IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 10, 2019, accepted July 18, 2019, date of publication July 24, 2019, date of current version August 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2930936

Analysis of the Impact of the Evolution Toward 5G Architectures on Mission Critical Push-to-Talk Services

AITOR SANCHOYERTO, RUBEN SOLOZABAL, BEGO BLANCO, AND FIDEL LIBERAL
Networking Quality and Security, University of the Basque Country, 48013 Bilbao, Spain
Corresponding author: Aitor Sanchoyerto (aitor.sanchoyerto@ehu.eus)

ABSTRACT This paper analyzes the impact of the evolution of mobile broadband networks from 4G architectures toward 5G on mission critical push-to-talk (MCPTT) key performance indicators (KPIs). This paper focuses on how the deployment of MCPTT over these architectures affects the service quality. We carry out a comprehensive analysis of the call flows and the contribution of each network segment to the service KPIs defined by the Third Generation Partnership Project (3GPP). A long-term evolution (LTE) architecture will be considered as the baseline scenario; from there on, different core network proposals are evaluated: the LTE multi-access edge computing (MEC) and 5G. We analyze how these strategies for data and control plane distribution affect the service, identifying current performance bottlenecks and exploring latency reduction techniques. Throughout this paper, we show that the current implementation of the MCPTT service according to the standards defined by the 3GPP remains a key challenge in terms of the KPI compliance. Still, the evolution toward 5G architectures, particularly those leveraging the MEC, further improves the KPIs of the MCPTT service, especially the KPI 3 (mouth-to-ear latency) since the services are deployed closer to the end users. These results strengthen the commitment to mobile broadband networks for the deployment of mission critical communications.

INDEX TERMS Public safety comunications, MCPTT, 5G, MEC, CUPS, NFV.

I. INTRODUCTION

In the last years, the limitations of legacy Mission-Critical (MC) Public-Safety Communications have opened the discussion about considering commercial mobile broadband standards to support MC applications. From the very beginning, LTE attracted the attention of the research community as the reference technology for mission-critical communications. However, LTE was not designed to comply with reliability, confidentiality and security standards required in these Mission Critical Services (MCS). For that reason, and considering the increasing weight of the PS sector, the 3GPP has encouraged the evolution of LTE specifications to address these requirements. As a result, MCPTT is the first in a series of MCS standards pushed by the 3GPP's SA6 working group. The standardization of MCPTT over LTE began in Release 13, and new functionality has been added in Releases 14 and 15 [1].

The MCPTT service supports the communication between pairs of users (private calls) and multiple users (group calls), where each user has permission to speak (transmit voice / audio) in specific time slots. The system allows participants to request the token to speak, traditionally by pressing a button on the user's device. It also provides a priority mechanism to arbitrate the participation of the users in the communication. When multiple users request the token simultaneously, the choice of which user request is accepted, and which requests are rejected or queued is based on a number of features (including the priorities of the users in dispute). Additionally, the MCPTT service provides a method for a user with a higher priority (for example, in an emergency situation) to overrule the current speaker. Finally, this service also implements a mechanism to limit the time a user speaks. In this way, when a user in a call runs out of time, the users of the same or lower priority have the opportunity to win

The associate editor coordinating the review of this manuscript and approving it for publication was Zeeshan Kaleem.

the authorization to talk. Additional features of an MCPTT service include the monitoring of the active group calls a user participates in, late call entry and the identification of the current speaker in each of the active group calls.

MCPTT arises from classical Push-to-talk Over Cellular (POC) services, but presenting an integration with commercial mobile phone networks that opens up new possibilities. In order to enable this coexistence, the mobile network should be assisted with key features such as prioritized and preemptive access and Quality of Service (QoS) guarantees for media flows. For this purpose, 3GPP created service specific Quality Class Indicators (QCI) that the LTE network has to assume to host MC services. Currently, network operators are evolving their infrastructure to support these QCIs signalled over the Policy and Charging Control (PCC) interfaces. The support of QCIs enables operators to offer MC services sharing the network infrastructure among mission critical and commercial users. In addition to QCI compliance, MCPTT also has to satisfy its own KPIs, mainly related to latency. The fulfillment of these KPIs depends heavily on the network architecture as well as on the service implementation.

In conclusion, legacy Private Mobile Radio (PMR)/Land Mobile Radio (LMR) narrowband networks rely on standards that were developed 25 years ago. By contrast, commercial networks have evolved towards broadband. Current public safety users have new demands, such as video and data transmission, making the transition to broadband ineluctable. However, commercial networks were not designed to cope with the tight requirements of public safety mission critical services. In spite of this, taking LTE as a reference, the latencies measured presently remarkably improve those of the previous network generations, getting closer to the requirements of MC communications. Thus, the main contribution of this work is the assessment of the suitability of current LTE deployments to provide mission critical services, in particular, MCPTT. To this end, we perform a comprehensive analysis of the components involved in the call setup and call latency times, as the most relevant KPIs of MCPTT. We start with an assessment of the latencies involved in current commercial real-world LTE networks to prove that MCPTT over LTE is presently a challenging option that barely achieves the required KPIs in favourable conditions, but cannot guarantee the required latencies in heavily loaded scenarios. Then, we further analyze the impact on such KPIs when evolving from an LTE baseline scenario to LTE MEC and eventually 5G architectures. We demonstrate that the fullfilment of KPIs will even improve in the near future, making the network sharing model a promising candidate for future broadband MC communications.

With this aim, this paper is organized as follows: next section provides a brief overview on the current state of network performance measurements towards an LTE mobile network. Then, Section III describes the system model, including the measurement methodology. Next, Section IV describes the baseline LTE scenario, Section V the MEC alternatives and Section VI the reference non-standalone 5G architecture used in the experiment. Then, the details of the experiments are presented in VII. Finally, Section VIII presents the discussion of the obtained results and Section IX summarizes the main conclusions of the article.

II. RELATED WORK

The Public Safety Agencies (PSA) have detected the need to update their traditional narrowband networks (TETRA, TETRAPOL, P25) to broadband networks in order to support the services demanded by the emergency services nowadays [2], [3]. These services impose higher bandwidth requirements for data communication than the currently granted bandwidth for PMR networks. [4] In this context, the migration of legacy narrowband networks to LTE is a promising solution. As a matter of fact, the Electronic Communication Committee (ECC) [5] has proposed a roadmap up to 2025 for the transition to broadband communication in mission critical systems. However, the existing literature shows relatively limited studies that present a comprehensive survey and/or comparison on PMR and LTE systems. Among the few available works, [6] presents a discussion on the performance of voice over LTE as an important aspect of PSC. [7] and [8] explore the ability of LTE to meet the requirements of the Public Safety Network (PSN) and possible future improvements to LTE in order to provide PSC.

It must be noted that PMR networks are specifically designed to serve PS users, who often work in groups, and therefore inherently implement specialized services such as group calls. These networks are geared to low data rate voice communications, maximizing network capacity in terms of users [9]. Conversely, LTE networks are mainly designed for one-to-one communication, with large payloads of voice and video communications. In order to fill in this gap and provide broadband capabilities to public safety services, the 3GPP is developing new standards that will enable LTE to provide mission critical capabilities. As a result, MCPTT arises as the first specification in a series of mission critical services [1]. In [10] we find an exhaustive analysis of legacy and emerging technologies for PS communications that briefly comments the features of mission-critical push-to-talk over LTE. However, this study does not go into performance details of the service. In this regard, the MCPTT standard also defines a set of KPIs [11], mainly oriented to latency. In contrast, the use of KPIs to assess the network performance and the Quality of Service (QoS) in PMR networks is not very common. The work in [12] introduces a set of measurable KPIs considered necessary in order to allow TETRA operators to be aware of the whether provided services meet the QoS requirements.In any case, most of the experimental results about performance are based on simulations, not on real-world deployments.

On a separate issue, other studies consider the importance of coverage in this discussion. Traditionally, legacy PS networks have optimized their infrastructure to the geographical area where their users are located. Outside that zone, the coverage is bad or nonexistent. [13] indicates the possibility of using the already deployed commercial broadband network for Public Safety agencies when the radioelectric coverage has ‘dark zones’ (without radioelectric coverage). Likewise, [14] indicates that Mobile Personal Cell (mPC) can be utilized for PS to provide the network in the disaster area. However, these considerations about coverage are beyond the scope of this work.

One key element of the deployment of an end-to-end MCPPT service is the Mobile Broadband (MBB) network that supports the service. As aforementioned, the reference standard nowadays is LTE. Fortunately, the analysis of MBB network performance is one of the main points of study of the networking research community and has significantly contributed to the feasibility assessment of MC services over commercial networks. There are three main approaches for measuring the performance of MBB networks: (i) crowd-sourced results from a large number of MBB users [15], [16], (ii) measurements based on network-side data such as [12], [17], [18], and (iii) measurements collected using a dedicated infrastructure [19]–[21]. Part of the experimental results shown later in this work is extracted from the MONROE project infrastructure. [22], [23] MONROE is a EU project that provides a transnational open measurement platform for performance evaluation over a cellular access network. It has a dedicated infrastructure with fixed and mobile nodes distributed over Norway, Sweden, Spain, Italy and Greece. The platform provides a flexible infrastructure to launch performance-related measurements and an accurate assessment of a variety of features in 4G cellular networks. MONROE has been used to perform latency measurements in the real-world over different commercial operators in different countries and network conditions.

There are several studies that analyze the latency and Quality of Service (QoS) on LTE networks using the MONROE platform. Among them, [24] is specially relevant for our study, since the authors identify the latency contribution of the components of the overall LTE architecture and estimate the lower achievable latency limits. Based on these results, this paper presents the values of the KPIs that have been obtained when deploying the service on different network architectures: (IV) Baseline LTE scenario, (V) MEC Deployment in 4G (A Distributed Evolved Packets Core (EPC), B Distributed Serving and PDN Gateway (S/PGW) and (VI) Deployment in non-standalone 5G.

In addition to the nodes provided by MONROE, we have also used another server in France to deploy the MCPTT ASs. After measuring the current path latencies in a wide variety of scenarios the conclusion is that, in a realistic situation, over the 90% of the communications experienced end-to-end latencies close to 50ms to reach a centralized MCPTT server located behind a conventional EPC core (Round Trip Time). On the other hand, the latency experienced between the end-user and the PGW (also in the 90% of the tests) has been around 30ms. Therefore, we have considered a RAN uplink delay and a RAN downlink delay of 30ms for the calculation of the KPIs.

TABLE 1. Threshold and QCIs defined for MCPTT KPIs (TS 23.203).

MCPTT KPIs	Threshold	Likelihood	QCI involved	LTE packet Delay Budget
KPI 1	<300 ms	95%	QCI 69	<60 ms
KPI 2	<1000 ms	N/A	QCI 69	<60 ms
KPI 3	<300 ms	95%	QCI 65	<75 ms

The current related literature shows no study where the KPIs of an MCPTT service are measured. The closest research work is [25], which accomplishes a delay analysis of Push-to-talk Over Cellular (POC) service over LTE networks. As previously commented, PoC is considered the precursor of the MCPTT standard. Functionally and conceptually they are very similar, yet MCPTT needs also to meet the strict requirements demanded for MC services. However, it is an important reference for our analysis of the impact of the network architecture on the performance of an MCPTT service.

III. SYSTEM MODEL

The service requirements for MCPTT defined in 3GPP TS 22.179 [11] include the following performance indicators:

- **KPI 1 - Access Time:** It is defined as the time between the moment an MCPTT user requests to speak (by pressing the MCPTT button on the MCPTT User Equipment (UE)) and the moment this user receives the token granted for the MCPTT Server. This time does not include confirmations from receiving users or the time to affiliate to the group.
- **KPI 2 - End-to-End Access time:** It is defined as the time between the moment an MCPTT user requests to speak and the moment this user gets a signal to start speaking, including MCPTT call establishment (if applicable) and the acknowledgement from the first receiving user before voice can be transmitted.
- **KPI 3 - The Mouth-to-ear Latency:** It is the time between an utterance by the transmitting user, and the playback of the utterance at the receiving user’s speaker.

MCPTT KPIs have a set of strict performance requirements that are summarized in Table 1. As it is shown, each KPI has a time threshold and a likelihood ratio that must be met. The fulfilment of these requirements depends on the network architecture as well as on the quality of the service implementation. The objective in this work is to analyze the impact on the KPIs that an MCPTT service presents when deployed along the following network architectures. We will start measuring the KPIs obtained in a traditional LTE network. The traditional LTE network serves as a baseline scenario to later address the introduction of MEC strategies and, finally, a standalone 5G network.

As introduced before, in order to provide MC services network operators must support a new range of QCIs defined for these services. Usually, during normal operation, the network

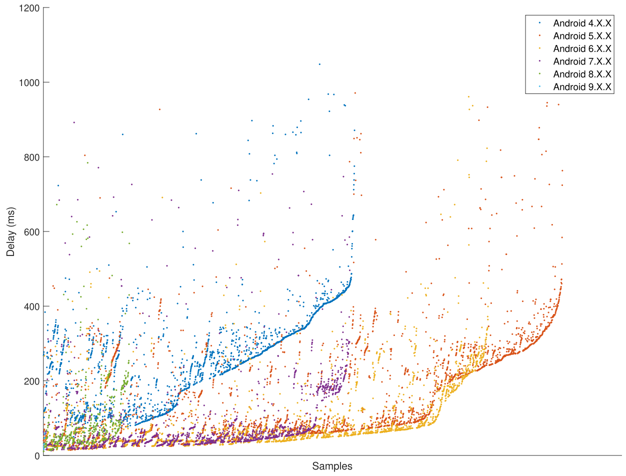


FIGURE 1. Mouth-to-ear latency calculated in an Android device.

conditions are much more favorable than the QCIs guarantee. However, this QCIs become crucial in case of congestion for the service continuity. For this reason, we measure the performance of the service int the following scenarios:

- **Average network load:** represents LTE under normal circumstances. Measured latency is considerably lower than the packet delay budget granted.
- **The worst-case scenario of network load:** MC specific bearer traffic is scheduled with an appropriate Quality of Service (QoS). According to standard 3GPP LTE QoS Classes, the tenant will assign a QCI = 69 for signaling plane bearers and a QCI = 65 for user plane bearers. This QCI guarantees a packet delay budget of 60 ms for the control plane and 75 ms for the user plane.

Another aspect of great significance is the implementation on the user equipment. The operating system of the user's device has a big impact in the service performance. The tests executed for this purpose have been carried out using different versions of Android operating system. As Figure 1 shows, latest releases of Android have a major impact in UE's audio latency reduction. The presented results have been obtained from [26], and suggest that close collaboration with mobile phone manufacturers is key to obtain low latency implementations. A more direct access to the hardware by the audio functions of the operating system should reduce processing times and decrease the latency of the phone's audio production.

In order to analyze which are the main contributions of the different network components and segments into the KPI, a comprehensive analysis of the communication flows needs to be carried out. Figure 2 shows the flow diagram of a private call with automatic commencement, as an example of the

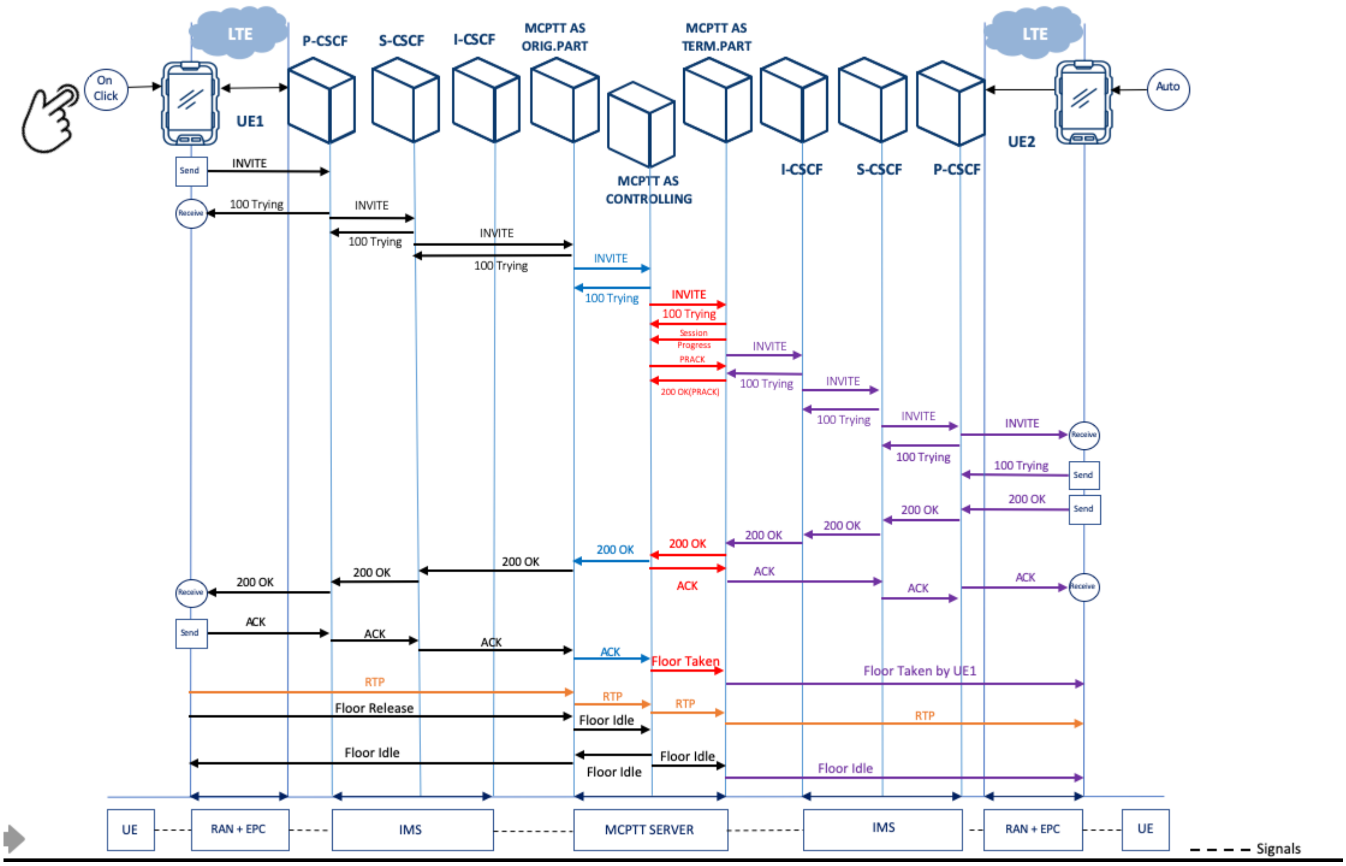


FIGURE 2. On-demand on-network private call with automatic commencement.

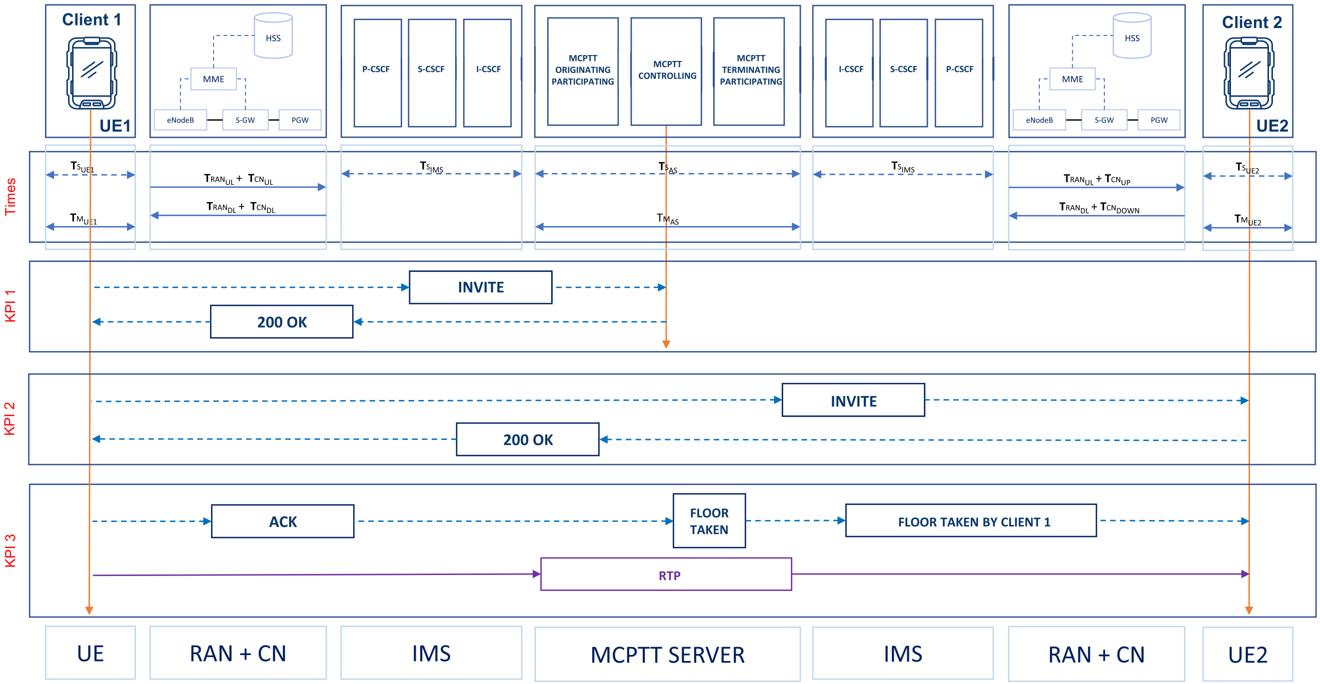


FIGURE 3. Functional architecture, control procedures and processing time for an MCPTT call.

TABLE 2. Definition of the processing times.

$T_{S_{UE}}$	Signaling processing time on the client
$T_{M_{UE}}$	Media processing time on the client
$T_{RAN_{UL}}$	RAN uplink delay
$T_{RAN_{DL}}$	RAN downlink delay
$T_{CN_{UP}}$	Core network uplink delay
$T_{CN_{DN}}$	Core network downlink delay
$T_{S_{IMS}}$	Signaling processing time on the IMS
$T_{S_{AS}}$	Signaling processing time on the AS
$T_{M_{AS}}$	Media processing time on the AS

traffic flows involved in MCPTT communications. The overall delay contributions associated to each KPI are depicted in Figure 3. It presents the trace analysis of an MCPTT call in the baseline LTE scenario. The signaling is represented by broken lines and the media flows in solid lines. The latency time involved in each KPI is calculated as the sum of the processing times corresponding to the implied components. The processing times that have been taken into consideration are defined in Table 2.

A. MEASUREMENT METHOD

The lack of a standardized method to measure KPI 1 and KPI 2 latencies, led us to define a suitable measure method. Both indicators are related to the signaling of the call, so, the methodology followed has been to capture all the traffic generated in the calls and to sort the packets according to the logic of a call establishment defined by the 3GPP in [11] and depicted in 3. In order to automate and simplify this computation we have developed an analysis tool that allows us to obtain the KPI 1 (1) and KPI 2 (2) for each call based on the call signaling messages.

$$\begin{aligned} KPI1 = {} & 2T_{S_{UE1}} + T_{RAN_{UL}} + T_{CN_{UP}} + 2T_{S_{IMS}} \\ & + 2T_{S_{AS}} + T_{CN_{DN}} + T_{RAN_{DL}} \quad (1) \\ KPI2 = {} & 2T_{S_{UE1}} + 2T_{RAN_{UL}} + 2T_{CN_{UP}} + 4T_{S_{IMS}} \\ & + 2T_{S_{AS}} + 2T_{CN_{DN}} + 2T_{RAN_{DL}} + T_{S_{UE2}} \quad (2) \end{aligned}$$

The calculation of the KPI 3 (3) implies the measurement of the latency between the moment the caller speaks in the microphone until the callee listens it through the loudspeaker of his device.

$$\begin{aligned} KPI3 = {} & T_{M_{UE1}} + T_{RAN_{UL}} + T_{CN_{UP}} \\ & + T_{M_{AS}} + T_{CN_{DN}} + T_{RAN_{DL}} + T_{M_{UE2}} \quad (3) \end{aligned}$$

This KPI is identified as a key component of the quality of experience in an MCPTT communication. For this indicator, we have adopted the specific method developed by NIST's Public Safety Communications Research (PSCR) division to measure and quantify the KPI 3 of any communications system transmitting audio, with specific emphasis on Push-to-Talk (PTT) devices [27]. This measurement method is the first step in establishing QoE key performance indicators for

TABLE 3. Number of times that the signaling and data plane crosses each component over a baseline LTE scenario.

SIGNAL.	$T_{S_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{S_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{S_{UE2}}$
KPI 1	2x	1x	1x	2x	2x	1x	1x	N/A
KPI 2	2x	2x	2x	4x	2x	2x	2x	1x

MEDIA	$T_{M_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{M_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{M_{UE2}}$
KPI 3	1x	1x	1x	N/A	1x	1x	1x	1x

mission critical voice (MCV) and a measurement system to quantify these KPIs.

Once the system model is established and the measurement methods are described, next sections get into the analysis of the KPI performance over the proposed scenarios.

IV. BASELINE LTE SCENARIO

The first analyzed scenario corresponds to a traditional LTE architecture. The processing times involved in the computation of the KPIs are outlined in Table 3. This table represents, for each KPI, the processing time corresponding to each stage of MCPTT end-to-end call, as depicted in Figure 3. As it can be observed, KPIs 1 and 2 only consider signaling messages, while the calculation of KPI 3 only contemplates media messages.

Based on this formulation, we have measured the performance of an LTE testbed, whose functional architecture is detailed in Section VII. The average delay measured in radio access under normal non-congested operation is 30 ms. This latency time is the same that can be measured using MONROE nodes [22], [23] as described in Section II of this article. This latency is assumed as the delay budget for T_{RAN}. In the worst-case scenario, under high traffic conditions and congested operation, the tenant is guaranteed with a packet delay budget of 60 ms for the signaling and 75 ms for the media, due to the use of QCIs.

V. MEC DEPLOYMENTS IN 4G

Multi-Access Edge Computing (MEC) [28], is a European Telecommunications Standards Institute (ETSI) standard that aims at moving applications from the core network to the mobile edge in close proximity to end-users. As a result, service latencies get reduced thanks to the closeness to the end user and the core network is relieved from the portion of traffic that is now processed at the edge. From a high-level point of view, MEC itself can be considered a cloud computing entity with access and capabilities to perform tasks at the edge of mobile networks. MEC has the ability to apply traffic rules to program the data path and redirect the traffic to the corresponding service provider, whether it's local or remote, to lower the latency. In this way, the service can be handled by the corresponding server to improve the perceived user experience in a totally transparent manner.

This work presents two MEC scenarios [29], which are further described in the following subsections. In both scenarios, the MCPTT service is deployed in the edge, including IMS and MCPTT Servers. The difference between them is the deployment of the EPC. In the first one, the MCPTT service is collocated with the complete EPC of the operator also deployed at the edge. In the second one, only the SGW and PGW entities are deployed at the edge site, whereas the control plane functions such as the Mobility Management Entity (MME) and Home Subscriber Server (HSS) are located at the operator's core site.

As it is going to be described next, bringing the service near to the end-user will decrease the processing times involved in media and/or signaling flows, improving the performance of the service in terms of latency. The data used to demonstrate the latency reductions are the same as in the baseline scenario. In this testing platform, we have measured the delay to reach a service located behind a traditional LTE EPC core network, under a wide variety of traffic load circumstances. As we depict in Figure 5, we have obtained that, in a realistic situation, over a 90% of the communications has experienced end-to-end latencies close to 50ms to reach a centralized MCPTT server located behind a conventional EPC core. On the other hand, the latency experienced also in the 90% of the tests between the end-user and the PGW has been about 30ms. Consequently, it is expected that latencies for the signaling and data of services deployed applying the MEC paradigm will be reduced at least a 40% in each way, compared to the baseline scenario case.

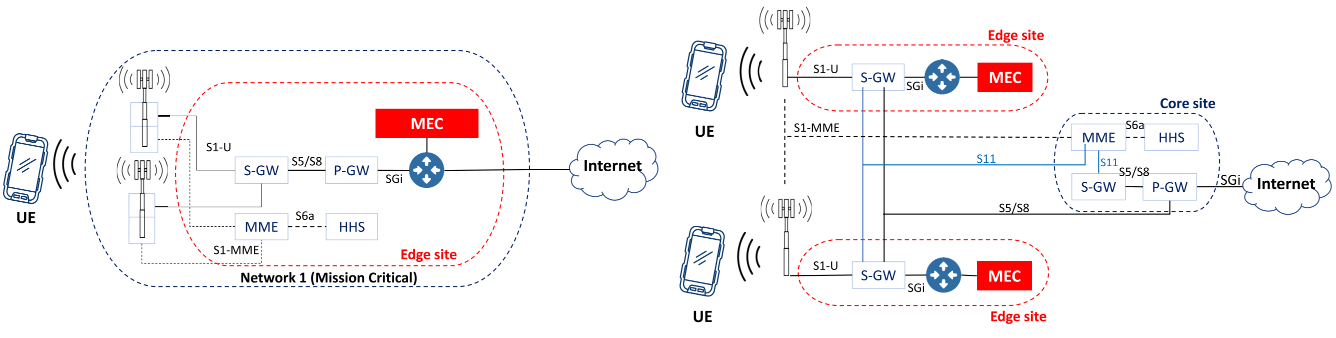


FIGURE 4. MEC deployment (distributed EPC on the right and distributed S/PGW on the left).

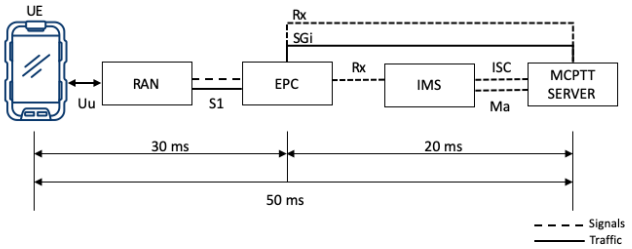


FIGURE 5. Medium delay values from the UE to a MCPTT server located behind an EPC core measured in the MONROE platform.

TABLE 4. Number of times that the signaling and data plane crosses each component over MEC (Distributed EPC).

SIGNAL.	$T_{S_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{S_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{S_{UE2}}$
KPI 1	2x	1x	N/A	2x	2x	N/A	1x	N/A
KPI 2	2x	2x	N/A	4x	2x	N/A	2x	1x
MEDIA	$T_{M_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{M_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{M_{UE2}}$
KPI 3	1x	1x	N/A	N/A	1x	N/A	1x	1x

A. DISTRIBUTED EPC

The first proposed alternative corresponds to an EPC service fully deployed on the MEC platform. Under these circumstances, the signaling and media data must not cross to the core infraestructure, which obviously leads to a reduction of the latency.

The experienced reduction of the latency comes from the shorter path of the transmissions now that the service has been brought to the edge. Table 4 shows the new formulation for the processing times through the network components, which saves two hops in the case of KPI 1 and KPI 3 and four hopes in the case of KPI 2.

B. DISTRIBUTED S/PGW

In the second EPC edge deployment proposal, only the SGW and PGW entities are deployed at the edge site, whereas the control plane functions such as the MME and HSS are located at the operator's core site. MEC hosts connect to the PGW over the SGi interfaces so that data media may not cross the core network of the operator's site.

The deployment with the SGW and PGW co-located at the network edge requires the operator to extend the S5 interface to the MEC site. This type of deployment makes it possible for the operator to retain full control over the MME. In this deployment, the signaling crosses the core network in a way that KPI 1 and KPI 2 are not improved compared to the baseline scenario. On the contrary, the mouth-to-ear latency (KPI 3) gets a reduction, as the data plane has been moved to the edge. Table 5 shows the contribution to latency of the different components of this deployment proposal. As it can be observed, in this case T_{CN} in the media plane is passed over.

VI. DEPLOYMENT IN NON-STANDALONE 5G

Finally, we also explore the deployment of the service over a non-standalone 5G architecture and analyze the expected improvement of the defined KPIs.

TABLE 5. Number of times that the signaling and data plane crosses each component over MEC (Distributed S/PGW).

SIGNAL.	$T_{S_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{S_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{S_{UE2}}$
KPI 1	2x	1x	1x	2x	2x	1x	1x	N/A
KPI 2	2x	2x	2x	4x	2x	2x	2x	1x
MEDIA	$T_{M_{UE1}}$	$T_{RAN_{UL}}$	$T_{CN_{UP}}$	$T_{S_{IMS}}$	$T_{M_{AS}}$	$T_{CN_{DN}}$	$T_{RAN_{DL}}$	$T_{M_{UE2}}$
KPI 3	1x	1x	N/A	N/A	1x	N/A	1x	1x

3GPP is also in charge of developing the new 5G core architecture (5GS) [30], which involves a new Radio Access Network - 5G New Radio (NR), available from Release 15 onwards, and a new 5G core network (5GC) defined as of Release 16. The analysis performed in this section considers a 5G non-standalone architecture, which involves 5G New Radio but maintains an LTE core network. In this scenario, we propose an MCPTT service deployment in an edge cloud platform. The edge location is suited to 5G because of the extreme low-latency requirements of some services (such as public safety services) and the scalability offered to meet growing traffic demands in specific locations.

The procesing times and associated KPI calculations of this deployment are represented in Table 5. The latencies here obtained are similar to the case described in Section V-A, but, in this case, benefiting from the low latency that 5G radio links will offer. 5G-NR is expected to achieve link connections between 1-2 ms, while in current enhanced Universal Terrestrial Radio Access Network (eUTRAN) this values are 10 times higher [31].

VII. EXPERIMENTAL DETAILS

The functional architecture of the testbed has the following components (as depicted in Fig. 2): eNodeB LTE, EPC, IMS (IP Multimedia Subsystem), MCPTT Servers and Android MCPTT clients. During the test, we establish 1000 sequential calls. The hardware platform on which the IMS and MCPTT servers have been deployed is a barebone with an INTEL i7 processor. The operating system installed in the barebone is Ubuntu 16.4 Server. The IMS / HSS and MCPTT Servers are deployed as virtual machines. Finally, the eNodeB and EPC have been installed in a separated barebone. The local network connection between the barebones has been established through a 1Gb switch. The KPI values presented in this work have been calculated according to the methodology defined by the NIST's PSCE division [27]. All the experimental parameters are listed in Table 6.

VIII. DISCUSSION OF THE RESULTS

After presenting the evolutive scenarios under study and the analytical approach to the delay formulation, we now compare the obtained results and their impact on the defined KPIs. To this effect, Figure 6 summarizes the impact of the evolution of the presented mobiles architectures on an MCPTT service. The figure exposes the level of compliance of each deployment option with the KPI thresholds imposed by the 3GPP for the MCPTT standard. The results show

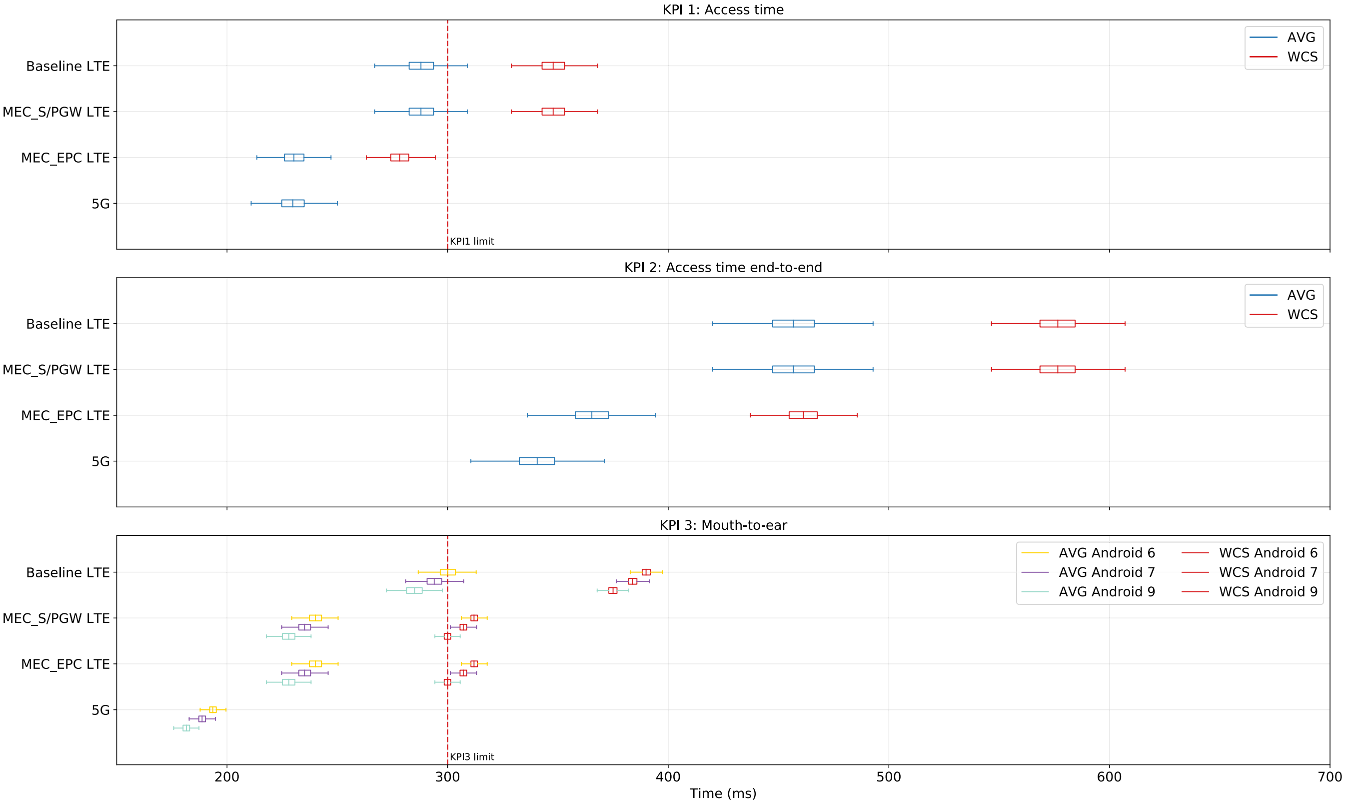


FIGURE 6. Impact of the evolution towards 5G architectures on mouth-to-ear latency.

TABLE 6. Experimental parameters.

Functional Architecture	
MCPTT AS compliant	3GPP Rel. 13 (v13.3.0 and above)
LTE Network EPC	Compact core network. MME/S-GW/P-GW/HSS/PCRF. Support for all FDD and TDD bands. LTE bearers with MCPTT QCIs
Operating System	Linux Ubuntu 16.04
Server Hardware	Barebone Mini-PC
Linux clients	
MCPTT compliant	3GPP Rel. 13 (v13.3.0 and above)
Operating system	Linux Ubuntu 16.04
Number of Clients	2.000 pre-registered clients
Number of Calls	1.000 simultaneous calls
Call rate	50 calls per second
Type	Private calls
Android clients	
MCPTT compliant	3GPP Rel. 13 (v13.3.0 and above)
Android version	8
N Clients	20 pre-registered clients
N Calls	10
Type	Private calls

that, under normal operational conditions, current LTE networks are compliant with the performance indicators defined in Table 1. Nevertheless, if we take into account the delay budgets guaranteed by the operator, KPIs 1 and 3 are slightly over the established limits. On the other hand, KPI 2, which measures the time a user takes to establish an end-to-end communication, is satisfactory in all scenarios.

The numerical results corresponding to Figure 6 are condensed in Table 7. The first column of this table illustrates the degree of compliance with the KPIs of the MCPTT service using eUTRAN and an LTE core under both normal operational conditions and the worst-case overloaded scenario. As it is shown, only Access Time KPI (KPI1) is not compliant with the 300 ms established by the 3GPP. However, this breach is barely for some milliseconds. This means that an MCPTT service deployment, though still challenging with current 4G technology, is getting remarkably closer to the requirements for Public Safety communications.

Furthermore, novel network architectures evolving towards 5G can benefit the MCPTT service performance in terms of reducing the latencies even more. When the service is deployed along with the EPC on the edge, the signaling and the user data is not uploaded to the core network, which implies a 40% time reduction in both directions. Moreover, if we only distribute the SGW and PGW to the edge, KPI 1 and KPI 2 will take similar values regardless the service is located in the edge or not, while KPI 3 will be significantly improved.

At another level, throughout the study we have detected that one of the most determining elements in the service performance is the user device. In this paper, we have graphically represented the improvement in the KPI 3 values obtained by using mobiles manufactured by operating system developers or manufacturers that are building their hardware to be able to take better advantage of the operating system.

TABLE 7. Impact on Mission Critical Push to Talk Service of the evolution from eUTRAN towards 5G architectures (ms).

	Radio Access Network	eUTRAN						5G
	Core Network	Centralized EPC		Distributed EPC		Distributed S/PGW		Standalone
	Network Load	Average	Congested	Average	Congested	Average	Congested	Average
KPI1	Access time	301 ms	359 ms	241 ms	287 ms	301 ms	358 ms	193 ms
KPI2	Access time end-to-end	479 ms	594 ms	384 ms	475 ms	479 ms	594 ms	267 ms
KPI3	Mouth-to-ear	293 ms	379 ms	235 ms	304 ms	235 ms	304 ms	185 ms

To sum up, these results show how the evolution of the architectures and the adaptation of the MC services to them will allow a significant reduction of the latency in PS communications.

IX. CONCLUSION

The 3GPP has done a great effort of standardizing the MCPTT service. This effort aims to leverage existing 4G/LTE and future 5G broadband networks, to provide push-to-talk voice communications that get closer to the performance of PMR/LMR voice. The final objective is to offer a cost-effective, open, interoperable alternative to legacy MC networks, while paving the way to the upcoming data and video services with broadband requirements.

Our results illustrate how the MCPTT service deployed over a current LTE network is not yet able to guarantee the compliance with the KPIs defined by 3GPP under any circumstances. Still, this work envisages the promising future of (near-to) 5G networks regarding the performance improvements expected for low latency services, such as mission-critical communications. The transition to a fully capable 5G network is expected to happen by gradually taking over the 4G equipment. Hereupon, as one of the main technology enablers for 5G, MEC can be used in combination with current mobile networks to target the support for MC services from now on, even before 5G is rolled out.

Extending the study to 5G, the estimated results are even better, demonstrating the beneficial impact of the evolution of 4G/5G commercial broadband networks on MC services. Therefore, our next steps will be focused on extending this KPI assessment to the future MC services and on studing the impact of new network paradigms, such as MEC and NFV, on the performance of commercial broadband networks.

ACKNOWLEDGMENT

The research leading to these results has been supported by the European Commission under the H2020 5G-PPP project ESSENCE (grant agreement no 761592), by the US NIST under the Public Safety Innovation Accelerator Program (PSIAP), and by Spanish MINECO under the project 5RANVIR (no. TEC2016-80090-C2-2-R).

REFERENCES

[1] Y. Lair and G. Mayer, "Mission critical services in 3GPP," Tech. Rep., 2017.

[2] M. Peltola, "Socioeconomic benefits as criteria in the valuation of broadband mobile services for public safety and security," Aalto Univ., Espoo, Finland, Tech. Rep., 2018.

[3] R. Ferrüs and O. Sallent, "Future mobile broadband ppdr communications systems," in *Mobile Broadband Communications for Public Safety: The Road Ahead Through LTE Technology*. Hoboken, NJ, USA: Wiley, 2015, pp. 81–125.

[4] C. Gruet, X. Pons-Masbernat, and P. Force, "The LTE evolution: Private mobile radio networks," *IEEE Veh. Technol. Mag.*, vol. 8, no. 2, pp. 64–70, Jun. 2013.

[5] CEPT, "Harmonised conditions and spectrum bands for the implementation of future European broadband PPDR systems," CEPT, ECC Rep. 218, Oct. 2018.

[6] A. Paulson and T. Schwengler, "A review of public safety communications, from LMR to voice over LTE (VoLT E)," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 3513–3517.

[7] T. Doumi, M. F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore, "LTE for public safety networks," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 106–112, Feb. 2013.

[8] R. Ferrus, O. Sallent, G. Baldini, and L. Goratti, "LTE: The technology driver for future public safety communications," *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 154–161, Oct. 2013.

[9] M. Stojkovic, "Public safety networks towards mission critical mobile broadband networks," M.S. thesis, NTNU, Trondheim, Norway, 2016.

[10] A. Kumbhar, F. Koohifar, I. Güvenç, and B. Mueller, "A survey on legacy and emerging technologies for public safety communications," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 1, pp. 97–124, 1st Quart., 2016.

[11] *Universal Mobile Telecommunications System (UMTS); LTE; Mission Critical Push to Talk (MCPTT) over LTE; Stage 1*, document TS 22.179, 3GPP, Jan. 2016.

[12] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. Mao, S. Sen, and O. Spatscheck, "An in-depth study of LTE: Effect of network protocol and application behavior on performance," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 363–374, 2013.

[13] D. del Rey Carrión, L. Juan-Llácer, and J.-V. Rodríguez, "Radio planning considerations in tetra to LTE migration for PPDR systems: A radioelectric coverage case study," *Appl. Sci.*, vol. 9, no. 2, p. 250, 2019.

[14] Z. Kaleem, Y. Li, and K. Chang, "Architecture and features for 5G mobile personal cell," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2015, pp. 164–166.

[15] M. Welsh. *MobiPerf*. Accessed: Jul. 8, 2019. [Online]. Available: http://mobiperf.com

[16] A. Nikravesh, D. R. Choffnes, E. Katz-Bassett, Z. Mao, and M. Welsh, "Mobile network performance from user devices: A longitudinal, multidimensional analysis," in *Proc. Int. Conf. Passive Act. Netw. Meas.*, Mar. 2014, pp. 12–22.

[17] E. Halepovic, J. Pang, and O. Spatscheck, "Can you GET me now?: Estimating the time-to-first-byte of HTTP transactions with passive measurements," in *Proc. Internet Meas. Conf.*, Nov. 2012, pp. 115–122.

[18] M. Zubair Shafiq, L. Ji, A. X. Liu, J. Pang, S. Venkataraman, and J. Wang, "A first look at cellular network performance during crowded events," in *Proc. ACM SIGMETRICS/Int. Conf. Meas. Modeling Comput. Syst.*, vol. 41, Jun. 2013, pp. 17–28.

[19] Z. Koradia, G. Mannava, A. Raman, G. Aggarwal, V. Ribeiro, A. Seth, S. Ardon, A. Mahanti, and S. Triukose, "First impressions on the state of cellular data connectivity in India," in *Proc. 4th Annu. Symp. Comput. Develop.*, Dec. 2013, pp. 1–3.

[20] D. Baltrunas, A. Elmokashfi, and A. Kvalbein, "Measuring the reliability of mobile broadband networks," in *Proc. ACM SIGCOMM Internet Meas. Conf. (IMC)*, 2014, pp. 45–58.

[21] Ö. Alay, A. Lutu, R. García, M. Peón-Quirós, V. Mancuso, T. Hirsch, T. Dely, J. Werme, K. Evensen, A. Hansen, S. Alfredsson, J. Karlsson, A. Brunstrom, A. S. Khatouni, M. Mellia, M. A. Marsan, R. Monno, and H. Lonsethagen, "Measuring and assessing mobile broadband networks with MONROE," in *Proc. IEEE 17th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Jun. 2016, pp. 1–3.

[22] A. Safari Khatouni, M. Mellia, M. Ajmone Marsan, S. Alfredsson, J. Karlsson, A. Brunstrom, O. Alay, A. Lutu, C. Midoglu, and V. Mancuso, "Speedtest-like measurements in 3G/4G networks: The monroe experience," in *Proc. 29th Int. Teletraffic Congr. (ITC)*, Sep. 2017, pp. 169–177.

[23] Ö. Alay, A. Lutu, M. P. Quirós, V. Mancuso, T. Hirsch, K. Evensen, A. F. Hansen, S. Alfredsson, J. Karlsson, A. K. Brunstrom, A. S. Khatouni, M. Mellia, and M. A. Marsan, "Experience: An open platform for experimentation with commercial mobile broadband networks," in *Proc. MobiCom*, 2017, pp. 70–78.

[24] A. Kurian, "Latency analysis and reduction in a 4G network," M.S. thesis, Delft Univ. Technol., Delft, The Netherlands, 2018. [Online]. Available: http://resolver.tudelft.nl/uuid:e1badd8d-a384-49a1-b958-a0c1e499c539

[25] C. Lu, "Delay analysis of push-to-talk over cellular (PoC) service solutions for public safety communications over LTE networks," M.S. thesis, Univ. Politècnica Catalunya, Barcelona, Spain, 2012.

[26] (2019). *iOS and Android Audio Latency*. [Online]. Available: https://superpowered.com/

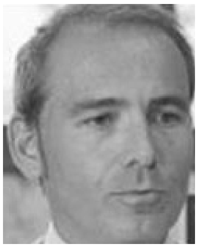
[27] J. Frey, J. Frey, J. Pieper, and T. Thompson, "Mission critical voice QoE mouth-to-ear latency measurement methods," NIST, Washington, DC, USA, Tech. Rep. NISTIR 8206, 2018.

[28] *Mobile Edge Computing (MEC); Framework and Reference Architecture*, document ETSI GS MEC 003, V1.1.1 Release, Mar. 2016.

[29] F. Giust, G. Verin, K. Antevski, J. Chou, Y. Fang, W. Featherstone, F. Fontes, D. Frydman, A. Li, A. Manzalini, D. Purkayatha, D. Sabella, C. Wehner, K.-W. Wen, and Z. Zhou, "MEC deployments in 4G and evolution towards 5G," ETSI, Sophia Antipolis, France, ETSI White Paper 24, 2018.

[30] *System Architecture for the 5G System*, document 3GPP TS 23.501, 3rd Generation Partnership Project, Dec. 2018.

[31] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, "A survey on low latency towards 5G: RAN, core network and caching solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3098–3130, 4th Quart. 2018.



AITOR SANCHOYERTO is an Engineer for more than 20 years of experience in telecommunications projects in public safety sectors such as defense, national security, and transport. He is specialized in interoperability protocols between organizations with different levels of security and radio technology whose objective is to be able to establish voice and data communications in a safe way. In recent years, he has been actively working on convergence projects from narrow-band radio communications to broadband communications based on the 3GPP-defined standard for mission-critical services. Recently, he has joined UPV/EHU to pursue the Ph.D. thesis on MC services management in 5G.

RUBEN SOLOZABAL received the M.Sc. degree in telecommunications engineering from the University of the Basque Country, in 2015, where he is currently pursuing the Ph.D. degree with a particular interest in public safety over the 5G architecture. He was involved in automation industry for three years focusing on industrial communications.

BEGO BLANCO received the B.S., M.S. and Ph.D. degrees in telecommunications engineering from the University of the Basque Country, Spain, in 2000. She is currently a Lecturer and a Researcher with the Faculty of Engineering, University of the Basque Country, Bilbao. Her research interests includes PQoS/QoE/QoS assessment and multicriteria optimization in 5G networks.

FIDEL LIBERAL received the Ph.D. degree from the University of the Basque Country (UPV/EHU), in 2005. He has been the Project Coordinator in the FP7-SEC-GERYON Project, from 2012 to 2014, dealing with public safety interoperability problems. He is currently a Lecturer and a Researcher with the Faculty of Engineering of Bilbao, University of the Basque Country, where he cooperates in different international projects. He has coauthored over 75 conferences and journal papers. His research interests include next generation emergency networks and 5G.