users watching the same tile might be envisioned. In addition, the benefit of using D2D communications to relieve the uplink channel of multiple end-users' access every time a tile report is required will be evaluated. End-users may be grouped in clusters and certain end-users may be elected as cluster head to handle the tile reporting on behalf of the other group members. Such end-user cooperation is expected to reduce the uplink channel load, as most of the users will communicate with the cluster head via D2D links.

5.1.1.4 Multi-RAT IoT

In this 5G-CORAL use case the focus has been on the EFS development for showcasing the feasibility of softwarizing IoT communication stacks in the Edge and sharing network data between softwarized network functions. In this way, a multi-RAT IoT network is transformed to a more data-driven network where a huge amount of data can be stored and exchanged. This would enable to utilize ML and AI techniques to further automate and optimize the network. This would be very useful for massive IoT applications where reliable connectivity to billions of devices needs to be managed cost-effectively. Therefore, the future direction of this use case is to explore the possibility of integrating the capabilities of ML and AI to the system and investigate the use cases of network automation and optimization. For example, optimizing Edge resource utilization given traffic load, improving radio resource management (especially regarding interference management) as well as link quality estimation, localization accuracy, etc.

5.1.1.5 Software Defined Wide Area Network (SD-WAN)

In the current implementation of the SD-WAN use case, the allocation of resources for federation is done dynamically under the demand of a new connection by the end-user terminals. As a direct consequence of this, the end-user will experience some delay during the migration from one AP to another due to the need to instantiate a new application process in the federated EFS when a connection change is detected. However, this event could be further predicted from the data collected in relation with the user activity and the quality of the connection. By leveraging AI and the processing of a series of KPIs such as the jitter, the RTT delay of the connection and the end-user activity (detecting that the end-user is currently employing the application and has not yet ended the transaction), the SD-WAN orchestrator could predict that the end-user will soon move to a different AP while still requesting the use of the applications. This prediction would then be used to instantiate and configure the connection to a new process in the other federated EFS before this action takes place. Other inputs that are exploited for this purpose could be the time of day or the day of the week/year to predict the end-users' activity/load.

5.1.2 Connected Cars scenario – Safety use case

Vehicular crashes and congestions became two of the main socio-economic problems in developed countries due to the fast development of metropolitan areas, city roads and highways. The global status report on road safety launched by the World Health Organization (WHO) in December 2018 reports that the number of annual road traffic deaths has reached 1.35 million: road traffic injuries are now the leading killer of people in the range of 5-29 years old, with the highest price in terms of deaths paid by pedestrians, cyclists and motorcyclists. Furthermore, traffic jams also cause a tremendous waste of time and fuel, which therefore contribute to huge economic losses. The above-mentioned report suggests that the price paid for mobility is too high and drastic action is needed to put measures in place in order to save as many lives as possible: among these measures, provisioning of timely information to drivers and/or vehicles may allow for a significant reduction of this high price in terms of human deaths [60]. Focusing mainly on road safety, efforts in reducing road traffic deaths vary significantly among the different countries in the world: with an average rate of 27.5 deaths per 100.000 inhabitants, the risk of a road traffic death is more than 3 times higher in low-income countries than in high-income ones, where the average rate is 8.3 deaths per 100.000 inhabitants; from another perspective, it should be noted that although only 1% of the world's motor vehicles are in low-income countries, 13% of deaths incurs in these countries. This suggests that connectivity technologies – working jointly with distributed analytics platforms as the 5G-CORAL solution – which allows to exchange driving information among vehicles that are not natively provided with embedded connectivity for safety purposes may be a market segment to focus on. This might also be motivated even further when considering that a considerable number of existing cars can still come "online": there are millions of cars on roads today (almost all carrying a mobile device inside) and although connected cars with embedded connectivity will dominate the market in the future, used cars today are lasting longer than ever and, as a result, drivers holding their vehicles longer also want to get more from the vehicles themselves. Even though car manufacturers strive to bring consumers the newest advances in vehicles' connectivity technology (e.g. in-car Wi-Fi), they require an additional price for connectivity rather than including it in the cost of the vehicle. This represents a barrier to consumers, who will not buy the connectivity unless they deem it worth the cost [61]. In order to recover their investments in providing an embedded wireless connection to their products, car makers are trying to directly monetize vehicle connections by means of subscription-based services or marketing platforms (e.g. General Motors (GM)'s Marketplace), but what car companies need to do is to make connectivity an intrinsic element of operating a car: with its native flexibility, 5G technology will surely help in making this a reality. Although car manufacturers such as GM, BMW, Audi (accounting for almost 80% of cars with embedded connectivity [62]) rely on mobile operators to deliver connectivity, telecom operators are currently failing to derive significant revenues in this market, with analysts estimating an average monthly revenue from car connectivity subscriptions to be in the $1 to $2 range per vehicle [63]. The lack of revenues with in-car connectivity has pushed several telecom operators out but others, such as Verizon and AT&T, are engaging in partnerships with major car and truck makers in order to ingrain themselves in the Connected Cars business, hence playing a larger role than just mobile-service supplier, and are looking for new revenues from safety applications as well as selling data mined from drivers/vehicles and entertainment [63], pushed by the growth opportunity offered by the in-car connectivity, as 380 million connected cars are estimated to be on the road globally by 2021 [62].

With this estimation in mind, it is up to mobile operators to capitalize on the growing demand for advanced car safety technologies, many of which remain out of reach for the average driver who is expected to keep its car for many years (far beyond the 2021). So far, the used car market has not been considered by operators, mainly interested in the natively-connected cars market. Operators might establish new revenues in the used cars market not only via aftermarket solutions such as OBUs installed within cars and operating over the cellular network (e.g. LTE) or On-Board Diagnostics (OBD) dongle-based solutions [64] – even if it is difficult for operators to drive subscriptions' costs down, because hardware devices are expensive to manufacture – but also via the exploitation of the smartphone(s) each driver owns nowadays, in both mature and developing countries. For the safety use case as intended in the context of the 5G-CORAL Project, a short-term future improvement might be the implementation of direct communications, e.g. IEEE 802.11p (also known as Dedicated Short Range Communications, DSRC) or 3GPP LTE-V2X (i.e. V2X communications over PC5 interface), in addition to wireless technologies currently used within OBUs (i.e., LTE and Wi-Fi). But another future direction of this use case might involve a deeper exploitation of resources available within smartphones.

Cars are getting smarter, but smartphones are getting smarter at a much faster rate: the increasingly sophisticated range of sensors (GPS, accelerometers, gyroscopes) has reached a level of maturity that allows them to capture a detailed profile of a driver's behaviour. Moreover, smartphones are capable of capturing aspects of driving behaviour that OBUs cannot, e.g. whether the driver is on the phone or texting while driving. A smartphone-based safety solution might measure driver caution, control and focus while on the road; collisions might be automatically detected by measuring a variety of factors such as sharp acceleration, mobile usage, length and time of trip, speeding and hard braking. In addition to this, a distributed analytics system realized by exploiting the flexibility and ubiquity of the 5G-CORAL platform allows for data generated by smartphones to be processed locally so to provide the driver suggestions for continuous monitoring and correction of dangerous driving habits. In case a collision occurs, drivers and passengers might be kept safe by collecting and sending valuable crash data to emergency rescuers timely (e.g. within seconds), drastically increasing the probability of saving the lives of people involved in the incident.

There are strong business opportunities for a telecom operator when offering a smartphone-based safety application to its customers: for instance, parents might be willing to pay for premium services allowing them not only to know whether their sons and relatives are safe (69% of parents monitor their teen drivers by using a smartphone application [65]) but also to be supported in the event of an incident. Beyond this, a driver safety proposition can also form the core of a wider ecosystem of services related to driving, e.g. roadside assistance, Pay How You Drive (PHYD) insurance, city parking locators and so on, not only provided by a single player in the ecosystem but, on the contrary, exploiting the services and resources made available by a plethora of actors.

5.1.3 High-Speed Train scenario

With this use case, a holistic Edge and Fog service management system to provide mobile service continuity was investigated. Two issues arose in regards to mobile service continuity, namely control signalling storm towards the core network and backhaul capacity impacts. To solve these issues, a group handover scheme which exploits Edge and Fog capabilities by deploying vMMEs at the Edge and Fog tiers to lower signalling storm. Then, a container-based pre-copy migration scheme has been developed to relocate applications in an Edge and Fog environment to mitigate the impacts on the capacity of backhaul links towards the core network. Future directions for this use case include the following items:

- *Improved user classification*: this operation needs to be executed prior the end-users migration from the train to the train station. Currently, end-users are classified based on the QoS levels of the services they are requesting. In the future it is expected to also classify end-users based on their applications to provide application-tailored service continuity.
- *Dedicated containerized application*: currently, a single application container is instantiated to a group of end-users. To improve the system performance, each end-user should be accommodated with a dedicated containerized application.
- *Smart service continuity*: ML- and AI-based techniques could be used in the future to enable the system to fully self-recognize and self-act without operator involvement.

With future improvements, this holistic Edge and Fog system will open doors for new applications not only for moving networks but also for smart cities. In the case of moving networks, new applications for the High-Speed Train such as the driverless train might be envisioned: for instance, France aims to have driverless trains travelling at high speeds by 2023 [66]. The trains will be able to tightly coordinate among each other and will be equipped with sensors that allow them to detect obstacles and automatically brake, if necessary. This represents an important economic impact as it will be possible to increase both the speed and frequency of train journeys. When coming to the smart city, further news services associated with the High-Speed Train might include provisioning of city information, 3D interactive offline city maps, points of interest indications and local information to all the train passengers based on users destination.

5.2 On-device Artificial Intelligence (AI) and Machine Learning (ML)

Over the past few years, research interest on Artificial Intelligence (AI) has significantly grown boosted by the increasing number of software projects launched by Google. AI popularity is associated with a number of factors, including the large accessibility to affordable neural networks, the widespread support to Cloud computing infrastructure and the massive use of research tools and datasets. In [67], *Global AI as a Service Market* is expected to reach $20.2 billion (USD) by 2024 from $1.25 billion (USD) in 2016 at a Compound Annual Growth Rate (CAGR) of 48.9%. Figure 5-I lists the top-100 startups which make use of AI to drive transformation in several Verticals.
Within the 5G ecosystem, the general perception is that AI could supply more reliable services and make optimized decisions, for example by improving the communication system. Authors in [69] introduce AI-enabled use cases as well as anticipate that the rise of mobile 5G networks will have a tremendous positive impact on AI and robotics research areas. [70] shows various AI implementations to solve 5G communication system issues, such as automated out-of-coverage regions discovery, energy efficient network management, cognitive network resource management, context aware cross-layer optimization and agile PHY transmission optimization. ITU-T established ML5G focus group [71] to investigate the impact of AI on networks and

published a recommendation document [72] to describe the architectural framework for ML in future networks.

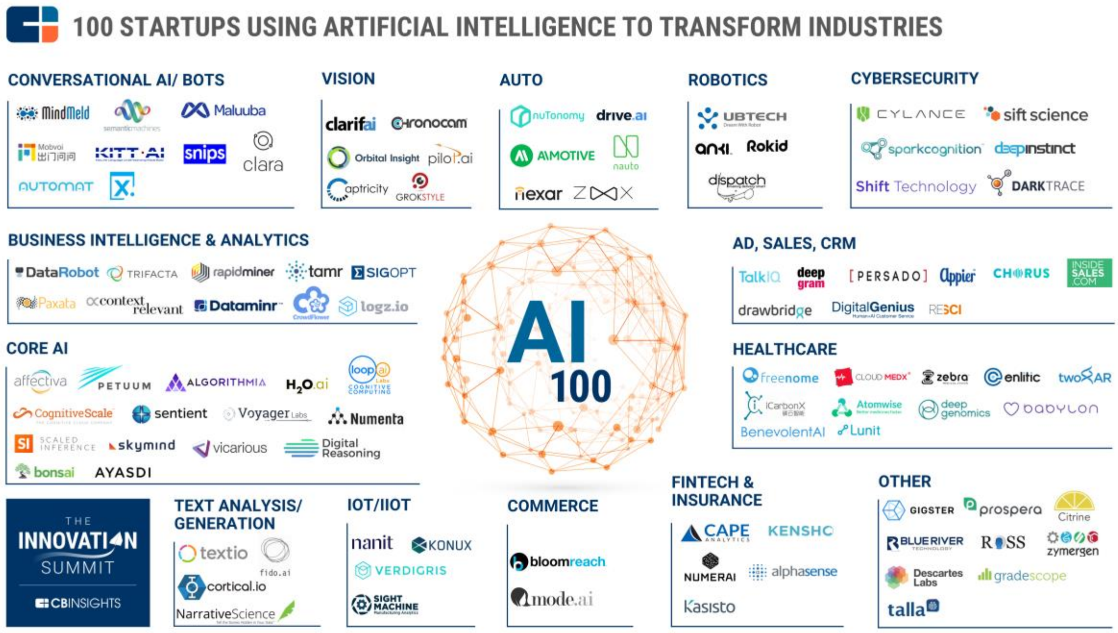

FIGURE 5-I – EXEMPLARY AI-BASED VERTICAL SERVICES [68].

Recently, AI-based cognitive radio resource utilization has been considered as an important research topic. The Federal Communication Commission (FCC) announced the adoption of AI tools for smart spectrum sharing [73]. In this research topic, multi-RAT convergence is a key aspect since the basic usage scenario for spectrum sharing assumes that a device is able to detect (or anticipate) and avoid the simultaneous resource occupancy by using a different RAT [74].

5.2.1 AI from Cloud to Edge and Fog

With AI technology, a machine (or software agent) determines a solution for a problem based on data collected from the *things* without any human intervention. ML is a prominent method to enable AI, where machines can infer a mapping rule from pairs of input and output data. Furthermore, Deep Learning (DL) is gaining interest as a promising ML technique, by making inferences (IF) with multi-layer perceptron of neural networks. Figure 5-II shows the relationship among AI, ML and DL.

Conventional ML is enabled through Cloud servers, since it requires huge processing power for iterative and complex computations as well as massive storage for big data. However, the Cloud-based ML comes with limitations:

- *Service latency*: a client will undergo the inevitable network delay to take an action, being it generally located far away from the Cloud server. This clashes with low latency service constraint dictated by specific use cases, such as immersive entertainment and automated vehicle.
- *Centralized architecture*: a client should always be connected to a Cloud server since the whole ML and inference processes are handled by the Cloud server. If the client loses connectivity, the ML service would be immediately halted.
- *Public data exposure*: raw data should be exposed to the Cloud server, yet privacy must be ensured. If the Cloud server gets compromised, data would be publicly disclosed.

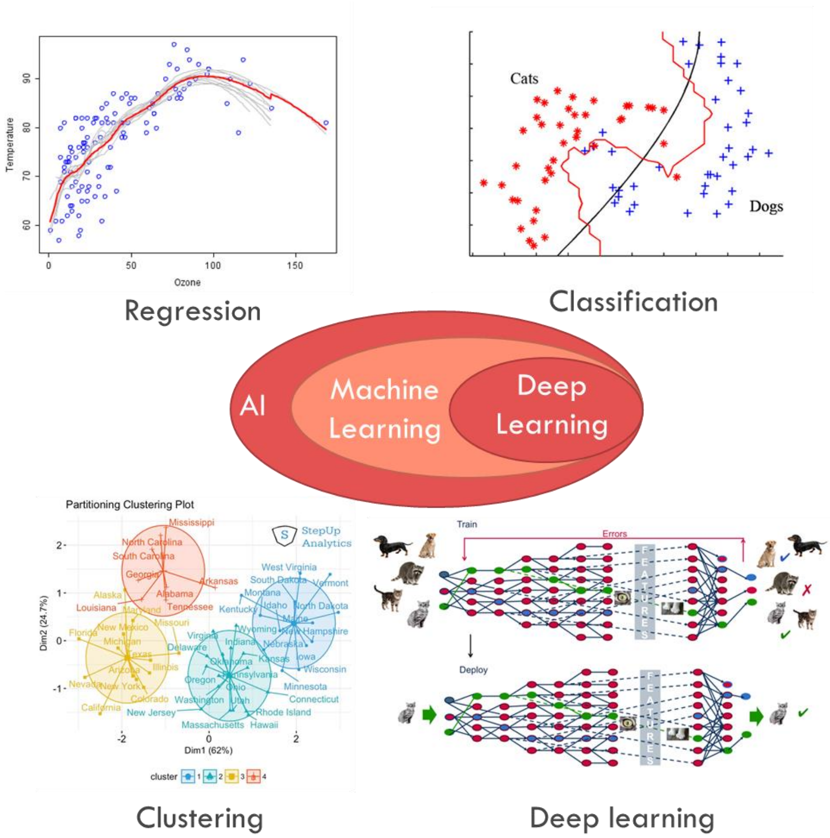

FIGURE 5-II – ARTIFICIAL INTELLIGENCE/MACHINE LEARNING/DEEP LEARNING.

Distributed learning is proposed to mitigate the restriction of Cloud-based ML. In distributed learning, learning and inference tasks are moved from centralized Cloud servers to Edge servers or Fog devices (see Figure 5-III). For instance, Amazon designed a distributed ML inference system with the AWS Greengrass platform [75], where ML inference is performed by AWS Greengrass Core on device and the input/output data is fed back to AWS SageMaker in the Cloud server and trained, with significant end-to-end latency reduction.

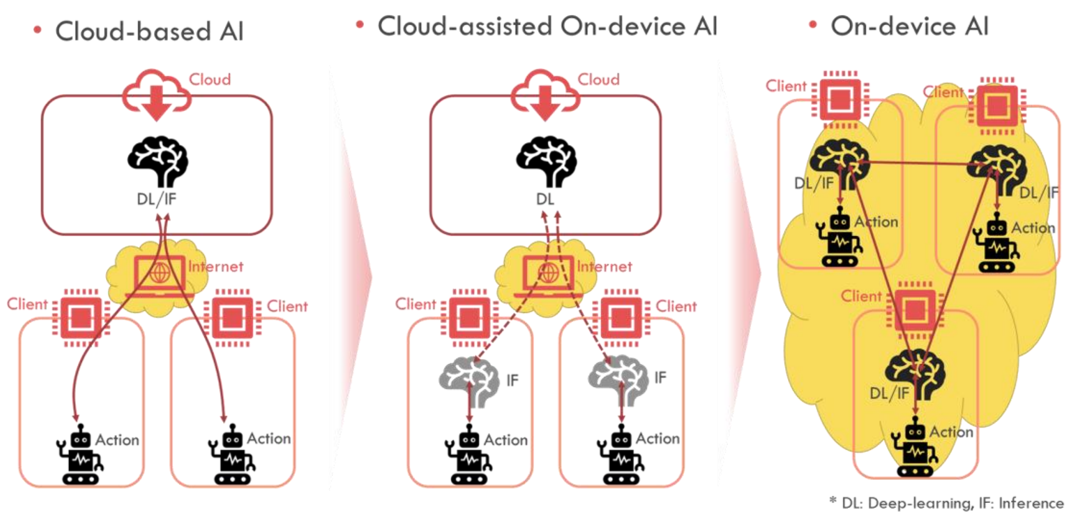

FIGURE 5-III – FROM CLOUD-BASED AI TO ON-DEVICE AI.

In [76], a distributed federated learning mechanism is proposed as a solution to cope with the service latency induced by Cloud-based AI in vehicular communications. Each vehicle determines an optimized Resource Blocks (RBs) configuration and transmission power level with the estimated queue distribution and signal quality information which will be inferenced based on own local ML model. Chipset vendors, e.g. Qualcomm and NVIDIA, and Cloud service providers, e.g. Amazon and Google, also stress the importance of *On-device AI* (or distributed AI) [77][78][79]. The key advantages of distributed AI are the following:

- *Low latency*: since ML inference is carried out near a device, either on an Edge server or Fog device, the action determined by the ML agent will take less time in comparison with Cloud-based ML.
- *Less strict dependency on centralized server*: distributed ML inference does not always require connectivity with Cloud server since each device has its own ML agent running.
- *Data filtering*: all raw data is not reported to a Cloud server. The system will use filtered metadata, thus consuming less bandwidth and improving the transport network utilization efficiency.

Table 5-I summarizes the comparison between Cloud-based and On-device AI.

TABLE 5-I – COMPARISON BETWEEN CLOUD-BASED AI AND DISTRIBUTED AI (ON-DEVICE AI).

	Pros	Cons
Cloud-based AI	• High computing power and large storage in a Cloud server • Big data base • Centralized learning	• High RTT for execution (inappropriate to low latency service) • Require always connection with Cloud server via Internet • Consume transport resources for interaction with Cloud server such as data report • Private data could be exposed
Distributed (On-device) AI	• Low latency (low RTT) for execution • Scalability • Less sensitive to intermittent Internet connectivity • Reduction of bandwidth costs for data report by client and control by Cloud server • Privacy and security (data are exchanged within a restricted area)	• Limited computing power • Small data base • Contention among multiple agents (semi-optimized)

5.2.2 AI leveraged by 5G-CORAL

The current AI development trend fits well with the 5G-CORAL platform evolution. As mentioned in the previous subsection, the AI tasks move to the Edge node or Fog devices for enabling low latency services as well as ensuring privacy and security. The main target of the 5G-CORAL platform is to optimally utilize computing substrates and networking resources of Edge nodes and Fog devices in a distributed manner, on top of virtualization and orchestration principles.

The 5G-CORAL platform enables a client device to use virtualized computing resources hosted by Edge servers and Fog devices within the EFS service area such that the client device can perform more complex computing tasks than by own computing resource. As distributed (On-device) AI is assumed, a client device could execute ML inference as well as heavy DL tasks through the computing resources provided by 5G-CORAL platform (see Figure 5-IV).

The 5G-CORAL platform can provide networking resource information via the EFS Service Platform. AI applications would be deployed optimally considering both computing and networking resources, which could make more precise Edge/Fog computing substrate deployment decisions. Regarding a specific AI-based function relevant to wireless communication, such as AI-based Self Organized Network (SON) and AI-based cognitive spectrum sharing, the networking resource information provided by the EFS Service Platform would enable these AI-based functions. Specifically, the multi-RAT convergence envisioned in 5G-CORAL is a crucial requirement of AI-based cognitive spectrum sharing technology, where a device detects and avoids the active resource occupancy through a different RAT.

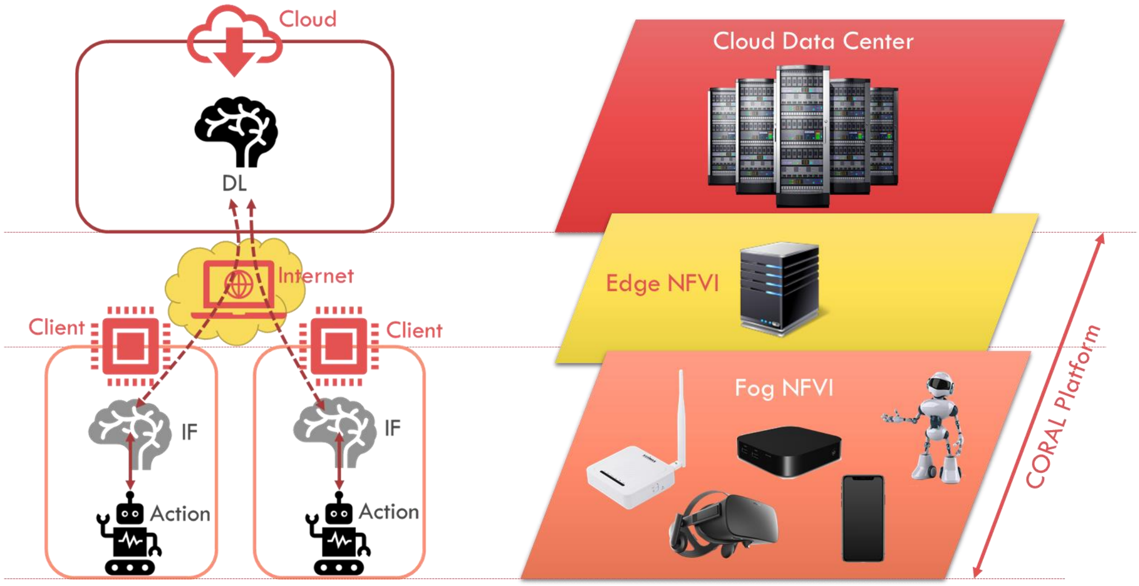

FIGURE 5-IV – AI FACILITATED BY THE 5G-CORAL PLATFORM.

The 5G-CORAL platform is not tied to a static server or to a specific network technology. The EFS NFVI can be dynamically adjusted and on the move, not regionally restricted to a specific location within the EFS service area, hence the 5G-CORAL platform is foreseen to bring very flexible deployment capabilities for AI-based applications.
Table 5-II shows the level of importance (High/Medium/Low) of each 5G-CORAL key benefit when taking AI into account, giving insights into how the 5G-CORAL platform can be exploited in the context of AI.

TABLE 5-II – 5G-CORAL KEY BENEFITS EXPLOITABLE FOR AI.

ID	5G-CORAL key benefit	Description	Level of importance
KB-01	De-centralization in terms of	proximity (lower latency)	High
		offloading (power saving, faster execution)	High
		processing data locally (congestion avoidance and security of data)	High
KB-02	Maximized utilization of local computing, networking and storage resources	Virtualization, abstraction and pooling of locally-available RAN resources are the enablers	Medium/High
KB-03	Improved AI-based network function (e.g. SON, cognitive spectrum sharing)	Achieved via optimized interworking among the multiple RATs in the RAN	High
KB-04	Cooperative on-demand networking and computing	This can also be provided in real-time and on the move	Low/Medium
KB-05	Improved flexibility and scalability of application deployment	AI applications could not be regionally restricted and could be delivered on top of various network technology	High
KB-06	Open value chain for including more stakeholders	This includes also end-users to participate in the solution as providers (towards incentives)	High

5.3 Private wireless networks

Businesses often depend on wireless systems for handling (i.e. communicating and controlling) automated systems and gathering data, by relying on licensed narrowband radios or unlicensed Wi-Fi networks. However, those types of networks are not able to support the characteristics of today's workplaces since there is a need to securely connect hundreds of people, to precisely control complex/critical machinery and to analyze data from thousands of sensors, e.g. monitoring production processes. To meet these high demands, enterprises are starting to consider new technologies which encompass the so called *private networks*.

A private (wireless) network can be defined as a local network that uses dedicated radio equipment to service a premise with specific applications and services. By focusing on specific requirements, the private network can be tailored for more optimized performance (e.g. low latency). Furthermore, the use of dedicated equipment allows the network itself to be independent of traffic fluctuation in the public, wide-area macro network [80]. Private networks can be deployed in several flavors, basically depending on the enterprise's requirements: industrial markets (and related business-critical market applications) are highly fragmented and characterized by customers who are conservative adopters of new technologies as a consequence of their operationally-intensive activities (i.e. failure or disruption of devices, services or systems will negatively impact business operations), whose choices are also driven by implementation costs and Return On Investment (ROI). As a result, business-critical markets put requirements on the network characteristics including low latency, high reliability and high security.

5.3.1 Private LTE networks

A private LTE network is an LTE standard-based network which is designed to serve specific enterprise business, government or educational purposes, leveraging on the same technology which is used by communications' carriers to connect today's smartphones. In [82], a private LTE network is defined as a *dedicated network for consumers, businesses and IoT deployed either in licensed, unlicensed or shared spectrum*. A recent study from Harbor Research indicated that the private LTE network market in the context of IoT could reach $31billion (USD) by 2022 [81]; different types of industrial market segments, in fact, are showing increasing interest in this networking solution, including diverse resource-based industries, manufacturing and infrastructure-driven segments covering a wide range of domains, e.g. [80] (see Figure 5-V):

- Mining and oil/gas exploration and delivery
- Chemicals, pharmaceuticals, consumer goods, automotive, aerospace and electronics
- Power generation, transmission and distribution
- Transportation venues such as airports and railway stations
- Healthcare delivery and,
- Supply chain including warehouses, distribution and container facilities.

When referring to IoT use cases, it has been observed that existing technologies have proven cumbersome and costly to apply due to many conflicting protocols and poor integration capabilities among them: many IoT networks, in fact, are Wi-Fi, Bluetooth or Zigbee-based and, while this might be enough for certain enterprise applications, enhanced network capabilities are required for integrating connectivity into industrial, business-critical market segments. This can be achieved by private LTE networks since their design do not only take sensors and connectivity into account but also an intelligent data management system being in charge of handling the sensors' inputs properly.

Private LTE networks could be operated not only by public, traditional mobile operators but also by private, 3rd-party network providers. Unlike public operators – typically requiring carriergrade solutions along with robust network management services – private network operators generally include large utilities and industrial companies which are more focused on specific applications and use cases that can be provided by deploying simpler network solutions with a straightforward user experience. When designing a private LTE network, two main aspects have to be defined first: *equipment* and *spectrum*. Several equipment vendors in the marketplace already produce LTE devices working in the bands most likely to support private LTE networks.

More critical is the choice of spectrum: individuals and businesses may decide to use any of the already licensed commercial mobile spectrum bands for private LTE systems; while most of that spectrum has already been licensed to mobile network operators over broad geographic regions, "secondary" market activities could be encouraged by spectrum regulation bodies, including offering private LTE networks by leasing spectrum and license assignment agreements that involve partitioning (assigning spectrum over a specific geographic area) and disaggregation (assigning specific spectrum) or a combination of the two.

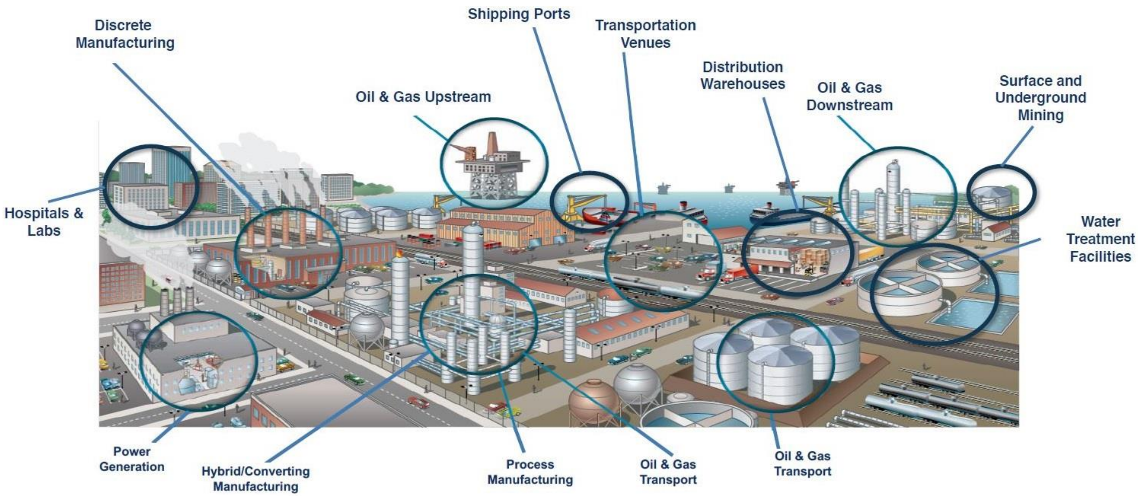

FIGURE 5-V – INDUSTRIES ADDRESSABLE BY PRIVATE LTE NETWORKS [81].

Recently there has also been significant progress around other spectrum options, i.e. unlicensed and shared spectra, which significantly lower the barrier for new entrants hence opening private LTE network deployments to a wider ecosystem as stated in [83]. In addition to private LTE, alternative spectrum types are also attracting mobile network operators being interested in low-cost network densification options in indoor and/or venue environments (due to power restrictions as typically seen in unlicensed and shared spectrum).

Initial efforts have mainly been focused on unlicensed spectrum, e.g. resulting in the MulteFire Alliance ecosystem [84] allowing for deployment of a private LTE network in the globally-available, unlicensed 5GHz band. Not only enterprises but also traditional operators are showing interest in using MulteFire to go beyond the boundaries of frequency ownership: in fact, they are evaluating the possibility to cover corporate's international customers located in different countries where the operator does not own spectrum, eventually offering them the value of LTE even overseas while ensuring its inherent security, quality of service and mobility features [86].

Lately, also shared spectrum was considered, an example of which is the CBRS Alliance [85] – including members such as all four major US mobile operators, i.e. AT&T, Verizon, T-Mobile US and Sprint, as well as companies such as Google, Intel, Nokia, and Qualcomm – which developed a private, TDD-based LTE network deployed in the US Citizen Broadband Radio Service (CBRS) band – a 150MHz shared spectrum band around the 3.5GHz frequency. CBRS is a technology-neutral, sharing framework where spectrum can be available for use either on a Priority Access

Licensee (PAL) basis (via auction) or on a General Authorized Access (GAA) basis to any interested party.
The efforts of the CBRS and MulteFire alliances represent the initial steps towards the establishment of a collaborative ecosystem necessary to develop a new generation of wireless networks to be used in specific private contexts: for instance, the partnership among Nokia, Qualcomm and General Electric around industrial deployments of private LTE over the CBRS band represents a tangible evidence of how private LTE can be adopted in the enterprises' industrial wireless networks [87]. Also, Huawei has effectively deployed private LTE solutions across shipping ports, mine sites and oil exploration platforms achieving better coverage, capacity and availability that improves the overall security, efficiency and sustainability of operations [80]. These examples basically confirm the need to establish strategic partnerships to enable means for significant penetration of private wireless network solutions, provided that these networks will ultimately need to be interoperable and easily integrated into existing enterprises' infrastructures and systems.

5.3.2 Private networks in upcoming 5G systems

Based on the previous subsection, it can be stated that private wireless networks today are mainly 4G-based, where LTE is one of the main technological enablers for the next industrial revolution. With the latest progress in 3GPP, it is expected that these networks will evolve into private 5G networks in the future. In this sense the availability of a brand-new, cloud-native and virtualized core network (i.e. 5GC) – characterized by a high degree of modularity of its functionalities – allows to achieve enough flexibility to be exploited for realizing one of the main innovations of 5G systems, that is, *Network Slicing*. The ability to support several services, each with specific performance requirements, on a common network infrastructure is a powerful idea with interesting commercial opportunities, with Network Slicing being essential as it represents the technological enabler allowing for mobile operators to configure end-to-end (i.e. including RAN, core network and service platforms) logical virtual network instances – each optimized to the specific functional requirements of a customer or application – more quickly and at lower costs than designing and deploying traditional dedicated networks. Network slices can be at individual user/service-level (e.g. a video streaming slice) or at a company-/industry-level (e.g. an automotive slice, a utilities slice, etc.) and they can be seen as a network being able to adapt itself to the needs of the specific application. Network Slicing is of particular interest for operators aiming at diversifying their business activities across different industry verticals, hence expanding their customer base and increasing their revenues massively and sustainably.
3GPP extensively worked on developing Network Slicing on both core network and RAN sides during Rel-15 and produced a set of technical specifications addressing the topic [88][89]. At RAN side, it is noteworthy to recall that network slices can be provided to UEs not only via NR radio access but also by exploiting already deployed LTE base stations (Rel-15 compliant eNBs) since it is possible to logically connect Rel-15 LTE eNBs to the 5GC.
Actually, it is already possible for operators to offer commercial "slicing" propositions in the near term by exploiting today's commercial 4G mobile networks, without the need to wait for the 5GC to be widely deployed and completely specified in terms of features availability. These "pre-5G" solutions basically provide a commercial and technical bridge to 5G and exploit techniques such as APN routing, Multi-Operator Core Network (MOCN) and DEdicated CORe network (DECOR). As an example, Verizon provides a private LTE service which securely extends corporate networks to employees' mobile devices or branch offices, also offering wide area connectivity over the public shared RAN. A dedicated per-enterprise network gateway is used to allow the enterprise itself to control user policies, data caps, traffic management, QoS, etc. [90].

5.3.3 5G-CORAL applicability to private wireless networks

Once the definition, addressed use cases and possible implementations of a private wireless network have been discussed, it needs to be understood how this application fits the 5G-CORAL concept with related business considerations. Within the distributed computing infrastructure foreseen by 5G-CORAL the focus is on the Edge and Fog tiers (with related interactions), which are also integral parts of the 5G infrastructure as they represent the enablers for achieving not only end-to-end low latency (via proximity), high bandwidth (via exploitation of multiple, convergent RATs) and increased reliability (via consumption of data locally) but also security, as risks can be limited up to the Edge tier. Security is one of the key requirement especially for business-critical markets' customers (i.e., enterprises), which require networks being able to securely connect hundreds of company's employees and to properly handle confidential business information (i.e. information whose disclosure may harm the business, that is, trade secrets, sales and marketing plans, new product plans, notes associated with patentable inventions, customer and supplier information, financial data, etc.). The deployment flexibility of the 5G-CORAL solution allows the solution itself to be easily applied to several use cases (often having contradictory sets of requirements when compared to each other) encompassing also the private networks. This is basically due to the overall distributed system composed by Edge and Fog nodes whose characteristics in terms of computing/storage/networking capabilities, heterogeneity, and ownership can be exploited for deploying a private network fulfilling the needs of a specific application (e.g. a business-critical production process within an enterprise). Table 5-III reports the level of importance (High/Medium/Low) of each 5G-CORAL key benefit when taking private networks into account, basically giving an insight of how the 5G-CORAL technical solution can be exploited in the context of private networks.

TABLE 5-III – 5G-CORAL KEY BENEFITS EXPLOITABLE FOR PRIVATE NETWORKS.

ID	5G-CORAL key benefit	Description	Level of importance
KB-01	De-centralization in terms of	proximity (lower latency)	High
		offloading (power saving, faster execution)	Low
		processing data locally (congestion avoidance and security of data)	High
KB-02	Maximized utilization of local computing, networking and storage resources	Virtualization, abstraction and pooling of locally-available RAN resources are the enablers	Medium/High
KB-03	Improved 5G RAN area spectral efficiency	Achieved via optimized interworking between the multiple RATs in the RAN	High
KB-04	Cooperative on-demand networking and computing	This can also be provided in real-time (e.g. in different areas of the implant served by the private network) and on the move	Low/Medium
KB-05	Lower deployment cost and time	Achieved via densification of constrained Fog devices and automated orchestration and control	High
KB-06	Improved flexibility and scalability for supporting different network topologies	Achieved via pluggable/detachable Fog nodes based on the needs of the specific use case requiring a private network (e.g. enterprise)	High
KB-07	Relax dependencies on sophisticated infrastructures	Achieved by leveraging local (already-deployed) networks and resources	High
KB-08	Open value chain for including more stakeholders	This includes also end-users to participate in the solution as providers (towards incentives)	High

Players acting within the 5G-CORAL ecosystem as well as the key roles in the ecosystem itself have been identified and extensively described in 5G-CORAL deliverable D1.2 [2]. Also, a mapping between key players and possible role(s) was also provided in deliverable D1.2, reported in Table 5-IV for the sake of completeness. When thinking of private networks, it is straightforward to link them to the so-called Verticals, e.g. enterprises. Hence, a player acting as an Edge and Fog system provider such as the telecom operator will be able to integrate 3rd-party resources – provided by both the Vertical company (i.e. the facility owner mapped into the

Edge and Fog site owner) and the hardware vendor – as part of the private network designed and deployed for the Company being able to fulfil Company's requirements, by also leveraging and integrating the already-deployed corporate DCs and IT environment. Private networks can then be seen as a type of telecom services and it is easy to understand that the telecom operator can run this business to increase its revenues by leveraging its experience and knowledge in network operation, sales and customer relationships. The case where the telecom operator acts not only as the connectivity provider but also as the Edge and Fog system provider is considered in Figure 5-VI: basically, the Vertical company (i.e. the enterprise, which is the Edge and Fog application/service end-user) requests a private network service to the operator which, consequently, can be considered also as the Edge and Fog application/service provider.

TABLE 5-IV – MAPPING BETWEEN KEY PLAYERS AND ECOSYSTEM ROLES [2].

Ecosystem role	Telecom operator	Software vendors	OTT service providers	Cloud providers	Hardware vendors	Vertical Companies
Edge and Fog system provider	X		X		X	
Edge and Fog site owner						X
Edge and Fog hardware vendor					X	
Edge and Fog system software vendor		X		X		
Edge and Fog application/service provider	X		X			
Edge and Fog application/service software Developer		X				
Edge and Fog application/service end-user						X
Cloud provider				X		
Connectivity provider	X					

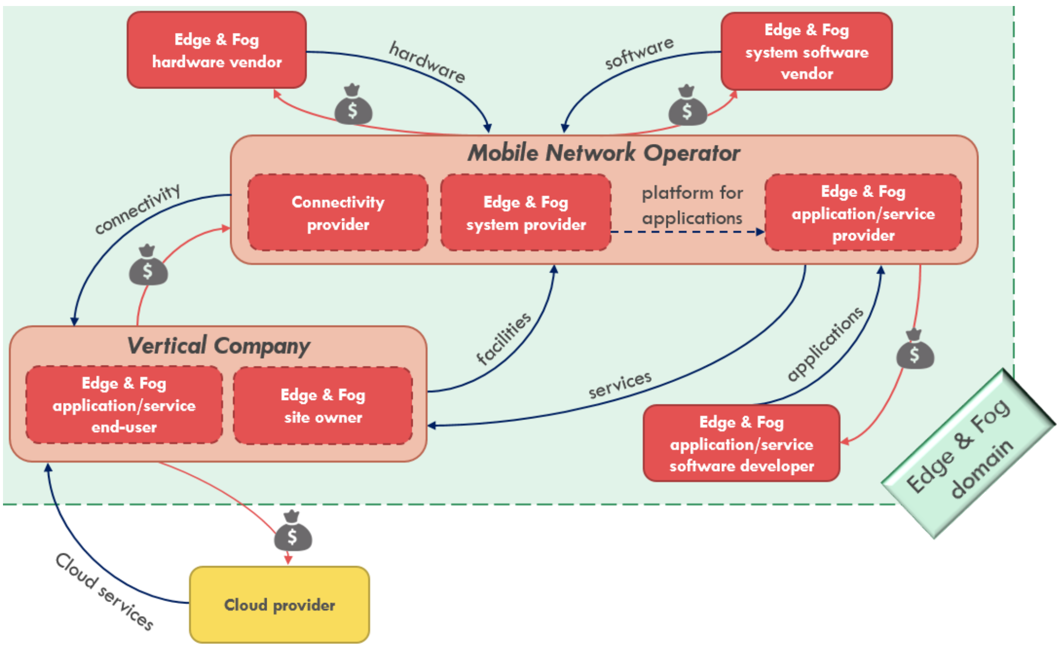

FIGURE 5-VI – EXEMPLARY RELATIONSHIPS AMONG 5G-CORAL ROLES WHEN CONSIDERING PRIVATE NETWORKS.

As a system provider, the operator might properly exploit existing Company's network infrastructure and facilities to provide all the means (e.g. hardware and software) needed to establish the private network within Company's premises. It should be noted that such private network services can also be provided by new players, e.g. network equipment vendors when using technical solutions based on unlicensed/shared spectrum operation (e.g. MulteFire) as well as by the Vertical company itself when partnering with telecom operators by means of spectrum sharing agreements or in case local spectrum licensing models will be allowed by spectrum regulation bodies.

6 Conclusions

This third WP1 deliverable concludes the work on the system-level design of the 5G-CORAL technical solution and identifies related future directions. Summarizing the work conducted in WP1, the first deliverable D1.1 [1] focused on the identification and prioritization of the most promising use cases – from both the technical and economical point of view – which were used to motivate the need of a virtualized, integrated Edge and Fog platform exploiting multiple RATs. Both the functional and non-functional requirements needed for the design of this Edge and Fog platform were also listed in D1.1, so to derive a baseline system architecture being able to fulfil the requirements of the considered use cases. The architecture, composed by two main building blocks, i.e. Edge and Fog computing System (EFS) and Orchestration and Control System (OCS), was then revised and updated in deliverable D1.2 [2] following the feedback from the work conducted in Project's technical Work Packages, i.e. WP2 (for the EFS) and WP3 (for the OCS), also considering the activities related to demonstrations and PoCs in the scope of WP4. Along with the revised architecture, the deliverable D.12 was mainly focused on the analysis of the business perspectives of the 5G-CORAL technical solution and, in that context, the concept of federation among different administrative domains was introduced as a mean to allow for the different players of the 5G-CORAL ecosystem to create business relationships aiming at extending or creating new end-to-end services by importing external features or resources (being used for other services) from already deployed Edge and Fog systems, without the need of deploying dedicated infrastructure or building new complex architectures.

Departing from the above findings, this deliverable D1.3 first proposed in Section 2 a list of the business requirements for each of the key players in the 5G-CORAL ecosystem in order to complete the analysis on the need of resource federation mechanisms in 5G-CORAL, which was initially drafted in deliverable D1.2. The State-of-the-Art of both the performance metrics and technical features considered in each of the 5G-CORAL use cases is then reported in Section 0, which can be used as a baseline for the assessment of performance improvements and/or system enhancements achieved when adopting the 5G-CORAL platform. The system-level refinement of the architecture, detailing how each use case makes use of multiple RATs and how the EFS and the OCS interact with each other, was in the scope of Section 4. In addition, the capability of the OCS to manage heterogeneous virtualization platforms has been assessed for each of the 5G-CORAL use cases. Section 0 discussed the procedures and techniques for the incremental deployment of the 5G-CORAL solution into existing networks by taking the viewpoints of both the telecom operator and the Vertical company into account. Finally, future directions of the 5G-CORAL solution when applied to the Project's use cases were reported in Section 5, along with the possibility to adopt the 5G-CORAL solution for other promising use cases, in particular the On-device AI and ML as well as private wireless networks.

For what concerns future works on the Edge and Fog platform as developed in 5G-CORAL, from the system-level and business-oriented viewpoint of WP1 it can be stated that the platform itself is already well positioned in each tier of a distributed computing infrastructure, i.e. Cloud, Edge and Fog. The industry perfectly understood that the Edge and Fog are needed along with the Cloud: their inclusion and, more important, their integration will be a big part of Companies' IT strategies in 2019, with a market size being expected to value at $3.24 billion (USD) by 2025 [91]. This is driven by the following key aspects [91][92], with respect to which the 5G-CORAL technical solution is fully compliant:

- *Both Edge and Fog computing are likely to work in tandem with the Cloud, not completely replace it* – Cloud will still play an important role, e.g., for the Connected Cars scenario, vehicles might send the data they have collected to the Cloud at the end of a day of driving. This will allow for cars' manufacturers to use the data from cars in order to train and refine their software, so vehicles will be able to offer smoother and safer rides. For other use cases, e.g. the ones within the Shopping Mall scenario, devices at the Edge and Fog tiers are in charge of the initial and/or immediate processing, but they will rely on Cloud intelligence to improve how they work.

- *Reduction of bandwidth budgets* – the adoption of Cloud solutions has led to a bandwidth problem due to the need to send and receive applications' data from the Cloud. By exploiting the pervasiveness and the computing capabilities of an Edge and Fog platform, both data aggregation and processing happen locally, meaning that a large portion of the data does not need to be sent to the Cloud, only the processed results do. Therefore, customers can lower their bandwidth costs and the corresponding savings might be used for placing more compute and storage resources at the Edge and Fog tiers.
- *Businesses are continuously increasing the amount of generated IoT data generated year-over-year* – it has been estimated that, at the end of 2018, there were over 23 billion connected IoT devices across the Globe [91]. Therefore, it is likely that this trend will continue also in the incoming years, with organizations moving toward the IoT space to increase and improve the intelligence they can offer and the automation they can achieve. This will be done not only via a unique IoT-specific RAT but potentially all the available RATs for IoT – NB-IoT, LTE-M, Sigfox, LoRa, etc. – could be considered equally: in this sense, the Multi-RAT IoT use case is an example of how multiple IoT technologies might be integrated within a single platform which can be used by different companies based on their needs.
- *Microservices will facilitate Edge-like developments* – with the incoming 5G systems and the introduction of the Service Based Architecture, it will be easier to push rules, processes and policies into discrete packages that can be aligned with various business needs, easing the ability to scale, deploy new services quickly and integrate with other platforms. All these benefits are foreseen also for the 5G-CORAL platform, which has been designed with a microservices approach, hence making it possible to work properly within the 5G systems.
- *Artificial Intelligence will continue to drive business development* – Artificial Intelligence (AI) is often dependent on the Cloud, as it requires massive computing power. But with the need to develop AI-based services characterized by low latency and high reliability, AI needs to be hosted far away from the Cloud, i.e. in the Edge and Fog, giving it capabilities to learn and deliver a highly effective result in real-time. It has been discussed how the 5G-CORAL platform is able to offer AI-based services as part of its future directions, and these services might be different based on the considered player in the ecosystem. For instance, when considering the telecom operator, the networking resource information provided by the EFS Service Platform would enable AI-based functions relevant for operating a mobile network, such as AI-based Self Organizing Network (SON) and AI-based cognitive spectrum sharing. For the latter, in particular, the multi-RAT convergence envisioned in 5G-CORAL is a crucial requirement, where a device detects and avoids the active resource occupancy through a different RAT.

7 References

[1] "*D1.1 – 5G-CORAL initial system design, use cases, and requirements*". [Online]. Available: http://5g-coral.eu/wp-content/uploads/2018/04/D1.1_final7760.pdf.

[2] "*D1.2 – 5G-CORAL business perspectives*". [Online]. Available: http://5g-coral.eu/wp-content/uploads/2018/09/D1.2-final12828.pdf.

[3] "*D2.1 – Initial design of 5G-CORAL edge and fog computing system*". [Online]. Available: http://5g-coral.eu/wp-content/uploads/2018/06/D2.19803.pdf.

[4] "*D2.2 – Refined design of 5G-CORAL edge and fog computing system and future directions*". [Online]. Available: http://5g-coral.eu/?page_id=37.

[5] "*D3.1 – Initial design of 5G-CORAL orchestration and control system*". [Online]. Available: http://5g-coral.eu/wp-content/uploads/2018/06/D3.19802.pdf.

[6] "*D3.2 – Refined design of 5G-CORAL orchestration and control system and future directions*". [Online]. Available: http://5g-coral.eu/?page_id=37.

[7] "*D4.2 - 5G-CORAL proof of concept and future directions*". [Online]. Available: http://5g-coral.eu/?page_id=37.

[8] "*Innovating the Telco Business Model – Drivers and Emerging Trends*", Capgemini, 2011. [Online]. Available: https://www.capgemini.com/wp-content/uploads/2017/07/Innovating_the_Telco_Business_Model_Drivers_and_Emerging_Trends.pdf.

[9] "*ABB and Deutsche Telekom to develop smart grid solutions*", ABB Press Release, April 2010. [Online]. Available: https://new.abb.com/news/detail/13375/abb-and-deutsche-telekom-to-develop-smart-grid-solutions/.

[10] "*AT&T wireless scraps flat-rate Internet plan*", C. Kang (The Washington Post), June 2010. [Online]. Available: http://voices.washingtonpost.com/posttech/2010/06/att_wireless_scraps_flat-rate.html.

[11] "*Internet Policy's Next Frontier: Usage-Based Broadband Pricing*", D. A. Lyons (Boston College Law School), April 2013. [Online]. Available: http://www.fclj.org/wp-content/uploads/2014/01/66-1-Lyons.pdf.

[12] "*Farmers: dig your own trench for NBN access*", IT News, May 2009. [Online]. Available: https://www.itnews.com.au/news/farmers-dig-your-own-trench-for-nbn-access-145126.

[13] *Blyk* Mobile Virtual Network Operator. [Online]. http://www.blyk.com/.

[14] "*Customer Success Stories – How Telecom Operators worldwide embrace success with Comarch BSS/OSS*", Comarch, 2019. [Online]. Available: https://www.comarch.com/files-com/file_6/Comarch-Customer-Success-Book.pdf.

[15] "*Sprint inks $5B network outsourcing deal with Ericsson*", M. Dano (FierceWireless), July 2009. [Online]. Available: https://www.fiercewireless.com/wireless/sprint-inks-5b-network-outsourcing-deal-ericsson.

[16] "*Connect to Planet*", AT&T, August 2018. [Online]. Available: https://about.att.com/ecms/dam/csr/issuebriefs/IssueBriefs2018/environment/energy-management.pdf.

[17] PoC 12 MEC enabled Over-The-Top business – ETSI https://mecwiki.etsi.org/index.php?title=PoC_12_MEC_enabled_Over-The-Top_business.

[18] E. Datsika et al, "*Software Defined Network Service Chaining for OTT Service Providers in 5G Networks*", IEEE Communications Magazine, Vol. 55, pp. 124-131, November 2017. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8114562.

[19] Switch Abstraction Interface (SAI), https://github.com/opencomputeproject/SAI.

[20] ARCore, Google, https://developers.google.com/ar/.

[21] "*Recognize and Augment Images*", Google, August 2018. [Online]. Available: https://developers.google.com/ar/develop/c/augmented-images/.

[22] W. Zhang et al., "*CloudAR: A Cloud-based Framework for Mobile Augmented Reality*", Proceedings of the Thematic Workshops of ACM Multimedia 2017, co-located with MM 2017, pp. 194-200, October 2017. [Online]. Available: https://arxiv.org/abs/1805.03060.

[23] T. Skeie et al., "*Ethernet in substation automation*", IEEE Control Systems Magazine, Vol. 22, pp. 43-51, August 2002. [Online]. Available: https://www.researchgate.net/publication/3206774_ETHERNET_in_substation_automation.

[24] Z. Huang et al., "*CloudRidAR: A Cloud-based Architecture for Mobile Augmented Reality*", Proceedings of the 2014 Workshop for Mobile Augmented Reality and Robotic Technology-Based Systems, pp. 29-34, 2014. [Online]. Available: http://www.cse.ust.hk/~panhui/papers/mars2014_cloudridar.pdf.

[25] G. Mohanarajah et al., "*Rapyuta: A cloud robotics platform*", IEEE Transactions on Automation Science and Engineering, Vol. 12, pp. 481-493, July 2014. [Online]. Available: http://roboearth.ethz.ch/wp-content/uploads/2013/03/gajanIROS-Tokyo.pdf.

[26] L. Agostinho et al., "*A cloud computing environment for supporting networked robotics applications*", IEEE 9th International Conference on Dependable, Autonomic and Secure Computing (DASC), December 2011.

[27] H. J. Hong et al., "*Dynamic module deployment in a fog computing platform*", IEEE 18th Asia-Pacific Network Operations and Management Symposium (APNOMS), October 2016.

[28] S. V. N. Kotikalapudi, "*Comparing Live Migration between Linux Containers and Kernel Virtual Machine – Investigation study in terms of parameters*", Master of Science in Computer Science at Blekinge Institute of Technology, February 2017. [Online]. Available: http://www.diva-portal.org/smash/get/diva2:1085809/FULLTEXT02.

[29] A. A. Tabassam et al., "*Fast and seamless handover for secure mobile industrial applications with 802.11r*", IEEE 34th Conference on Local Computer Networks, October 2009.

[30] A. Machen et al., "*Live Service Migration in Mobile Edge Clouds*", IEEE Wireless Communications, Vol. 25, pp. 140-147, August 2017. [Online]. Available: https://core.ac.uk/download/pdf/156963730.pdf.

[31] M. Hosseini et al., "*Adaptive 360 VR video streaming: Divide and conquer*", IEEE International Symposium on Multimedia (ISM), December 2016. [Online]. Available: https://arxiv.org/pdf/1609.08729.

[32] T. Kämäräinen et al., "*Dissecting the end-to-end latency of interactive mobile video applications*", Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications, February 2017, pp. 61-66. [Online]. Available: https://arxiv.org/pdf/1611.08520.

[33] O. Liberg et al. "*Cellular Internet of Things: Technologies, Standards, and Performance*", Elsevier Science Publishing Co Inc, EAN: 9780128124581, 2017.

[34] Amazon Web Services Network Test, http://cloudharmony.com/speedtest-latency-for-aws-limit-1.
[35] Amazon Web Services Network Test, http://cloudharmony.com/speedtest-downlink-uplink-latency-for-aws-limit-20s.
[36] 3GPP TS 22.185 "*Service requirements for V2X services*", Rel-15. [Online]. Available: http://www.3gpp.org/ftp//Specs/archive/22_series/22.185/22185-f00.zip.
[37] 3GPP TR 22.885 "*Study on LTE support for Vehicle-to-Everything (V2X) services*", Rel-14. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/22_series/22.885/22885-e00.zip.
[38] 3GPP TS 32.203 "*Policy and charging control architecture*", Rel-16. [Online]. Available: http://www.3gpp.org/ftp//Specs/archive/23_series/23.203/23203-g00.zip.
[39] I. Widjaja et al, "*Comparison of MME Signaling Loads for Long-Term-Evolution Architectures*", IEEE 70th Vehicular Technology Conference Fall, September 2009.
[40] H. T. Larasati et al., "*Performance evaluation of handover association mechanisms in SDN-based wireless network*", 3rd International Conference on Wireless and Telematics (ICWT), July 2017.
[41] F. Laassiri et al., "*An Improvement of Performance Handover in Worldwide Interoperability for Microwave Access using Software Defined Network*", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 9, January 2018. [Online]. Available: https://thesai.org/Downloads/Volume9No12/Paper_13-An_Improvement_of_Performance_Handover.pdf.
[42] M. R. Hines et al., "*Post-Copy Live Migration of Virtual Machines*," ACM SIGOPS Operating Systems Review, Vol. 43, pp. 14-26, July 2009. [Online]. Available: https://kartikgopalan.github.io/publications/hines09postcopy_osr.pdf.
[43] V. Medina et al., "*A Survey of Migration Mechanisms of Virtual Machines*", ACM Computing Surveys (CSUR), Vol. 46, p. 30, January 2014. [Online]. Available: https://romisatriawahono.net/lecture/rm/survey/computer%20engineering/Medina%20-%20Migration%20Mechanisms%20of%20Virtual%20Machines%20-%202014.pdf.
[44] J. Nider et al., "*Cross-ISA Container Migration*", Proceedings of the 9th ACM International on Systems and Storage Conference, p. 24, June 2016.
[45] R. Morabito et al., "*Evaluating Performance of Containerized IoT Services for Clustered Devices at the Network Edge*", IEEE Internet of Things Journal, Vol. 4, pp. 1019-1030, August 2017.
[46] European Telecommunications Standards Institute (ETSI) GS MEC 003, "*Multi-access Edge Computing (MEC); Framework and Reference Architecture*", January 2019. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/02.01.01_60/gs_MEC003v020101p.pdf.
[47] European Telecommunications Standards Institute (ETSI) GS MEC 009, "*Multi-access Edge Computing (MEC); General principles for MEC Service APIs*", January 2019. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/009/02.01.01_60/gs_MEC009v020101p.pdf.
[48] European Telecommunications Standards Institute (ETSI) GS MEC 011, "*Mobile Edge Computing (MEC); Mobile Edge Platform Application Enablement*", July 2017. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/011/01.01.01_60/gs_MEC011v010101p.pdf.
[49] Kernel Virtual Machine (KVM), https://www.linux-kvm.org.

[50] Infrastructure for container projects, https://linuxcontainers.org/#LXD.
[51] Enterprise Container Platform for High-Velocity Innovation – The Docker Enterprise Difference, https://www.docker.com.
[52] NVIDIA Virtual GPU Technology – Accelerate Virtual Desktops with NVIDIA vGPUs, https://www.nvidia.com/en-us/data-center/virtual-gpu-technology/.
[53] NextEPC Open Source, https://nextepc.org/.
[54] "*5G Network & Service Strategies Survey Report*", Heavy Reading, February 2019. [Online]. Available: https://www.lightreading.com/lg_redirect.asp?piddl_lgid_docid=749589.
[55] MobiledgeX, https://mobiledgex.com/.
[56] "*Does Amazon's purchase of Whole Foods redefine "Edge Compuing"?*", D. Bubley (Disruptive Analysis), July 2017. [Online]. Available: https://www.netmanias.com/en/post/blog/12526/5g-iot-mec/does-amazon-s-purchase-of-whole-foods-redefine-edge-computing.
[57] "*Vested Interests Drive Divergence in Edge Computing*", Tolaga Research, June 2018. [Online]. Available: http://www.tolaga.com/home/tolagaDocument?doc=edgeComp2018.
[58] "*MEC Deployments in 4G and Evolution Towards 5G*", ETSI, White Paper No. 24, February 2018. [Online]. Available: https://www.etsi.org/images/files/etsiwhitepapers/etsi_wp24_mec_deployment_in_4g_5g_final.pdf.
[59] "*MEC in 5G networks*", ETSI, White Paper No. 28, June 2018. [Online]. Available: https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp28_mec_in_5G_FINAL.pdf.
[60] "*Global status report on road safety 2018*", World Health Organization, December 2018. [Online]. Available: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.
[61] R. Lancelot, "*Connected Cars: High Tech, Low Demand*", November 2018. [Online]. Available: https://www.strategyanalytics.com/strategy-analytics/blogs/automotive/infotainment-telematics/infotainment-telematics/2018/11/08/connected-cars-high-tedh-low-demand.
[62] B. Jonston, "*Connected Car Market to Grow 270% by 2022 GM Bests BMW, Audi & Mercedes-Benz*", April 2018. [Online]. Available: http://www.autoconnectedcar.com/2018/04/connected-car-market-to-grow-270-by-2022-gm-bests-bmw-audi-mercedes-benz/.
[63] S. Moritz, "*Telecom Giants Fear Missing the Money as Cars Go Online*", April 2018. [Online]. Available: https://www.bloomberg.com/news/articles/2018-04-10/can-telecom-giants-find-ways-to-make-real-money-on-smarter-cars.
[64] Frost & Sullivan (Research and Markets – The World's Largest Market Research Store), "*Opportunity Analysis of the OBD II Automotive Aftermarket – Forecast to 2023*", December 2017.
[65] Allstate Newsroom, "*Allstate Survey Reveals Two-Thirds of Parents Think Their Teen Drivers are Safe Behind the Wheel – Despite Concerns About Distracted Driving*", October 2018. [Online]. Available: https://www.allstatenewsroom.com/news/allstate-survey-reveals-two-thirds-of-parents-think-their-teen-drivers-are-safe-behind-the-wheel-despite-concerns-about-distracted-driving/.
[66] AFP (The Guardian), "*SNCF to launch driverless trains in mainland France by 2023*", September 2018. [Online]. Available:

https://www.theguardian.com/world/2018/sep/12/sncf-to-launch-driverless-trains-in-mainland-france-by-2023.

[67] Maximize Market Research, "*Global artificial intelligence as a service market : Industry Analysis and Forecast (2017-2024) – by Software Tool, by Service Type, by Technology, by Organization Size, by Vertical and by Geography*", 2019. [Online]. Available : https://www.maximizemarketresearch.com/market-report/global-artificial-intelligence-service-market/5754/.

[68] Nanalyze, "*Artificial Intelligence (AI) as a Service and "Core AI"*", October 2017. [Online]. Available: https://www.nanalyze.com/2017/05/artificial-intelligence-service-core-ai/.

[69] GTI (China Mobile, Softbank, CloudMinds, Huawei, Skymind), "*GTI Enabling AI Vision with 5G Networks*", White Paper, version V1, February 2018. [Online]. Available: http://www.gtigroup.org/d/file/Resources/rep/2018-02-22/a97b3970bb0111fa348e833074251de8.pdf.

[70] Wireless World Research Forum (WWRF), "*Wireless Big Data of Smart 5G*", White Paper 2, November 2017. [Online]. Available: https://www.wwrf.ch/files/wwrf/content/files/publications/outlook/White%20Paper%202-%20Wireless%20Big%20Data%20of%20Smart%205G.pdf.

[71] ITU-T Focus Group on Machine Learning for Future Networks including 5G (ML5G FG), https://www.itu.int/en/ITU-T/focusgroups/ml5g/Pages/default.aspx.

[72] ITU-T Y.3172 (formerly Y.IMT2020-ML-Arch), "*Architectural framework for machine learning in future networks including IMT-2020*", March 2019. [Online]. Available: https://www.itu.int/itu-t/workprog/wp_item.aspx?isn=15020.

[73] Pai A. (FCC Chairman), "*Remarks Of FCC Chairman Ajit Pai At FCC Forum On Artificial Intelligence And Machine Learning, Washington, DC*", November 2018. [Online]. Available: https://www.fcc.gov/document/chairman-pai-remarks-fccs-artificial-intelligence-forum.

[74] R. Mennes et al., "*Deep Learning-based Spectrum Prediction Collision Avoidance For Hybrid Wireless Environments*", IEEE Access Vol.7, pp. 45818 – 45830, April 2019. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8684944.

[75] AWS IoT Greengrass ML inference, https://aws.amazon.com/greengrass/ml/.

[76] Samarakoon S. et al., "*Distributed Federated Learning for Ultra-Reliable Low-Latency Vehicular Communications*", Globecom 2018, Abu Dhabi (UAE), August 2018. [Online]. Available: http://www.laneas.com/sites/default/files/publications/4223/camera_ready_sumudu.pdf.

[77] Qualcomm Technologies Inc., "*Making on-device AI ubiquitous*", March 2018. [Online]. Available: https://www.qualcomm.com/media/documents/files/making-on-device-ai-ubiquitous.pdf.

[78] ST Microelectronics – Innovation and Technology, "*Artificial Intelligence*". [Online]. Available: https://www.st.com/content/st_com/en/about/innovation---technology/artificial-intelligence.html.

[79] Venture Beat (VB), "*The rise of on-device AI – and why it's so important (VB Live)*", March 2019. [Online]. Available: https://venturebeat.com/2019/03/21/the-rise-of-on-device-ai-and-why-its-so-important-vb-live/.

[80] "*The Private LTE Opportunity for Industrial and Commercial IoT*", Harbor Research, July 2017. [Online]. Available: https://www.multefire.org/wp-content/uploads/HRI_Paper_Private-LTE-Network-Paper_20-July-2017_Final.pdf.

[81] "*The Market Opportunity for Private IoT Leveraging LTE/MulteFire*", Harbor Research, February 2017. [Online]. Available: http://harborresearch.com/wp-content/uploads/sites/8/2017/03/HRI_MutleFire-Alliance_February-2017.pdf.

[82] "*Private LTE: Making the Enterprise future-ready*", RCR Wireless, July 2018. [Online Webinar]. Available: http://content.rcrwireless.com/private_lte_webinar.

[83] "*Network Evolution in Unlicensed and Shared Spectrum*", ABI Research, 2017.

[84] MulteFire Alliance, https://www.multefire.org/our-members/.

[85] CBRS Alliance, https://www.cbrsalliance.org/about-us/.

[86] "*Understanding MulteFire for industrial IoT*", S. Kinney, May 2018. [Online]. Available: https://www.rcrwireless.com/20180529/network-infrastructure/lte/multefire-industrial-iot-tag17-tag99.

[87] "*GE, Nokia and Qualcomm Unveil First Private LTE-based Trial Network Customized for Industrial IoT*", Nokia Press Release, February 2017. [Online]. Available: https://www.nokia.com/about-us/news/releases/2017/02/22/ge-nokia-and-qualcomm-unveil-first-private-lte-based-trial-network-customized-for-industrial-iot/.

[88] 3GPP TS 38.300 "*NR; Overall description; Stage-2*", Rel-15. [Online]. Available: http://www.3gpp.org/ftp//Specs/archive/38_series/38.300/38300-f40.zip.

[89] 3GPP TS 23.501 "*System architecture for the 5G System (5GS)*", Rel-15. [Online]. Available: http://www.3gpp.org/ftp//Specs/archive/23_series/23.501/23501-f40.zip.

[90] "*4G & 5G-Ready Network Slicing*", Heavy Reading, August 2017. [Online]. Available: https://affirmednetworks.com/wp-content/uploads/2017/12/Affirmed-Networks-Network-Slicing-Whitepaper.pdf.

[91] F. Cittadino (Forbes Technology Council), "*Four Predictions For Edge Computing In 2019*", January 2019. [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2019/01/17/four-predictions-for-edge-computing-in-2019/#798d12004f06.

[92] M. Weinberger (Business Insider), "*One of the biggest trends in tech over the last decade has been the shift to cloud computing – and we're seeing the first signs of what might be next*", February 2018. [Online]. Available: https://www.businessinsider.com/edge-computing-and-fog-computing-explained-2018-2?IR=T.