UNIVERSIDAD DE MÁLAGA

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE INFORMÁTICA

TESIS DOCTORAL

COMUNICACIONES MÓVILES DE MISIÓN CRÍTICA SOBRE REDES LTE

PROGRAMA DE DOCTORADO EN TECNOLOGÍAS INFORMÁTICAS

AUTOR: CÉSAR AUGUSTO GARCÍA PÉREZ

DIRECTOR: PEDRO MERINO GÓMEZ

MÁLAGA, 2020

AUTOR: César Augusto García Pérez

http://orcid.org/0000-0001-9390-7872

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga

Authorship Statement

DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR

D. *César Augusto García Pérez*

Estudiante del programa de doctorado *Tecnologías Informáticas* de la Universidad de Málaga, autor de la tesis, presentada para la obtención del título de doctor por la Universidad de Málaga, titulada: *Comunicaciones de Misión Crítica sobre Redes Móviles*

Realizada bajo la tutorización y dirección de *Pedro Merino Gómez*.

DECLARO QUE:

La tesis presentada es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, conforme al ordenamiento jurídico vigente (Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), modificado por la Ley 2/2019, de 1 de marzo.

Igualmente asumo, ante a la Universidad de Málaga y ante cualquier otra instancia, la responsabilidad que pudiera derivarse en caso de plagio de contenidos en la tesis presentada, conforme al ordenamiento jurídico vigente.

En Málaga, a 1 de MARZO de 2020

Fdo.: CÉSAR AUGUSTO GARCÍA PÉREZ

Acknowledgements

To Ada,
may you enlighten this world
as you have enlightened mine's.

César

The work described in this thesis has received funding from the European Horizon 2020 Programme under grant agreement 688624 (Q4Health project). The validation of the approach has been made thanks to the collaboration of Polaris Networks Inc.

Contents

Authorship Statement **iii**

Acknowledgements **v**

Acronyms **xiii**

Preface **1**

1 Introduction **3**
1.1 Mission Critical Mobile Communications 5
1.2 Motivation 7
1.3 Contributions 9
1.4 Rest of this book 10

2 Experimentation Tools **13**
2.1 Introduction 15
2.2 State of the art 16
2.3 PerformNetworks 19
2.4 Contributions to PerformNetworks 21
2.4.1 Improving Automation 21
2.4.2 Tools 24
2.4.3 UE Selection methodology 31
2.5 Conclusions 33

3 Standard LTE architecture **35**
3.1 Introduction 37
3.2 Standard Architecture 38
3.3 LTE Performance 40
3.3.1 Throughput 40
3.3.2 Latency 41
3.3.3 Prioritization and Reliability 44
3.3.4 Security 46
3.4 Conclusions 46

4 Mobile Mission Critical Communications Architecture **49**

4.1 Introduction 51
4.2 Requirements 52
4.3 Early Enablers 54
4.4 MCX Architecture 57
4.5 Conclusions 60

5 Low Latency Architectures 63
5.1 Introduction 65
5.2 Analysing the latency in LTE networks 67
  5.2.1 LTE architecture latency overview 67
  5.2.2 Qualitative overview of latency contributions 68
  5.2.3 Experimental evaluation of the latency 70
5.3 Fog Gateway 73
  5.3.1 Architecture 73
  5.3.2 Emulated Results 77
5.4 GTP Gateway 79
  5.4.1 Architecture 79
  5.4.2 Latency Emulated Results 81
  5.4.3 Comparison with the Fog Gateway 82
5.5 Third Party Exposure 83
5.6 Conclusions 86

6 Fog Gateway Implementation 89
6.1 Introduction 91
6.2 Implementation 92
  6.2.1 Implementation alternatives 92
  6.2.2 Software architecture overview 94
6.3 Latency analysis 96
  6.3.1 Validation Scenarios 96
  6.3.2 Results 98
6.4 Conclusions 103

7 Conclusions 105
7.1 Summary of our contributions 107
  7.1.1 Publications 107
  7.1.2 Tools 108
  7.1.3 Projects 109
7.2 Discussion on the results 110
  7.2.1 Mission Critical Mobile Networks 110
  7.2.2 Experimental Testbeds 111
7.3 Future work 112

Bibliography 133

List of Figures

1.1 3GPP Mission Critical Standardization 8

2.1 PerformNetworks architecture 19
2.2 XML-SCPI Integration Architecture [GPRPRG+16] 22
2.3 Remote Impairments Architecture [MGM+18] 23
2.4 EPC Automation Architecture [MGM+18] 23
2.5 PingAnalyzer dissected protocols 25
2.6 Dissectors UML Diagram 25
2.7 NAS and S1AP Libraries 26
2.8 Time consumed by signalling procedures [GPDZR+17a] 27
2.9 S1DatabaseGenerator state machine [GDM+18] 28
2.10 Tunnel Tester [GP17] 29
2.11 Methodology Setup[GP17] 31
2.12 Normalized Comparison of two User Equipment (UE) 33
2.13 Setup for the validation of the Fog Gateway [GPDZR+17a] 34

3.1 LTE architecture overview 38
3.4 Data plane latency measurements [GPDZR+17a] 42
3.5 Control plane latency measurements [GPDZR+17a] 43
3.8 Mean Opinion Score (MOS) Measurements under heavy background traffic [GP15] 45

4.1 LTE architecture to support railway communications 54
4.2 Updated MCC Architecture 59

5.1 LTE basic architecture [GPM17] 67
5.2 GTP header [GPM16] 69
5.3 Experiment setup to characterize the latency split [GPM16] 70
5.4 LTE RTT baseline under different radio conditions [GPM16] 72
5.5 Architecture of the Fog Gateway [GPM17] 73
5.6 Message Sequence Chart of the Fog Gateway [GPM17] 74
5.7 Fog Gateway ARP interactions 77
5.8 Standard network and Fog Gateway RTT comparison 77
5.9 Emulated Fog Gateway RTT under different radio conditions [GP17] . 78
5.10 Architecture of the GTP gateway [GPM17] 79

5.11 GTP Gateway MSC for the user plane [GPM17] 80
5.12 Experiment Setup to Evaluate the GTP Gateway [GPM17] 81
5.13 GTP Gateway Performance Under Different Channels[GPM17] 81
5.14 RTT CDF Comparison between Fog Gateway and GTP Gateway [GPM17] 82
5.15 End-to-end architecture 84
5.16 Comparison of traffic with and with QoS enforcement [GPDZR+17a] . 85
5.17 GTP Gateway Stacks 86

6.1 Prototype implementation methodology [GP17] 93
6.2 Software architecture diagram [GP17] 94
6.3 Packet Parsing Module Class Diagram [GP17] 95
6.4 Queue Module Class Diagram [GP17] 95
6.5 Development setup [GP17] 96
6.6 Experiments setup [GP17] 97
6.7 Fog Gateway platform contribution characterization [GP17] 99
6.8 Fog Gateway fog services comparison [GP17] 100
6.9 Fog Gateway cloud services comparison [GP17] 101
6.10 UDP downlink throughput [GP17] 102
6.11 UDP uplink throughput comparison [GP17] 102
6.12 Summary of the results obtained with the Fog Gateway [GP17] 103

7.1 RTT CDF Comparison standard LTE network, Fog and GTP Gateways 111

List of Tables

1.1 Cost comparison GSM-R, TETRA and LTE 8

3.1 Summary of the latency measurements 44

4.1 GSM-R QoS Requirements 52
4.2 Early-enablers standards for railway communications [DZGPMG14a] . 56
4.3 Updated summary of the standards for MCC 58

5.1 T2010A Configuration 71
5.2 RTT split time comparison 72
5.3 Summary of the results obtained with the different latency solutions . . 82

6.1 Summary of the platform characterization results 99
6.2 Summary of the Fog Gateway and standard EPC comparison 100
6.3 Cloud Services RTT when using the standard EPC or the Fog Gateway 101
6.4 Summary of the UDP throughput with and without the Fog Gateway . 103

7.1 Summary of the publications 107
7.2 Summary of the tools 108
7.3 Summary of the research projects 109

Acronyms

5GC	5G Core
ACE	Adaptive Communication Environment
AKA	Authentication and Key Agreement
ATCA	Advanced Telecommunications Computing Architecture
ANDSF	Access Network Discovery and Selection Function
AP	Access Point
APN	Access Point Name
ARP	Address Resolution Protocol
ASN.1	Abstract Syntax Notation One
AWGN	Additive White Gaussian Noise
BGCF	Breakout Gateway Control Function
BLER	Block Error Rate
BM-SC	Broadcast Multicast Service Centre
BW	Bandwidth
CAPEX	Capital Expenditure
CBC	Cell Broadcast Centre
CBE	Cell Broadcast Entity
CBG	Code Block Group
CDF	Cumulative Distribution Function
CDN	Content Distribution Networks
CDR	Charging Data Record
COTS	Commercial-Off-The-Shelf
CS	Circuit Switched
CSCF	Call State Control Function
DL	Downlink

DMRS	Demodulation Reference Signal
DNS	Domain Name System
DoS	Denial of Service
DPI	Deep Packet Inspection
DRA	Diameter Routing Agent
DRB	Dedicated Radio Bearer
DUT	Device Under Test
EC	European Commission
eNB	Evolved Node B
eMBMS	Evolved Multimedia Broadcast Multicast Services
EPA5	Extended Pedestrian A 5Hz
EPC	Evolved Packet Core
ePDG	Evolved Packet Data Gateway
EPS	Evolved Packet System
ERA	European Union for Railways
E-RAB	E-UTRAN Radio Bearer
ERTMS	European Rail Traffic Management System
e-SMLC	Evolved Serving Mobile Location Centre
ETCS	European Train Control System
ETSI	European Telecommunications Standards Institute
E-UTRAN	Evolved Universal Terrestrial Radio Access Network
EVA70	Extended Vehicular A 70Hz
FD	File Distribution
FGW	Fog Gateway
FIRE	Future Internet Research Experimentation
FRMCS	Future Railway Mobile Communication System
GBR	Guaranteed Bit Rate
GENI	Global Environment for Network Innovations
GPP	General Purpose Processor
GPRS	General Packet Radio Service
GSA	Global mobile Suppliers Association
GSM	Global System for Mobile Communications
GSM-R	Global System for Mobile Communications Railway

GTP	General Packet Radio Service Tunnelling Protocol
GUI	Graphical User Interface
HARQ	Hybrid Automatic Repeat Request
HO	Handover
HSS	Home Subscriber Server
ICMP	Internet Control Message Protocol
I-CSCF	Interrogating Call Session Control Function
IDE	Integrated Development Environment
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IOPS	Isolated Operation for Public Safety
IoT	Internet of Things
ISDN	Integrated Services Digital Network
ISUP	ISDN User Part
ITU	International Telecommunication Union
IWF	Inter-Working Function
JCR	Journal Citation Reports
KPI	Key Performance Indicator
LCS	Location Services
LMR	Land Mobile Radio
LMU	Location Measurement Unit
LTE	Long Term Evolution
LTE-U	LTE Unlicensed
MAC	Medium Access Control
MAD	Median Absolute Deviation
MANO	Management and Orchestration
MBMS	Multimedia Broadcast Multicast Services
MBMS-GW	Multimedia Broadcast Multicast Services Gateway
MBR	Maximum Bit Rate
MCC	Mission Critical Communications
MCE	Multi-cell/multicast Coordination Entity
MCData	Mission Critical Data
MCPTT	Mission Critical Push to Talk

MCVideo	Mission Critical Video
MCX	Mission Critical X with X standing for Video, Data and PTT
MEC	Multi-access Edge Computing
MGCF	Media Gateway Control Function
mHealth	Mobile Health
MME	Mobility Management Entity
MC-MTC	Mission Critical Machine Type Communications
M-MTC	Massive Machine Type Communications
MNO	Mobile Network Operator
MORSE	Mobile Networks and Software Reliability
MOS	Mean Opinion Score
MRF	Media Resource Function
MSC	Message Sequence Chart
MT	Mobile Terminated
MTC	Machine Type Communication
MTU	Maximum Transfer Unit
NAS	Non-Access Procedure
NAT	Network Address Translation
NFV	Network Function Virtualization
NR	New Radio
OAI	Open Air Interface
OEDL	OMF Experiment Description Language
OMEC	Open Mobile Evolved Core
OMF	ORBIT Management Framework
OML	OMF Measurement Library
OPEX	Operational Expenditure
OS	Operating System
OVS	Open vSwitch
PCC	Policy Charging Control
PCRF	Policy and Charging Rules Function
P-CSCF	Proxy Call Session Control Function
PDCP	Packet Data Convergence Protocol
PDN	Packet Data Network

PDU	Protocol Data Unit
PER	Packed Encoding Rules
PESQ	Perceptual evaluation of speech quality
PGW	PDN Gateway
PMR	Professional Mobile Radio
PoC	Push to Talk over Cellular
PPDR	Public Protection and Disaster Relief
ProSe	Proximity Services
PS	Packet Switched
PTP	Precision Time Protocol
PTT	Push to Talk
PWS	Public Warning System
QCI	Quality Class Indicator
QoE	Quality of Experience
QoS	Quality of Service
RAT	Radio Access Technology
RF	Radio Frequency
RLC	Radio Link Control
RoHC	Robust Header Compression
ROI	Return of Investment
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RTP	Real-time Transport Protocol
RTT	Round Trip Time
SCPI	Standard Commands for Programmable Instruments
S-CSCF	Serving Call Session Control Function
SET	SUPL Enabled Terminal
SDN	Software Defined Network
SDR	Software Defined Radio
SDS	Short Data Service
SGW	Serving Gateway
SIP	Session Invitation Protocol
SLA	Service Level Agreement

SLP	SUPL Location Platform
SNR	Signal-to-noise Ratio
SUPL	Secure User Plane Location
SON	Self Organizing Networks
SPS	Semi-Persistent Scheduling
SSH	Secure Shell
TAS	Telephony Application Server
TEID	Tunnel Endpoint Identifier
TFM	Trabajo Fin de Master
TFT	Traffic Flow Template
TSN.1	Transfer Syntax Notation One
TTI	Transmission Timer Interval
UDP	User Datagram Protocol
UE	User Equipment
UIC	International Union of Railways
UL	Uplink
UMA	Universidad de Málaga
UMTS	Universal Mobile Telecommunication System
UPF	User Packet Function
URLLC	Ultra-Reliable Low Latency Communication
V2X	Vehicle-to-Everything
VM	Virtual Machine
VoLTE	Voice over LTE
XSLT	Extensible Stylesheet Language Transformations
XML	Extensible Markup Language

Preface

Andábamos sin buscarnos pero sabiendo que andábamos para encontrarnos
Julio Cortazar

Thanks for reading this. Uh no wait, let's start from the beginning. I wanted to be a bio-engineer. I purchased some slides for my bio-engineering class in the University copy centre and, surprise, there were some extra (non-free) pages. They were a dissemination article about TCP and the problems it has faced since its conception. I just read the introduction disappointed for the stranger in my copies. By the time I have reached the bus stop, I already have read the full article and decided that wanted to become a protocol engineer (not the ones that sit people in royalty events but the ones that design and implement very cool communication stacks). And this is why this thesis is about mobile communications.

Ok, some things happened in the middle. I wanted to write my dissertation on TCP, so I asked the teacher of my Protocol Engineering class, who was, by the way, Pedro Merino, if he knew someone that wanted to lead a master final dissertation about TCP. And this is how Pedro became the director of my PhD. Yes, more things happened in the middle. Pedro offered me to collaborate in a private-public collaboration project about LTE. However, I had already decided that I wanted to go to Switzerland to work on vehicle to vehicle communications so I said no. Honestly, I mainly said no because I did not know by the time what LTE means and when I discovered what it was (thanks to Wikipedia, yes I am a donor thanks for asking Mr Wales) I begged him to reconsider me for the position, and so he did, giving me a profession that I love. And this is why LTE appears a few times in this document.

Then another project lead to another project, and I forgot about TCP (eventually I completed the dissertation a few years later), and I was always involved in private sector transfer projects. That was until Almudena Diaz suggested that we could publish something about railway communication using the research with the results obtained for one of the projects. Best idea ever, I loved that, I enjoyed writing the article, discussing with my colleges and the reviewers and the very fresh look of the issue in which it appeared. So I decided I was going to try to do the PhD and started a master focused on research. And this is why my first paper with Almudena (thanks!) is separated from the rest of the articles.

Just a few more things to recap. I had a wonderful son that kept me locked long enough to complete some publications with a smile in my face. I got married to Eva, who is the most amazing woman I have ever met (not to speak about how great teacher-/mother/racooner/traveller/disserter she is). I moved to the private sector to a pretty cool company. I had an incredible daughter that always cheers me up even when she is in destruction mode. And for all these reasons you will find a slight gap between the dates of the papers that support this thesis and the moment you are reading it.

It has been a long way that I have enjoyed and, for a while, this will be my last research work. I hope you will enjoy this document at least half as much as I have enjoyed all this adventure. So, for now, I think it is time to say... Thanks for reading this!

César

Chapter 1

Introduction

Contents

1.1 Mission Critical Mobile Communications 5
1.2 Motivation 7
1.3 Contributions 9
1.4 Rest of this book 10

Synopsis

In this chapter, we introduce the topic of this dissertation, which is the provision of Mission Critical Communications (MCC). First, we provide a brief introduction to the topic, followed by an analysis of the requirements for mission critical communications. Then an overview of the related work is provided, followed by the objectives and contributions of this thesis. Finally, an overview of the rest of this book is done.

1.1 Mission Critical Mobile Communications

Mission Critical Communications (MCC) span across many different types of services, which usually are very demanding in terms of reliability but also in terms of latency or bandwidth. Some of the typical scenarios involving MCC are blue light services communication (ambulances, police, etc.), critical infrastructure surveillance or crisis management. Still, new scenarios are appearing in the market such as augmented reality, robotics, vehicle to vehicle communications or artificial vision systems. Whereas traditional services have been typically supported with voice and reliable data communications [TNW13], the new applications require more complex systems with higher requirements on the latency, data rates, network functionality and availability.

MCC are typically provided by the use of niche Professional Mobile Radio (PMR)[1] technologies, such as TETRA [DDT^{+}13], which have been designed around the provision of voice services. But the use of proprietary technologies hardens the adoption of improvements as the size and characteristics of these markets harden the Return of Investment (ROI). On the other hand, mobile technologies are being deployed massively [Int19], they provide a much better performance[2], and they are generally cheaper than PMR, not only in terms of Capital Expenditure (CAPEX) and Operational Expenditure (OPEX) but also because they enable other deployment/ownership strategies.

This demand for more services and the improvement of the existing ones is apparent in the mission critical market. Some of these new services were analysed in [GPDZR^{+}17b], public safety is an important one but there are others such as mHealth, vehicle to vehicle or next generation industries. To analyse their requirements, we selected railway communication as a driver use case. From a performance perspective, it is very demanding (not only its Key Performance Indicators (KPIs) but also the scenarios), and additionally, it has many functional requirements for the network (such as group communications or location dependant addressing).

Currently, railways are managed by European Rail Traffic Management System (ERTMS), a standard management system defined by European Union for Railways (ERA). ERTMS communications are supported by the Global System for Mobile Communications Railway (GSM-R) system, which is an extension of Global System for Mobile Communications (GSM) specifically developed to support railway communications by adding the network functional requirements that were not provided by GSM. ERA is indeed looking for replacements of GSM-R, so we explored the provision of an LTE architecture to support railway communications in [DZGPMG14a], where we concluded that the majority of the requirements demanded by mission critical scenarios could be already fulfilled with the LTE Commercial-Off-The-Shelf (COTS) technology that was standardized by then.

---

[1] We will use equally both PMR and Land Mobile Radio (LMR) as the only difference will be the focus or not on public safety sectors.

[2] For instance TETRA release 2 is providing less than one megabit per second, in the 2008 Long Term Evolution (LTE) release 8 75Mbit/s data rates were already standardized.

But, the development of this thesis was done at the same time than the standardization efforts of mission critical dedicated standards and the development of 5G [ABC$^+$14][CFH14], so in the meantime, new standards and technologies emerged. Standardization bodies have now introduced reference architectures targeting the requirements coming from the different stakeholders of MCC. For instance, ERA has already provided some studies and requirements for Future Railway Mobile Communication System (FRMCS), which is the new standard to replace GSM-R and 3GPP has defined a new architecture, which provides specific network elements for MCC.

Analysing how mobile networks could fulfil the functional network requirements of MCC is essential, but it is also important to characterize their performance. Networks are becoming more complex and more difficult to characterize and, for this reason, the use of experimental platforms has been increasing, for instance, in [Aut16] we described our platform PerformNetworks[3] along with more than other 50 platforms devoted to 5G across Europe. The use of these platforms eases the characterization of the technologies for different purposes, reducing the cost of the evaluation and pilots by sharing the testing infrastructure[PJR$^+$20]. Additionally, they can improve the development and evaluation of prototypes by providing reference implementations, portable tools to generate signalling scenarios and components with different characteristics to increase the number of integration tests.

After using our testbed to characterize LTE networks, we concluded that end-to-end latency is an aspect that could be improved[GPM16], primarily to support new applications such as self-driving cars, robotics or haptic technologies[LLM$^+$17]. In this area, the Multi-access Edge Computing (MEC)[4] paradigm plays a vital role, it proposes the location of the services closer to the users to reduce the end-to-end path.

To summarize, the global objective of this thesis is the support of Mission Critical Communications using standard mobile technologies, as we think they will accelerate the adoption of innovations and reduce the cost of deploying and operating them. To do so, we defined more concrete objectives:

- Evaluate if mission critical communications could be supported with standard mobile networks, both from functional and performance perspectives.
- Improve the existing experimental platforms to increase the possible evaluation scenarios and their characterization.
- Reduce the end-to-end latency of mobile networks to better support future mission critical services.
- Improve the end-to-end performance of MCC over mobile networks.

---

[3] Currently known as TRIANGLE `https://www.fed4fire.eu/testbeds/triangle/`(indoor part) and 5Genesis`https://5genesis.eu/malaga-platform/`(outdoor part).

[4] In this document we use MEC to refer to a general methodology and as a synonym of fog computing. To refer to the ETSI standard MEC architecture we use ETSI-MEC.

To accomplish these objectives, we have done a qualitative and quantitative analysis of mobile networks and their suitability to support a complex MCC scenario such as railway communications. For the experimental platforms, we have developed different tools that improve the interconnection of instrumentation, prototypes and standard equipment, as well as a methodology to improve the end-to-end performance by identifying the best UE for a given application. We have developed several prototypes of a new design to support MEC on standard LTE networks and have evaluated a more efficient solution for future networks.

1.2 Motivation

The provision of mission critical communications with mobile networks has been gaining importance over the years. PMR systems are challenging to maintain both from an economical (OPEX and CAPEX are high to be sustained by a single user or agency) and from a technical perspective (the arrival of new use cases and technologies is increasing the requirements for these applications). The MCC market is expanding, according to [Hil18] its size will grow from 12.6 billion dollars in 2017 to 20.1 billion dollars by 2023, with a compound annual growth of 8% per year. More than 30 countries have already started either assessment, design or deployment of MCC networks with the primary focus of replacing legacy systems[GSM14a].

As stated, many mission critical communications currently rely on PMR solutions, which are frequently exploited, operated and maintained by a single agency. This exploitation model has been questioned in some studies. For instance, in [FPSB13], the authors analyse the provision of the techno-economic drivers for future Public Protection and Disaster Relief (PPDR) systems. The authors from [FGMM16] identify the deployment costs and times, the spectrum availability and coverage, and the resilience and prioritization capabilities as the main issues to be considered when evolving current PPDR networks.

Governmental bodies have also been working on the provision of mission critical systems with mobile networks. The US government has made a considerable investment (7 billion dollars[5]) to support FirstNet [Kru17], an independent group that is implementing a national public safety network, with dedicated frequencies and currently based in LTE. The European Commission (EC) carried out a study [SFB14] to assess if commercial cellular networks could be used to support mission critical broadband. The study concludes that mobile communications could support mission critical broadband and also that, although it will be feasible to use a dedicated network, the main barrier will be the initial investment.

MCC users are also researching the use of mobile networks. For instance, the study [IDA15] compares the prices of GSM-R with TETRA, LTE and LTE rugged[6], the results

---

[5] Considered the starting point, experts estimated the total cost in tens of billion dollars[SFB14].

[6] It is a version hardening LTE, the main difference is an increased power redundancy to last several days according to the study.

	CAPEX (per site)	CAPEX (per user)	OPEX (% of CAPEX)
GSM-R (ETCS + Voice)	30K per Km of line	45K per cabin	10%
TETRA	1400K	900 handset 2000 per car	9%
LTE	76.38K	<100	20%
LTE Rugged	125.38K	<100	20%

Table 1.1: Cost comparison GSM-R, TETRA and LTE

are provided in Table 1.1, as can be noticed, LTE provides the lowest costs. Additionally, 3GPP has been standardizing some of the functionality that was missing in LTE, such as group communications or Mission Critical Push to Talk (MCPTT).

Standardization bodies were also targeting the inclusion of new functionality to support MCC [LM17]. For instance, during the development of this thesis 3GPP introduced some enablers (group communications, device to device communications, etc.) and then in Release 13 started to focus on the application domain in the standards with MCPTT and later on Release 14 Mission Critical Data (MCData) or Mission Critical Video (MCVideo). The 3GPP standardization track, for mission critical communications, during the last years, is depicted in Figure 1.1, we will discuss it in Section 3.

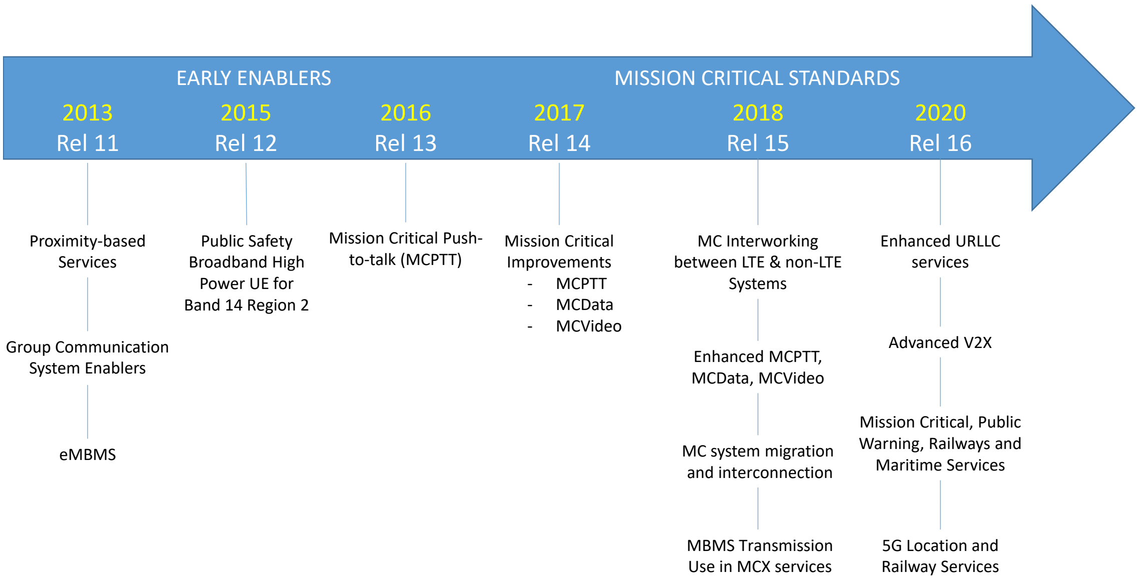


Figure 1.1: 3GPP Mission Critical Standardization

The arrival of the new 5G standards also brings some relevant functionality such us Ultra-Reliable Low Latency Communication (URLLC) or standalone mode in the base stations. Indeed, these new standards have common requirements with many critical systems. 5G is targeted to provide very low latency, improved reliability and availability, high bandwidths, etc.

Finally, there are also many research studies being conducted around the different aspects of MCC [BCD$^+$11][PLG$^+$18a][MMA19]. Some of them are focused on network services such as [SSA$^+$18], where the authors discuss on the allocation of an MCPTT service in the edge of the network as a way of improving service and increasing reliability by enabling deployment on base stations without access to the core network. Others are focused on end-to-end architectures, for instance, the authors of [PLG$^+$18b] proposes a holistic approach to support network slicing, multi-connectivity and end-to-end quality to support mission critical traffic reliability in 5G networks. We will provide more specific related work in each of the chapters of this book.

1.3 Contributions

We started our research by analysing the LTE network to evaluate its feasibility to support MCC. To do so, we picked a driver use case, railway communications, which has complex requirements both in terms of functionality and performance. We carried out an analysis that consisted of identifying services, requirements, and the existing solutions to replace GSM-R with a COTS mobile technology such as LTE[DZGPMG14a]. We proposed feasible architectures employing COTS LTE technologies to support all the gathered requirements, also identifying gaps that should be provided with external applications. Finally, we have provided a qualitative analysis of the most important KPIs employing standard equipment. These results are employed as a baseline for some of our results and methodologies and also as metrics to assess the feasibility of using mobile technologies to support MCC.

We have made contributions on the network design, prioritizing the use of standard technology, as a way of accelerating the adoption of the technology advances on mission critical scenarios. Our first proposal to reduce latency was based on a new component to be added to the standard architecture in order to improve its behaviour. This new component, the Fog Gateway, is fully compatible with standard LTE networks, reduces the end-to-end latency and reduces the signalling load in the networks[GPM16]. We have implemented several prototypes[GP17] and evaluated them using COTS equipment and instrumentation.

Nevertheless, we have also explored pathways to evolve existing networks and further improve their reliability and latency figures; we defined the GTP Gateway[GPM17] a component to be used in future mobile networks (or in current ones modifying the standards) to reduce unnecessary overhead in the provision of fog services. A network architecture has also been proposed, focused on closing the loop between applications and the network by providing an API [GPDZR$^+$17b] that could be used to setup application-defined Quality Class Indicators (QCIs) and to deploy fog services in our solutions.

To validate some of the proposed solutions, several improvements to a European research platform[GPMMR17] have been developed aiming to support the combination of different equipment, including standard network components, instrumentation or prototypes. We have developed new automation techniques to ease the integration of new Standard Commands for Programmable Instruments (SCPI) compliant equipment[GPRPRG+16]; measurement tools to obtain accurate results on the end-to-end Round Trip Time (RTT) and the split between segments; and deployments tools to roll-out Evolved Packet Core (EPC) on-demand and setup IP impairments in any of the interfaces of the network.

Due to the complexity of the full architecture, it is essential to include all the elements, including services and user equipment, so we have provided a methodology to select the best UE for a given application[GPDZR+17a]. We have also modified conformance testing instrumentation to support standard S1 interfaces, so we have connected the instrumentation with commercial core networks, but we still have full control over the radio channel and the base station stack.

In Chapter 7, we will provide a detailed discussion of the results and our contributions, but to provide an overview here, we can identify the following:

- Scientific publications, we have published three journals, indexed in the Journal Citation Reports (JCR), four book chapters, presented in six congresses and completed two master thesis.
- Tools, we implemented several tools to improve our experimentation platform and several prototypes to evaluate our low latency proposals.
- Projects, we have contributed to four European research projects and one national, cooperating with relevant stakeholders of the mobile communication market.

1.4 Rest of this book

Experimental platforms are covered in Chapter 2. To validate some of our proposals we combine COTS equipment with research prototypes. This combination is done to deal with the increasing complexity of the networks and stacks, the effect of any proposed change in them is more difficult to assess using simulations in each generation, so more complex testbeds have to be employed. We will provide an overview of existing testbeds, focusing in PerformNetworks[7] (formerly PerformLTE and now known as Triangle[8]), which is the testbed that we have employed to generate most of the empirical results of this thesis. We will describe the methodology and architecture of the testbed as well as our contributions to improve automation and the tools to support triggered network procedures, measurement analysis and development. Finally, we will outline our methodology to select the best UE for a given application.

Chapter 3 provides an overview of the basic LTE architecture to support voice and data services. We will start with a review of the standard architecture, analysing the role

---

[7] `http://performnetworks.morse.uma.es/performnetworks`

[8] `https://www.triangle-project.eu/`

of its components at the end-to-end service. Then, a characterization of its performance by analysing different KPIs is provided and the results are obtained highlight the relevant places where contributions could be made.

Different MCC architectures are described in Chapter 4, where we will analyse a use case, railway communications, to gather requirements for MCC. Then we propose a standard LTE architecture based on Release 12, and then we provide an updated version of the architecture, which takes into account the latest mission critical standards defined by 3GPP.

After analysing COTS architectures to support MCC, we also research possible latency improvements of the mobile networks for these communications in Chapter 5. We will first analysis qualitatively and quantitatively the latency provided by a commercial mobile network. Then, we will describe the Fog Gateway, a MEC network element that can be employed on standard architectures. The GTP Gateway, which is an evolution of the Fog Gateway that requires modifications of the standard network elements, is discussed later. For both solutions, we provide some figures based on emulated scenarios and an architecture to expose the functionality to third-party users.

Chapter 6 covers the implementation of a Fog Gateway prototype. We will first discuss on the different implementation approaches, to identify advantages and limitations that could be useful for other researchers. Then, we will provide details on the implementation of the prototype. Finally, we describe the results obtained with these prototypes, including the end-to-end latency of fog and cloud services and the effect on the rest of the traffic.

On the final Chapter 7, we will analyse the results obtained by this thesis, providing more details about our contributions and about the future work that can follow up this thesis.

Chapter 2

Experimentation Tools

Contents

2.1 Introduction **15**
2.2 State of the art **16**
2.3 PerformNetworks **19**
2.4 Contributions to PerformNetworks **21**
  2.4.1 Improving Automation 21
  2.4.2 Tools 24
  2.4.3 UE Selection methodology 31
2.5 Conclusions **33**

Synopsis

The first stage of this thesis consisted of the analytical evaluation of existing mobile networks to determine if they were able to support mission critical services and also to identify where could they be improved. In the context of the thesis, we have developed several tools, methodologies and prototypes. One of the fundamental pillars of our approach has been the validation of our proposals by empirical evaluation. To do so, we have used PerformNetworks, which is an experimental platform that combines COTS, with research elements and telecommunication instrumentation. We have also expanded this testbed, improving its automation as well as its capabilities to communicate with standard equipment. Several measurements tool have been deployed, enabling the possibility of obtaining more detailed data from the elements of the platform. In this chapter, we provide an overview of the existing experimental platforms for mobile communications, we describe the testbed and our experimentation methodology and provide some details on our contributions to experimentation platforms.

2.1 Introduction

One of the distinct characteristics of our approach is the verification of our proposals by evaluating realistic scenarios. The core platform of our methodology has been PerformNetworks[1], which is a testbed building and maintaining an experimentation eco-system, managed by the Mobile Networks and Software Reliability (MORSE) group at Universidad de Málaga (UMA). The primary objective of PerformNetworks is the provision of a cutting-edge and realistic environment for all the stakeholders of mobile communications, from researchers to operators or developers. Its unique combination of COTS equipment, instrumentation and open-source platforms provides the ideal environment to validate some of our proposals.

PerformNetworks is part of the Future Internet Research Experimentation (FIRE) community, an initiative by the European Commission aiming to enable experimentally driven research for the European industry. The testbed has been part of several research projects, which have served as a driver of this thesis contributions. For instance, in Tecrail[2], a national research project, we explored the provision of ERTMS communications based on LTE. In the context of the project, we proposed an architecture, different tools to evaluate realistic railways conditions and a pilot to evaluate the feasibility of the use of LTE to support high-speed trains communications. Fed4Fire and Fed4Fire+[3] along with FLEX and Triangle[4] are European research projects that bring third parties to the testbed to exploit it, and several tools developed for this thesis were exposed in PerformNetworks to be available for these projects.

Various combinations of the elements of the testbed were used to support the different stages of our research. A baseline of the current behaviour of mobile networks was established by combining conformance testing emulators, which were extended to support communication with a standard core network, with an LTE core network and UEs, the obtained results are described in Chapter 3. The emulators and open-source tools were employed to develop and validate some of the tools that are described later in section 2.4. Finally to validate some of our proposals to improve latency and reliability several prototypes were developed, guaranteeing their conformance to the standards using the elements of the testbed, and combined with them to generate the different results.

Along our research on MCC, we also generated several contributions to the testbed itself. For instance, to improve automation, we designed an architecture for integration were we allow the testbed patron to integrate new SCPI based instrumentation by merely define an Extensible Markup Language (XML) that is used to map experimenter functionality into the appropriate instrument commands. Using this architecture we extended the testbed introducing new components. We implemented elements to provide

---

[1] `http://performnetworks.morse.uma.es/`

[2] `https://www.uma.es/tecrail/`

[3] `https://www.fed4fire.eu`

[4] `https://www.triangle-project.eu/`

artificial IP impairments in any of the testbed interfaces, which can be used to emulate the behaviour of different transports. We also implemented an SCPI compliant control software that enables the automatic deployment of the EPC network elements and the trigger of signalling procedures on demand.

In [GPDZR^{+}17a], we proposed a methodology to select the best UE for a given application, which is based on generating several KPIs (designed to identify the most important scenarios for different type of applications). We do so by running end-to-end tests to fully characterize the UE data and control planes. We then combine the information obtained with other characteristics such as cost or price, which can be used to compare different UEs to help decide which one will provide better performance in the target scenario.

Also, different tools to generate or analyse measurements were implemented. For instance, we created tools to analyse ping traces to estimate RTT across all the elements of the EPC or tools to analyse control plane captures in order to calculate the time consumed by the different procedures, which were triggered by our control software. Also elements to ease the development of prototypes were develop, such as a tunnel encapsulation tool which can capture traffic in real-time and transport it over General Packet Radio Service Tunnelling Protocol (GTP) tunnels.

In this chapter, we will discuss the state of the art of experimental platforms for mobile communications. Then we will describe the PerformNetworks testbed, providing an overview of its architecture and the methodology. Finally, we will discuss our contributions to the testbed, which includes architectures to improve automation, some measurement tools and a methodology to select the best UE for a given application.

2.2 State of the art

Evaluation of mobile networks can be done with simulations, emulations or commercial elements depending on the maturity of the technology. Simulations are typically used to support experimentation on new technologies. These approaches have some advantages as they act as inexpensive early enabler. However, it also has some limitations, mainly simulations work well to analyse small parts of the network, e.g. concrete technologies on the physical layer, scheduling algorithms, etc. but fails to simulate system behaviours, as it is difficult to introduce all the interactions on a network. One clear example is on the MAC scheduling algorithms, there is much different research covering the issue, but the actual implementations are limited to a few algorithms, mainly because the scheduler allocations have to be done in less than 1ms, which poses a hard constrain to be filled.

Currently, one of the most popular simulation tools is Matlab, which provides a 5G toolbox[5]. This software can be used to process and analyse data such in [SAHI18], to validate designs as in [SJK^{+}18] or in system-level simulations such in [IRB15]. The

---

[5] `https://www.mathworks.com/products/5g.html`

Riverbed Modeler[6] (formerly OPNET) was very popular to support LTE evaluation at system-level, as it was designed in conformance to the 3GPP Release 8 specifications and provided a statistics toolset. The ns-3 emulator is also widely used by the research community, and there are several modules focused on LTE/5G, being the most popular LENA [MBR$^+$18], which provides an eNB-EPC emulator. This open-source emulator has been used to developed modules to support other technologies such as visible light communications [ARK$^+$17] or satellite links [CL16].

Another aspect to be considered is the usage of instrumentation, which is generally used for conformance and design verification, this equipment provides typically cutting edge technology that works with (future) COTS UEs. Some frequently used vendors are Keysight[7], for instance, in [CKC$^+$17] or [RDM16], or Rohde & Schwarz[8] as in [LMRZ16] or [KDR$^+$17]. These vendors provide test sets for mobile technologies from 2G to 5G, which can be potentially used to evaluate mobile communications. The main issue with this approach is that the test sets usually do not support integration with EPCs, their cost is high, and the latest UE might not be available commercially, having to be obtained by collaboration agreements with vendors, which are ordinarily challenging to sign for researchers.

The use of Software Defined Radio (SDR) is very relevant as they enable early prototyping as well as full implementations. One of the most popular open-source implementations for mobile networks is Open Air Interface (OAI)[9], which provides SDR implementations of UE and Evolved Node B (eNB), as well as support for a small EPC. OAI is now also used in the Mosaic5G an initiative targeting the development of 5G open-source networks, for instance in [HNS$^+$17a] they describe their low latency MEC implementation or in [KNH18] they cover their orchestrator for network slicing. SRS[10] also provides an open-source implementation of a base station and a UE with support for the VOLK[11] acceleration libraries to improve performance, some examples of usage can be found in [GZW18] and [GSN$^+$17], which provides results for a system to combine LTE Unlicensed (LTE-U) with Wi-Fi. The authors of [GWL$^+$17] provide a comparison of the performance of both platforms in terms of memory usage, time consumed by common procedures and throughput.

Experimentation with real platforms is frequently based in testbeds built ad hoc for a specific scenario, the main advantage is that this approach can be used to address a particular research problem, but it burdens the reusability of the infrastructure as well as the access by other researchers. For instance, the authors of [SKCK16] deployed several LTE nodes in a railway line and evaluated the performance of different services when compared to GSM-R. In [HMP$^+$16] an eMBMS field trial is employed to characterize the performance of multicast services both with COTS UEs and SDR solutions.

---

[6] `https://www.riverbed.com/gb/`

[7] `https://www.keysight.com`

[8] `https://www.rohde-schwarz.com`

[9] `https://www.openairinterface.org/`

[10] `https://www.softwareradiosystems.com/`

[11] `http://libvolk.org/`

On the other hand, there are also other experimental platforms as PerformNetworks, which provide access to sophisticated equipment and consultancy services to third party researchers and companies. These platforms expose functionality that can be used by third parties, being the main limitation that the functionality present can be too general for specific experiments or require too much development efforts. We contributed to [Aut16], which provides an overview of the existing experimental platforms for 5G mobile technologies. Many of the testbeds that are described there are part of the FIRE initiative, funded by the EU. One valuable testbed in the context of the FIRE platforms is ORBIT, as it provided ORBIT Management Framework (OMF), which is a framework, adopted by many FIRE projects, to offer experimenters a standard interface.

There are three essential testbeds, besides PerformNetworks[12], covering the mobile domain experimentation in FIRE: w-iLab.t[13], FUSECO playground, and NITOS. The w-iLab.t testbed is operated by IMEC and offers SDR two commercial EPCs and femtocells, cognitive radio implementations and a Wi-Fi deployment. An example of usage can be found in [VVDM17] where the authors evaluate an indoor position strategy on an industrial environment. The FUSECO[14] playground testbed provides several toolkits for mobile experimenters such as the OpenIMS, Open5GCore or Open5GMTC. These two later toolkits were employed in [CQC^{+}18] to evaluate the performance of a virtualized Open5GCore, using different hypervisors, under machine type traffic. The NITOS[15] testbed provides a combination of multiple technologies, such as Wi-Fi, WiMAX, LTE, SDR or SDN, their approach is described in [MZK^{+}15].

The key difference of PerformNetworks with others testbed is the possibility of using customized instrumentation, which can be used in full end-to-end deployments, and that can be combined with any standard equipment. Most of the testbeds feature either commercial base station or open SDR solutions on the radio side. The use of commercial base stations limit the quality of the results, it is difficult to provide a stable environment and generating channel conditions is hard, as it requires the use of channel emulators along with equipment to limit the output power of the base station. SDR solutions can be connected to channel emulators more easily but the existing implementations are still limited in terms of coverage of the standards and working scenarios. On the other hand, using the testbed modified instrumentation can produce very realistic results as it can be combined with standard equipment but allowing the use of channel emulation, the configuration of the full radio stack, and the use of triggered procedures. All these increase the number of possible scenarios and improves reproducibility of the tests. More details on the testbed and its functionality will be provided in the following section.

---

[12] `http://performnetworks.morse.uma.es/`

[13] `https://doc.ilabt.imec.be/ilabt/`

[14] `https://www.fokus.fraunhofer.de/go/en/fokus_testbeds/fuseco_playground`

[15] `https://nitlab.inf.uth.gr/NITlab/nitos`

2.3 PerformNetworks

The PerformNetworks testbed is an experimental platform that is designed to support realistic experimentation on mobile networks. The idea behind the platform is to support the different stages of research by combining different equipment. On the first phases of research, the testbed can be used to gain insights on the behaviour of mobile networks under different channel conditions and configurations, frequently employing the SDR solutions or the conformance testing equipment. Additionally, these components can be used to support the design and verification of prototypes or modified elements on the network. On maturer stages, the indoor deployment can be used to have a controlled and functional testing environment as well as to support interoperability testing with COTS equipment, which will lead to the final stage where experimenters can test their solutions on field deployments.

Figure 2.1 depicts the PerformNetworks architecture; we have grouped the different components of the testbed in blocks:

- User Equipment, which includes the different elements that can be used as UE as well as instrumentation and network clients.
- Channel, which comprises components that can be used to introduce Radio Frequency (RF) channel impairments.
- Radio Access, which consist of different Radio Access Technologys (RATs) equipment and platforms.
- Transport and Services, which include the servers and switching technology to interconnect the elements of the network.

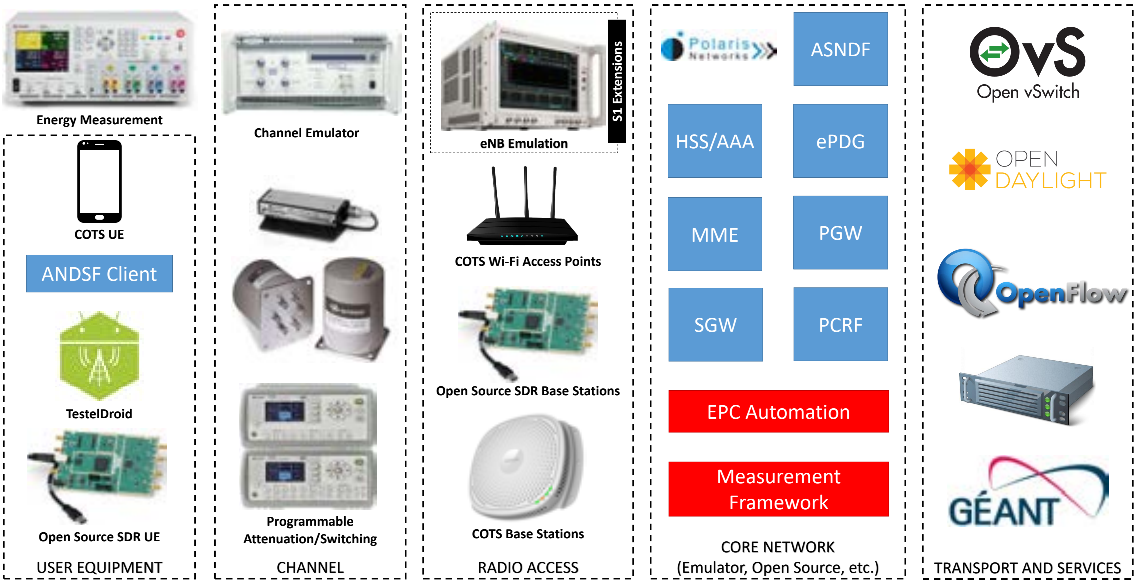


Figure 2.1: PerformNetworks architecture

The testbed allows different UE configurations from COTS UE of different vendors, which can be USB dongles or phones, to SDR UE, which will be generally implemented with srsUE or OAI. The TestelDroid tool is a custom tool that can be used to monitor the UE. The most relevant functionalities that it can provide are the sniffing capabilities, the Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ) measurements and the estimation of MOS for voice calls. More details on this tool can be found in [ÁlvarezDMR12b] and [ÁlvarezDMR12a] and Section 3.3.3.

Channel emulation instruments can be used between the UE and the base station. A Spirent equipment featuring fading and Additive White Gaussian Noise (AWGN) is available as well as programmable attenuators which can introduce controlled attenuation on demand, for instance, to force the signal strength from one of the cells become higher than other. Additionally, energy measurement equipment can be used to monitor the energy consumption of the UE.

On the radio side, the testbed features Wi-Fi access points, LTE and LTE-A conformance testing equipment, commercial base stations and SDR devices. The conformance test sets are the Keysight T2010 and UXM, which have been modified to implement an S1 interface, so they allow communication with standard EPCs. This equipment also includes channel emulation, which includes fading and AWGN generation. The SDR units usually are used with OAI eNB, but it allows other implementations such as srsLTE or OpenBTS[16]. There are also available some standard COTS LTE eNB on band 7 (2.6GHz) from Athena Wireless (now part of Google Fiber) and Nokia.

For the core network, the experimenter can use the Polaris EPC emulators, which comprise Mobility Management Entity (MME), PDN Gateway (PGW), Serving Gateway (SGW), Home Subscriber Server (HSS), Policy and Charging Rules Function (PCRF), Access Network Discovery and Selection Function (ANDSF) and Evolved Packet Data Gateway (ePDG). The emulators provide standard LTE functionality, implement the 3GPP Release 13 and enable on-demand deployment as well as triggered procedures while still maintaining carrier-grade performance. General Purpose Processor (GPP) compatible implementations, such as the open-source solutions OAI and Open Mobile Evolved Core (OMEC)[17] can also be deployed on the testbed. There are also a set of measurement and automation tools that can be used to obtain different KPIs, which will be described later.

The testbed also features GPP servers where services can be deployed; for instance, experimenters could exploit the existing VoIP service or deploy IP Multimedia Subsystem (IMS) such as ClearWater[18] or Kamailio[19]. The interconnection between the backhaul and the EPC and the services can be done using different Software Defined Network (SDN) switching deployments, normally Open vSwitch (OVS) with a controller

---

[16] `http://openbts.org`

[17] `https://www.opennetworking.org/omec`

[18] `https://www.projectclearwater.org/`

[19] `https://www.kamailio.org`

such as OpenDaylight[20] or ONOS[21]. We provided more details on the equipment and methodology available on the testbed in [DZGPMG15] and [GPMMR17].

Currently, PerformNetworks has evolved into two platforms, TRIANGLE[22] described in [DZPB$^+$18], which holds the components to setup experiments indoor (such as the conformance test sets or the power analyser), and 5Genesis [23] [KTG$^+$18], which support outdoor experimental deployments.

2.4 Contributions to PerformNetworks

In this section, we describe some of the tools that have been generated during the execution of this thesis. We first provide an overview of the tools to increase the usability of the platforms; then we describe the implementations done to generate measurements; finally, the methodology to select a UE for a given application is covered.

2.4.1 Improving Automation

To enable automatic deployments and triggering procedures from an application, we provided several automation tools to the testbed, being the first one an SCPI integration tool. The SCPI standard defines messages (command and responses) of a language to control compliant instruments. These commands are text-based and allow configuring and querying the device, providing a simple way of creating text scripts to handle the instrument more comfortably. To ease the integration with the automation architecture of the testbed, we implemented an XML abstraction tool.

PerformNetworks employs the OMF[ROJS10] architecture to offer a standard interface to experimenters. OMF has been deployed in many of the FIRE testbeds and in Global Environment for Network Innovations (GENI)[24], allows the design and execution of experiments using a standard interface. The definition of the experiments is done using OMF Experiment Description Language (OEDL), and the measurement and results can be retrieved with the OMF Measurement Library (OML). The OEDL scripts are interpreted by the experiment controller, which will talk with the resource controllers of each of the connected resources. To support new equipment integration, we designed a generic SCPI Resource Controller that was implemented by the research team and that can be used to expose in OEDL any SCPI compliant instrument.

Figure 2.2 depicts the architecture; the central element is the XML definition, which provides a mapping between functionality, SCPI commands and aliases. This definition serves as input to an OEDL generator that will generate a reference OEDL script. OEDL is the language used by OMF to define the experiment execution, the measurements

---

[20] `https://www.opendaylight.org/`

[21] `https://onosproject.org/`

[22] `https://www.fed4fire.eu/testbeds/triangle/`

[23] `https://5genesis.eu/malaga-platform/`

[24] `https://www.geni.net/`

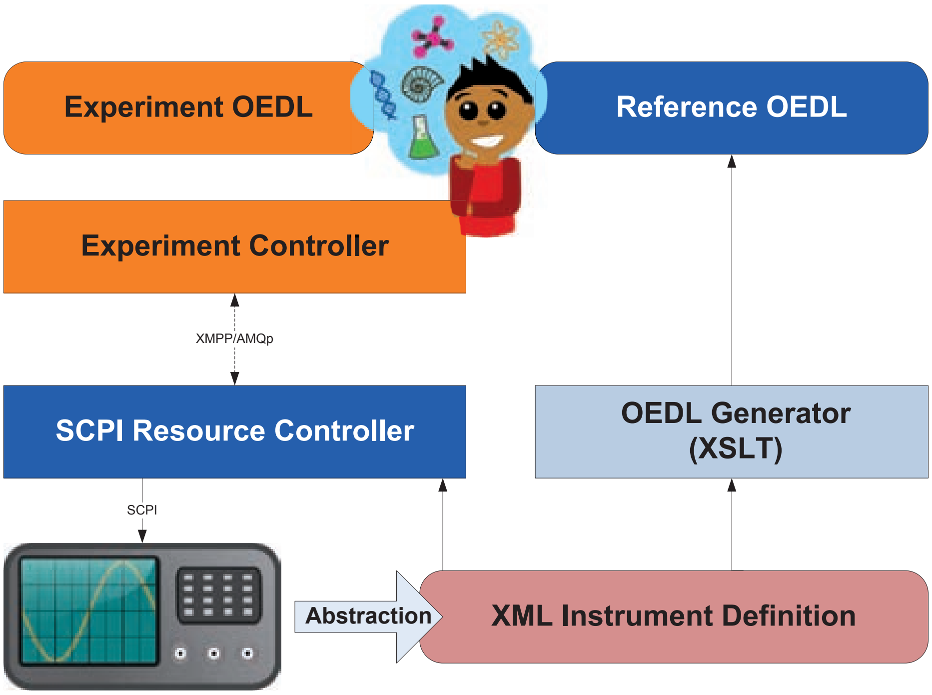


Figure 2.2: XML-SCPI Integration Architecture [GPRPRG$^+$16]

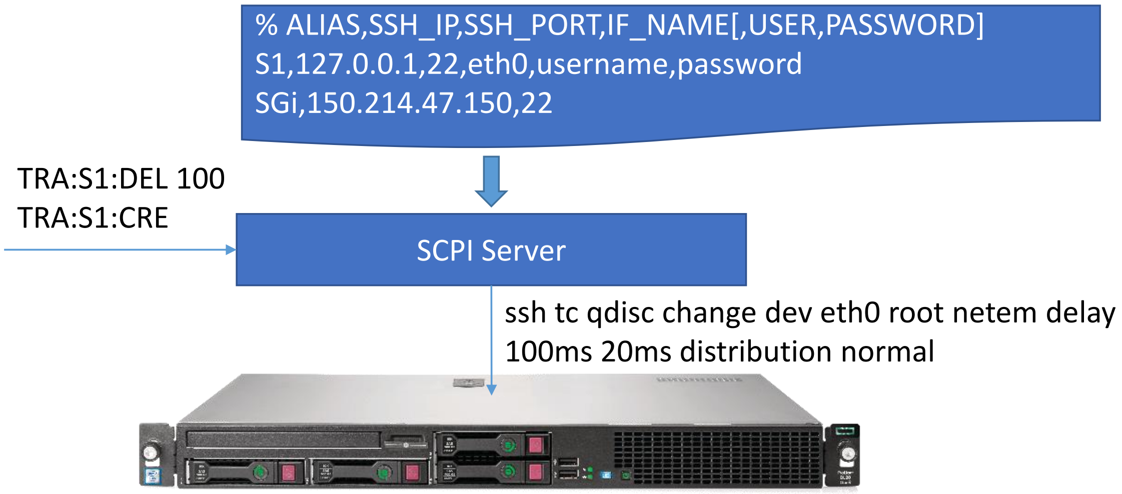
and the plots. The definition is also used by the SCPI interpreter that will translate functionality aliases into the appropriate SCPI commands. Experimenters can then use the reference OEDL to define their experiments in an OEDL script, which can be then fed to the Experiment Controller to be executed. Finally, the results can be extracted from the instruments using an OML collection server. We provided more details about this architecture in [GPRPRG$^+$16] and [DZRPGPM16].

In addition to the XML-SCPI framework, we also provided SCPI interfaces for part of the functionality in the testbed. For instance, a tool to automatically deploy the Polaris EPC emulators was implemented, along with a framework to activate artificial IP impairments in network interfaces remotely. The design and implementations were done by us in the context of a European project, TRIANGLE[25] [CMD$^+$16] [DSM$^+$17], which was an initiative focused on the provision of tools to benchmark and evaluate mobile applications on 5G networks.

One of the tools that was integrated with the XML-SCPI framework was the artificial impairment tool, which allows the experimenter to add latency in any of the interfaces of the test. This functionality could be used to emulate the affect of different components not present in the testbed, such as the IPSec encryption, the backhaul technologies, the domain changes in the cloud services, etc. The tool can be configured to generate different statistical distributions to match the experimenter requirements.

---

[25] `https://www.triangle-project.eu/`



```
% ALIAS,SSH_IP,SSH_PORT,IF_NAME[,USER,PASSWORD]
S1,127.0.0.1,22,eth0,username,password
SGi,150.214.47.150,22
```

```
TRA:S1:DEL 100
TRA:S1:CRE
```

```
ssh tc qdisc change dev eth0 root netem delay
100ms 20ms distribution normal
```

Figure 2.3: Remote Impairments Architecture [MGM+18]

Figure 2.3 depicts the workflow for the Impairment SCPI Server. The tool is based in a Linux kernel module named netem[26], which is characterized in [JLHW11]. The server accepts an interface definition file, which contains a list of interface aliases and the Secure Shell (SSH) connection details (IP, port, interface name and, optionally[27], user and password). The SCPI commands are sent specifying an alias interface, and when received, the SCPI server will open an SSH connection to the server and execute the necessary netem commands. The system provides support to enable and disable artificial latency, errors, packet duplication and packet reordering in any of the network interfaces of the testbed.

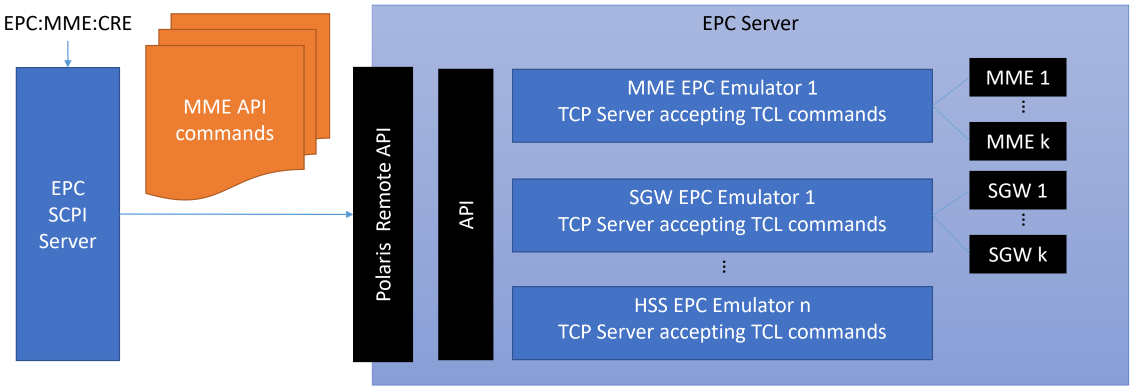


Figure 2.4: EPC Automation Architecture [MGM+18]

The architecture employed to support an SCPI interface to trigger the deployment of the EPC elements and the network procedures is depicted in Figure 2.4. The EPC SCPI Server accepts commands addressed to one type of emulator, e.g. *EPC:MME:CRE*, *EPC:MME1:*, will create a default configuration for the MME, and will generate and

---

[26] https://wiki.linuxfoundation.org/networking/netem

[27] The use of ssh keys is preferred.

send the appropriate commands to the emulators. The interface will also expose some triggered network procedures for the created network elements. The server controls the emulators using a proprietary Polaris TCL interface and is designed to support multiple instances of the emulators by indexing them.

The EPC SCPI Server supports commands to deploy MME, SGW, PGW, HSS, PCRF, ePDG and ANDSF. On the activated signalling procedures, the user can also trigger the following:

- MME Detach with cause 0 and reattach required indication, which can be used to trigger a new attach procedure after the detach.
- MME UE Context Release, used to trigger Service Request procedures.
- MME Paging.
- PCRF Dedicated Bearer Creation, with the specified Quality of Service (QoS) parameters.
- PCRF Dedicated Bearer Release.

For all the procedures, the International Mobile Subscriber Identity (IMSI) of the user has to be provided. The commands of the MME are sent with a delay of 100ms; this is done to allow the measurement framework to be prepared to capture data.

2.4.2 Tools

We design and implemented several tools to generate more measurements from the testbed and the implemented prototypes. The tools, named PingAnalyzer, ControlPlaneAnalyzer and TunnelTester, allow us to extract information from the signalling links and ease the development of new network elements. PingAnalyzer was implemented to help us characterize the user plane latency split on the different segments of the network. We also provided ControlPlaneAnalyzer, which is a component to obtain the time consumed by the control plane procedures. S1Database is a prototype to build a bearer information database from the S1 interface and to help on the development of components based on GTP tunnels was also developed. Finally, TunnelTester can be used to encapsulate IP packets into GTP tunnels, which removes the need for using actual base stations to test some of our network elements.

PingAnalyzer is an Internet Control Message Protocol (ICMP) trace analyser that can be used to evaluate the data plane latency split. The tool requires a PCAP trace file and it can extract the ICMP information from IP packets but also from GTP and a proprietary transport protocol on top of Packet Data Convergence Protocol (PDCP) employed by the conformance testing equipment. The involved protocols are depicted in Figure 2.5, we implemented dissector for the ones in orange (IPv4, UDP, Proprietary Transport, GTP and ICMP).

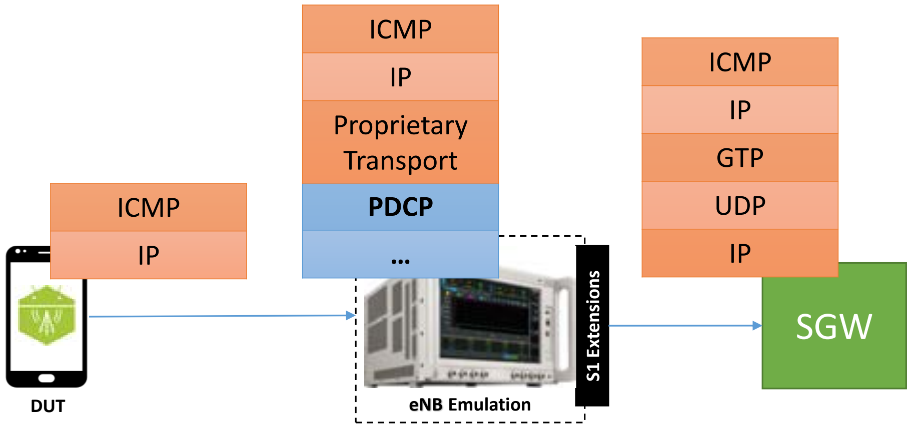


Figure 2.5: PingAnalyzer dissected protocols

Figure 2.6 provides a diagram of the dissector library. There is a basic class to support different encoding/decoding methods, encoding support depends on the protocol, typically we do the XML encoding for ASN.1 based protocols while for the rest we use binary only. There are other dissectors, not depicted such as the proprietary dissector.

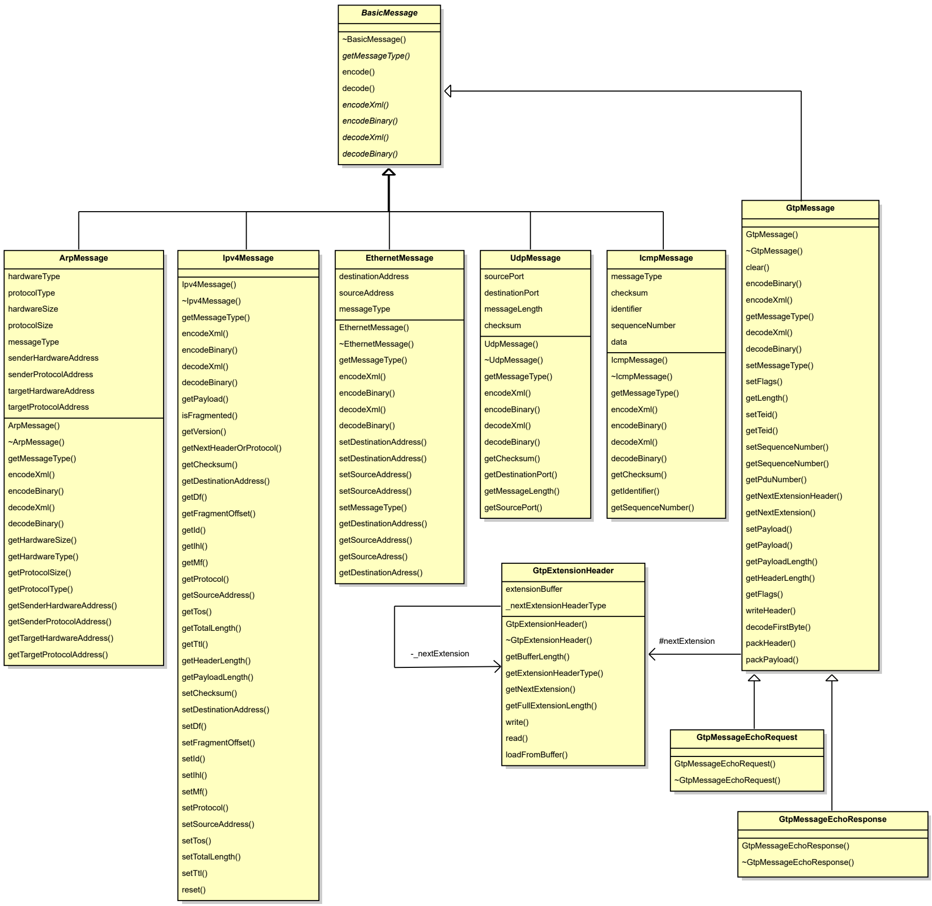


Figure 2.6: Dissectors UML Diagram

The tool estimates the RTT by comparing the PCAP timestamp of the ICMP echo and response packets. To use the tool we generate a continuous ICMP flow of packets, we capture the packets on the relevant interfaces of the network and pass them to PingAnalyzer, which will provide an estimated RTT of each of the segments.

For the control plane procedures, ControlPlaneAnalyzer, which supports S1 and Non-Access Procedure (NAS) dissectors, was implemented. The tool will accept PCAP traces and will look for the S1 messages and for the any NAS linked information. The implementation relies in different libraries, one hand the previously described dissector library, which is used as base, a NAS dissection library and a S1 library. Figure 2.7 provides an overview of the libraries employed in ControlPlaneAnalyzer. We have depicted in blue the ones developed by the author, in green the ones written in collaboration with other members of the team and in orange the external ones.

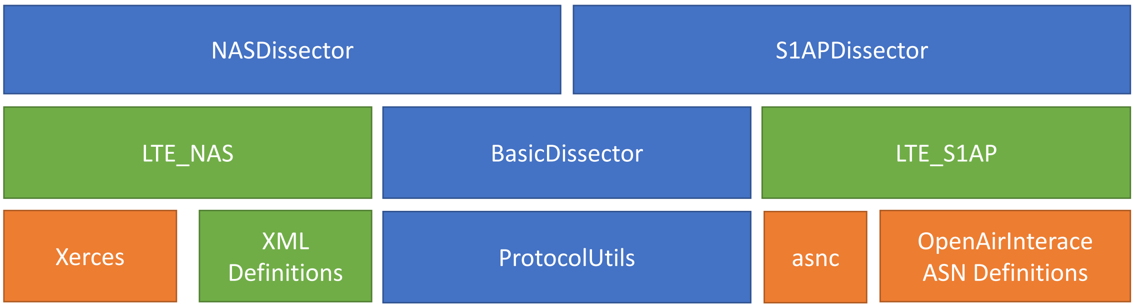


Figure 2.7: NAS and S1AP Libraries

The NAS library was implemented as part a collaboration agreement with Keysight Technologies by Jose Alberto Fernandez Prat and the author of the thesis. The tool relies on XML definitions that represent each of the messages of the protocol, which are handled with the Xerces libraries[28] along with an in-house implementation of binary coding/encoding tool. There are other ways of implementing the NAS protocol, for instance, the Transfer Syntax Notation One (TSN.1)[29] approach is interesting. TSN.1 is a formal notation focused on a transfer syntax, this is on definitions of protocols that are made in binary, such as IP, TCP or NAS. The main issue is that the notation is copyrighted, so the only implementation available is the one sold by Protomatics. Open source alternatives are available in the previously mentioned SDR implementations (srsUE and OpenAirInterface), being their main drawback that their coverage of the specifications is limited.

Finally, we also have an S1 dissection library, which is based on the LTE_S1AP library, which was implemented by Leticia Lavado Muñoz, who also did modifications in the asn1c [30] Abstract Syntax Notation One (ASN.1) compiler, and the author of this

---

[28] `http://xerces.apache.org/`

[29] `http://www.protomatics.com/`

[30] This is the most extended open source ASN.1 compiler `https://github.com/vlm/asn1c`. Originally developed by Lev Walkin offers a compiler able to support code generation and coding/decoding of messages.

thesis. The main challenge of implementing the S1AP protocol is that it uses some characteristics that were[31] not supported in asn1c, namely Information Objects and aligned Packed Encoding Rules (PER) encoding. To overcome this issue we employed the OpenAirInterface approach, on one hand they provided a patch to support aligned PER, on the other the modified slightly the definitions replacing the information objects. There are also commercial tools witch full support of the ASN.1 specification and dedicated modules for the 3GPP definitions, some vendors are OSS Nokalva[32] and Objective Systems Inc.[33]. The S1APDissector is implemented on top of an abstraction of LTE_S1AP to allow replacement of the underlying ASN.1 compiler.

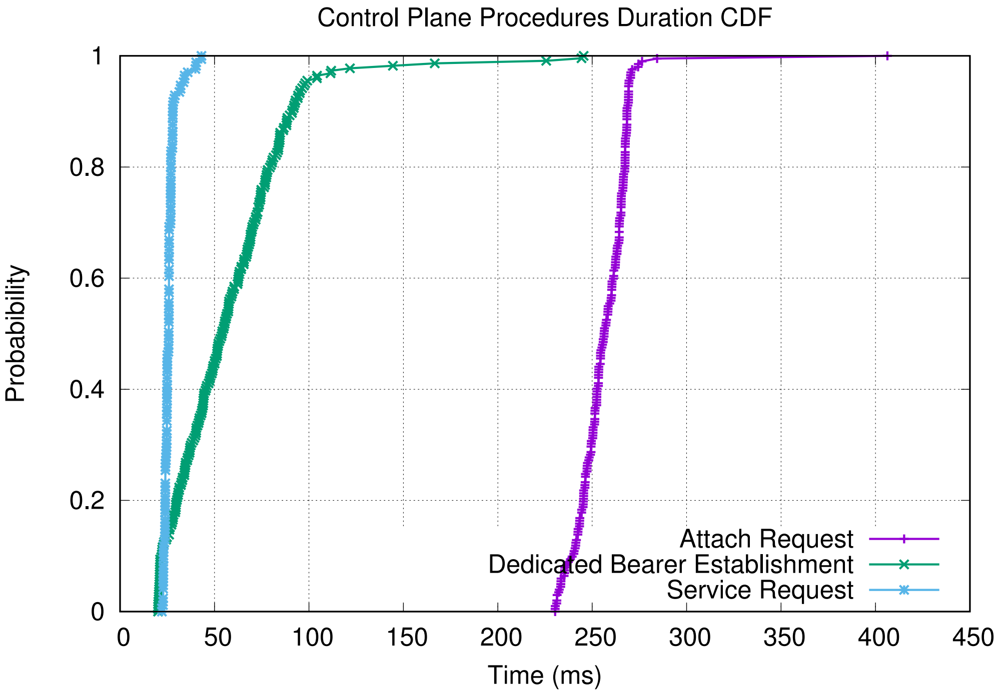


Figure 2.8: Time consumed by signalling procedures [GPDZR+17a]

In order to generate sufficient samples, to produce statistically relevant results, the EPC automation tools described in the previous section was employed. We first setup the end-to-end setup which consisted of the standard LTE EPC architecture, the T2010 emulator that acted as the base station and different UE. Then we use the tooling to trigger signalling procedures to generate a large number of network samples. The supported procedures are the following:

- Service Request. The time is measured between the NAS Service Request Message and the S1 Initial Context Setup Response message. The messages are generated triggering UE Context Release Request and then Paging messages.

---

[31] In 2020 there is an ongoing pull request in the project to support these functionalities.

[32] `https://www.oss.com/`

[33] `https://obj-sys.com/`

- Attach Procedure. The time is measured between the NAS Attach Request and the NAS Attach Complete messages. The procedures are generated triggering Detach messages with the reattach flag enabled.

- Dedicated Bearer Establishment. The time is measured between the NAS Activate Dedicated EPS Bearer Context Request and the NAS Activate Dedicated EPS Bearer Context Accept. The procedures are generated by triggering Dedicated Create Bearer Request and Release Dedicated Bearer.

The required messages will be triggered while capturing on the S1 interface in the MME. Then the tool to calculate time will be used over the PCAP file, and the median time will be generated. Figure 2.8 depicts an example that was generated during one of the pilots organized in the testbed.

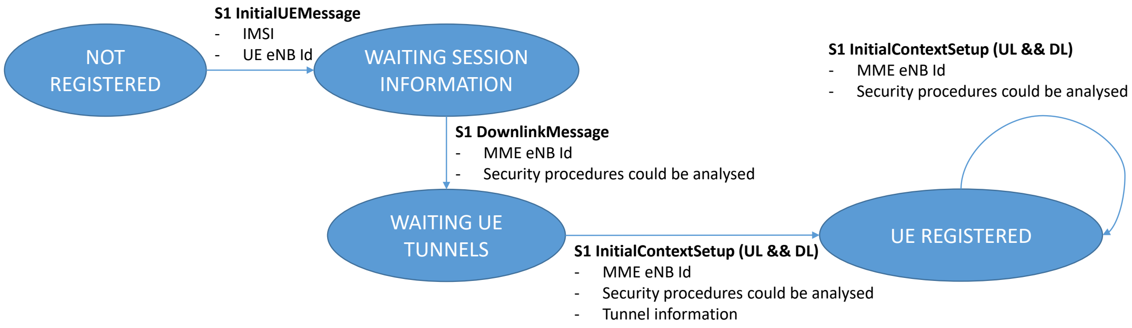


Figure 2.9: S1DatabaseGenerator state machine [GDM+18]

The same dissectors employed for the ControlPlaneAnalyzer are employed in the S1DatabaseGenerator. The tool was a proof of concept that was created to implement data plane acceleration in the context of the European project Q4Health. The objective of the tool was to generate a database of the user with their GTP settings so a central controller could be used to prioritize the tunnel traversing different IP domains. The generator analyses real-time traffic from a remote interface (using the remote capture capabilities from Winpcap[34]) and analyses the S1 messages. S1DatabaseGenerator state machine is depicted in Figure 2.9; the tool will look for:

- Identity messages (transporting the IMSI)

- Initial Context Setup messages (to get the default bearer information)

- Context Release Messages (to determine when a context has been released)

- E-UTRAN Radio Bearer (E-RAB) Setup (to get the dedicated bearers)

- E-RAB Release messages (to know when bearer has been released)

---

[34] The remote capture capabilities were introduced in Winpcap (`https://www.winpcap.org/`) but can be compiled for Linux systems.

For each of the bearers, the database will store information such as the GTP Tunnel Endpoint Identifier (TEID), the transport address, the direction, E-RAB id, priority level, Access Point Name (APN), etc. This information can be used to identify the default and any dedicated bearers of a particular UE inside the GTP flows. More details on this approach can be found in [GDM+18].

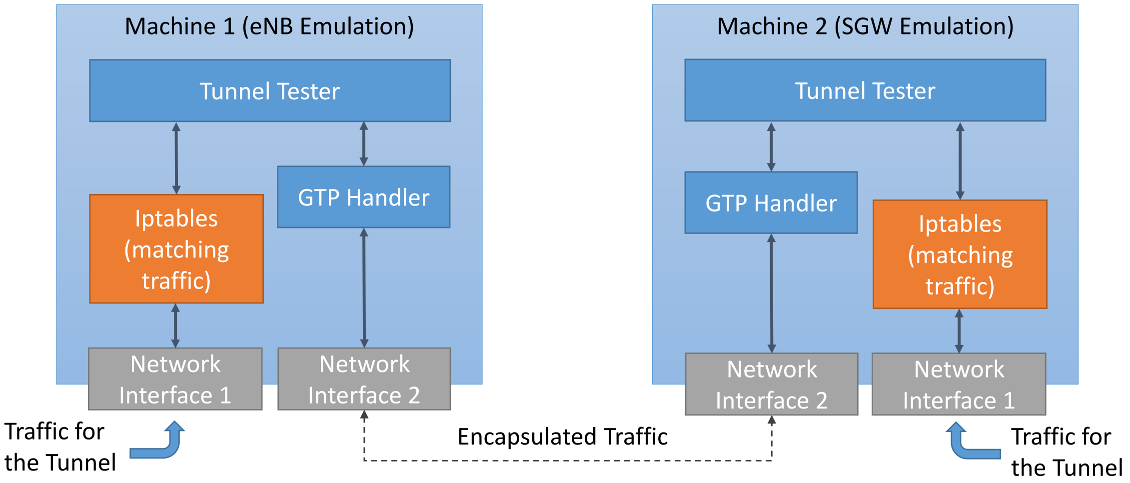


Figure 2.10: Tunnel Tester [GP17]

To ease the development of data plane acceleration prototypes, we implemented a tool to generate GTP traffic from real-time user IP packets, named TunnelTester. TunnelTester is implemented in python and based on the use of libnetfilter_queue[35], which is a library that can be used to capture packets queued from the kernel by iptables. Figure 2.10 depicts a diagram of how the tool works.

An iptables rule has to be defined to decide which traffic is going to be encapsulated, e.g. listing 2.1 defines a rule that will capture all the traffic going to 10.10.0.128/25 on interface eth1 and put it in queue id 1.

Listing 2.1: IPTables example

```
#!/bin/bash
iptables -A FORWARD -i eth1 -p all -d 10.10.0.128/25 -j NFQUEUE --
    queue-num 1
```

Every traffic matched by the rule will be sent to the TunnelTester library that will add GTP headers and send it over the GTP Handler. Listing 2.2 has the implementation of the loop called each time libnetfilterqueue send a packet to our application. The application will add the GTP header, send it over a UDP socket towards the other peer and drop the packet in iptables. In case of error the application will accept the packet, which means that the copy stored by iptables will be processed by the operating system.

---

[35] `https://netfilter.org/projects/libnetfilter_queue/`

Listing 2.2: IPTables Packet processing

```
def packet_processor(self, pkt):
   try:
      # First get the data from the packet
      data = pkt.get_payload()
      # Append a GTP header and send the data over an UDP socket
      gtp = self.tunnel_parser.append_tunnel_tester(data)
      if gtp is not None:
         # Send over UDP packet
         self.udp_socket.sendto(gtp, (self.peer_ip, self.peer_port))
         for i in range(0, self.copies):
            self.udp_socket.sendto(gtp, (self.peer_ip, self.peer_port))
         # If the packet is sent drop the iptables copy
         pkt.drop()
      else:
         # If a GTP header cant be appended send the packet as it is
         pkt.accept()
   except Exception as e:
      # On exception we accept the packet
      pkt.accept()
```

Listing 2.3 provides the calls when a GTP packet is received. First we receive data from the UDP socket, then we remove the GTP headers to obtain the payload. Then we check if we have payload[36], and, if so, we send the payload to ourselves using a raw socket.

Listing 2.3: IPTables Packet processing

```
def read_gtp_packet(self):
   # Get the data from the UDP socket
   data, addr = self.udp_socket.recvfrom(4096)
   # Remove the GTP header
   payload = self.tunnel_parser.remove_tunnel_header(data)
   # Inject the data to do so check if the packet is GTP
   if payload:
      self.raw_socket.sendto(payload, (self.local_peer, 0))
```

This application allow us to quickly emulate end-to-end GTP communication with real user traffic, which enables the possibility of extracting realistic measurements, e.g. we could send an iperf flow over the GTP tunnel to measure the maximum throughput or a continuous flow of ICMP packets to estimate the delay.

---

[36] In case we receive an Echo Request or Response message there will be no payload.

Following the same approach, we also developed a tool to introduce redundancy at IP level, named Redundancy Tester. The tool was installed on the UE and per each IP packet generated a configurable number of copies of the packet were generated. On the base station side, only the first copy of the packet (based on IP id) was delivered, the rest were dropped silently. With the tool, we were able to determine how many copies, were required to maintain a certain level of latency.

2.4.3 UE Selection methodology

In [GPDZR+17a], we explored the support of wearable mHealth applications following a holistic approach, trying to optimize all the elements of the network end-to-end. One of our proposals was the definition of a methodology that can be used to identify the best UE for an application. This methodology can help as different UEs provide different KPIs, even when they are on the same device category. It is not restricted to mHealth applications and is useful to evaluate the behaviour for any scenario, including mission critical communications. The methodology and the latency tools were implemented by the author while the radio measurements were implemented by Almudena Díaz Zayas and Pilar Rodríguez Pinos.

The required KPIs can be obtained by combining the different elements of the testbed. The setup that is employed to generate them is depicted in Figure 2.11. The power consumption measurements are generated by using a power analyser connected to the battery of the Device Under Test (DUT), if the DUT is an Android UE we install the Testeldroid tool. The DUT is connected to the LTE emulator and the channel emulator using RF wires connected the antennas[37]. The uplink and downlink are separated using splitters and filters. The uplink is connected to a channel emulator and from there is connected to the LTE emulator. Finally, the LTE emulator is connected to the EPC emulators, where impairments to emulate the backhaul and cloud services are introduced, and the service can be deployed in GPP machines.

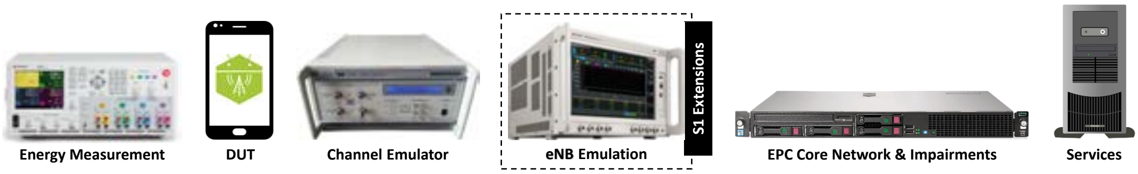


Figure 2.11: Methodology Setup[GP17]

In order to provide realistic data that could be used to assess the performance of a UE we defined a set of basic measurements, but these can be extended with specific end-to-end scenarios for the given application. For instance, if we are evaluating a voice communication system, we could analyse end-to-end voice traces to determine which UE is providing better results under the same network conditions. the set of measurements that we identified are the following:

---

[37] For some UEs the RF connectors are not available (either they cannot be easily exposed or their design is proprietary), in these cases contact antennas could be used. These antennas will introduce looses but could serve to compare two UEs.

- Cell Edge Measurements: Maximum Output Power and Receiver Sensitivity. The Maximum Output Power determines how good the uplink connection is in the worst RF conditions (at the cell of the edge, where we have the most reduced coverage). The Receiver Sensitivity is used to characterize the behaviour of the downlink under the same circumstances. These parameters are essential for applications on scenarios of high mobility, as they will determine how the UE will behave in the worst-case scenario, the cell edge where there will be the minimum received power and signals from other cells.

- Throughput. Although differences on maximum throughput might not be very relevant, it is interesting to analyse the behaviour of the throughput under different mobility patterns, as the behaviour of the UE might also change (as it is also influenced by the retransmission implementation, the cell edge measurements, etc.). To support these measurements, iperf[38] is employed, so the jitter and packet loss can also be obtained and compared.

- Data Plane Latency. To obtain the data plane latency, we estimate the RTT based on ICMP samples. The latency is obtained per each of the segments of the network, so we can isolate the contribution from the radio access, which is usually consuming the more significant part of the latency.

- Control Plane Latency. The control plane procedures also influence the behaviour of the data plane. For instance, the attach procedure determines how much time will be consumed by the UE the first time it registers on the network. The service request is the procedure that is triggered when there is data available for a UE that is in power-saving mode. Finally, the dedicated bearer establishment determines the time to establish a high priority bearer, for instance, to transport critical data.

- Power Consumption. To characterize power consumption, we analyse the instantaneous power consumption under different network scenarios. This KPI is relevant for systems that work with a battery to determine which one will run out of energy later.

We applied this methodology to two different UE with the same category (3). To compare them, we used a normalized spider diagram that is depicted in Figure 2.12. We included all the previously described measurements plus the cost. However, we could have introduced specific ones of interest for the scenario. For instance, instead of considering all the control plane procedures aggregated in a single metric, we could have used the service request and the power consumption for applications with long energy-saving periods. Mechanical characteristics such as size or weight for applications limited in space or specific metrics for the scenario being evaluated.

According to the diagram, DUT A provides better latency behaviour and power consumption while DUT B provides better throughput, cell edge behaviour and cost.

---

[38] `https://iperf.fr/`

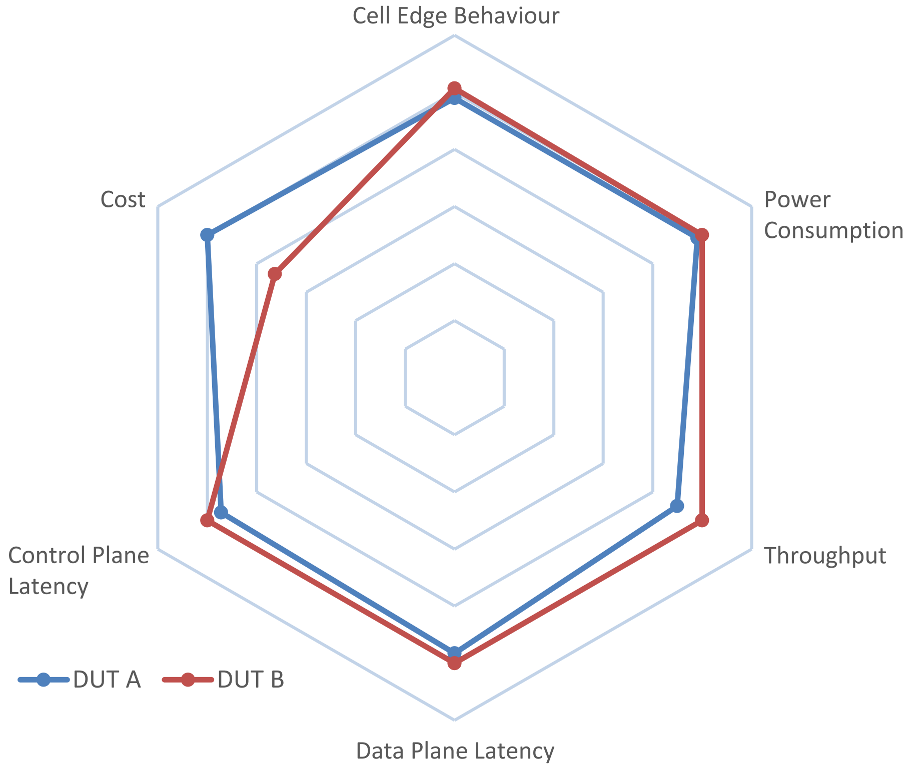


Figure 2.12: Normalized Comparison of two UE

Depending on the scenario and on the target KPI for the application, it might be better to select one UE or the other. More detail on the concrete procedures to obtain the measurements can be found in [GPDZR+17a].

2.5 Conclusions

In this section, we have provided an overview of the experimentation platforms for mobile networks, highlighting the differences with PerformNetworks, which is a European testbed that we have used to generate most of the results of this thesis and to which we have contributed in several ways.

A critical improvement to the methodology of the testbed has been enabling the configuration of scenarios mixing open-source, emulator and commercial solutions. This has been achieved by the design and development of the S1 extensions for the T2010 as well as by the integration of a wide range of equipment. Additionally, we have designed a methodology to select the most appropriate UE for a given application, establishing ways of extracting measurements to be correlated with other information such as the cost or availability.

On the tooling side, we have designed a way to easily integrate SCPI compliant instruments into our testbed and also developed several SCPI solutions. The remote Impairments tool is one example that allows users to introduce artificial IP impairments in any of the interfaces of the testbed, increasing this way the realism of the results. The EPC deployment tool has also been integrated with SCPI and allows the experimenter the deployment of EPC components in the testbed.

Figure 2.13: Setup for the validation of the Fog Gateway [GPDZR+17a]

We have also added tools to assist during the development of network solutions. Tunnel Tester can be used to encapsulate/decapsulate any type of traffic into GTP, which is the development of solutions that require this traffic by generating realistic traffic over the protocol. The replication tool can be used to generate redundancy at IP level to evaluate resiliency mechanisms.

Additionally, tools to generate measurements from the data of the testbed has been developed. For instance, PingAnalyzer can provide RTT measurements from PCAP ICMP traces. The ICMP packets can be transported over IP as usual but also over GTP or PDCP. ControlPlaneAnalyzer is focused on the estimation of time consumed by the signalling procedures of the S1 interface.

PerformNetwork testbed has evolved to support the integration of commercial equipment and to combine it both with instrumentation, allowing the definition of particular configurations, and open-source solutions, which can support the implementation of custom elements. On Chapter 7, we will provide some insights on possible future work around this experimental platform.

Chapter 3

Standard LTE architecture

Contents

3.1 Introduction **37**
3.2 Standard Architecture **38**
3.3 LTE Performance **40**
  3.3.1 Throughput 40
  3.3.2 Latency 41
  3.3.3 Prioritization and Reliability 44
  3.3.4 Security 46
3.4 Conclusions **46**

Synopsis

In this Chapter, we analyse the standard LTE network. First, we provide an overview of a basic LTE architecture with voice services, analysing the role of each of its elements. Then, using the tools developed in the previous chapter, we obtain results for the most frequent KPIs present in the basic architecture and provide a qualitative analysis of some network characteristics relevant for MCC.

3.1 Introduction

Mission critical services have been supported with dedicated proprietary networks, as it could be the case of TETRA. As stated in the introduction, the use of proprietary networks frequently raise the CAPEX and OPEX of their operators and harden the adoption of network improvements for new services. The adoption of standard mobile networks to support MCC could accelerate the introduction of improvements and ease the exploitation models, allowing the shared use of the infrastructure and its lease to third parties.

LTE is a good candidate to support MCC. It has been introducing functionality to support mission critical markets since Release 8 [LM17], where the first early enablers for mission critical appeared, up to these days where there are still ongoing activities to improve different mission critical services. Furthermore, according to the study [GSA19] from the Global mobile Suppliers Association (GSA), 791 operators were investing in LTE worldwide in December 2019 and 4.9 billion subscribers in Q2 2019, which are the 53% of the total of mobile subscribers. The availability of specific mission critical functionality and the spread of the technology worldwide make LTE an excellent candidate to support MCC.

The LTE basic end-to-end architecture, without any modification for MCC, can already support different use cases, most of its KPIs are higher than the existing MCC supporting technologies. In this chapter we will provide an overview of KPIs that could be relevant for different MCC technologies, employing the methodology that we have described in Section 2.4.3. Other studies are covering this characterization, the authors from [KKGM17] provide a survey on legacy and emerging technologies, including LTE, for Public Safety systems. Simulations are frequently used, for instance in the previously mentioned study the authors also provide some simulation results employing LENA ns3, another example is [JKY^{+}17] whose authors employed a modified LENA ns3 implementation to evaluate split bearer approaches. There are also studies employing live networks, such as [BRT^{+}14], where the end-to-end performance on live networks was evaluated, but these approaches are less frequent as they are normally more expensive and can be limited. In our case we have employed the modified conformance testing equipment combined with commercial core networks and UEs, this allows us to have controlled RF channel conditions, full configurable radio stacks but still maintaining compatibility with COTS equipment, increasing the realism of the results.

In the rest of this Chapter, we will provide an overview of the basic standard architecture of LTE, describing the basic functionalities of the different components of the network. Then we will evaluate this architecture in terms of network functionality and KPI, covering throughput, latency, prioritization, reliability, security and network services. Finally, we will discuss the different results obtained, identifying the elements to improve.

3.2 Standard Architecture

Figure 3.1 depicts the elements for the basic LTE functionality for voice and data services. There are three main blocks of functionality the eNB, which is the base station in LTE, the EPC, which provides the LTE core network, and the IMS, which provides multimedia services to the network.

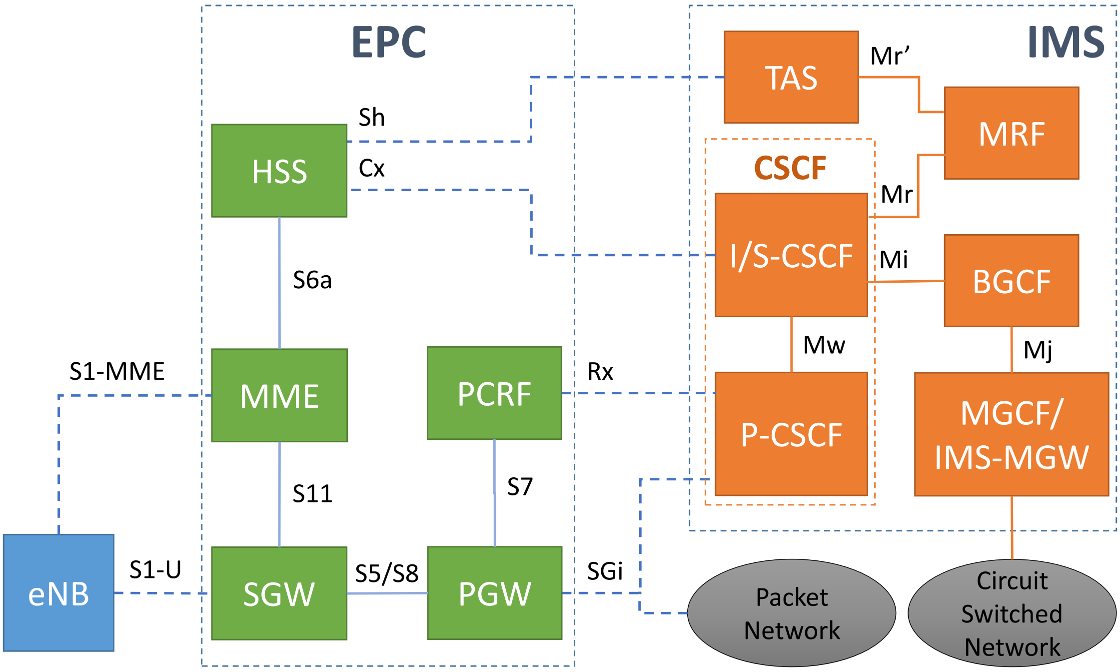


Figure 3.1: LTE architecture overview

The EPC's basic components are the MME, the SGW and the PGW (sometimes deployed together), the PCRF and the HSS. These components provide the end-to-end IP transport functionality, in order to support voice IMS needs to be employed. IMS comprises the Call State Control Function (CSCF), which can be divided in the Interrogating Call Session Control Function (I-CSCF), the Proxy Call Session Control Function (P-CSCF) and the Serving Call Session Control Function (S-CSCF), the Telephony Application Server (TAS), the Media Resource Function (MRF), the Breakout Gateway Control Function (BGCF) and the Media Gateway Control Function (MGCF).

The control plane procedures are handled by the MME, which exchanges signalling information with the UE employing the NAS procedures. The MME responsibilities are session and mobility management. The UEs register on the network employing the attach procedure, which consists of the exchange of identification, security, capabilities and data plane configuration. The MME keeps track of the location of the UEs across the different base stations and acts as the control plane mobility anchor.

Data plane is handled by the SGW and the PGW, sometimes are deployed in a single element. The SGW, which is responsible for the maintenance of the data plane tunnels with the eNB, and also acts as the anchor for the data plane during mobility procedures, for non-connected base stations a tunnel between the source eNB and the SGW and the SGW and the destination eNB will be established (indirect tunnel). The SGW will also establish tunnels toward the PGW, which provides the routing the packets towards the Internet. The tunnels employed on the core network to transport and route the subscriber data are based on GTP. The data goes over a small GTP header, which is transported over UDP, this way routing is done more easily.

The subscription information is stored in the PCRF and the HSS. The PCRF provides rules that determine the QoS of the users as well as an application interface, which can be used to trigger Service Level Agreement (SLA) on the network. For instance, when a voice call is going to be established using IMS, the CSCF will try to establish dedicated bearers for the voice traffic. The HSS is in charge of storing the information about the users, mainly their security keys and some service configuration. When the users are attaching to the network the MME will get their security information from the HSS in order to authenticate and identify them.

The control system in IMS is the CSCF, which does session control, communicates with the EPC to setup the path, to generate Charging Data Records (CDRs) and generate subscriptions. The element can be divided in three subsystems, P-CSCF, S-CSCF and I-CSCF. The P-CSCF is the element receiving the Session Invitation Protocol (SIP) messages from the UE and then forwarding them towards the IMS core. The element communicates with the PCRF in order to trigger any necessary dedicated bearer in the mobile core. The S-CSCF is in charge of session handling, generating CDRs and of enquiring the HSS about the applicable subscriber profiles. The I-CSCF is in charge of routing the registration request to the S-CSCF and also in Mobile Terminated (MT) calls of interrogating the HSS to find the S-CSCF where the UE is located.

IMS needs to communicate with the Circuit Switched (CS) domain, to do that we have two different components, the MGCF and the BGCF. The control/media plane interworking between CS and Packet Switched (PS) is done by the MGCF. The BGCF is a SIP proxy that routes the SIP messages, for calls terminated in the CS domain it will do the selection of MGCF, for calls terminated in the IMS domain it will select the appropriate interconnect.

We also have some components to handle the services itself such as the TAS or the MRF. The TAS provides support for the 3GPP mandatory telephony services, i.e.: supplementary services such as call forwards or line identification. The MRF provides media plane processing, such as transcoding or multi'party conferencing.

The basic architecture depicted in Figure 3.1 supports end-to-end data and voice communications for the LTE UE. In the following section, we will analyse qualitatively which capabilities it provides off-the-shelf and how it could be used to support MCC.

3.3 LTE Performance

The KPIs that we should assess to support MCC depends on the application that we are considering. From a network point of view, some of the most common indicators are:

- Throughput, most of the current MCC do not need very high throughput, at least compared with the current mobile network standards, however, this might change with the arrival of new services such as robotics or augmented or virtual reality.
- Latency, especially critical in systems that require real-time behaviour such as railways, vehicle to vehicle, etc.
- Prioritization and Reliability, some services require prevention of data loss, and sometimes this data has to be delivered with minimum latency.
- Resiliency, many MCC needs service continuity at any cost; the network has to provide this connectivity.
- Security, it is mandatory in most of MCC.

In the following sections, we will provide some reference figures for the measurable indicators and some related work for the non-quantifiable ones.

3.3.1 Throughput

Mobile communications can provide very high throughput, for the case of LTE, combining 3 cells in ideal conditions and with modulations 256QAM modulation it can reach 1Gbps [ZSTM17]. In [GPDZR$^+$17a] we provided some realistic figures of the throughput in three different scenarios ideal, pedestrian (3GPP Extended Pedestrian A 5Hz (EPA5)[3GP19r]) and vehicular (3GPP Extended Vehicular A 70Hz (EVA70)[3GP19r]) fading models. We considered uplink throughput which is more demanding for the UE. The setup employed is depicted in the previous section in Figure 2.11. We have two different UE that we put inside a shielding box, to avoid external interferences. We separate uplink and downlink, uplink is connected to a channel emulator where we setup the fading profiles. The uplink channel from the channel emulator and the downlink from the UE are connected to a modified UXM conformance testing equipment, which is connected to the core network.

The objective of the experiment was to set the devices on a very demanding scenario so we setup an estimated Signal-to-noise Ratio (SNR) of 20dB. The experiments were repeated five times per each of the devices as recommended in [3GP15c]. With this setup, we obtained high Block Error Rate (BLER)[1], it was close to 70% for the vehicular fading scenario and it was around 20% in the pedestrian.

---

[1] Normally, a BLER lower than 10% is considered as favourable radio propagation conditions.

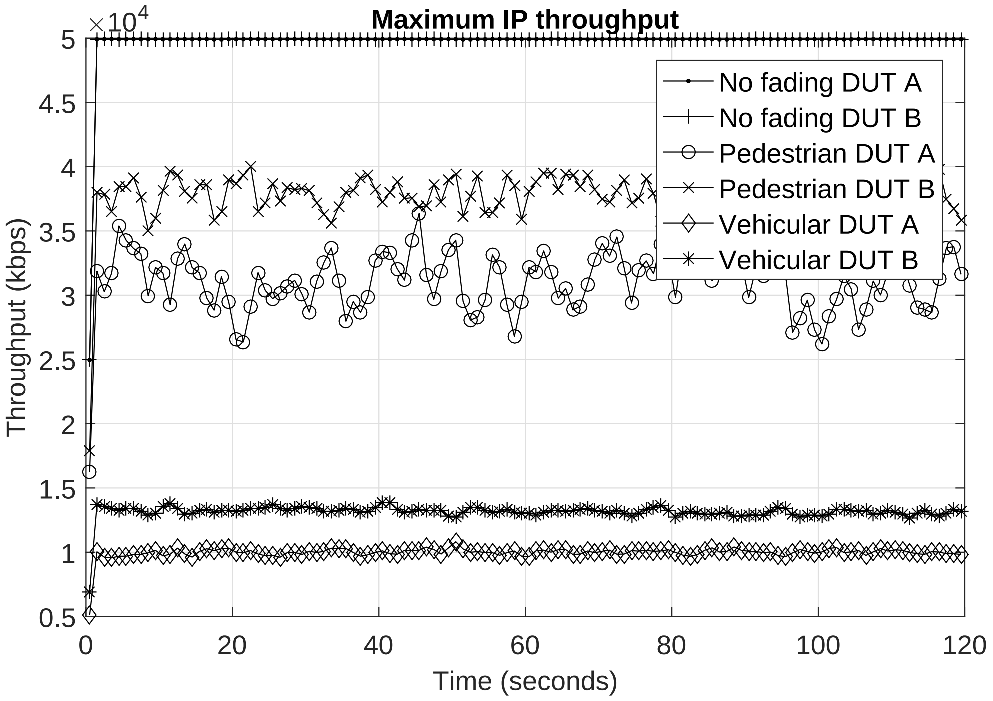


Figure 3.2: Uplink IP traffic measurements [GPDZR+17a]

Figure 3.2 depicts the obtained uplink results with the ideal scenario; both devices behave similar and are close to their uplink announced limit, which is 50Mbps. In the worst-case scenario, which is the vehicular one, the devices are still able to provide more than 10Mbps. It can be noticed that although both devices have the same LTE category (3), they provide different end-to-end performance.

3.3.2 Latency

There are two different aspects to consider about the latency. On one hand, we will have the latency on the data plane, which is the one that affects the end-to-end traffic of the final user. On the other hand, we could consider the control plane latency, which is the one related to the signalling procedures necessary to setup and establish the connection. The control plane latency is relevant as it affects the data plane of the UE, for instance, how fast it can connect after a restart or how fast it comes back from sleep mode.

In [GPDZR+17a], we obtained some measurements, using the tools described in Section 2.3 and the two same UEs that were used for the throughput measurements. Figure 3.3 depicts the setup employed, we used the UE connected to a modified T2010 conformance unit with a fixed Medium Access Control (MAC) scheduling allocation. The equipment was connected to the Polaris EPC, where we setup impairments in the backhaul (normal distribution mean 30ms) and the transport(normal distribution mean 40ms) networks. We considered two RF channel scenarios: ideal (no fading, no AWGN) and non-ideal (fading EVA70, AWGN to have SNR of 6).

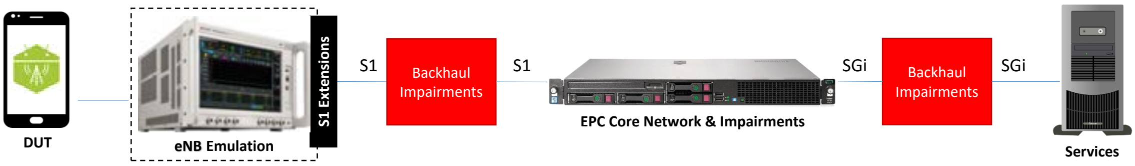
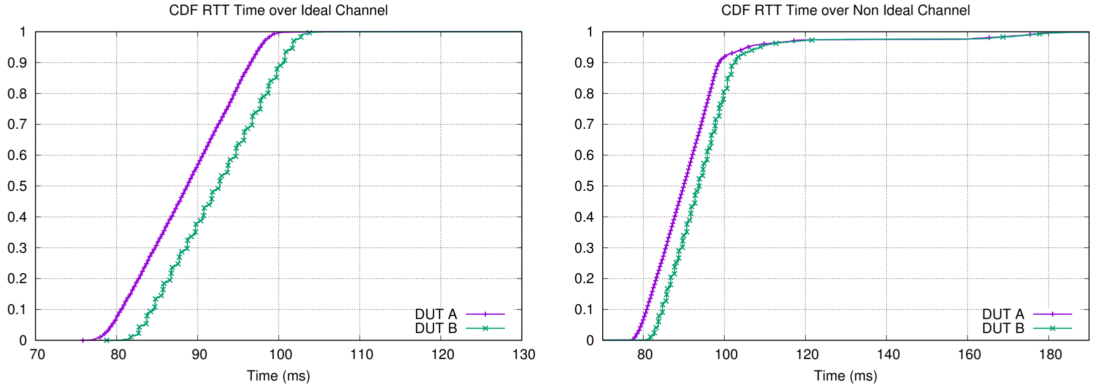


Figure 3.3: Setup for the latency measurements

Figure 3.4 depicts the Cumulative Distribution Function (CDF) of the LTE end-to-end latency on both scenarios. Again we notice different behaviour between the two devices, offering DUT a better performance in both scenarios. In the non-ideal channel, it can be observed how the RTT degrades as there is a higher number of retransmissions due to the errors introduced by the channel. The figures of the latency might not be enough for certain scenarios, using cloud services to support MCC harden the fulfilment of low latency requirements. A more detailed qualitative and quantitative analysis of latency is provided in Section 5.2.3.



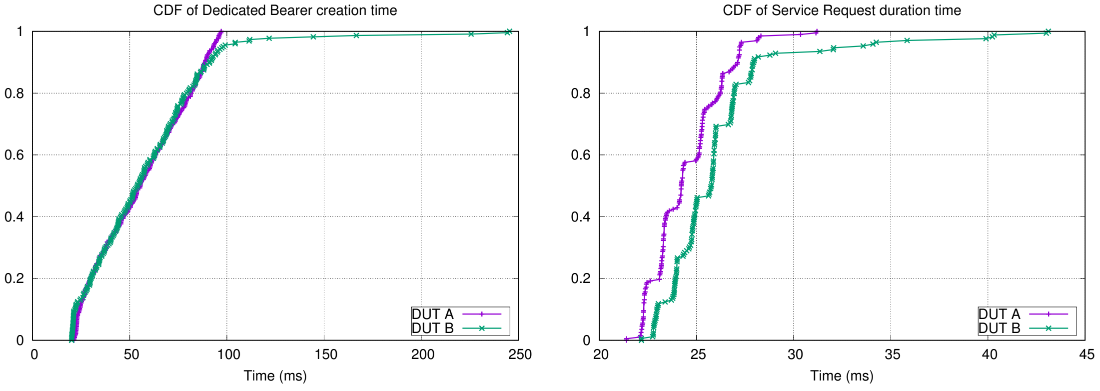
(a) Ideal channel

(b) Non Ideal channel

Figure 3.4: Data plane latency measurements [GPDZR+17a]

In the case of the control plane latency, we have analysed three procedures: attach, dedicated bearer creation and service request. These procedures have an impact on the behaviour of the UE in certain scenarios. The attach procedure is invoked by the UE to register on the network, so it will indicate how much time does it need to reconnect after a restart or initialization. The dedicated bearer creation is used to establish a prioritized channel, this is typically done by the voice over LTE procedures or, as we propose in Section 5.5, when a third party asks for a dedicated bearer transport. Finally, the service request is called when the modem exits from sleep mode, so it will provide an idea on the time consumed by the modem due to wake up, this is relevant for scenarios requiring low battery consumption.

The setup that we have employed to obtain the measurements is similar to the one in Figure 3.3, except that we use ideal backhaul and transport networks. We used the EPC automation tool to trigger the different procedures, as explained in Chapter 2.



(a) CDF DRB Establishment

(b) CDF Service Request

Figure 3.5: Control plane latency measurements [GPDZR+17a]

Figures 3.5 and 3.6 depict the CDF for the Dedicated Radio Bearer (DRB) establishment and service request, and the attach procedure respectively. Again there are slight differences between the two employed devices, being one of them faster in all the procedures under analysis. For devices in sleep mode wanting to send data, we have to add a mean 25 ms to the latency. The attach procedure is the procedure that takes longer, a mean of 240 ms, which is normal as it involves several security procedures and the establishment of the default bearer.

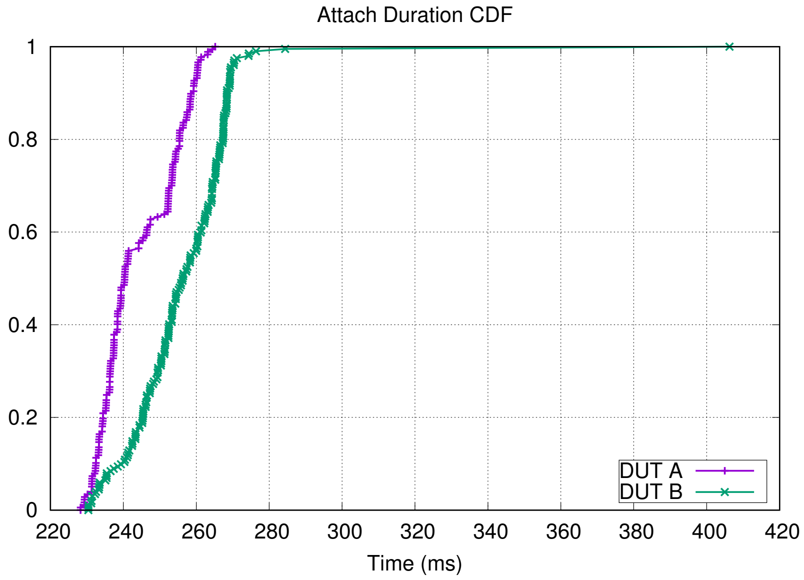


Figure 3.6: CDF Attach procedure [GPDZR+17a]

Table 3.1 provides a summary of the obtained Median and Median Absolute Deviation (MAD) RTT measurements. For instance, the service request time might be too high for critical devices. The time consumed by the attach procedure is high as well, which means that the device restart will have a high cost in terms of timing.

Table 3.1: Summary of the latency measurements

Measurement	DUT A Median (MAD)	DUT B Median (MAD)
Dataplane ideal channel RTT	88,710 (4,936) ms	92,714 (5,004) ms
Dataplane bad channel RTT	89,942 (5,464) ms	93,733 (5,061) ms
Attach Time	240.35 (9.290) ms	256.34 (10.390) ms
DRB Establishment Time	54.060 (20.56) ms	55.490 (23.60) ms
Service Request Time	24.225 (1.500) ms	25.780 (2.076) ms

As stated, the latency of LTE networks might be insufficient for some MCC services. In particular, data latency might be too high if we think on robotics, augmented reality or vehicular to everything communications. We will explore how to reduce these figures in Section 5, where we will discuss on new architectures to reduce this data plane latency. The control plane latency might pose a challenge for standard LTE networks, as to reduce it changes in the architecture and in the radio will be required.

3.3.3 Prioritization and Reliability

Mobile technologies provide mechanisms to prioritize the traffic since Universal Mobile Telecommunication System (UMTS), although they have not been traditionally exposed by Mobile Network Operators (MNOs) [DMPR07]. In LTE the QoS provision has gained importance, IMS is an important part of the network and requires dedicated bearer for the signalling, so providing prioritization for certain traffic can be done with a standard network.

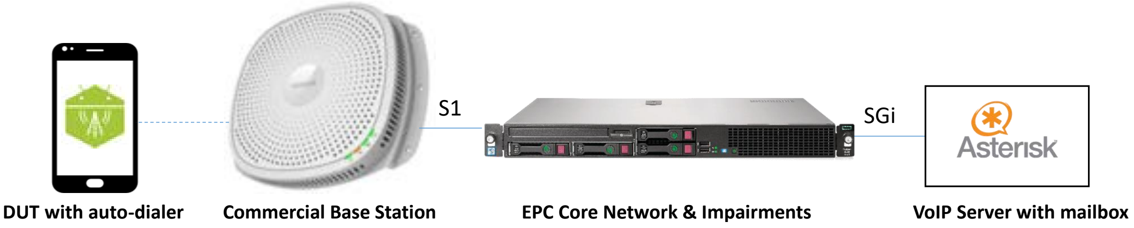


Figure 3.7: Setup employed for the priority measurements

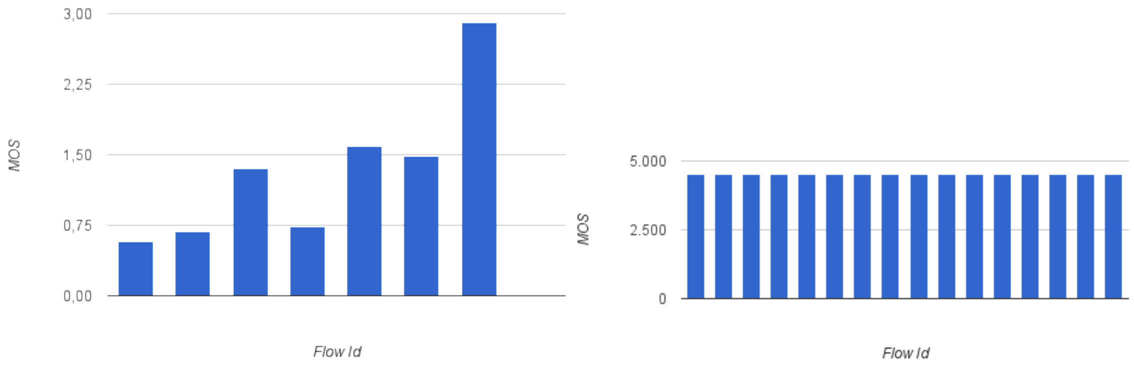
In [GP15], we carried out some QoS measurements on voice services under high background load with and without prioritization. The setup employed is depicted on Figure 3.7 and consists of an Android phone, a Nokia base station, the Polaris core network and an Asterisk[2] VoIP server.

We configured the UE with Testeldroid [ÁlvarezDMR12c] and an automatic dialer. The automatic dialer will place calls to a voice mail that plays back a recording. We then capture the voice traffic in the UE and the server and proceed to estimate the MOS using the Testeldroid post-processing tool. To generate errors we introduced some background services generating more than 90Mbit/s towards the UE. The eNB provides a

---

[2] `https://www.asterisk.org/`

maximum bitrate in the downlink of 50Mbit/s so the background traffic overloads the links generating drops and delays affecting voice.



(a) Voice on QCI 9 (b) Voice on QCI 5

Figure 3.8: MOS Measurements under heavy background traffic [GP15]

Figure 3.8 depicts two examples obtained using different bearers, one where the voice traffic was transported on the default bearer (QCI 9) and the other transported over a QCI 5 dedicated bearer. While the traffic on the default bearer ranked shallow values and even experienced called drops, the prioritized traffic MOS was the maximum that the estimator could provide 4.5[3].

The reliability has to be considered both on the network and service levels. In the network the LTE network, there are two mechanisms to prevent data loss one at the PHY/MAC layers and another at Radio Link Control (RLC). At the PHY/MAC layers, the Hybrid Automatic Repeat Request (HARQ) mechanism is employed, all the frames send through PHY has to be confirmed upon reception, if no ACK is received then a retransmission will be requested. The data received is also compared with any previously received data to improve error detection, and the sender will not process any data until an ACK/NACK is received. To prevent blocking the buffers due to a packet waiting there are eight HARQ processes, so if one gets blocked the others will keep on sending, only one process can send data in each Transmission Timer Interval (TTI) (in LTE 1ms) so each retransmission will introduce a delay of at least 8ms. In [GPM16] we analysed the latency under different conditions obtaining an additional mean latency of 14 ms when we introduced channel conditions that resulted in a BLER of 50% at the MAC layer.

In summary, mobile networks provide mechanisms to protect the data both against congestion and channel effects, the main problem will be for MCC services requiring very low latency as channel effect will introduce delays for them. To overcome this challenge sticking to standard networks we can introduce redundancy at IP level in the UE, but this generates a huge overload, we are not only multiplying the traffic but also

---

[3] The PESQ estimation is based on the reference algorithm provided by International Telecommunication Union (ITU) in [(IT01], which provides values between 0.5 and 4.5 (rather than between 1 and 5 as the MOS).

the underlying layers of the stack to transport it. Reducing the TTI will be the way to go, and indeed is one of the measures taken by the new 5G designs.

3.3.4 Security

In 4G security on the mobile networks was improved, in respect to the previous generations, by introducing Evolved Packet System (EPS) Authentication and Key Agreement (AKA), an improved mutual authentication between the UE and the MME, however, there are still some vulnerabilities in that networks [CML+14]. The authors from that research and from [HYA18], provide a very detailed survey on all the possible issues, some of them are:

- The flat IP architecture makes it easier to compromise the network, a compromised eNB can compromise all the network, and the network is more vulnerable to traditional IP attacks.
- The EPS AKA algorithm allow IMSI disclosure, Denial of Service (DoS) attacks and relies on mutual trust between roaming operators
- The handover mechanism can compromise all the security due to lack of backward security, is vulnerable to replay attacks and to de-synchronization attacks.
- The Radio access is susceptible to jamming attacks by centralized attacks to the control channel.
- The use of SIP in the IMS architecture makes easier capturing, forging and injecting messages.

Some solutions to these problems are also identified in the previously mentioned papers. There are also works focused on the security of mobile mission critical systems, for instance, [MRHR17] provides an analysis of the resilience of mission critical portable LTE systems. The authors provide some figures on the degradation of a MCC system under different interference conditions. Commercial technology is more attractive for attackers, as it is more widespread than niche technologies. This is both and advantage and a drawback, there is more information on the possible attacks and solutions, but there are also more known vulnerabilities.

3.4 Conclusions

In this chapter, we have discussed the feasibility of employing the standard LTE network to support MCC services. The use of standard technologies can reduce the CAPEX and the OPEX and accelerate the adoptions of innovations. In this sense LTE is a good candidate, it is deployed by many operators in the world and it is in constant evolution adding new features, including support for MCC scenarios, as we will discuss on Chapter 4.1.

The standard LTE architecture can provide voice and data services. In this chapter we have described the architecture and the role of each of the involved components. Additionally we have provided an analysis of the most features required for MCC.

For instance, for the throughput, we have provided some KPIs for a very demanding scenario, the uplink transmission under different (bad) channel conditions. The figures that we have obtained show that we can achieve the maximum throughput in static scenarios (in our case it was 50Mbps but can be higher) and throughput higher than 10Mbps in a bad vehicular scenario (that caused a 70% BLER at the MAC layer). This figure are much higher than many existing MCC technologies, for instance, TETRA Release 2 has a maximum throughput of 1Mbit/s.

On the latency, we have provided figured for both the control and the data plane. This is one of the aspects that need to be improved. For instance, the maximum call setup time for SIP calls is 50ms, which, in bad scenarios, cannot be fulfilled easily by LTE networks. For the data plane we will propose approaches to reduce the latency on standard networks but to provide better results or to improve the control plane behaviour changes in the base station and the architecture are required. We will discuss the latency in detail in Chapter 5.

Prioritization is also important and has been there for a while in the 3GPP standards, in 3G there were already QoS definitions. The main challenge here is the exposure of this prioritization to third parties willing to use the network over-the-top, we described our proposal in chapter 5. The commercialization of these services is also something to discuss both in terms of commercial strategy and of the technical realization of the CDRs. The existing mechanisms for reliability in LTE work but they have an impact on the latency, so if the application requires both latency and reliability then LTE present some limitations, mainly due to its TTI and the lack of low overhead transports for replicated data. In 5G URLLC is one of the target scenarios.

Finally, on the security side, there are still aspects to improve. For the author of this thesis, which is not an expert on the subject, the jamming attacks are maybe the harder to prevent. The availability and price of SDR solutions make easier the implementation of tools to block the control channels, which can leave a full base station out of service.

Although LTE can provide the required performance for voice and data services there are other functional requirements that need to be fulfilled. For instance, many MCC requires group communications or location dependant addressing. In the next chapter, we will discuss them.

Chapter 4

Mobile Mission Critical Communications Architecture

Contents

4.1 Introduction 51
4.2 Requirements 52
4.3 Early Enablers 54
4.4 MCX Architecture 57
4.5 Conclusions 60

Synopsis

In this chapter, we explore the provision of mission critical functionality employing mobile standard networks with a focus on functional requirements for the network. To gather functional requirements we will use the ERTMS as the driver use case, as it is very demanding in terms of network functionality. We will then analyze two architectures, an LTE architecture based in LTE, which we propose as a reference architecture for railways and an updated architecture using the Mission Critical X with X standing for Video, Data and PTT (MCX) standards.

4.1 Introduction

In the previous chapter, we have analysed the performance of standard mobile networks to support mission critical technologies over the top. This approach can fulfil the KPIs of some applications but there are scenarios were improvements on different KPIs, such as the latency or the security, are required. Additionally, MCC applications present functional requirements that were not addressed initially by the standardization bodies. However, some of these functionalities can be implemented over the top of mobile services and others are currently being tackled by the standardization bodies.

For the functional requirements, railway communications are a good use case, as they also demand other network features to support the end on the end service. To analyse MCC functional requirements, we will perform a qualitative evaluation of the standards to see if we can provide the network functionality to support complex MCC systems. To do so, we have chosen ERTMS, a railway control system, as our driver use case.

ERTMS is a signalling and speed control system to supports management and interoperation of railways in the European Union. Its standards are defined by the ERA and comprise the operating rules, European Train Control System (ETCS), which is the signalling system, and GSM-R [PTMK09], a proprietary mobile technology based on GSM that is employed to transport the signalling messages. GSM-R is based on GSM Release 99, but it also supports specific functionality for voice broadcast services, voice calls, functional addressing and location dependant addressing.

Other scenarios could have been analysed [GPDZR^{+}17b], but railway communications cover most of the requirements demanded by other MCC applications. From a functional point of view, railway requires many functions that were not natively supported by standard mobile networks, such as group and broadcast communications, location-dependent addressing, push to talk, etc. At the same time, the railway communications are demanding in term of performance requirements, they require support for high and fast mobility, latency, prioritization or even throughput[1]. Additionally, the analysis of mobile alternatives to GSM-R is also interesting because ERA wants a replacement before 2025[MMP16].

We will first provide an overview of the different requirements of GSM-R both in terms of KPIs and functionalities. Then we will describe the first proposal that we made [DZGPMG14a], which was based on the Early Enablers standards, a Release 12 LTE architecture to fulfil the requirements. But, many standards have emerged since we described this architecture. Indeed, some authors are pointing to 5G as the new FRMCS such as [PC19], where they suggest to skip LTE, or [CLML18], that points as 5G if it is mature enough by the time of standardization.

---

[1] Indeed current GSM-R standards cannot provide performance enough to support the most advanced self-driving scenarios.

From the standardization point of view, there is also an upcoming architecture study carried out by the ETSI Technical Committee for Rail Telecommunications [ETS20a] that have already carried out several performance studies for both 5G [ETS20b] and LTE [3GP19x]. And 3GPP has evolved the Early Enablers architecture, providing the MCX standards, which support technologies for mission critical communications, covering some of the gaps that we found.

In this section, we will describe the railway requirements, focusing on the ones described by the ERTMS standards, and then we will analyse how mobile networks have been used to support these communications. We will first describe an architecture based on the Early Enablers and then we will provide an update of the architecture, including the recent MCX standards that include specific functionality to support MCC.

4.2 Requirements

We consider GSM-R good use case for mission critical communications because it has been in production for a long time and it has many functional requirements that were not present in mobile networks. Its main advantage is that it already has standardized frequency bands across Europe to support it. However, nowadays, GSM-R presents severe limitations such as constrained bandwidth (the system was designed for voice services so maintaining ETCS links with trains can limit capacity) and higher OPEX and CAPEX, mainly because that is a niche technology only used on railways in some countries.

Table 4.1: GSM-R QoS Requirements

QoS Parameter	**Required Value**	**Specification**
Minimum signal level (95% of probability)	-92 dBm	EIRENE SRS
Connection establishment delay of mobile originated calls	<5 s (95%) <7.5 s (99%)	EIRENE FRS
Connection establishment radio	$<10^{-2}$	Subset 093
Bit error rate for transparent 4.8kb/s channel	$<10^{-4}$	GSM 05 05
Maximum end-to-end transfer delay (30 bytes data block)	<0.5 ms (99 %)	Subset 093
Connection Loss Rate	$\leq 10^{-2}$ /h	Subset 093
Transmission interference period	<0.8 s (95%) <1 s (99%)	Subset 093
Error free period	>20 s (95%), >7 s (99%)	Subset 093
Network registration delay	$\leq 30s(95\%)$, $\leq 35s(99\%)$, $\leq 40s(100\%)$	Subset 093
Maximum break during handover	<500 ms	

QoS requirements for GSM-R might not be very representative, as an old standard designed for voice it is not very demanding. However this is not an issue for our purposes as we have already characterized the most frequent requirements in the previous section. Anyway, we will provide an overview of the ones of GSM-R, there are some general QoS requirements (see Table 4.1) defined in the standards. Comparing the values depicted in the table with the ones that we obtained in the previous Section, we can conclude that most of them can be easily fulfilled by LTE, some of them by the KPIs of the network, others by the deployment strategy.

Still, some requirements are difficult to be supported by GSM, for instance, ETCS layer 3 introduces significant improvements, position and speeds of the trains is reported every 5 seconds by the radio system, instead of relying on the infrastructure of the train tracks. An outage of the radio will make the trains stop, so it requires a communication system with high availability and reliability. To overcome this challenge, there is a hybrid level 3 defined which combines both strategies, but a modern communication system should be able to support this saving money on the infrastructure.

The functions required by railway were not covered by commercial mobile communications before LTE and were introduced in the GSM-R standards, extending the functionality of GSM. The EIRENE specification [Gro15] provides a list of the ones to be supported, which are the following:

- Voice services, including point-to-point voice calls, public emergency voice calls, broadcast voice calls, group voice calls and multi-party voice calls.
- Data services, which covers text messages, and bearer services for general data applications, automatic fax and train control applications.
- Call related services, multi-level priority and pre-emption, advanced call handling (e.g.: call hold, call transfer, call queuing, etc.), auto-answer service, barring incoming or outgoing calls, call supervisory indications and charging information.
- Railway specific applications, support for functional addressing by train, engine or coach number or functional number, call specific persons depending upon user location, a specific mode for shunting operations providing a link assurance signal, multiple driver communications within the same train and railway operational emergency calls.
- Railway specific features, set-up of urgent or frequent calls through single keystroke or similar display of functional identity of calling/called party, fast and guaranteed call set-up, seamless communication support for train speeds up to 500 km/h, automatic and manual test modes with fault indications.

The specification also defines the call setup times based on the type of call, for instance, a railway emergency call has to be established in less than 1 second, while the low-priority calls admit 10 seconds.

There are additional aspects that have to be considered to replace GSM-R, such as security or deployment considerations. 3GPP has introduced security designs at different levels, but still, some threads have to be addressed. It is essential to analyse the cost of deploying and operating the network. This analysis should include the evaluation of the coexistence of both GSM-R and its replacement and the assessment of any potential legal requirements, such as spectrum requirements. There are also ongoing discussions on the type of deployment either having a dedicated one, which will increase the cost of the solution, or having a shared one, which can introduce more risks in term of security and safety.

In the following sections, we describe the architecture proposals to support railway services. We will start with the Early Enablers architecture that we proposed for LTE Release 12 and then we will analyse the most recent MCX standards.

4.3 Early Enablers

In [DZGPMG14a], we provided an analysis of the specifications at the time, and we proposed a standard architecture, which could be used to support railway communications. This architecture was based on LTE Release 12, which was being enhanced with a set of components that are now known as early mission critical enablers, as they were meant to enable network functionality to support MCC.

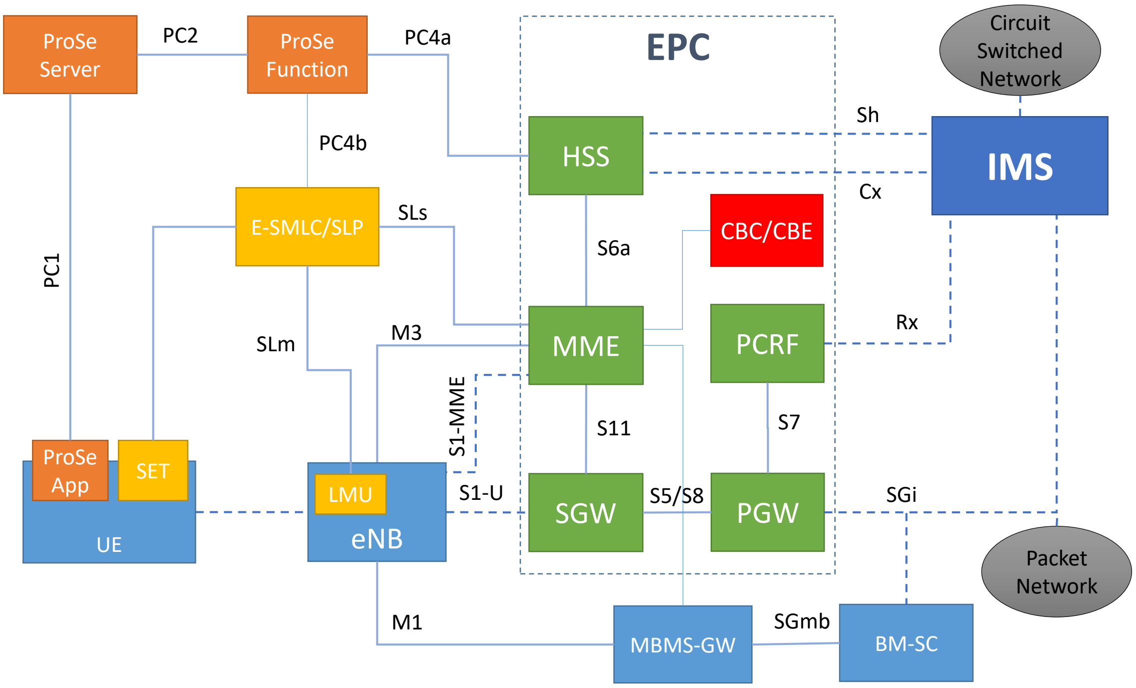


Figure 4.1: LTE architecture to support railway communications

Figure 4.1 depicts the architecture that we proposed in [DZGPMG14a]. We have included Proximity Services (ProSe) elements (ProSe Server, ProSe Function and ProSe App), positioning services (collocated Evolved Serving Mobile Location Centre (e-SMLC) with the SUPL Location Platform (SLP), SUPL Enabled Terminal (SET) device and Location Measurement Unit (LMU) base station), broadcast/multicast services (Multimedia Broadcast Multicast Services Gateway (MBMS-GW) and Broadcast Multicast Service Centre (BM-SC)), emergency services (Cell Broadcast Centre (CBC) and Cell Broadcast Entity (CBE)) and multimedia services (represented by the IMS block, fully depicted in Figure 3.1).

For voice communications scenarios, the use of Voice over LTE (VoLTE) [GSM14b] was foreseen, the main issue that we found by the time was the establishment time of VoLTE calls for emergency calls. EIRENE standards required this time to be lower than one second, which is higher than most of the results obtained for VoLTE in empirical evaluations, for instance, the authors from [LSRH15] only measure call setup times under the limit for the VoLTE to VoLTE calls of devices already in connected state (0.9 s) with the rest of the values doubling the target value.

Emergency calls for the IMS architecture are standardized in [3GP19q] and they require positioning services, which are obtained by the UE and the network. As positioning services are also required for location dependant addressing (which is not supported currently by the network) and for device to device communications, we proposed the use of e-SMLC [3GP17b] combined with LMU base stations [SDM14], which can provide accurate positioning data. The LPP protocol can be used on LTE to exchange positioning information between the UE and the LTE network but also Secure User Plane Location (SUPL) on SET devices, which is also available on 2G/3G, can be used [TZ13].

To increase reliability, we proposed the use of device to device communications, which is standardized in 3GPP as ProSe [3GP14b], which defines new channels that can be used for side link discovery and configuration. ProSe was used to enable communications between trains and track-side operators during the event of network failure as it enables discovery and communication of UEs close to others. In [KS18] we can find some of the challenges for these technologies.

The use of the Public Warning System (PWS), which was used for earthquake and tsunami alerts, was also foreseen to support emergency messages between the different railway stakeholders. PWS is defined at [3GP19a], which describes the technical realization of the CBS, which sends the messages received from the CBE (outside the scope of the standard) to the affected cells.

From a functional perspective, the main gaps to support railway communications were the provision of Push to Talk over Cellular (PoC) and of group communications. The standard [OMA11] that provided the requirements of the PoC enablers, which were yet to come, did not fulfil the EIRENE requirements (for instance, right-to-speak times, which are the equivalent to call setup time targeted 1.8 s). Furthermore this would an over-the-top solution, without integration to the operator network, so additional functions to trigger dedicated bearers using the PCRF Rx interface or to communicate with others elements of the mobile network should be developed.

For group communications there was not a standardized solution, broadcast and multicast services can be supported by the Evolved Multimedia Broadcast Multicast Services (eMBMS) architecture [3GP19b] but just for the downlink. The architecture introduces the MBMS-GW and BM-SC for the data plane and the interface M3 between the MME and the base station, we assume that the base station implements the Multi-cell/multicast Coordination Entity (MCE) functionality. To correctly support group communications, application level solutions will need to be implemented, however, eMBMS could be used to optimize downlink communications of a group of users.

Table 4.2: Early-enablers standards for railway communications [DZGPMG14a]

Requirement	*LTE Implementation and Challenges*		*Related Spec/s*
Call Communications	Might be provided using PoC; performance should be improved for emergency calls		OMA-RD-PoCv2 [OMA11]
	Point -to-point calls	Voice over LTE	GSMA IR 92 [GSM14b]
		Device-to-device	3GPP TR 23.703 [3GP14b]
	Emergency Calls	IMS emergency sessions	3GPP TS 23.167 [3GP19q] 3GPP TS 23.869 [3GP20k]
	Group Calls	eMBMS	3GPP TS 22.468 [3GP19p]
	Emergency group calls		3GPP TS 23.768 [3GP14a]
Broadcast services			3GPP TS 23.246 [3GP18c]
Priority Management	Bearer-established models		3GPP TS 23.107 [3GP18e]
	Service-specific access control		3GPP TS 22.011 [3GP20j] 3GPP TS 24.173 [3GP18b] 3GPP TS 27.007 [3GP20f]
	Multimedia priority service		3GPP TS 22.153 [3GP18d] 3GPP TR 23.854 [3GP11]
Location-dependant addressing	LTE provides standard positioning services, but the location-dependent functions must be standardized		3GPP TS 23.167 [3GP19q] 3GPP TS 36.355 [3GP20g] 3GPP TS 36.455 [3GP20a]
Security	Two levels of security, intermediate security layers might be necessary, as well as radio jamming mitigation techniques to protect control channels		3GPP TS 33.210 [3GP20b] 3GPP TS 33.401 [3GP20c] 3GPP TS 33.310 [3GP20i]
Spectrum allocation	Selection of standard railway frequencies for LTE studies of coexistence between mission-critical and public networks		3GPP TS 36.101 [3GP19r] 3GPP TS 36.816 [3GP12b]
Coverage	Transmission Time Interval (TTI) bundling		3GPP TS 36.824 [3GP12a]
	Adaptive beamforming		3GPP TS 36.912 [3GP18a]
	Standardization of the higher UE power class in the selected spectrum and analysis of the impact on other services and band		3GPP TS 36.837 [3GP13b]
	Carrier aggregation techniques		3GPP TS 36.823 [3GP13a]

Table 4.2 provides a summary of the functionality that we foresaw to enable railway communications. In the next section we will discuss about the most recent standards, which could be used to better support MCC requirements.

4.4 MCX Architecture

As mentioned before, since the Early Enablers standards there has been many improvements for MCC. 3GPP has worked in many items to support MCC applications, the process started with the early enablers, which are described in the previous section, and continue with the standardization of MCX. In [3GP19l] 3GPP has defined a detailed list of requirements for MCX, such as group, broadcast and private communications, traffic prioritization, functional aliases, location, security, interworking, etc. for ProSe, unicast and multicast scenarios.

A common functional architecture has also been defined [3GP19d] covering the control and application planes with group communications [3GP19i] or ProSe [3GP19n] and location with Location Services (LCS) [3GP19h]. Security for MCX services is defined in [3GP19o], covering signalling with interfaces with IPSec and HTTP interfaces with TLS. There are also other ongoing specifications and studies covering related aspects, such as operation in isolated environments [3GP20d], which provide the requirements of the Isolated Operation for Public Safety (IOPS) mode, or the provision of multicast over 5G [3GP20e].

The new standards cover one of the gaps of the early enablers architecture, which was the support of Push to Talk (PTT) calls, with the MCPTT[3GP19f] architecture. It was the first MCC application standardized by 3GPP to provide a PTT solution to public safety and other MCC entities. The service supports peer to peer and group communications and provides priority mechanisms to decide which user has the right to speak if multiple users have requested it, to override a call in case of emergency or to allowing low priority users getting the right to speak by implementing time budgets [SSBL19]. The authors from that paper and [SSA^{+}18] provide an analysis of the performance of the system using different technologies, including standard LTE, MEC with collocated EPC or collocated SGW, and 5G, achieving the minimum KPIs defined by 3GPP in all the scenarios except for standard LTE.

3GPP has also standardized a similar service for video, MCVideo [3GP19g], which defines a set of functions to be supported by the application such as pulling or pushing video from other clients or servers, transmission control, location information or codec adaptation. The architecture follows the one described for MCX services in [3GP19d], specific QCIs are assigned depending on the video mode and architecture requirements are also defined.

Finally, the other MCX service that has been standardized is MCData [3GP19e]. The standard defines the capabilities to be supported by the MCData application, which include Short Data Service (SDS) (maximum 1000 bytes in the payload sent over signalling or data planes), File Distribution (FD) and data streaming, all of them support one-to-one and group communications. The system should also provide an MCData message store, that shall be used to maintain the messages and their meta-data of a conversation securely.

Another aspect that we did not cover was the interworking of other MCC solutions with the architecture provided by the standards. Indeed, the MCX systems depict Inter-Working Functions (IWFs) in their architectures, as they are an important part to consider. The study [3GP17a] identifies some key issues when working with different MCC systems, including interconnection and migration issues with proposed solutions. The standard [3GP19k] provides a reference architecture so MCX systems can interact with LMR solutions. It is also important to mention the non-3GPP access architecture[3GP19c] that could be used to support technologies such as Wi-Fi (indoor), or other non-trusted LMR solutions. The authors from [LCC17] did a study analysing some the existing MCPTT solutions from an interworking point of view, including the solution that was awarded to support MCPTT, which according to them does not fulfil the interworking requirements.

Mission Critical Services can be supported by LTE with some MEC modifications, according to our research it can provide latencies around 10ms with good radio conditions. The MCX architecture is compatible with New Radio (NR) and the proposals that we will describe in Chapter 5 too, so, if lower latencies are required, then NR would be the preferred choice as it supports functionality to considerably reduce the time consumed by the radio (faster retransmissions, mini-slots, etc.).

Requirement	Mobile Implementation	Related spec
Call communications	PTT provided with MCPTT	TS 23.379 [3GP19f] TS 24.380[3GP19s]
	Point to Point supported with VoLTE or VoWiFi	TS 23.228 [3GP19j] TS 23.402 [3GP19c]
Common MCC Architecture	Functional Architecture	TS 23.280 [3GP19d]
	Group Calls	TS 23.468 [3GP19i]
	Device to Device, covered without network assistance	TS 23.303 [3GP19n]
	Emergency Calls	TS 23.167 [3GP19q]
	Security	TS 33.180 [3GP19o]
Data	MCData	TS 23.282 [3GP19e] TS 24.581 [3GP19t]
Video	MCVideo	TS 23.281 [3GP19g] TS 24.582 [3GP20h]
Initiator Context Dependent Addressing	Location provided LCS, context might be shared with MCData	TS 23.271 [3GP19h]
Low Latency and reliability	URLLC, ongoing studies. Modifications for NR.	TR 38.824 [3GP19y]
	Ongoing IOPS mode	TS 23.180 [3GP20d]
Interworking	Reference architecture to interact with LMR	TS 23.283 [3GP19k] TS 23402 [3GP19c]

Table 4.3: Updated summary of the standards for MCC

An updated summary of the standards to support MCC is provided in Table 4.3. In this table, we have not included the standards related to the basic functionality (except for the case of those either related or referenced by MCC). Additionally, we are not covering the 5G standards in there as they are still ongoing[2].

Figure 4.2 provides an updated view of a mission critical network. For the sake of simplicity, the end-to-end interfaces have been hidden, and we only depict the mobile network interface transporting them, i.e. we have not depicted the links toward the IMS or a SIP core but the interfaces that will be used to interconnect with it). We have also omitted the end-to-end reference points with the UE, the IWF or other MCX systems. For the ProSe architecture, we have included the full enabler, but the common MCX architecture only contemplates the communications without network assistance using the PC5 reference point. The MCData interfaces have also been collapsed, and the particular interfaces per capabilities are represented with a single MCData-cap1,n.

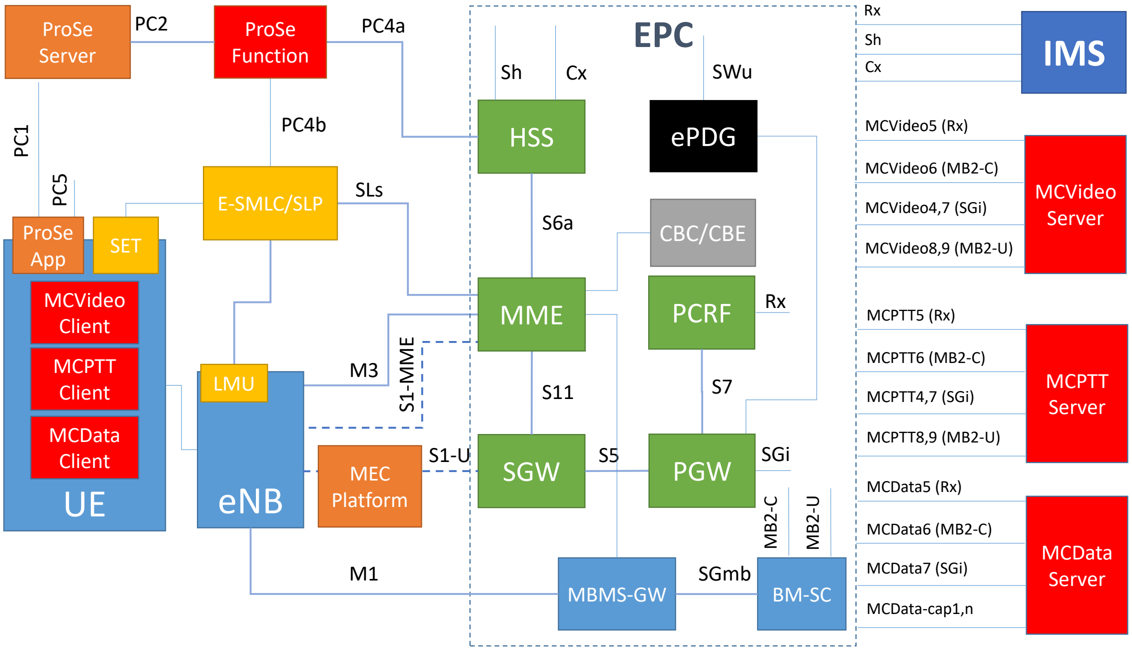


Figure 4.2: Updated MCC Architecture

In the figure we have the ProSe enablers (ProSe Server, App and Function), the function part should be implemented in the MCX service (in case the user of network-assisted device to device is required). For location services, we have used the standard LTE LCS (e-SMLC and in our scenario a collocated SLP), interconnection functions might be developed to interact with any other legacy positioning system[3]. We have also included a MEC platform, which will to reduce the latency but also to reduce the traffic towards the core and to assist on off-network scenarios. For the EPC we have considered the standard elements plus the Multimedia Broadcast Multicast Services (MBMS) (MBMS-GW

---

[2] At the time this was written, radio standards where already available but many of the ones defining the 5G Core (5GC) functions were yet to appear.

[3] e.g.: existing track beacons in railways.

and BM-SC), the PWS (CBC/CBE) and the untrusted non-3GPP access (ePDG) architectures.

On the radio side, we have depicted a single LTE base station, but we could also have depicted other non-3GPP accesses such as Wi-Fi or others from the domain being integrated. The UE implements the required clients for MCX applications, ProSE and location services. Finally, we have represented several applications, namely IMS[4] as a single block, and all the MCX service reference architectures.

4.5 Conclusions

In this chapter, we have focused on the specific functional requirements demanded by MCC services. To do so, we have used railway communications as a driver use case. We were able to provide an Early Enabler architecture, covering most of the functional aspects required except for group communications, which required solutions at the application level and PTT, whose requirements did not comply with the ones for the EIRENE specification. Except for these issues, the proposal was entirely based on COTS solutions, which reduces both the CAPEX and OPEX of the system.

Between the proposals we made and the writing of this thesis, some exciting improvements have been made. For the particular use case of railway communications, International Union of Railways (UIC) has defined a new set of specifications for the FRMCS in [AT19]. They provide an overview of different communication applications categorised as critical, which are the ones essential for train safety and for legal obligations, performance, which are the ones improving efficiency and business, which are focused on business operations. Different use cases are described in [Gro19].

UIC is providing the use cases, and user requirements and the standardization is carried out by the ETSI Railway Telecommunications Technical Committee along with the 3GPP Technical bodies. These organizations have carried out several performance studies for LTE. We would like to highlight [3GP19x], made by the ETSI provides some simulation-based figures to support future railway communications in LTE from a radio perspective. The report points to some other aspects to improve such as the intercell interference in a frequency reuse schema (which depending on regulations could be a plausible scenario) or the performance on rural environments (with higher train speeds).

3GPP has also evolved the standards, focusing on the provision of mission critical functionality, and now they provide solutions to some of the challenges that we found. For instance, different studies are covering the requirements for MCC, focused on identifying gaps in the standards or specifications providing new functionality. For PTT 3GPP has standardized MCPTT [3GP19f] that target access times (lower than 300 ms according to [SSBL19]) much lower than the one foresaw by the PTT enablers spe-

---

[4] Many will have depicted this as part of the EPC, to make the figure clearer we have depicted outside the EPC.

cification. Other MCC functions such as high priority data (MCData [3GP19e]) or video (MCVideo [3GP19g]) have been also standardized.

To conclude the chapter we have provided an updated architecture for these enhancements, referenced as the MCX architecture. As already mentioned, the updated architecture provides many of the functionality that we discussed on the early enablers and also some of the elements have stayed there (e.g. positioning or broadcast communications).

The efforts carried out by the standardization bodies to support this type of communications are clear and that the provision of MCC over standard networks is an important feature for all the involved stakeholders is too. However, from our point of view, there are still challenges to face such as the interworking with existing LMR (standardization has been done but implementations yet to appear), the regulation on critical services (governmental requirements on communications, dedicated frequencies, etc.) and the barriers to newcomers willing to enter this type of markets (due to the legacy deployments, size of the existing actors, etc.).

In the next section we will focus on the provision of low latency services in mobile networks, which is one of the weakest points to enable MCC over LTE networks.

Chapter 5

Low Latency Architectures

Contents

5.1 Introduction 65

5.2 Analysing the latency in LTE networks 67

5.2.1 LTE architecture latency overview 67

5.2.2 Qualitative overview of latency contributions 68

5.2.3 Experimental evaluation of the latency 70

5.3 Fog Gateway 73

5.3.1 Architecture 73

5.3.2 Emulated Results 77

5.4 GTP Gateway 79

5.4.1 Architecture 79

5.4.2 Latency Emulated Results 81

5.4.3 Comparison with the Fog Gateway 82

5.5 Third Party Exposure 83

5.6 Conclusions 86

Synopsis

In this chapter, we describe our proposals to reduce the latency on mobile networks. First, we analyse the end-to-end latency of standard LTE networks, giving an overview of the time split between the different components. Our proposals consist of the provision of services located close to the base station, following the MEC paradigm. We will define two architectures, the Fog Gateway that is compatible with standard LTE networks and the GTP Gateway, an evolution that requires modifying the network. For both proposals, we will provide some emulated results and we will describe an architecture to exposure MCC functionality for third-party applications.

5.1 Introduction

Low latency communications are essential for many of the upcoming technologies, such as vehicle to vehicle communications, virtual/augmented reality or the tactile Internet. To support these technologies over mobile networks, we need to modify the architectures and the stacks. Current mobile networks are composed of different elements, which have to be passed through by the traffic of the users. The fog computing [BMZA12] and MEC [HPS+15] paradigms target precisely bringing the services closer to the users, reducing the number of elements to be traversed.

Fog computing and MEC span across many different research areas, as described by the authors of [MYZ+17b] that provides a detailed survey on the different aspects of the paradigms. For instance, an essential feature of the architectures is the distribution networks employed for the services. This is explored by some authors such as in [LYS16] that gives an overview of different algorithms to implement Content Distribution Networks (CDN) on wireless networks. The authors of [Zha16] explore the reduction of traffic on CDN using caching techniques. Others researchers focus on the service side like [Mak15], where a video streaming use case is provided analysing technology and service designs. On [ASS17] optimization algorithms for augmented reality applications are studied.

Computational offloading is also important on the MEC paradigm; indeed MEC was initially used to refer to local computational offloading capabilities, although now is used in the same sense than fog computing [MYZ+17b]. A survey on the different solutions to support offloading is provided in [MYZ+17a]. The allocation of resources on the fog is tackled by [WZ17], where the authors proposed several algorithms, taking into account delay and costs constraints. A possible set of instructions to support computational offloading is analysed in [SGK+17].

MEC can also be supported at network level, which is the approach that has been followed in this thesis. Different architectures to support MEC services on 5G networks are characterized in [LNPW14] in terms of data caching techniques and overall system performance. [PH16] proposes a framework is proposed to secure cloud/fog based applications. Security analyses were done during the design phase where different architectures were discussed. A similar approach is followed in [HPK+14], its authors analyse the data plane, providing a mechanism which can switch from a cloud base tunnel to fog one. The main difference with the Fog Gateway proposal is that the latter can run on GPP machines. An architecture based on femto-cells is proposed and analysed quantitatively in [LBP+14]. These types of architecture require the analysis of both the control and the data plane, which provide more resiliency, however, it is also a risk as it needs to break the NAS security in order to access to the messages. The Fog Gateway architecture is based on the analysis of the data plane, there is no control plane analysis involved, which simplifies considerably the integration a security of the solution and allows to deploy it on commercial networks. We will describe some of our experiments using emulation techniques [GPM16, GPM17] to validate this proposal.

To further reduce and support lower latencies or the URLLC paradigm, fundamental changes have to be done to the network. 3GPP standardization bodies are already targeting some of them. 3GPP made a study on URLLC in [3GP19y], exploring different options such as fix allocation grants, UE inter UE prioritization, mini slot level repetition, etc. The NR standards have introduced features such as flexible slot allocation [3GP19u] [PLGB18], reducing HARQ times by using Code Block Group (CBG) [3GP19v] [GFSH18] and front-loaded Demodulation Reference Signal (DMRS) allowing fast demodulation. Other approaches focus on the introduction of packet duplication so the latency is lower on bad channel conditions. For instance, in [RV18] an overview of the different layer of the stack and improvements proposal to support packet duplication is done, in this study, they also foresee the use of this strategy at cell edges and bad link detection to improve the error rate. The authors from [CLS+20] provide a system-level analysis where they exploit dual connectivity to support data replication on error detection at PDCP level, exploring different enhancements.

Our take on reducing latency is to remove unnecessary layers on the stack. The approaches provided by ETSI, in [KFF+18], which provides an overview of the different deployment options for MEC in 5G, are based on the use of a collocated User Packet Function (UPF), we think it will be better to modify the stacks depending on the target service. The use of GTP is necessary when the core network has to route the packets but not when these packets are going to be processed by a server close to the base station. We propose another architecture the GTP Gateway, which follows the same philosophy than the Fog Gateway but removing avoidable GTP encapsulations.

The exposure of the network functionality is also relevant for the MCC scenarios. Accessing the operator network to rollout services or setup specific SLAs its not realistic. For instance, in LTE networks, the entry point for QoS requests is the Rx interface in PCRF, which was initially designed to accept requests coming from IMS systems. According to our experience operators are not open to expose this interface to third parties, not even via a Diameter Routing Agent (DRA), so we need a way to expose the functionality to third parties outside the operator domain. To do so, we have proposed the use of an API to be located in a secure domain, offering functionalities to request and setup QoS and fog services. Additionally, adapting the lower layers configuration to the transport is also part of the API, as it should also improve the reliability by reducing overhead on specific scenarios. A typical example of this will be to disable the retransmissions that are higher than the latency budget of a real-time service.

In this chapter, we will first provide an overview of the latency on LTE networks, both from a qualitative and a quantitative perspective. Then we will describe our Fog Gateway architecture and provide some results based on emulation techniques. Then, the GTP Gateway solution is defined and also characterized. Finally, we will provide details on an API to allow access to the functionality and some conclusions.

5.2 Analysing the latency in LTE networks

5.2.1 LTE architecture latency overview

Figure 5.1 depicts an architecture for LTE that we have split into different segments, which are the ones that we have considered in terms of latency. The LTE base station is connected to the core network using the S1 interface. This interface is split into two interfaces the S1-MME, which is the control plane, and the S1-U that transport the data from the users. There is another interface which is used to interconnect base stations, namely the X2 interface. X2 interface is also divided into user and control planes, and it is used mainly to support fast handover and Self Organizing Networks (SON) procedures. The SGi reference point divides the operator and the Internet domains.

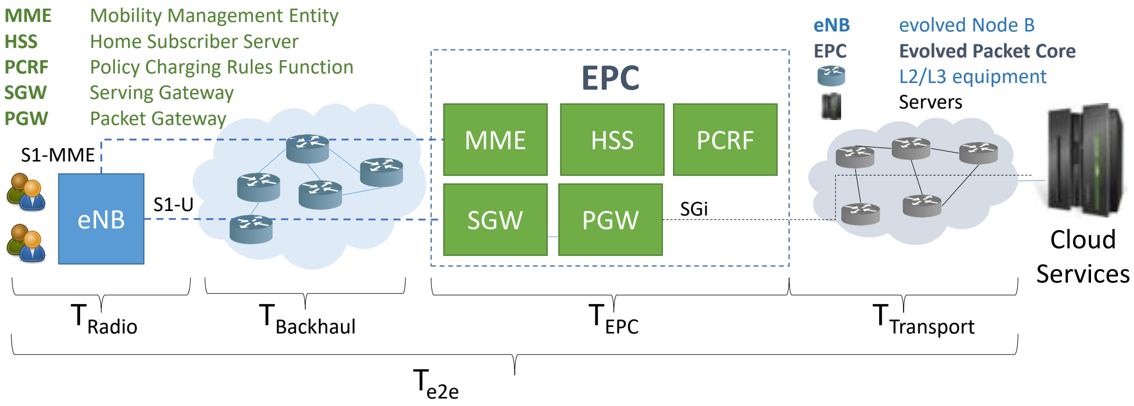


Figure 5.1: LTE basic architecture [GPM17]

We have represented two additional interconnection networks with the EPC, labelled as backhaul and transport networks. The backhaul is in charge of connecting the base stations with the core network while the transport network links the EPC with the external networks. User data is transported over GTP tunnels on the S1 and X2 interfaces.

In Figure 5.1 we have depicted the architecture with the different segments contributing to the end-to-end latency (T_{e2e}), which is the latency from the users to the services. We have divided the end-to-end path into the following segments:

- Radio (T_{Radio}): in this part, we have included both the base station and the UE.
- Backhaul ($T_{Backhaul}$): the backhaul interconnects the base stations with the core network.
- Evolved Packet Core (T_{EPC}): this contains the contributions from the elements previously described.
- Transport ($T_{Transport}$): in this segment, we have included the interconnection networks between the operator and the servers on the cloud.

The end-to-end delay (T_{e2e}) can be expressed as the sum of the previously described segments, as follows:

$$T_{e2e} \approx T_{Radio} + T_{Backhaul} + T_{EPC} + T_{Transport}$$

The latency doubles when we connect two users on the mobile network using an external server, and this will happen even if the users are camped on the same cell. Even if the connection between these two UEs is peer to peer we will see relevant contributions, we can define the peer to peer latency ((T_{p2p}) as:

$$T_{p2p} \approx 2 * (T_{Radio} + T_{Backhaul} + T_{EPC})$$

Looking at both expressions, it is clear that avoiding some of the elements of the communication the end-to-end latency will be reduced. To do so it is possible to deploy the servers very close to, or even in, the base stations. This way the traffic will not have to reach the transport, the core network nor even the backhaul. In Section 5.3, we described a solution that follows this approach on standard LTE networks.

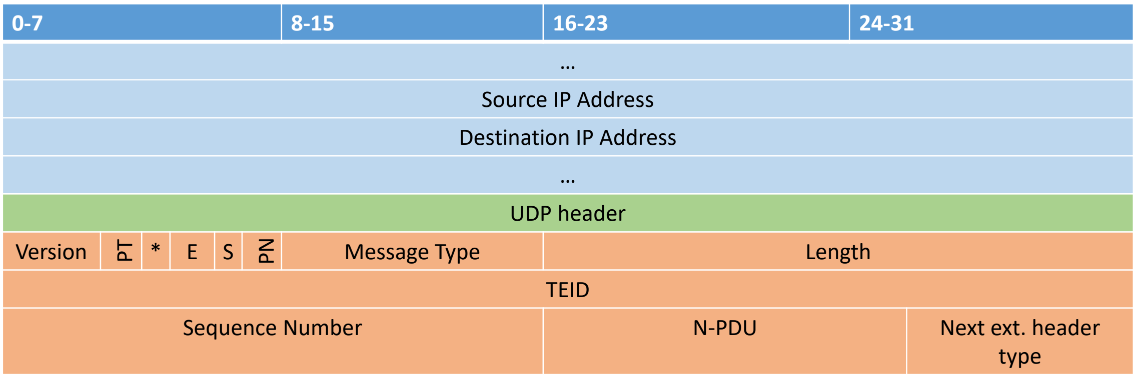
Before describing and characterizing our solution, we will provide some estimations on the contributions of the different segments of the network. There are several studies trying to analyse the latency on mobile networks, but the results vary a lot depending on where and how the characterization is done, so establishing our measurements will also serve as a baseline for the system performance. For instance, [Net09] provides an end-to-end latency of 20ms, half of which was consumed on the radio access. The authors from [LSRM^{+}12] obtain estimations for the end-to-end latency in live networks and found that was around 33ms.

5.2.2 Qualitative overview of latency contributions

In this section, a qualitative overview of the different sources of latency in each of the previously identified segments will be provided. The radio segment contribution is highly variable as it depends on different factors such as the equipment themselves and the RF conditions. If there are lousy channel conditions the coding scheme can vary (reducing the throughput) and there could appear HARQ retransmissions at RLC level, which will increase the contribution to the latency from this step considerably. When sending data from the UE in the uplink, the MAC scheduling procedures have a relevant contribution to the end-to-end figure. Every time that the UE has data to transmit it will have to ask the base station for a scheduling grant, and this procedure can take more than 4 ms [Net09]. In case the UE has data to send periodically, it could ask for a recurrent allocation employing the Semi-Persistent Scheduling (SPS) procedure.

The UE data in the base station will be prepared for transit to the core network at the PDCP layer. If the Robust Header Compression (RoHC) procedure is enabled there can be additional delays introduced in this stage[1]. The UE data will be matched against a Traffic Flow Template (TFT) in order to identify the bearer to which it belongs, then encapsulated using the GTP protocol with the information that has been negotiated with the MME for that bearer, and sent over the backhaul to the SGW.

GTP is used because simplifies the transport of the data carrying the IP packets of the UE over User Datagram Protocol (UDP). The GTP header, which is depicted in figure 5.2, is between 8 and 12 bytes long. Its most important fields are the TEID, packet length and message type, the rest of the fields are used to indicate the presence of other optional parameters, such as the sequence number of the NPDU. The main advantage of the protocol is that switching from one tunnel to other, for instance, due to a handover, can be done by changing the IP/s and the TEIDs that are part of the GTP header, the traffic from the UEs will remain unmodified.



0-7	8-15	16-23	24-31
...			
Source IP Address			
Destination IP Address			
...			
UDP header			
Version PT * E S PN	Message Type	Length	
TEID			
Sequence Number		N-PDU	Next ext. header type

Figure 5.2: GTP header [GPM16]

The base station is connected to the EPC by the backhaul networks. Different technologies can be used in the backhaul and selecting one or another will depend on the location of the base station; e.g. on cities, there will fibre available while in rural areas is more frequent to find microwave links. As higher frequencies are being standardized, more base stations will be required, so the variability introduced from the backhaul will increase. The authors from [ZQK$^+$16] provide an overview of the aggregated effect of different backhaul technologies. Additionally, many operators outsource this part of the network, so in addition to the variability due to the underlying technology, we can also find domain changes (from the operator to the third party supporting the links). The latency introduced by this segment varies a lot from a few milliseconds to the tens of seconds.

During the attach procedure there will be established at least two tunnels for the default bearer (one for the downlink and another for the uplink). The tunnels will be unequivocally identified by their TEID and the transport endpoints, which are typically the IP of the SGW and the eNB. Each of the tunnels will have associated certain quality

---

[1] RoHC is usually employed to reduce the overhead on the packets of the UE by compressing the headers from IP up to Real-time Transport Protocol (RTP). This usually is useful to better exploit the channel when transmitting small packets, for instance on VoIP

level, defined by the QCI, maximum and guaranteed bitrates, priority level and preemption capability and vulnerability. This information is extracted from the HSS and the PCRF. Once the data has arrived at the SGW it will be sent to the PGW over another tunnel, and the PGW will remove the GTP headers and route the traffic towards the Internet. The main advantages of using the tunnels are the routing simplification, the low overhead (8 bytes minimum) and an easier security application[2] but, as we will see, can introduce an unnecessary encapsulation for MEC applications.

The EPC elements are traditionally centralized, and the contributions from the latency normally come from this fact. The origin of this centralization is that current EPC networks run on dedicated hardware infrastructures (such as Advanced Telecommunications Computing Architecture (ATCA)), whose price along with the licenses is very high. This lead operators to own and deploy a few of this equipment. Again we will find variability depending on the country, and on how close and well connected to the central deployment the base station is.

Transport networks will usually include several domain changes and the contribution will be affected by the location of the services (and the core network). Some of the approaches to reduce latency proposed moving the cloud servers to the edge of the operator's network, while others explore the location of this services closer to the base stations, like our Fog Gateway solution.

5.2.3 Experimental evaluation of the latency

The end-to-end setup, employed for the characterization, is depicted in Figure 5.3. To measure the UE started a continuous ping against the server was launched and traces were captured on the UE (RTT_{UE}), the S1 interface (RTT_{S1}) and the SGi (RTT_{SGi}).

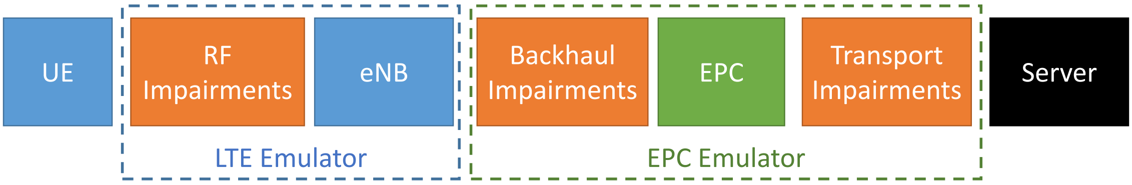


Figure 5.3: Experiment setup to characterize the latency split [GPM16]

The evaluation of the latency on LTE networks is carried out using different radio accesses. The extended conformance testing equipment (also referred as LTE emulator) T2010A (described in section 2.3) combined with the EPC emulators from Polaris Networks are used to provide a thorough characterization of the split of the latency on different network conditions. The basic configuration employed in the emulator is provided in Table 5.1.

---

[2] Although not covered by the standards, it is commonplace to secure the backhaul interfaces using IPsec that could be negotiated between the eNB and the SGW. If the traffic of the user is sent without tunnel, IPsec would have to negotiated between the UEs and the Servers.

Table 5.1: T2010A Configuration

Parameter	Configured Value
Frequency (Band 20 FDD)	DL 806MHz UL 847MHz
Bandwidth	10 MHz
Power	-61 dBm/15KHz
Uplink Modulation	22-64QAM
Downlink Modulation	22-64QAM
Antenna Configuration	SISO
Maximum HARQ Retransmissions	7

To produce realistic results the channel emulator of the unit was configured with different channel conditions, as follows:

- ideal channel with no fading and no noise, this resulted in an estimated MAC BLER of 0%.
- medium channel, fading profile EVA70[3] and channel noise -80dBm/KHz, which resulted in an estimated MAC BLER of 10%.
- bad channel, fading profile EVA70 and channel noise -70.5dBm/KHz, resulting in a MAC BLER of 50%.

The EPC is deployed with the basic setup (MME, SGW, PGW, HSS and PCRF). Impairments in the backhaul or the transport networks are not considered for the characterization of the different segments. The UE is a commercial LTE USB dongle that works both in band 7 and 20. In order to compare with COTS base stations, the LTE emulator was replaced with a small cell, and the measurements were repeated.

An initial baseline for the end-to-end latency was provided under different radio conditions. The setup in Figure 5.3 was employed to evaluate the effect of the channel on the RTT. Figure 5.4 depicts the obtained results; the latency increases considerably when the channel conditions get worst, mainly due to the effect of the HARQ retransmissions.

Then we characterized the latency split on the different segments of the networks. Besides the T2010, small cells were also combined with the Polaris EPC to have some figures using COTS equipment. The employed small cells worked on band 7 with a 10MHz bandwidth for both the uplink and downlink.

Table 5.2 summarizes the obtained results. The difference in the latency measurements between the small cells and the conformance testing equipment is due to the configuration of the T2010, which was configured to allocate all the available resources of

---

[3] EVA70 stands for Extended Vehicular A with a maximum Doppler frequency of 70Hz. This profile is defined on the LTE standards (see [3GP15b]).

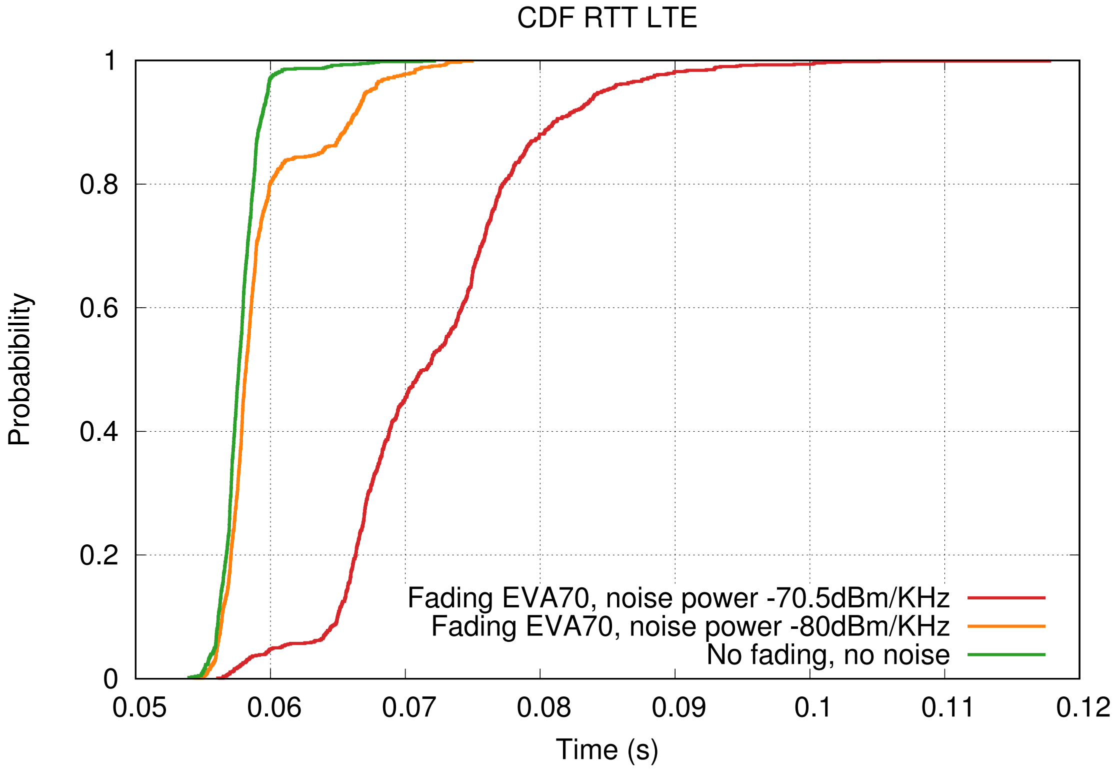


Figure 5.4: LTE RTT baseline under different radio conditions [GPM16]

the spectrum in every time slot. This configuration will be the equivalent to setup SPS on the system with all the resources of the base station for a user. It is important to take into account that these latencies are much lower than the ones that will be obtained on a live network. The main reasons are the lack of backhaul and transport networks, and the absence of more users camped on the cell. In these scenarios, most of the time is consumed on the radio access, the time consumed by the EPC is negligible.

Table 5.2: RTT split time comparison

Segment		COTS Small Cells		T2010 Equipment	
RTT(ms)	Obtained as	Median	MAD	Median	MAD
T_{e2e}	RTT_{S1} - RTT_{SGi}	28.775	4.887	11.830	0.253
T_{Radio}	RTT_{UE} - RTT_{S1}	28.223	4.882	11.577	0.247
T_{EPC}	RTT_{UE}	0.227	0.002	0.229	0.003

5.3 Fog Gateway

5.3.1 Architecture

Overview

The Fog Gateway architecture is fully compatible with standard LTE networks. It is designed to make possible the deployment of servers very close to the base station, in order to avoid the latency introduced by the backhaul, core and transport networks. Figure 5.5 depicts the proposed architecture. The essential elements of the LTE network remain as they are while a new component, the Fog Gateway is deployed to capture the X2-U and S1-U interfaces, which are in charge of transporting the user plane between base stations and the base station and the core network respectively.

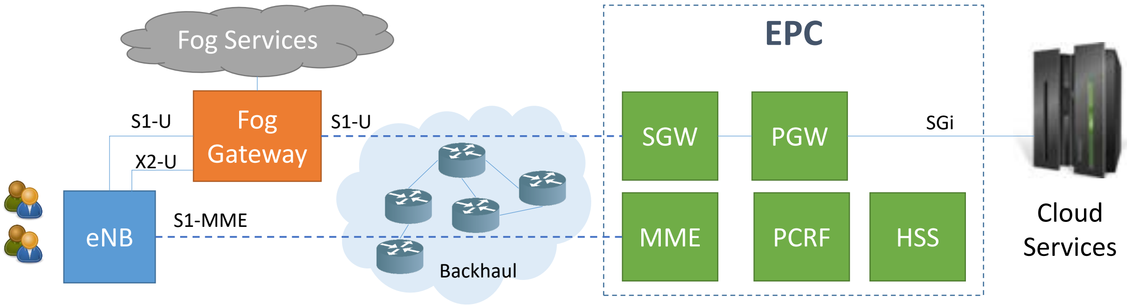


Figure 5.5: Architecture of the Fog Gateway [GPM17]

The X2-U interfaces are analysed to detect handover events, which might translate on tunnel identification changes. If the X2 interface exists between the source and destination eNB on a handover, the traffic for the user will be redirected from source to destination. If the X2 interface does not exist, an indirect tunnel will be established, and the source eNB will send the data via its SGW. If the source and destination SGWs are different then the source SGW will send to the destination SGW and this one to the destination eNB. All these tunnels will be based on GTP.

The Fog Gateway will receive all the tunnels between the different components. Identifying which tunnels are part of a direct or an indirect tunnel can be done based on the analysis of the inner IP header. For instance, if an uplink packet is received, but the inner destination IP is a known UE of the eNB from which the packet has been received, the gateway will identify the packet as part of a handover[4] procedure and will be ready to update the information about the tunnels on the database.

The identification of the packets, which have to be redirected to the fog, is based on the destination IP of the inner IP header. Two conditions will make the Fog Gateway process a packet locally:

- Uplink packet with inner destination IP of a service that is reachable from the Fog Gateway.

---

[4] This scenario occurs when the eNB is establishing an indirect tunnel to the new eNB.

- Uplink packet with inner destination IP of A UE and source inner IP another UE that belongs to the tunnel to which the packets belongs.

The MSC for the Fog Gateway is depicted in Figure 5.6 where the two main scenarios are depicted. On the first one, the packet is redirected to a local service while on the second the standard flow is followed.

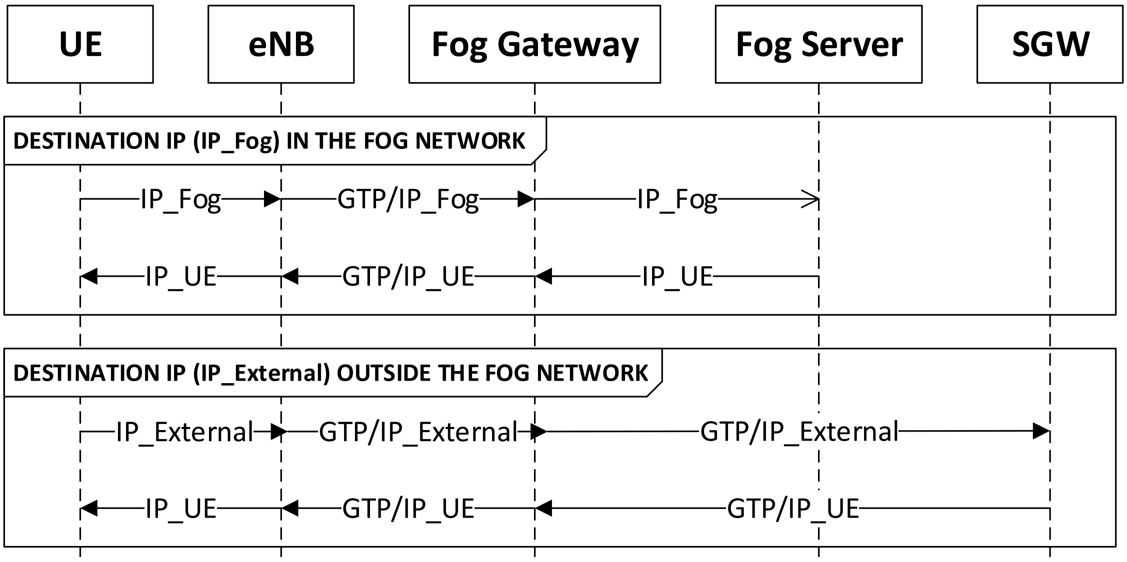


Figure 5.6: Message Sequence Chart of the Fog Gateway [GPM17]

The gateway can be split into two main functions: tunnel database building and packet routing. The tunnel database building is in charge of maintaining an updated database with the information of the UEs and their tunnels. The packet routing function has to decide whether a packet has to be processed locally, redirecting it to the fog network, or if it has to be forwarded to the SGW.

Tunnel Database Function

The tunnel database provides the bindings between the UE IPs and the information of their tunnels, which are defined by their TEID, eNB IP and SGW IP. Every GTP packet arriving at the gateway will be analysed in order to maintain an updated database.

The function will process the packets coming from the SGW in the downlink and from the eNBs in the S1 uplink and the X2 interfaces. There are different scenarios that will allow the gateway to identify the different situations in the S1 interface, which are the following:

- New UE (S1-U interface between the eNB and the SGW). This scenario is detected when a new TEID arrives and the UE IP is not present in the database. In that case, the database will be updated with the information and the packet will be transported normally.

- Handover with direct tunnel (X2 interface between the source and destination eNBs). This is detected because both outer IP addresses belong to an eNB and the inner destination IP address belongs to the UE camped on the eNB from which the packet comes.

- Handover with indirect tunnel (no X2 interfaces between source and destination eNB, forwarding tunnel established over the SGW/s). In the source eNB, the scenario is detected because the outer source and destination IPs belongs to that eNB and its SGW respectively, and the inner destination IP address belongs to a UE camped on the source. In the destination eNB, it is detected when the source IP is the SGW, the destination IP is the eNB, and the destination inner IP address is from an unknown UE.

- Dedicated bearer tunnel (S1-U interface between the eNB and the SGW). This situation is detected when a new TEID for an existing UE on the database is identified between the same endpoints (no change of eNB nor SGW). This scenario is the main limitation on the standard behaviour for the Fog Gateway, if a DRB is established to transport packets between the UE and the fog services the gateway will ignore any prioritization on the uplink (it is not possible to infer the traffic characteristics only by analysing the data plane). In the case of the downlink the traffic will not be transported on the dedicated radio bearer on the eNB, as it is not possible to deduce the TFT from the traffic. The TFT mask could only be estimated in case it consists of a single rule (it could be obtained by applying a mask to filter out the relevant fields and then apply between packets.

The tunnel database might not contain all the information for a given UE, for instance, if the UE has generated only uplink packets there will be no information about the tunnels of the downlink. To prevent that behaviour, the tunnel database will generate an ICMP packet to a known server (reachable from the SGi reference point), which we call the pinger server, in order to force downlink traffic. Once the ICMP response arrives at the gateway it will be automatically dropped so not extraneous traffic is sent to the UE.

Once the packet is processed to generate information for the database, it will be passed to the routing function, which will use the information of the database to decide where to route the packet.

Packet Routing Function

The packet routing function is in charge of deciding the destination of all the received packets. In most of the implementations, it will receive a parsed packet from the packet database building function. The uplink packets might also need to be reconstructed if they are fragmented. This can happen if the UE is using the Maximum Transfer Unit (MTU) and the GTP headers added by the eNB make the packets longer than the MTU of the S1 interface.

The uplink X2 packets will be forwarded directly to their destination eNB. The treatment of the S1 uplink packets is based on the analysis of the inner destination IP address, which can be:

- An IP belonging to a fog service. In this case, the function will remove the GTP header and will forward the packet to the appropriate service.

- An IP belonging to a UE that is on the database registered on an eNB which is registered on the Fog Gateway. This is the scenario where two UEs are communicating peer to peer. In this case, the outer IP and GTP headers have to be updated with the information on the database. The outer source IP address will be replaced with the SGW IP address, the outer destination IP address will be changed with the eNB address where the UE is camped, and finally the GTP TEID will be substituted by the destination UE downlink TEID on the database. The packet then will be sent to the destination eNB.

- An unknown IP. In the rest of the cases, the packets will be forwarded to their original destination with any further change.

The downlink packets coming from the SGW, as well as the ones from any X2 interface, will be forwarded to their destination eNB. The downlink packets from fog services also need to be processed based on their destination IP, which can be:

- The address of a UE whose downlink information is on the database. The gateway will craft a GTP header for these packets with the information from the database and will send the packet to the eNB.

- The IP of a UE with no downlink information on the database. This situation might happen when there have been no downlink packets coming from the SGW, and the pinger procedure has not been completed. These packets can be buffered until a downlink packet from the SGW arrives or a time out is reached, in which case they will be treated like in the next point.

- An unknown IP or timed out packet. This cases will depend upon the implementation that might drop the packet silently or send an ICMP destination unreachable.

In order to be transparent to the other network elements, the gateway does not have IP address, so it is necessary to provide Address Resolution Protocol (ARP) in all the interfaces. The gateway will have to send ARP replies in behalf of the UEs, towards the fog services. It will also reply to requests for the eNB address coming to the interface connected to the EPC and to requests for the SGW address in the interface conected to the eNB. We have depicted this behaviour in Figure 5.7.

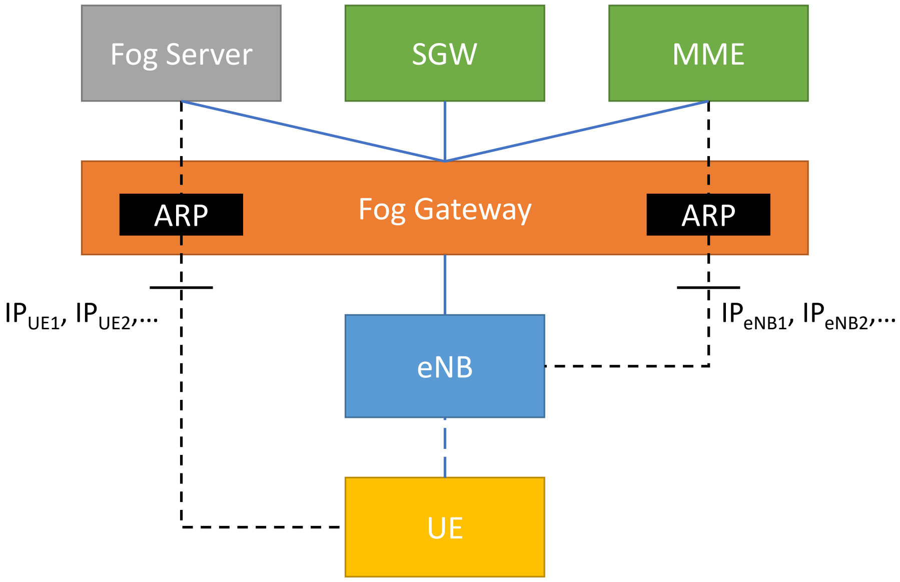


Figure 5.7: Fog Gateway ARP interactions

5.3.2 Emulated Results

For the emulated results we employed using the T2010 LTE emulator extended to support the S1 interface, the configuration employed for the experiments is provided in Table 5.1. The emulator was combined with the basic EPC emulators (see Figure 6.6) configured with some artificial delay impairments in the backhaul (the interface connecting the emulator with the S1-MME and S1-U interfaces) 15ms and in the SGi (the reference point that separates the operator from the external networks) 30 ms.

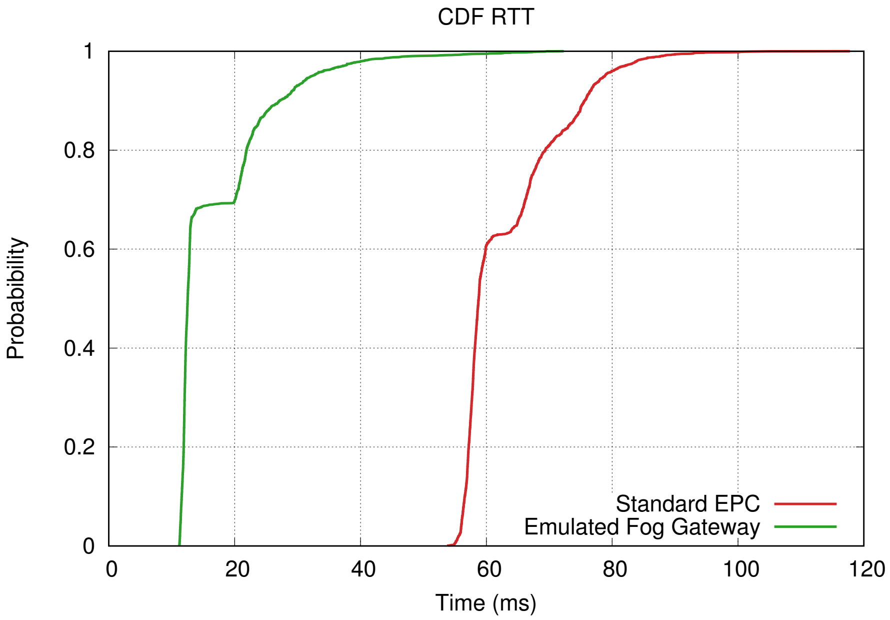


Figure 5.8: Standard network and Fog Gateway RTT comparison

We considered that the latency introduced by an EPC collocated with the base station would be similar to the one that could be introduced by the Fog Gateway, so to obtain the measurements we setup the emulators with different channel conditions and maintained the EPC without any impairments. Figure 5.8 depicts the CDF obtained for these two scenarios. The emulated Fog Gateway provided lower latency, as it will skip the latency introduced by the backhaul and transport networks. This gain will only be valid for services located in the fog, for the rest the standard EPC latency will be applied.

We also provided an analysis of the latency of the emulated Fog Gateway under different channel conditions, the results are depicted in Figure 5.9.

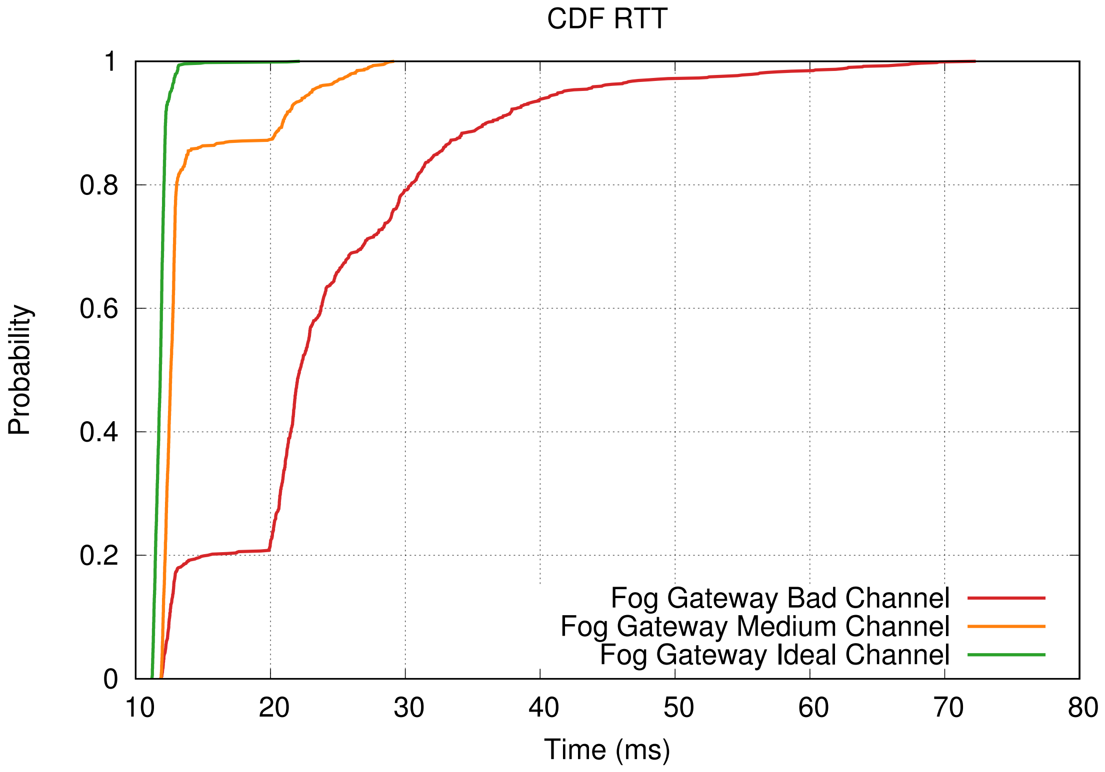


Figure 5.9: Emulated Fog Gateway RTT under different radio conditions [GP17]

In ideal conditions the Fog Gateway will be around 10 ms, on medium conditions, 80% of the users will be around 12 ms, a small percentage can go to 30 ms. Finally, on the bad channel conditions, 80% of the users will be under 30 ms, but the latency can go to 80ms. These values will be affected by the number of users, but taking into account that the system is designed to exploit geographical proximity we can estimate than less than a hundred stations will be connected to the gateway. More details on the emulated results are provided in [GPM16, GPM17].

5.4 GTP Gateway

In this section, we describe our proposals to reduce the latency by modifying the network architecture. We will describe the GTP Gateway, which is a proposal to remove unnecessary GTP encapsulations. We will also discuss the results that could be achieved by introducing these modifications on LTE networks by analysing the results obtained in an emulated scenario.

5.4.1 Architecture

In Chapter 5 we described the Fog Gateway, a MEC architecture that was fully compatible with standard LTE networks. The gateway decapsulates GTP packets when they are addressed to the fog, avoiding the latency of the backhaul, core and transport networks. This approach provides latency savings but has some trade-offs. For instance, to maintain an updated tunnel database for every UE, we introduced the pinger function, which generates unnecessary ICMP requests. We are also introducing computing overhead as we are analysing all the packets addressed or not to the fog, which introduces small increases on the end-to-end latency to the cloud.

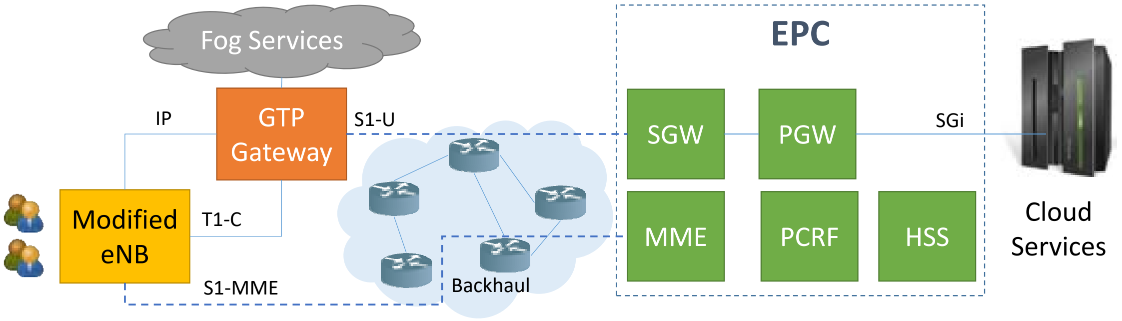


Figure 5.10: Architecture of the GTP gateway [GPM17]

Indeed the root of the trade-off introduced by our proposal is that we are doing some unnecessary GTP encapsulations on the base station (for the packets going to the fog) and checks to all the packets coming out of the base station. We can remove this overhead by modifying the elements of the architecture.

Figure 5.10 depicts our proposed architecture; we have decoupled the tunneling functionality from the base station. The new base station generates IP packets, taking the data directly from the PDCP layer and sending them to the GTP Gateway. The gateway will be in charge of deciding if the packets have to be addressed to the fog subnetwork or have to be encapsulated in GTP to be sent to the SGW using the S1-U interface.

In the architecture, we modify the data plane of the base station, which now send the user traffic without tunnels, but the control plane of the eNB and SGW remain unchanged. Nevertheless, the control plane information of the tunnels has to be shared with the GTP Gateway. There are two alternatives to this:

- the SGW sends the tunnel information to the GTP Gateway. The SGW received the tunnel information from the MME on the S11 interface. This same information could also be forwarded to the GTP Gateway. The main drawback of this approach is that SGW will have to wait for the confirmation from the GTP Gateway before confirming the message to the MME.

- the base station will send the tunnel information to the GTP Gateway, which is the option depicted in the figure with the interface T1-C. This presents the advantage of maintaining all the changes in the base station and not introducing delays on the rest of the control plane.

The interface T1-C (or the S11 if we go for the SGW modification) will share the TEIDs, the endpoint IPs, the TFTs and the bearer IDs. This interface could be implemented with a subset of the messages of the GTP-Cv2 interface like the S11 or S5/S8 interfaces. To create/modify/delete a bearer the base station could use the Create Bearer Request, Modify Bearer Request and the Delete Bearer Request respectively.

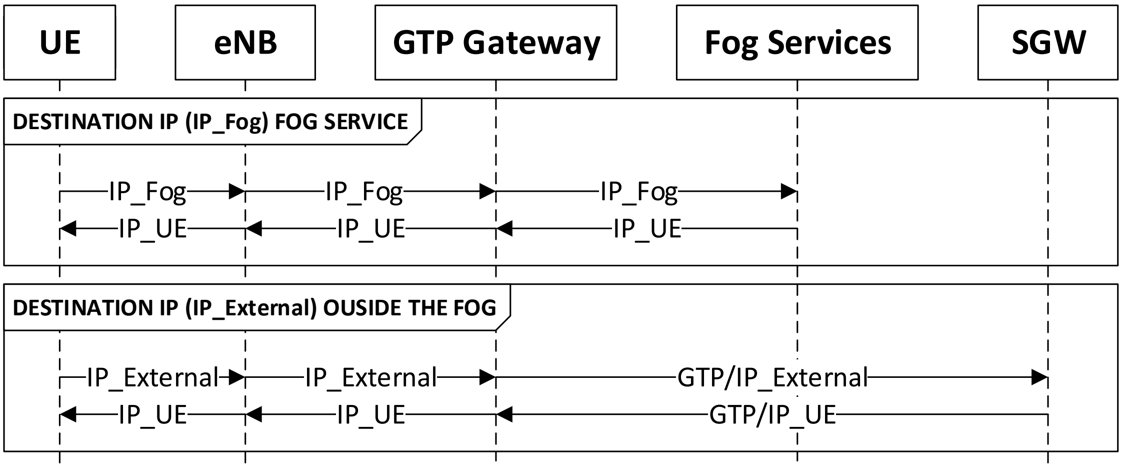


Figure 5.11: GTP Gateway MSC for the user plane [GPM17]

Figure 5.11 provides a Message Sequence Chart (MSC) of the two main scenarios covered by the gateway. When the destination IP address of the packets belongs to the fog the GTP gateway will forward the packet unchanged to the service, if the address does not belong to the fog the gateway will build the GTP header for the packet based on the source address IP and will send it to the SGW.

The GTP Gateway could act as an Access Point (AP), the IP addresses of the UE can be in the same subnetwork than the IPs of the subnetwork services. Any traffic going outside the subnetwork will go through the standard route, which will build a GTP header based on the information of the UE IP.

Following this approach, there is no unnecessary encapsulation/decapsulation, the GTP headers are only built when they are required. Additionally, there is no need for a pinger function as the GTP Gateway maintains an updated database of the established tunnels with the information sent from the base station (received either in the T1-C interface or in an S11 interface coming from the SGW). In any case, the control plane of the base station remains as it is.

5.4.2 Latency Emulated Results

We prepared the setup that is depicted in Figure 5.12 to provide an estimation of the latency that could be achieved with the GTP Gateway.

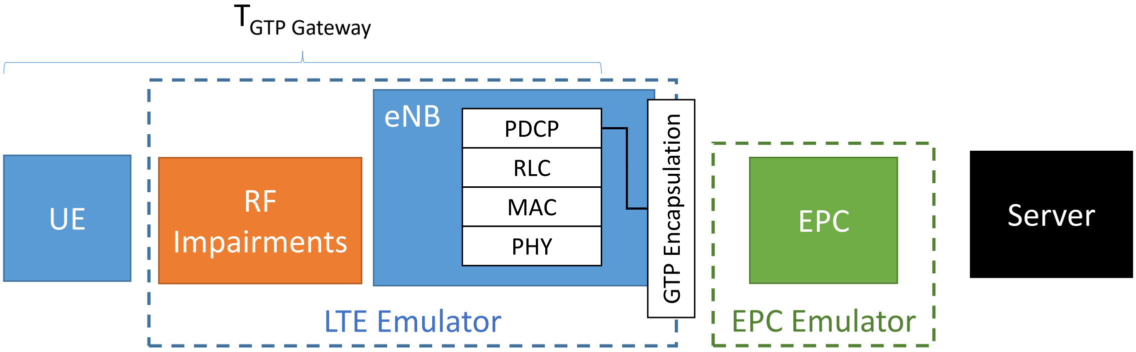


Figure 5.12: Experiment Setup to Evaluate the GTP Gateway [GPM17]

We implemented a tool that was able to capture and analyse ICMP packets at PDCP level, inside the LTE Emulator. PDCP is in charge of handling the packets received from the UE so it will be the layer from which we will forward the packets to the gateway in an actual implementation. We assumed that the bearers were established beforehand, so we did not provide a modelling of the T1-C interface, its contribution will only have an effect on the control plane during the attach or establish dedicate bearer procedures.

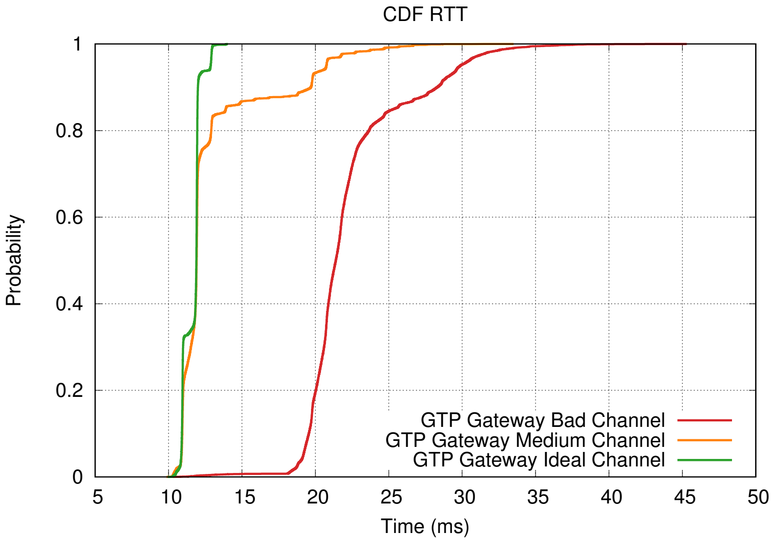


Figure 5.13: GTP Gateway Performance Under Different Channels[GPM17]

With this setup, we configured the T2010 in the same way than we did for the emulated experiments with the Fog Gateway in Section 5.3.2 and measured the RTT under the different channel conditions defined there.

Figure 5.13 provides the obtained results; we obtained a median value of 12ms for the RTT. The setup can provide a good idea on which will be the performance of a real implementation. It is missing the routing of the packets toward the fog network but this routing time would be in the order of microseconds.

5.4.3 Comparison with the Fog Gateway

The obtained results can also be compared with the emulated results obtained for the Fog Gateway in Section 5.3.2 and with the ones for the standard EPC network. Table 5.3 provides a summary of these results.

Table 5.3: Summary of the results obtained with the different latency solutions

Scenario	Median RTT (ms)	MAD RTT (ms)
Standard LTE network	58.8	2.1
Fog Gateway	12.5	0.7
GTP Gateway	12	1

As expected, both the Fog and GTP Gateway provide better performance than the standard architecture as they avoid the backhaul, core and transport networks. The performance they offer is very similar. The difference between the median RTT for the gateways is just half milliseconds, which is something that we could expect as we are just removing the encapsulation part, we think that this difference will be higher if we increase the number of subscribers in both the fog and the cloud networks.

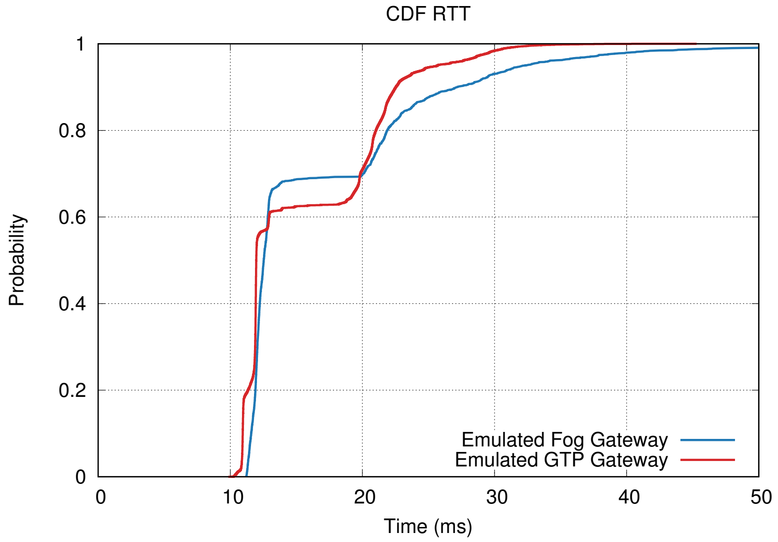


Figure 5.14: RTT CDF Comparison between Fog Gateway and GTP Gateway [GPM17]

Figure 5.14 depicts the RTT CDF of the two approaches. We have not provided an analysis of the control plane procedures, but we should expect some trade-off there, increasing the control plane times obtained in Table 3.1.

We expect the scalability of the GTP Gateway solution to be better than the one of the Fog Gateway as it does not have to infer the tunnel configuration based on the analysis of the packets and there is no need for the pinger functionality to maintain the databases updated. Also, the generation of CDR is easier with this approach, the GTP Gateway has the tunnel information so the consumption can be mapped to an IMSI easily. The main disadvantage of this approach is that it requires modifications on the base station, which will limit its use in existing deployments.

As we stated in [GPM17], we did not provide a comparison with other solutions because of the heterogeneity of the results in the bibliography. The figures between different research change considerably even when establishing a baseline or the time split per component, and in many cases, they are either based on qualitative analysis or focused on a different aspect of the solutions.

5.5 Third Party Exposure

In order to improve the reliability of the service we also proposed the use of the Rx interface of the PCRF to trigger the establishment of a dedicated bearer. This approach could be combined with solutions to setup certain QCIs of the base station to match the needs of the application layer. These features, along with the MEC capabilities either of the Fog or the GTP Gateways, have to be exposed securely to third parties. Figure 5.15 depicts an architecture for MCC, which was discussed in [GPDZR+17a] and [GPRM+17]. The architecture integrates our proposals for the network into a standard LTE network.

In this figure, we have divided the network into different domains. The most simple domain is the Internet domain, which is everything after the SGi reference point in LTE networks. We have the operator domain, where the owner of the network has all the elements required to support service[5], this domain is private and cannot be accessed. For instance, the PCRF, which offers the Rx interface to configure dedicated bearers, sits in this domain and is not accessible by third parties. There is another operator domain where the components to enable access to third parties are located (Operator Domain for Third-Party Network Applications). Interconnection between the two operator domains is made by dedicated secured links, sometimes even with intermediate elements (e.g. we could use a DRA for the Rx interface).

For the API we have considered three principal components from bottom to top the bearer configuration, MEC deployment and QoS. The bearer configuration is a pro-

---

[5] For this particular, and to maintain the Figure compact, we have only depicted elements to provide end-to-end data service, but at least IMS should be there and any other component to support required functionality by the specific MCC scenario.

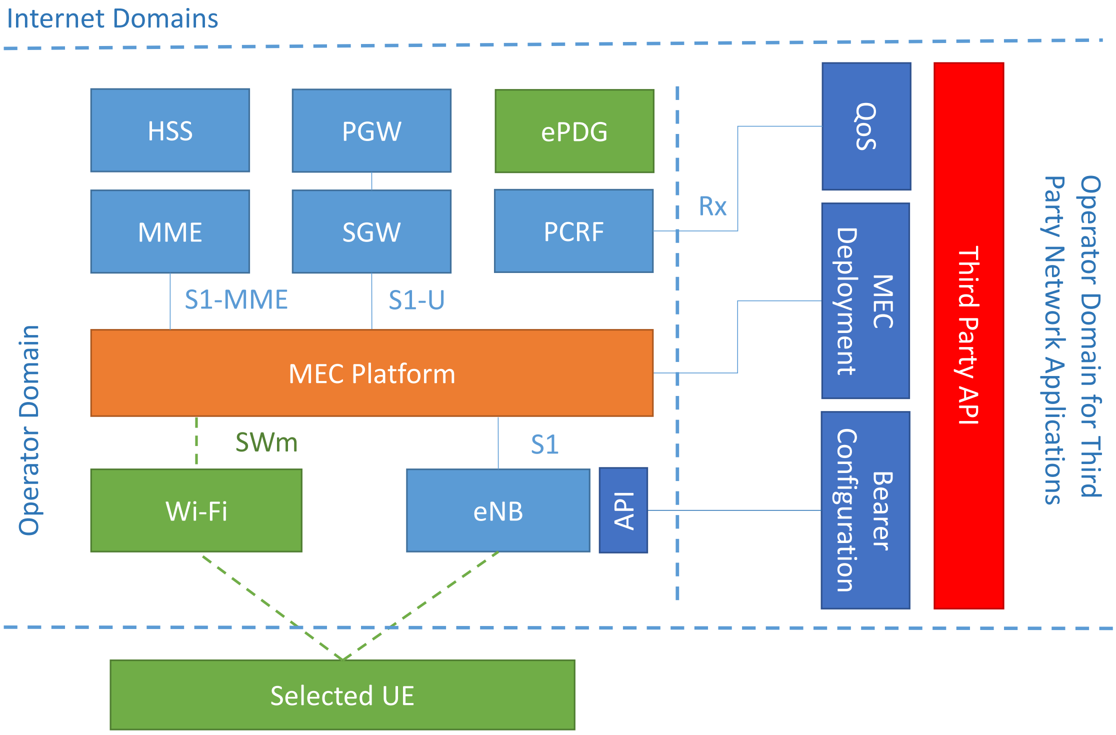


Figure 5.15: End-to-end architecture

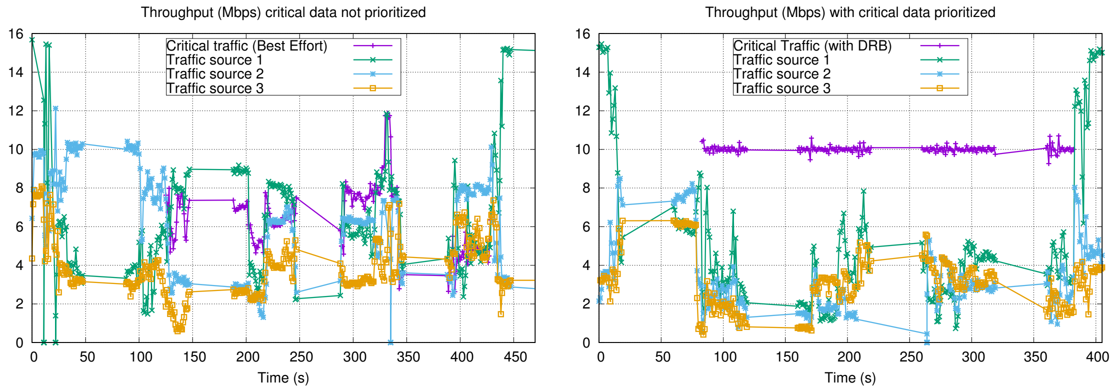
posal to have custom dedicated bearers. Current standards define a set of configurations identified by their QCI, which provides a priority, packet delay budget and packet error loss rate. The base stations then also allow to configure some aspects of these QCIs, we propose to have a range of QCI numbers that could be used to setup a dynamic configuration. We think that this is a clear need as the number of QCIs in the standard have grown from 9 in Release 8 to 15 in Release 14, and this is just focusing on network type applications. This dynamic QCIs will allow defining other aspects besides the end-to-end transport KPIs, such as the size of the buffers in the stack, the number of retransmissions or others. The API could expose these figures directly or also providing abstractions so applications can pick the correct configuration.

The MEC deployment API should offer all the necessary elements for the deployment of a service in the fog servers, the necessary rules to identify the users allowed to use the services and domains that will trigger the access to the fog. These are critical components that we have not considered in our research but poses security and flexibility problems.

Finally, the QoS API provides access to the functionality of the Rx interface. The API will trigger the necessary AA-Request messages, but rather than exposing a Media Component Description (which is composed of a list maximum and minimum bandwidths, which are later to be translated into a matching Policy Charging Control (PCC) rule by the PCRF) the API should offer a more comprehensive function, with a one to one mapping. This mapping could be based on the example services provided in

[3GP19m] or just definitions provided to third party applications with information on both charging and characteristics of the service.

We validated the QoS enforcement using the Rx interface approach in [GPDZR+17a]. To do so we setup one of the Nokia eNB with the our core network and the VELOX engine[6] We connected four terminals to a base station, we use three of them as background traffic generators and the other as a critical traffic generator.



(a) Ideal channel

(b) Non Ideal channel

Figure 5.16: Comparison of traffic with and with QoS enforcement [GPDZR+17a]

Then we proceeded to generate constant throughput from all the UEs as follows: 10Mbit/s from the critical traffic UE and 30Mbit/s from the background traffic generators. Figure 5.16 depicts the results obtained with and without triggering the dedicated bearer from the API. The critical traffic will achieve the required 10Mbit/s when the traffic is prioritized and it will stay around 6.3Mbit/s when it is not.

For the architecture, we have also depicted Wi-Fi APs, which are provided by the operator in order to improve the indoor coverage. Configuration of the APs can be done by a TR-096 interface, other remote management interface or a dedicated network element such as the ANDSF. The ePDG [3GP19c] is employed to allow untrusted access by establishing IPSec tunnels towards the API.

The MEC platform could be either the Fog Gateway or the GTP Gateway, we have not defined how access control will be done to this elements, one option could the use of a Domain Name System (DNS) server with the mapping of the allocated IPs with the IMSI, which will provide the fog IP for allowed subscribers and the cloud IP for not allowed ones. Also, control could be done at the application level, delegating access controls to third party applications.

We also considered the dynamic QoS enforcement over the different IP networks, e.g. on the transport and backhaul networks, which we are no longer considering. This enforcement was provided by employing SDN fabrics on different networks. However,

---

[6] An engine able to trigger Rx request, developed by Redzinc`https://redzinc.net/`.

we still see some limitations in this approach, mainly the effect on the end-to-end latency, the dependency of a central controller and the lack of support of GTP.

5.6 Conclusions

In this chapter, we have discussed the results on the different aspects of MEC and fog computing. To support our results, we have provided a qualitative and quantitative analysis of the contributions to the latency of the different elements of the network. A big part of the latency is consumed by the base stations but also by the backhaul, core and transport networks. We have proposed to deploy the services at the very edge of the network close to the base stations, removing all the contributions not coming from the radio.

The proposed Fog Gateway architecture is fully compatible with standard LTE networks, indeed we have implemented a prototype, which will be described in the next Chapter. Additionally, the architecture does not require the analysis of the control plane procedures to work, which ease the deployment of the tool in an operator environment as the security of the links does not need to be compromised.

Additionally, we have described some of the aspects that have to evolve to support MCC. We have provided our view on the evolution of the S1-U, which is similar to the new N3 interface in 5G, we think that stack could be modified to inject IP from the PDCP level and encapsulate in GTP only the packets going to the fog, Figure 5.17 depicts our proposal from a stack point of view.

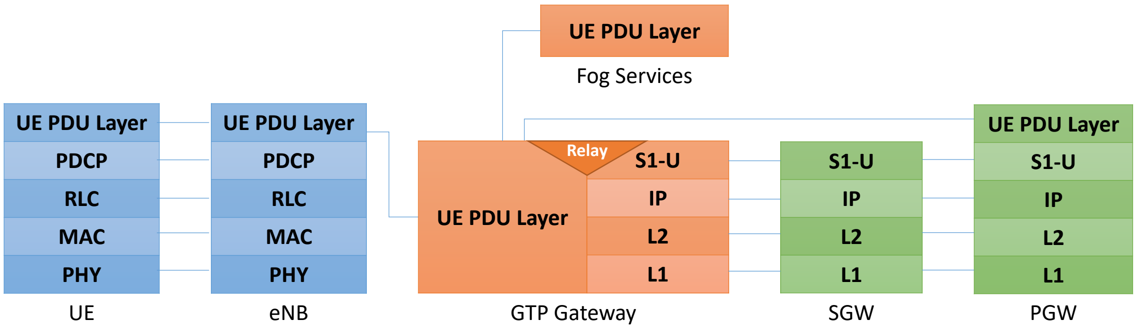


Figure 5.17: GTP Gateway Stacks

The behaviour of the GTP gateway will be similar to the one provided by a Network Address Translation (NAT) router, the IPs belonging to the fog subnetwork will be routed locally while the others will go to the default gateway, which will encapsulate the packets using GTP. The stacks for the user plane for 5G are so far similar to the ones in LTE so the solution could also be applied in the upcoming generations.

The median latency obtained with the emulation of the GTP Gateway is 0.5 ms lower than the one that we obtained with the Fog Gateway, this results will improve with the number of subscribers and offer similar benefits (removal of the core, backhaul and transport networks, reduction of the traffic in the core network, etc.) removing the overhead and the limitations with DRB traffic. The main disadvantages are the security that will need to be improved to prevent UE compromising the servers or other UEs and the requirement of an additional control interface so the base station and the gateway can negotiate the bearer configuration.

We have also proposed an API to make requests over the Rx interface in order to enforce certain QoS; this API could be combined with user-defined QCI so we can tailor the SLA to the service in place. This proposal is still valid but it might not necessary with the introduction in the standards of the MCData services. In this API we also foresaw the mechanisms to deploy service in the fog, which is still something that we need to define.

Our proposals could be used to support different use cases. The most obvious ones are the ones that require very low latency services, as it could vehicle to vehicle communications or augmented reality applications, or in general, applications that exploit the geographical location to provide real-time data. But there are less obvious use cases, for instance the gateway could help to reduce the traffic on the core network caused by massive deployments of IoT elements, as their gateways could be connected to the network. Also, multicast services could be improved, for downlink scenarios the eMBMS architecture could be combined but also for uplink the traffic towards the backhaul/core will be reduced. The resiliency of the services located in the fog is also improved as they will not be affected by errors in the backhaul and/or the core and transport networks.

In the next Chapter we will describe the implementation of a Fog Gateway prototype, as well as the experiments that we performed to characterize it.

Chapter 6

Fog Gateway Implementation

Contents

6.1 Introduction **91**
6.2 Implementation **92**
  6.2.1 Implementation alternatives 92
  6.2.2 Software architecture overview 94
6.3 Latency analysis **96**
  6.3.1 Validation Scenarios 96
  6.3.2 Results 98
6.4 Conclusions **103**

Synopsis

In this chapter, we describe the implementation of a Fog Gateway prototype. We will first discuss the different implementation alternatives. Then we will provide some details on the actual implementation. Finally, we provide some figures on the performance achieved by introducing the gateway, along with an analysis the associated trade-off.

6.1 Introduction

In the previous sections we have characterized, both by a qualitative analysis and some empirical evaluations based on emulations the performance of our proposed solutions. In order to have a more deep characterization and also to better evaluate the trade-off associated with the solution we have implemented a prototype of the Fog Gateway. With the prototype, we will be able to check if, indeed, the solution is fully compatible with standard mobile networks and also provide a better characterization in all the scenarios.

We can find some implementation of prototypes for MEC in the literature, but they are normally they are more focused on the service side. For instance, the authors of [MLZ+20] implement an autonomous Vehicle-to-Everything (V2X) prototype and evaluate with a commercial MEC+NR solution. In [LLL+20] a prototype based on OAI is implemented to evaluate a security design for MEC services. OAI is also employed in [ABF+19] to evaluate a virtual sensing service for vehicles. In our case, OAI was employed during the development phase, but the evaluation was carried out using conformance testing equipment and commercial base stations.

To implement the prototype we evaluated different strategies, which are the implementation for a GPP, the SDN/Network Function Virtualization (NFV) implementation. Many authors propose one or both of these paradigms to implement MEC services. Experimental NFV MEC deployments are characterized in terms of scalability, time to deploy and operational cost in [CCB+16]. An emulation of an NFV system is used to generate some results and some theoretical background in [YCPK16]. An NFV-MEC Wi-Fi demonstrator is implemented in [HLT+18], focusing on supporting different slices per type of peer. SDN is naturally employed, as the network functionality necessary to redirect the traffic can be implemented on SDN switches. An example of this approach is [HNS+17b], whose authors propose and evaluate a SDN-MEC architecture. The validation is done using a patched version of OVS, to support GTP, and OAI (described in chapter 2), showing savings both in latency and CPU load. Multi-domain fog services are characterized in terms of required rules in the switches in [BDLP17]. The discussion of the different implementation approaches is done in section 6.2.The security of these platforms is fundamental, for instance, in [LQNL17], a monitoring system to detect man in the middle attacks is designed.

However, we concluded that the existing SDN open source implementations were not enough to support our design with the required performance and went for a GPP implementation. We could have gone for a kernel implementation, tied to an Ethernet card but at the end we opt for something intermediate, using iptables and libnetfilter-queue[1] to capture packets in the kernel, this will cause an extra copy of the packets but it also improves the portability of the solution. To characterize our prototype, we first characterize the contribution of the underlying platform. Then, we carried out several campaigns to estimate the trade-off introduced in the packets going to the cloud and the

---

[1] `https://netfilter.org/projects/libnetfilter_queue/`

latency savings for the ones going to the fog services.

In this section we will first provide an overview of the different implementation strategies, then we will describe the architecture of our implementation. Then, the latencies will be characterized and also some basic tests for the throughput will be performed.

6.2 Implementation

The GPP and SDN/NFV implementation strategies will be discussed in this section, then we will provide an overview of the architecture of the prototype.

6.2.1 Implementation alternatives

GPP implementation

The most traditional approach for the implementation of network elements is the use of high-reliability ATCA platforms. For instance, there are vendors that provide stacks to support these implementations [WN10]. These solutions offer many advantages such as scalability, performance and pre-defined hardware functions. The main drawback is their price.

In our case, we decided to go for a GPP implementation with in-house stacks making use of open source libraries. This implementation runs over a regular computer and is based on a monolithic architecture. Raw sockets can be used to access the IP headers of the packets received in each interface and, after parsing them, pass them to the appropriate handler, which will decide whether to modify, forward or drop it.

For the implementation, carried out in [GP17], we provided two different versions of the gateway, one prototype implemented in python and another in C++. This one this way to ease the development, prototyping in python can be done fast and provide valuable inputs for the C++ design, which is normally more complex. Figure 6.1 depicts the approach that we followed, iterating over both implementations and using the feedback from one into the other.

In both cases, we made use of libraries providing high-level functionality such as libpcap[2] or libnetfilterqueue[3] for capturing packets and libnet[4] or scapy[5] to inject them. We could have used a kernel implementation to increase the performance, by avoiding a copy of each packet from kernel to user space, but at the cost of introducing dependencies with the underlying hardware platform, reducing the portability of the application.

---

[2] `https://www.tcpdump.org/`

[3] See footnote 1.

[4] `http://libnet.sourceforge.net/`

[5] https://scapy.net/

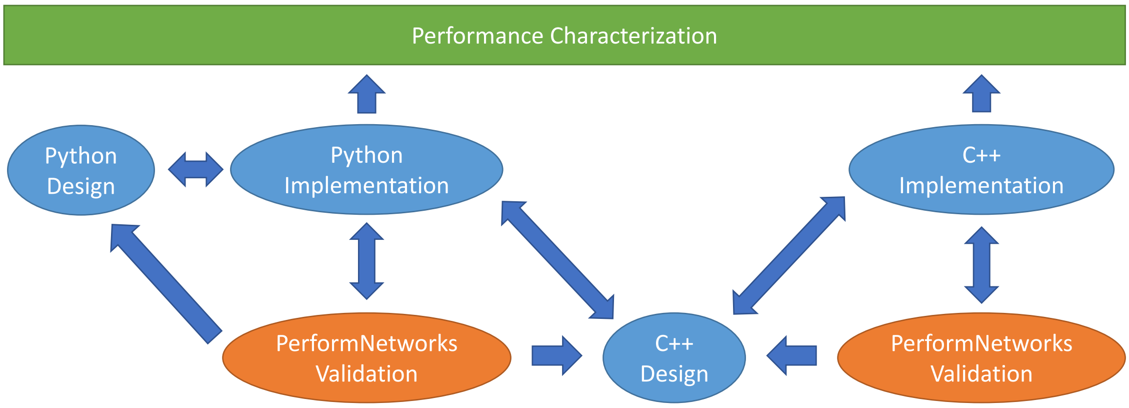


Figure 6.1: Prototype implementation methodology [GP17]

SDN/NFV Implementation

The SDN and NFV paradigms consist of a set of techniques aiming to separate the implementation for the underlying hardware and centralize the configuration to prevent dedicated low-level configuration per service. To achieve this, generic switches are introduced to support fast L2/L3 functionality managed by a central controller. The systems will be implemented as applications that will run on a central controller. The Fog Gateway architecture is suitable to be implemented using these paradigms but functionality to support high-performance fog was missing on the open source SDN switches implementations.

The main missing part to implement a high-performance SDN Fog gateway was the native support for GTP. Adding GTP support on the controller was feasible but impractical as it will mean that all the GTP packets will have to go to the controller. So to support GTP without affecting the performance, the switch will have to be modified to support parsing, addition, removal and modification of GTP headers. Additionally, the traffic that will go to the fog has to be identified and to do that we can use dedicated bearers or modify the switch header to support inspection of the IP headers transported by GTP.

The NFV paradigm could be considered a generalization of the SDN one, both paradigms are complementary, and they are frequently used together (normally abstracting network functions also require the use of abstracted L2 functionality). The paradigm consist on providing an architecture where the network behaviour is provided in functions that can be deployed separately with network servers on demand. 3GPP agreed on adopting the Management and Orchestration (MANO) architecture [ETS14], provided by European Telecommunications Standards Institute (ETSI) on the standards [3GP15a] and the new 5GC has introduced a new service architecture [3GP19z] and [3GP19w], but without the required support for SDN we decided not to explore NFV neither.

6.2.2 Software architecture overview

Figure 6.2 depicts the software architecture of the prototype that was implemented in [GP17]. The system was divided into several subsystems in order to reduce the complexity and to isolate external dependencies so they could be replaced when needed. The three main functionality blocks are packet parsing, which provides encoding and decoding of packets, including packet reassembly; packet injection, which provides functionality to send packets; and the iptables queue handling, which handles the packages received from iptables.

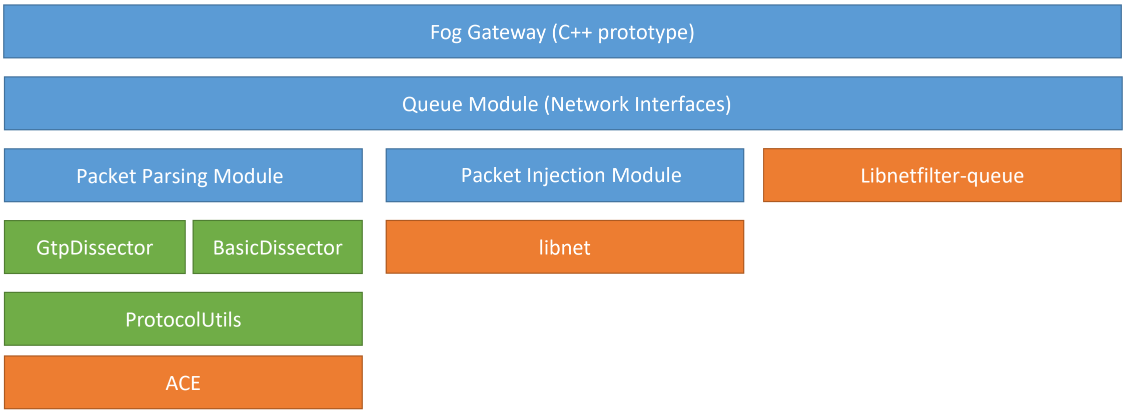


Figure 6.2: Software architecture diagram [GP17]

The external dependencies employed were the Adaptive Communication Environment (ACE) libraries[6], which provides multiple abstractions useful in software development, libnet[7] that was used to craft and inject packets and libnetfilterqueue[8], which is used to extract packets from the kernel using iptables.

The Packet Parsing module class diagram is depicted in Figure 6.3. This module is in charge of parsing different GTP and IP packets, including the ones transported over IP. The module has to support IPv4 reassembly[9] to forward the packets correctly to the fog layer. The parsing of the packets is based on the dissector library that we described in Section 2.4.2 as well as on a utility library, ProtocolUtils, implemented on top of ACE.

Figure 6.4 depicts the class diagram for the Queue Module. QueueInterface is an abstract class that provides the main loop in each of the interfaces of the system used in the prototype. The loop can be mapped to the logic that was previously described in 5.3. The first step is the parsing of the received packet (function parsePacket); in this phase,

---

[6] `http://www.dre.vanderbilt.edu/~schmidt/ACE.html`

[7] See footnote 4.

[8] See footnote 1.

[9] If the Ethernet layers are not configured to support Jumbo Frames and the UE is using the MTU for Ethernet, the packets received in the SGW will probably be fragmented (as with the GTP header the maximum MTU size will be exceeded).

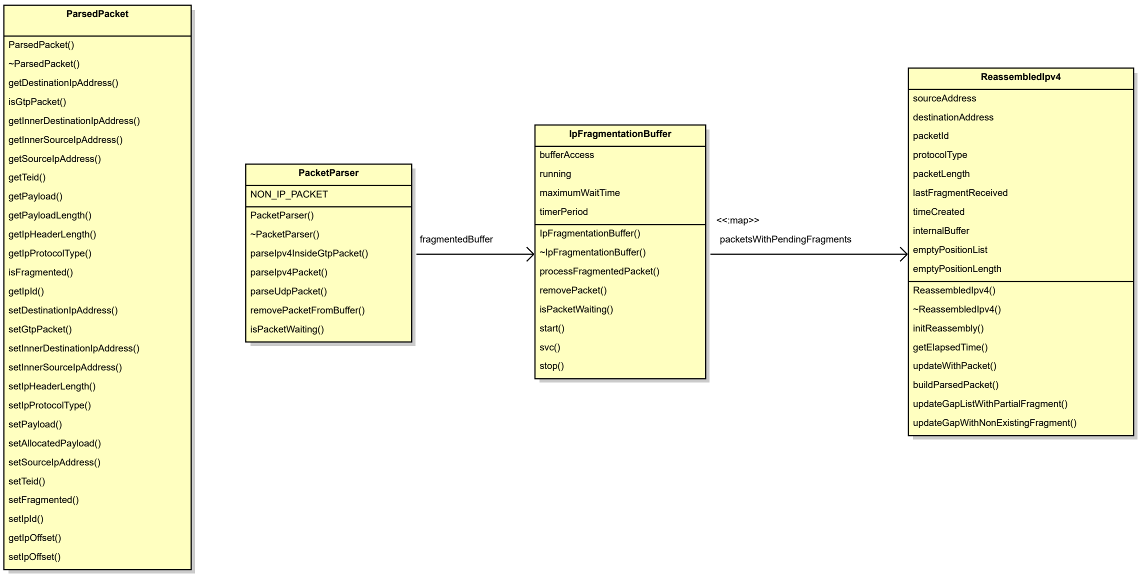


Figure 6.3: Packet Parsing Module Class Diagram [GP17]

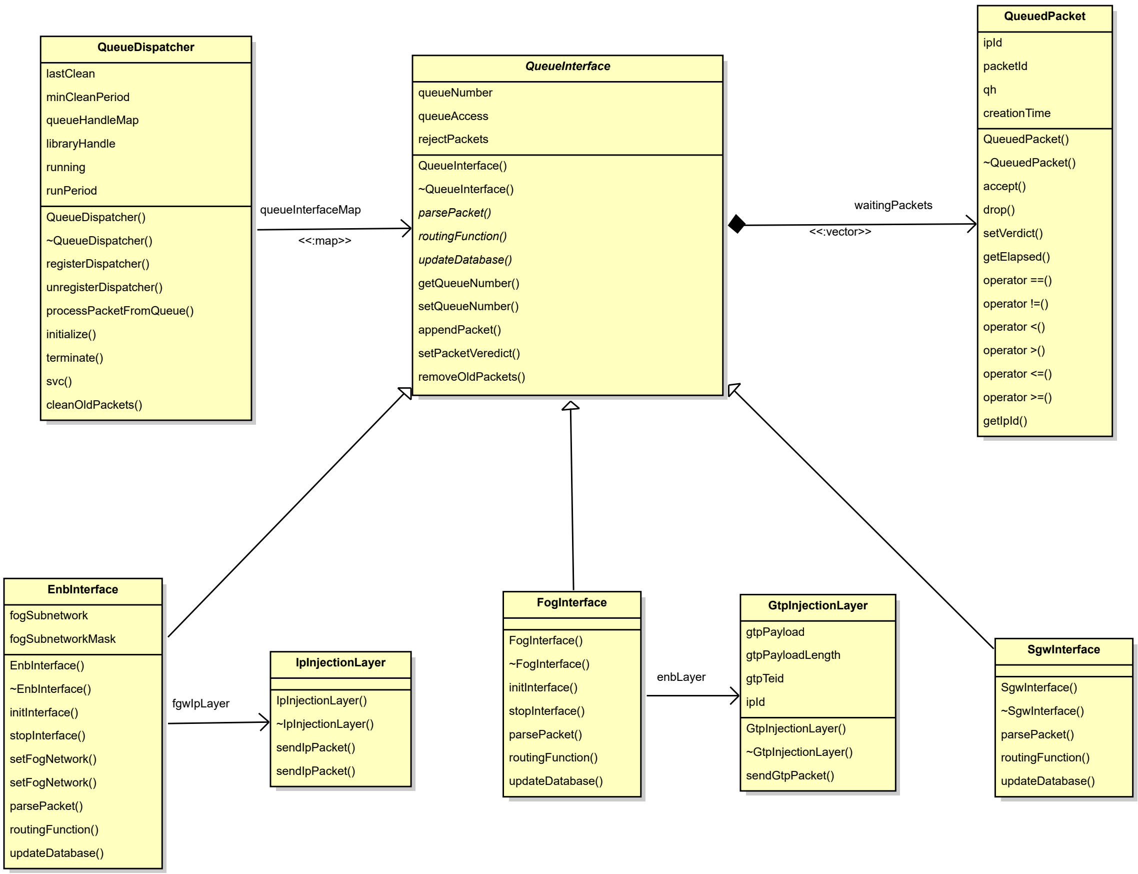


Figure 6.4: Queue Module Class Diagram [GP17]

the IP (for the interconnection with the fog) and the UDP and GTP (for the interconnections with the eNB and the SGW) headers are extracted. Then the information obtained from the packet is employed to update the corresponding database (function updateDatabase, downlink database for packets coming from the SGW and the uplink database for packets coming from the eNB). Finally, the gateway will decide (in the routingFunction) if it has to route the packet to a different entity, e.g. the fog subnetwork, or if it has to leave as it is and let the Operating System (OS) handle the packet normally. When a packet is routed to a different entity, a new packet will be crafted, and the received one will be dropped. All the interfaces of the Gateway implement this class.

Finally, the Injection Module provides two classes GtpInjectionLayer and the IpInjectionLayer. For GTP a normal UDP socket is employed, for IpInjectionLayer a raw socket could have been used but instead, we use libnet, which provides some packet handling functions, in C++ and Scapy, with a similar purpose, in Python.

6.3 Latency analysis

This section provides some of the results obtained with the prototype and compares them to different scenarios.

6.3.1 Validation Scenarios

The validation was carried out employing different toolsets. On the one hand, for the development process, an all software setup was designed. For the experimental results, different combinations of hardware and software were employed, following the methodologies described in Chapter 2.

Development setup

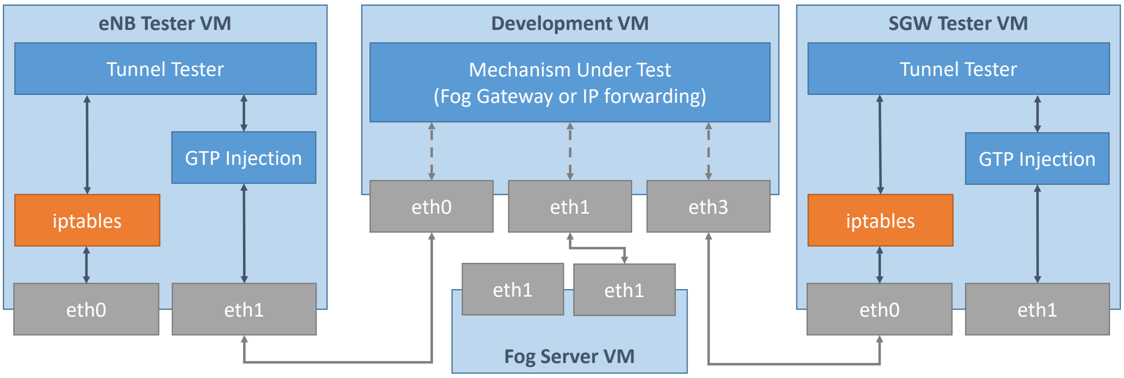


Figure 6.5: Development setup [GP17]

The development setup is depicted in Figure 6.6, it is entirely based in Virtual Machines (VMs) so it can be used in a regular computer. Four different virtual machines were employed on this setup, as follows:

- eNB Tester VM, which was employed for emulating the GTP traffic of an LTE base station. The GTP traffic is generated on real user traffic on one of the interfaces of the VM.
- SGW Tester VM, which generated the GTP traffic of an SGW and the GTP header removal and injection of the IP user traffic.
- Development VM, which ran the Integrated Development Environment (IDE) and the implementation under test.
- Fog Server VM that was in charge of running the different services.

The eNB and EPC emulation are based on the TunnelTester tool described in Section 2.4.2, which encapsulates user-defined traffic into GTP, and, in the other peer, it removes the GTP headers from the received responses and inject the transported IP packet back in the machine.

The development VM is where the software under development is deployed. It also serves to test other mechanisms in order to establish a baseline. In the Fog Server a regular service is deployed depending on the tests, for latency ping is used and for throughput characterization iperf.

After a development phase, we do an integration phase, which is normally done with OAI over and SDR card, as this configuration is portable and allow quick testing after fixing a bug. During the integration phases, we checked that the Fog Gateway is still working with standard network elements.

Experiments configuration

The setup employed for the experiments is depicted in Figure 6.6. Different radio accesses were used on different phases of the experiments. Commercial base stations were used for the compatibility and integration tests, the experiments under different radio channel conditions were run on the T2010, and OAI implementation over SDR was used for demonstration purposes.

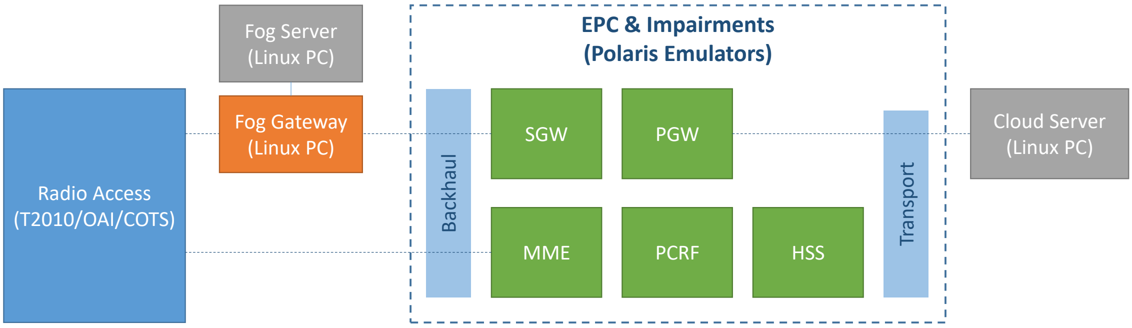


Figure 6.6: Experiments setup [GP17]

The Fog Gateway was deployed on a standard PC with multiple Ethernet interfaces; the employed OS was Linux. The services also run on a similar configuration, commercial computers with Linux on top. This setup is employed for the "‘Fog Gateway to Fog Service"’ and "‘Fog Gateway to Cloud Service"’ scenarios. The "‘Standard EPC to Cloud Service"’ scenario uses the same architecture but without deploying Fog Gateway.

The core network configuration provided access for a single operator with the basic elements to have end-to-end communication (MME, SGW, PGW, PCRF and HSS as depicted in the figure). Additionally, it was also used to introduce IP backhaul and transport impairments.

The configuration employed for the T2010 machines is provided in Table 5.1. For all the scenarios we introduce a normal distribution with mean 15ms in the backhaul link and a normal distribution with mean 30ms in the transport network.The channel configuration for the latency experiments is similar:

- Ideal channel, no fading and no noise.
- Medium Channel, fading profile EVA70 and noise power -80dBm, leading to a MAC BLER of 10%.
- Bad channel, fading profile EVA70 and noise power -70.5dBm, leading to a MAC BLER of 50%.

6.3.2 Results

The results obtained with the prototype were described in detail in [GP17]. The first experiment that was conducted with the prototype was designed to analyse the contributions from the underlying platform and the implementation. To do so, we analysed three different scenarios:

- A direct connection between the base station and the core network.
- A PC to interconnect the eNB and the core. In the PC forwarding is done with iptables, and artificial ARP is generated with arping.
- Fog Gateway prototype connecting the base station and the core network.

In these scenarios, we measured the UE RTT, Table 6.1 provides a summary of the Median and MAD results obtained. The introduction of a PC acting as a switch introduces an additional 0.8 ms to the RTT. If we use the Fog Gateway to evaluate services located in the fog using the same PC 1 ms is introduced. We can estimate the overall latency contribution from our implementation roughly to be 0.2 ms. Figure 6.7 depicts the CDF obtained in each of the connection scenarios.

After characterizing the platform, we used the T2010 under different channel conditions (defined in Section 6.3.1) to compare the Fog Gateway prototype and the standard

Table 6.1: Summary of the platform characterization results

Connection Type	**Median RTT (ms)**	**MAD RTT (ms)**
Fog Gateway	58.5	11.2
PC as switch	58.3	11.8
Direct Connection	57.5	11.6

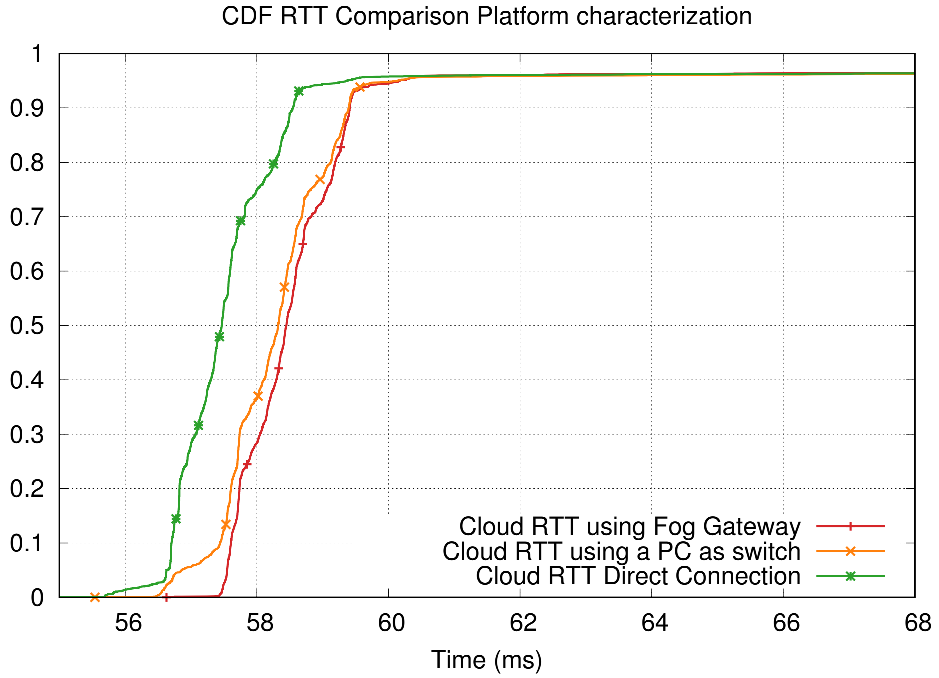


Figure 6.7: Fog Gateway platform contribution characterization [GP17]

EPC RTT. Figure 6.8 depicts the obtained CDF in each of the scenarios.

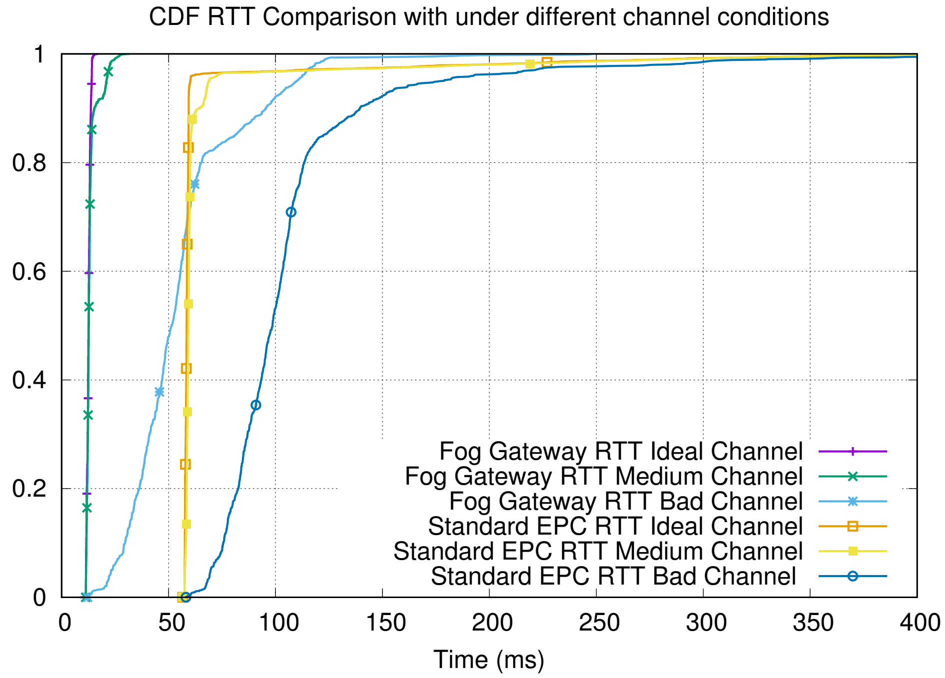


Figure 6.8: Fog Gateway fog services comparison [GP17]

The Fog Gateway provides better latency results in all the scenarios, although under bad channel conditions the performance will be similar to the one provided by services in the cloud with better channel conditions, so it is clear that if reliability is required additional mechanisms will have to be used. The RTT has been reduced around 45 ms in the ideal and up to 47 in the worst-case scenario, which is the gain that we would expect as the gateway removes the contributions from the backhaul, core and transport networks. Table 6.2 provides the median and MAD RTT on different scenarios.

Table 6.2: Summary of the Fog Gateway and standard EPC comparison

Scenario	Fog Gateway (median/MAD)	Standard EPC (median/MAD)
Ideal Channel	12.6 ms/0.63 ms	58.5 ms/11.2 ms
Medium Channel	12.7 ms/1.6 ms	59.3 ms/11.8 ms
Bad Channel	51.5 ms/1.83 ms	98.6 ms/24.9 ms

The gain when comparing fog services is apparent, but we also have to evaluate which is the trade-off of the platform, this is which is the increase of the latency for services not located in the fog. To do so, we analysed the RTT performance when comparing cloud services served by a standard architecture with others served using the Fog Gateway. Table 6.3 provides a summary of the comparison under different radio conditions, and Figure 6.9 depicts the RTT CDF in the different scenarios.

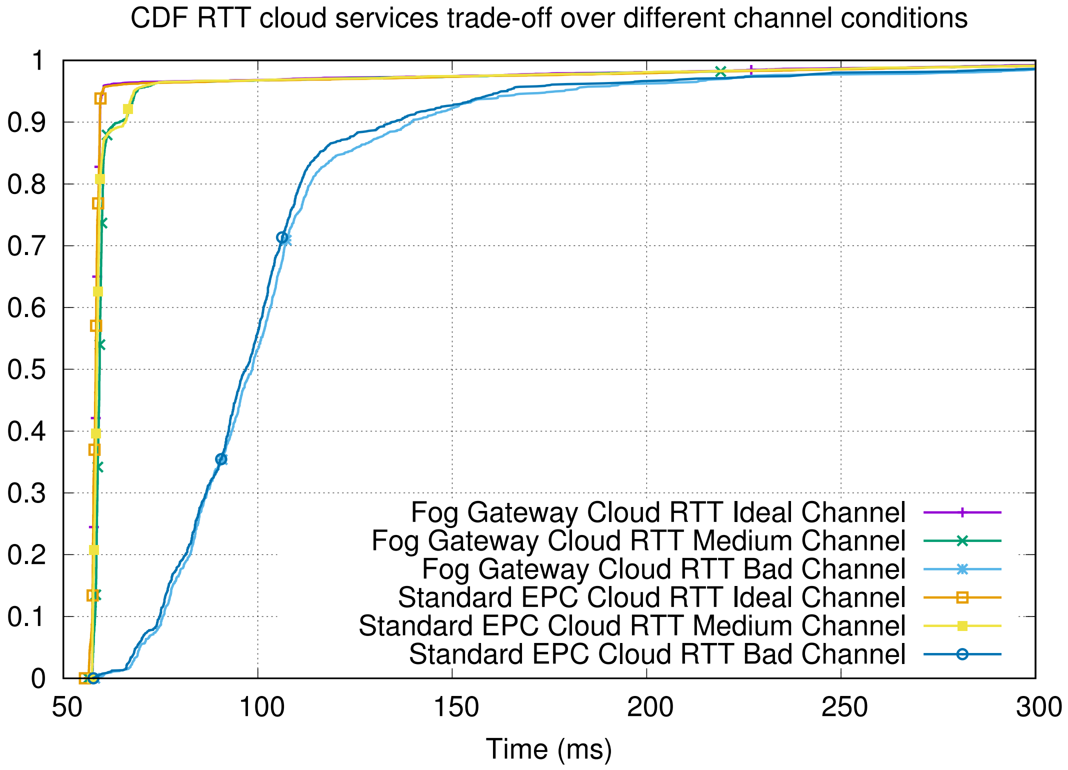


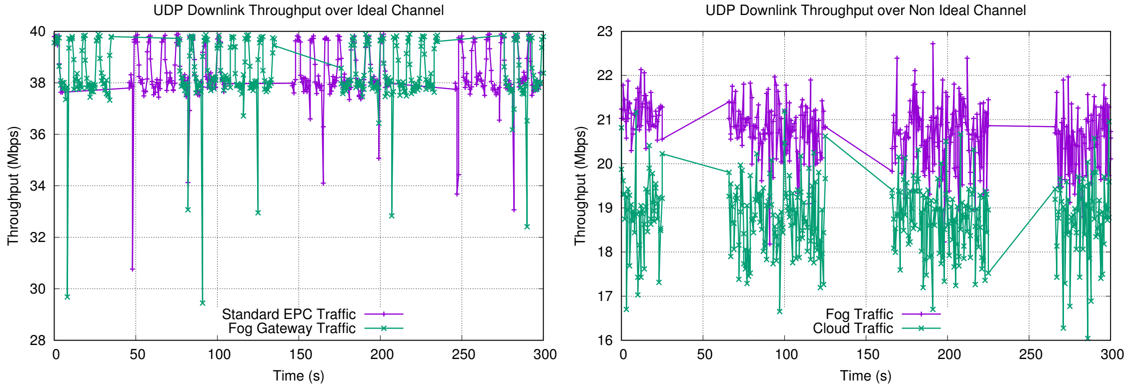
Figure 6.9: Fog Gateway cloud services comparison [GP17]

The latency introduced by the platform is approximately 0.2 ms in the ideal scenarios and increases when channel conditions get worst. These results are consistent with the ones that we obtained for the platform characterization in Table 6.1.

Table 6.3: Cloud Services RTT when using the standard EPC or the Fog Gateway

	Standard EPC (median/MAD)	Fog Gateway (median/MAD)
Ideal Channel	58.3 ms/11.8 ms	58.5 ms/11.2 ms
Medium Channel	58.6 ms/11.8 ms	59.3 ms/11.8 ms
Bad Channel	97.0 ms/23.2 ms	98.6 ms/24.9 ms

Another aspect of the trade-off, which has to be considered, is the throughput characterization. To evaluate it we employed the T2010 emulator, but only UDP was considered as the S1 interface implemented there introduces packet reordering in the links, affecting the performance of TCP. For the downlink we provided measurements for both ideal and non-ideal channel (in this EVA70 with noise power -75dBm) while for the uplink we only considered ideal channel measurements as the unit does not provide uplink RF impairments. We could have added external impairments with a channel emulator, but the receivers of this unit are not optimized to work under impairments so the results would have been limited. We evaluated the UDP performance for uplink and downlink with and without using the fog service.



(a) Ideal channel (b) Non Ideal channel

Figure 6.10: UDP downlink throughput [GP17]

Figure 6.10 provides outcomes for downlink on ideal and non-ideal scenarios, we have similar behaviours for both the standard EPC and the fog gateway. As mentioned before for the uplink we have only considered the ideal scenario. Figure 6.11 depicts and outcome for the uplink on an ideal scenario, the uplink measurements show some glitches, which are the result of the implementation of the emulator.

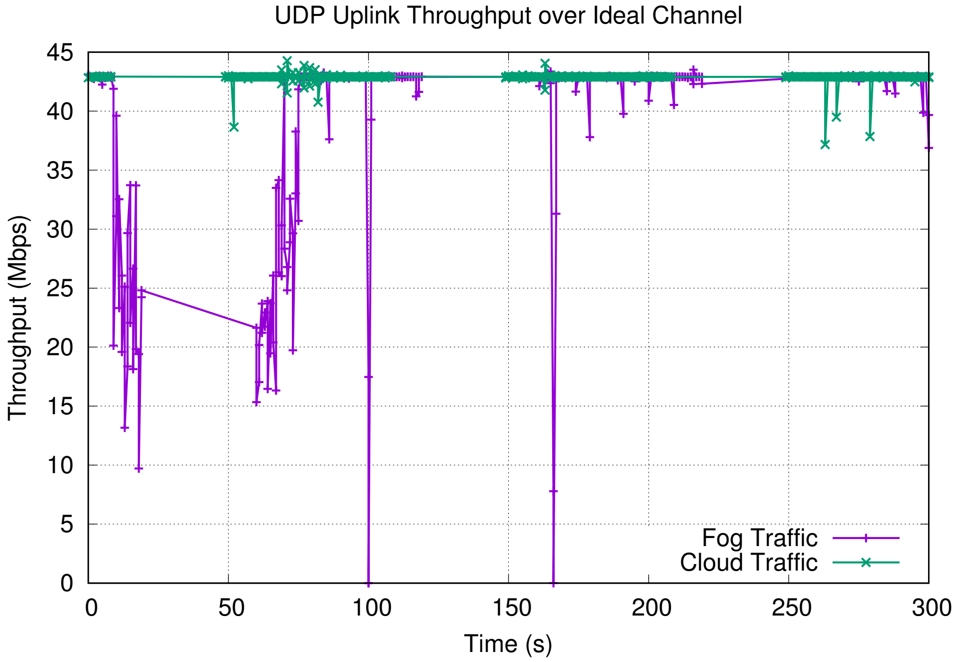


Figure 6.11: UDP uplink throughput comparison [GP17]

Table 6.4 provides a summary of the experiments that were carried out. The results obtained for the fog and cloud traffic are similar with minimal differences both in the uplink and the downlink, as expected the results for UDP are similar.

Table 6.4: Summary of the UDP throughput with and without the Fog Gateway

Scenario	Fog Gateway Throughput (Mbps		Standard EPC Throughput (Mbps)
	Fog Service	Cloud Service	Cloud Service
Ideal Downlink	40.036	38.376	38.4
Non-Ideal Downlink	20.073	18.78	18
Ideal Uplink	39.661	42.846	41.684

6.4 Conclusions

In this Section we have validated our Fog Gateway proposal, its architecture is fully compatible with standard LTE networks and indeed has been characterized with commercial base stations and also with emulators, to provide a better characterization of the solution under adverse channel conditions. Additionally, the Fog Gateway does not require the analysis of the control plane procedures to work, which ease the deployment of the tool in an operator environment as the security of the links does not need to be compromised.

A discussion on the implementation approaches has been carried out, we have discarded the SDN/NFV approach due to the limited support of the functionality required by the gateway. The ATCA platform has been also discarded, mainly due to the price and we have gone for a GPP implementation based on open source libraries. The architecture of this implementation and its main modules have been described.

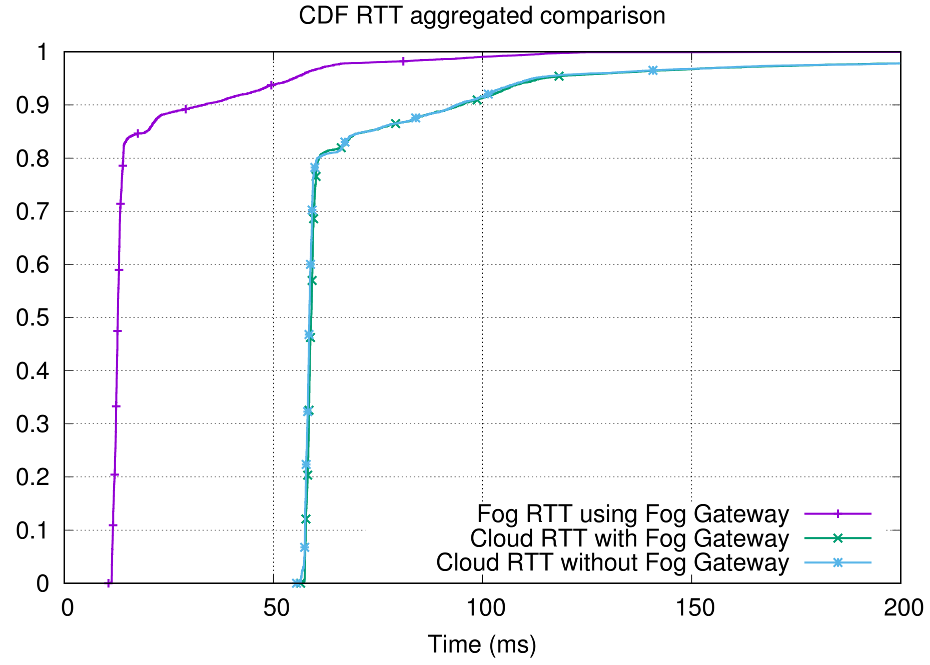


Figure 6.12: Summary of the results obtained with the Fog Gateway [GP17]

We have compared the performance of our prototype with the one obtained connecting the base station directly to the EPC or using a PC as a router (employing IP forwarding), showing a median difference of 0.2 seconds with the PC as router and 2ms with the direct connection. The trade-off of the solution for the traffic going to the cloud has also been evaluated, ranging from 0.2ms on the ideal scenario to the 1.6 ms on the bad channel conditions scenario. The latency reduction for services going to the fog servers is apparent and stays around 45 ms without affecting the throughput of the base station. The main feature that we missed was the characterization of the UE to UE communications that was attempted, but with many errors on the link a commercial base station. The improvement of the latency in this scenario should be better as it will benefit from removing twice the contributions from the backhaul and core networks. Figure 6.12 depicts the performance of fog and cloud services with and without using the gateway.

The obtained results are good, the main aspects missing from the evaluation are the tests at scale, to assess which is the performance of the prototype when handling hundred of subscribers, and the TCP performance, as it will be interesting to evaluate both the trade-off for cloud services and the gains, in particular in bad scenarios.

Chapter 7

Conclusions

Contents

7.1 Summary of our contributions 107

  7.1.1 Publications 107

  7.1.2 Tools 108

  7.1.3 Projects 109

7.2 Discussion on the results 110

  7.2.1 Mission Critical Mobile Networks 110

  7.2.2 Experimental Testbeds 111

7.3 Future work 112

Synopsis

In this section, we highlight the main contributions of this thesis. We start with a summary of them, covering the scientific publications, the tools and the research projects in the context of the thesis. Then we discuss the different results that we have obtained. Finally, we provide some insights on future research to follow up on the works of this thesis.

7.1 Summary of our contributions

In this section, we provide a summary of the contributions done in this thesis. First, we cover the scientific publications, then the different tools and finally the research projects and the private sector partnerships are described.

7.1.1 Publications

Table 7.1 summarizes the list of the relevant publications for this thesis. On these journals, we covered the provision of communications for railways, methodologies and architectures for eHealth applications and the provision of low latency communications over LTE networks. We have also presented works in six congresses, mainly on contributions to experimental platforms. Also, four book chapters have been published again covering experimental platforms, application and services for public safety networks and architectures to support IoT. Finally, we have also two master thesis one exploring video for mission critical applications and another on the implementation of one of our latency gateways.

Publications	Published in	Type[1]	Reference	Year
3GPP Standards for the Delivery of Critical Communications for Railways	IEEE Vehicular Technology Magazine	J(Q1)	[DZGPMG14b]	2014
Improving the Efficiency and Reliability of Wearable based Mobile eHealth Applications	Elsevier Pervasive and Mobile Computing	J (Q3)	[GPDZR$^+$17a]	2017
Experimental Evaluation of Fog Computing Techniques to Reduce Latency	Wiley Transactions on Emerging Telecommunications Technologies	J (Q1)	[GPM17]	2017
PerformLTE: A Testbed for LTE Testing in the Future Internet	Springer Selected Papers, Wired/Wireless Internet Communications (WWIC)	C	[DZGPMG15]	2015
3GPP Standards to Deliver LTE Connectivity for IoT	IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)	C	[DZGPRPM16]	2016
Q4HEALTH: Quality of Service and Prioritisation for Emergency in the LTE RAN Stack	European Conference on Networks and Communications (EuCNC)	C	[GPRM$^+$16]	2016
Extensive and Repeatable Experimentation in Mobile Communications with Programmable Instruments	International Conference on Remote Engineering and Virtual Instrumentation (REV)	C	[GPRPRG$^+$16]	2016
Remote Control and Instrumentation on Mobile Devices		C	[DZRPGPM16]	2016
Enabling Low Latency Services on LTE Networks	IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)	C	[GPM16]	2016
PerformNetworsk: A Testbed for Exhaustive Interoperability and Performance Analysis for Mobile Networks	River Publishers Building the Future Internet through FIRE	B	[GPMMR17]	2016
Q4Health: Mission Critical Communications over LTE and Future 5G Technologies		B	[GPRM$^+$17]	2016
Supporting new Applications and Services over LTE Public Safety Networks	Elsevier Wireless Public Safety Networks 3	B	[GPDZR$^+$17b]	2017
3GPP Evolution on LTE Connectivity for IoT	Springer Integration, Interconnection, and Interoperability of IoT Systems	B	[DZGPRPM18]	2018
Video for Mission Critical Applications over LTE	Universidad de Málaga	A	[GP15]	2015
Low Latency Communications for LTE networks		A	[GP17]	2016

Table 7.1: Summary of the publications

---

[1] The type has to be interpreted as follows: J=Journal, and includes also the JCR quartile; C=Congress; B=Chapter, in book; A=Academic.

Three journals indexed in the JCR have been published, two of them in the first quartile. The "Low Latency Communication for LTE networks" master thesis received the "Orange award to best master thesis in high-speed mobile communications" awarded by the Spanish Association of Telecommunication Engineers in 2018. According to ResearchGate[2] the publications provide an h-index (excluding self-citations) of six.

7.1.2 Tools

Many of the contributions to tools of this work have been done around the experimental platform PerformNetworks, where we have done contributions both at design and implementation level. Table 7.2 summarizes the list of contributions.

We have designed a tool to ease the integration of SCPI compliant instrumentation into OMF and implemented two SCPI compliant instruments, one to provide on-demand IP impairments on remote interfaces and on-demand deployment of EPC elements. Low latency prototypes of the Fog Gateway have been developed, as well as others to evaluate reliability, the Redundancy Tester, or different approaches to data plane acceleration (by analysing the control plane with the S1 Database Generator). To ease the development of these prototypes, we have used the Tunnel Testes, which supports encapsulating any traffic using GTP. Tools to characterize timing has been provided such as the Ping Analyzer (to extract RTT of ICMP samples transported over IP, GTP or PDCP) and the Control Plane Analyser, which provided functionality to measure the time consumed by S1 procedures. Finally, we have also developed a methodology to improve the selection of a UE for a given application.

Tool	**Contribution**[3]	**Area**
SCPI-Resource Controller	D	Experimental Platforms
Remote Impairments Configuration	D/I	Experimental Platforms
EPC on-demand rollout	D/I	Experimental Platforms
Ping Analyser	D/I	Measurements
Control Plane Analyser	D/I	Measurements
S1 Database Generator	D/I	Prototype
Tunnel Tester	D/I	Development Tool
Redundancy Tester	D/I	Prototype
UE Selection	D/I	Methodology
S1 Extensions for Wireless Test Sets	D/I	Experimental Platforms
Fog Gateway Implementations	D/I	Prototype

Table 7.2: Summary of the tools

---

[2] `https://www.researchgate.net/profile/Cesar_Garcia_Perez/scores`

[3] The contribution column has to be interpreted as follows: D=Development, I=Implementation, D/I=Development & Implementation.

7.1.3 Projects

The results of this thesis have also been actively used by research projects. We have contributed to them both during the proposal and implementation phases. Table 7.3 provides a summary of the key technical contributions to research projects.

Project	**Type**[4]	**Year/s**	**Key Technical Contributions**
Q4Health	E	2016-2017	Low latency communications, architectures for eHealth
Triangle	E	2016-2017	S1 Extensions for UXM, roll out tools
Fed4Fire/Fed4Fire+	E	2014-2017	Testbed integration
FLEX	E	2015-2016	S1 Extensions for T2010, interoperability
Tecrail	N	2014	Railway architectures and pilots

Table 7.3: Summary of the research projects

The initial idea of this work was conceived during the last year of the Tecrail project, in particular during the analysis of new standards to support railway communications. The Fed4Fire project was already active, in it and its continuation Fed4Fire+ we have made many contributions to experimental platforms and supported external experiments with some of the tools developed. On the FLEX project, we provided the implementation of the S1 extensions for the T2010 test set and carried out many interoperability tests with other experimental platforms. Triangle offers the possibility of continuing working in the UXM, which already provided pre-5G functionality, where we supported mainly work around the design of the extension and the implementation of the S1-C stack, along with other testbed extensions.

Most of the contributions of these project were performed in the context of the project Q4Health, where we developed our solutions to support low latency networks, designed proposals of network evolutions, methodologies to select UEs and more.

In the context of the projects, different collaborations with the private sector were done, to mention the most important ones:

- Polaris Networks [5] lend us some equipment that used to validate the proposals of this project and assisted us on the use of the emulators in different scenarios.
- Redzinc[6], a small company focused on the provision of solutions for blue light services. Several research initiatives were kicked off with them, including the Q4Health project. We have also collaborated in the realization of different showcases to final users involving their products and our prototypes.
- Keysight Technologies[7], providers of the UXM and T2010 test sets. We collaborated with them on extensions of these platforms (e.g.: the S1 interface) and the Triangle project.

---

[4] The type column has to be interpreted as follows: E=European, N=National.

[5] `http://www.polarisnetworks.net/`

[6] `https://www.redzinc.net/`

[7] `https://www.keysight.com/`

- Dekra[8] research department of the certification division. We have also participated in several initiatives with them including a pilot to evaluate LTE on railway environments.

7.2 Discussion on the results

In this section, we discuss the results obtained in this thesis. First, we discuss the different architectures described in this thesis and the new elements that have been proposed. Then the different contributions to external testbed are analysed to complete with use cases enabled by the obtained results.

7.2.1 Mission Critical Mobile Networks

The first architecture that we proposed to support MCC was based on 3GPP early enablers, before the developments of specific mission critical standards of the new 5G were done. Some of the missing aspects that we identified there were later covered by these new 3GPP standards, but in general, the architecture is a valid approach to support many different MCC scenarios with solutions developed at the application level. We also have provided an updated version of the architecture, where we have included the latest MCX standards that provide reference architectures for PTT, video and data services.

As mentioned, the existing standards are more appropriate to support MCC not only because of the dedicated specifications for MCC but also because of the performance of the ongoing 5G standards. For instance, low latency communications, which are also a focus on this thesis, have been improved with the arrival of the new radio stacks and the separation of the core network elements in smaller functions . However, our designs are still valid in 5G, GTP is still the preferred protocol in the N3 interface and MEC has been provided by collocating the UPF (or SGW/PGW in LTE) with the aggregation point of the base stations so that the performance will be slightly lower to our proposal due to the unnecessary encapsulation/decapsulation of fog traffic.

In terms of performance, both the Fog and GTP Gateways provide a significant reduction of the end-to-end latency, they both remove the backhaul, transport and core network from the path. The Fog Gateway has a trade-off for the packet going to the cloud, as all the traffic will have to analysed by the gateway, and will introduce additional ICMP packets for the pinger function to guarantee an updated uplink database. The GTP Gateway, on the other hand, does not has these limitations but requires the modification of the base station to generated only IP traffic. Both solutions exploit the geographical proximity of the user, for users with high mobility employing stateful fog services a context transfer between services is required. The IP address allocation of the UEs is done by the Packet Data Network (PDN) network so not conflict should be there; the address space has to allow traffic to the subnetworks of the fog services.

A comparison between both gateways and the standard LTE network is provided in

---

[8] `https://www.dekra.es/`

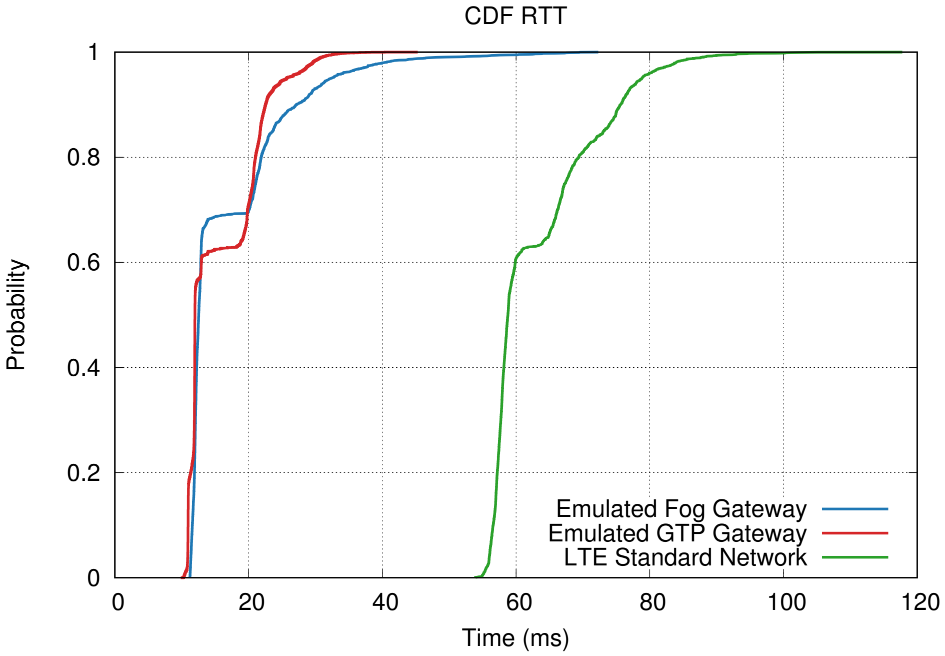


Figure 7.1: RTT CDF Comparison standard LTE network, Fog and GTP Gateways

Figure 7.1; both solutions provide a significant reduction on the latency. We also expect the GTP Gateway to introduce more gains under more substantial load conditions. We also described the use of an API to support QoS requests based on the use of dedicated bearers; the MCData service provides this QoS (besides other service functionalities), we also foresee the definition of several flexible QCI that could be configured by the end-to-end service to adapt them to its requirements. This approach could improve the usage of the underlying resources and the performance of the end-to-end service, but again, a proper API to be exposed to third parties will be required. Operators could also define their own QCIs for different types of service or transport protocols.

7.2.2 Experimental Testbeds

We have provided different tools and designs to have a platform that could combine commercial, prototypes and instrumentation solutions enabling different types of experiments in the area of mobile networks. The main limitation of the platform is that it offers a limited set of results when planning to evaluate under massive load conditions, these scenarios can only be reproduced by limiting the available number of resource blocks in our instruments.

Our framework to support SCPI compliant instruments provides an easy way of introducing new instrumentation in the platform just by defining an XML, which could later be used to generate more tools, such as Graphical User Interfaces (GUIs) or other control frameworks. Indeed for most of our developed tools, we provided an SCPI interface so they could be easily integrated into the platform. Some examples could be the IP Impairments tool, which allows introducing realistic errors in any of the IP links of the testbed or the EPC deployment tool, which enables the on-demand deployment of a full LTE core network.

The S1 Extensions to the conformance testing equipment enable combining these instrument with standard core networks. The instrument allows the usage of COTS UEs and EPC with controlled radio channel conditions and fully configurable radio station stacks. These extensions have allowed us to do a realistic evaluation of many of our proposals and also develop them more easily.

Finally, we have also provided a methodology to evaluate different KPIs that could be later used to select the best UE for a given application, which could be rollout in the testbed during these experiments.

7.3 Future work

In this section, we describe some possible follow-ups of the activities that have been carried out. For the MCC architecture, we will work on the evaluation of services using the new 5G standards. There are also opportunities in the evaluation of the MCData and MCVideo standards, where the creation of new prototypes will be required. The interworking with a common LMR such as TETRA could also be interesting; feasibility should be done first though as many of these solutions are proprietary. The 5GC service architecture is already available, but the standardization of the actual network functions was not (at the time of writing this), so analysing how they fit on the MCC architecture and how the MCX services could be implemented to follow the service architecture will also be of interest. Several issues identified in [3GP17a] have no proposed solution as they are considered out of the scope of the study, which might be interesting as future work, for instance maintaining functionality during migration could be of interest (groups, security, media routing, lawful interception, etc.).

On the field of experimental platforms, besides the natural evolution of the PerformNetworks platform integrating new technologies, the interfaces exposed to experimenters could be improved as well as the tooling to attract more users to the platform and to improve the capabilities to emulate large amounts of traffic. From our perspective, the main barrier of an experimental platform is the translation of all the configuration parameters into a language that an experimenter could employ. As the platform itself has been implemented and supported by researchers, the focus could look too scientific for non-academic customers and indeed that has been the case with many of the experiments that we have supported. On the other hand the exposed parameters are not always identified easily by researchers that have not work with commercial equipment. We could offer tools to configure the emulation as similar as possible to an operator network or even tools to attract researches more used to simulation.

To reproduce the operator network in the experiment platform, we could develop a UE able to extract the information from the SIB1 channels, for this purpose one of the open-source implementations such as srsUE [9] or OpenAirInterface[10] could be extended. If the tool is used for a drive-test we could also store the GPS coordinates (that could be later replayed with tools such as gpssim[11]) and the channel measurements. All this information could be later used to dynamically configure two of our emulators and have a very close-to-real scenario. Operators could even establish a secure link toward its test core network to debug and analyse conflictive scenarios in their pre-production environment. Providing functionality for legacy network and services will also attract more operators.

For researchers more used to simulation tools something similar could be done, we could provide an interpreter of ns3 able to emulate a scenario or part of a scenario to generate more realistic results. In this area, the evaluation of schedulers at MAC level is widespread, and indeed it is challenging to evaluate realistically using simulations as the implementation itself add a relevant contribution to the latency. Providing optimized frameworks to support MAC scheduling implementation and experimentation could be of interest, the framework should provide tools to ease the resource allocation and to generate artificial resource requests distributions. The generation of artificial traffic in the platform will also be of interest in these scenarios.

In general, the provision of replay tools, both at IP and radio levels, could be useful to assist on the development of new products. Experimenters could capture the outcome of their interoperability tests in the testbed and later reproduce them using an SDR platform (for radio scenarios) or a GPP to debug their products on their site and their own pace. In the core network side, the increase of the scale is also significant, the testbed currently holds a limited number of users, traffic generator tools could be useful, such as the ones of the Seagull project[12].

On the low latency solutions, different aspects could be improved, such as evolving the prototypes and creating new ones, exploring content distribution in the fog or providing secure-deployment tools that could be used on production environments. The implemented Fog Gateway has to be tested more thoroughly especially at terms of scale. We carried out some tests evaluating the effect of the regular traffic, but this has to be validated at scale to know how many users could be supported without affecting the overall performance. Indeed the implementation itself could be improved by porting it to a platform with dedicated programmable hardware to accelerate the handling of the packets or providing a kernel-based implementation (removing one the copies of the packets that have to be done from the kernel to the iptables queue).

To complete the fog platform, there are more aspects to consider, the distribution of the content is one of them; the use of CDN could be explored. Also, the identification of the users enabled in the platform has to be supported somehow, as their information is required to store CDRs. Another essential characteristic is the tooling for rolling out

---

[9] `https://github.com/srsLTE/srsUE`

[10] `https://github.com/OPENAIRINTERFACE`

[11] `https://github.com/osqzss/gps-sdr-sim`

[12] `http://gull.sourceforge.net/`

services on the fog, one could think of solutions based on Docker[13] or Kubernetes[14] but running on an isolated environment with the minimum traffic rules to enable service.

We can also improve the evaluation of the GTP Gateway by implementing a prototype compatible with one of the open-source base station implementations. The configuration of the stack to match the needs of the transport level and remove any redundant functionality already provided by the mobile stack could further reduce the latency. The first step will be to analyse different transport protocols modifying the stack to try to optimize the latency and the bandwidth available, and the next could be the design of the method to apply these modifications (e.g.: done automatically by the base station or triggered by an API). Definitely, the exploration of new link protocols could also contribute to latency reduction, especially to remove the redundancy introduced by the IP layer, that could be allocated to increase reliability.

---

[13] `https://www.docker.com/`

[14] `https://kubernetes.io/`

Bibliography

[3GP11] 3GPP. Enhancements for Multimedia Priority Service (MPS). Technical Report TR 23.854, 3GPP, 2011.

[3GP12a] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); LTE coverage enhancements. Technical Report TR 36.824, 3GPP, 2012.

[3GP12b] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Study on signalling and procedure for interference avoidance for in-device coexistence. Technical Report TR 36.816, 3GPP, 2012.

[3GP13a] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); Carrier aggregation enhancements; User Equipment (UE) and Base Station (BS) radio transmission and reception. Technical Report TR 36.823, 3GPP, 2013.

[3GP13b] 3GPP. Public safety broadband high power User Equipment (UE) for band 14. Technical Report TR 36.837, 3GPP, 2013.

[3GP14a] 3GPP. Study on architecture enhancements to support Group Communication System Enablers for LTE (GCSE_LTE). Technical Report TR 22.468, 3GPP, 2014.

[3GP14b] 3GPP. Study on architecture enhancements to support Proximity-based Services (ProSe). Technical Report (TR) 23.703, 3rd Generation Partnership Project (3GPP), 03 2014. Version 12.0.0.

[3GP14c] 3GPP. Study on architecture enhancements to support Proximity-based Services (ProSe). Technical Report (TR) 26.305, 3rd Generation Partnership Project (3GPP), 03 2014. Version 12.0.0.

[3GP15a] 3GPP. TR 32.842, Telecommunication management; Study on network management of virtualized networks. Technical report, Third Generation Partnership Project, 2015.

[3GP15b] 3GPP. TS 36.521 Evolved Universal Terrestrial Radio Access (E-

UTRA); User Equipment (UE) conformance specification; Radio transmission and reception; Part 1: Conformance testing. Technical report, Third Generation Partnership Project, 2015.

[3GP15c] 3GPP. TS 37.901 User Equipment (UE) application layer data throughput performance). Technical report, 3GPP, 2015.

[3GP17a] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on migration and interconnection for mission critical services. Technical Report TR 23.781, Third Generation Partnership Project, 2017.

[3GP17b] 3GPP. LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Stage 2 functional specification of User Equipment (UE) positioning in E-UTRAN. Technical Report TS 36.521, Third Generation Partnership Project, 2017.

[3GP18a] 3GPP. Feasibility study for Further Advancements for E-UTRA (LTE-Advanced). Technical Report TR 36.912, 3GPP, 2018.

[3GP18b] 3GPP. IMS Multimedia telephony communication service and supplementary services; Stage 3. Technical Report TS 24.173, 3GPP, 2018.

[3GP18c] 3GPP. Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description. Technical Report TS 23.246, 3GPP, 2018.

[3GP18d] 3GPP. Multimedia priority service. Technical Report TS 22.153, 3GPP, 2018.

[3GP18e] 3GPP. Quality of Service (QoS) concept and architecture. Technical Report TS 23.107, 3GPP, 2018.

[3GP19a] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of Cell Broadcast Service (CBS). Technical Report TS 23.401, Third Generation Partnership Project, 2019.

[3GP19b] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); General aspects and principles for interfaces supporting Multimedia Broadcast Multicast Service (MBMS) within E-UTRAN. Technical Report TS 36.440, Third Generation Partnership Project, 2019.

[3GP19c] 3GPP. 3rd Generation Partnership Project; Technical Specification

Group Services and System Aspects; Architecture enhancements for non-3GPP accesses. Technical Report TS 23.402, Third Generation Partnership Project, 2019.

[3GP19d] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Common functional architecture to support mission critical services; Stage 2. Technical Report TS 23.280, Third Generation Partnership Project, 2019.

[3GP19e] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Functional architecture and information flows to support Mission Critical Data (MCData);. Technical Report TS 23.282, Third Generation Partnership Project, 2019.

[3GP19f] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Functional architecture and information flows to support Mission Critical Push To Talk (MCPTT); Stage 2. Technical Report TS 23.379, Third Generation Partnership Project, 2019.

[3GP19g] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Functional architecture and information flows to support Mission Critical Video (MCVideo); Stage 2. Technical Report TS 23.281, Third Generation Partnership Project, 2019.

[3GP19h] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Functional stage 2 description of Location Services (LCS). Technical Report TS 23.271, Third Generation Partnership Project, 2019.

[3GP19i] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Group Communication System Enablers for LTE (GCSE_LTE); Stage 2. Technical Report TS 23.468, Third Generation Partnership Project, 2019.

[3GP19j] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2. Technical Report TS 23.228, Third Generation Partnership Project, 2019.

[3GP19k] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mission Critical Communication Interworking with Land Mobile Radio Systems; Stage 2. Technical Report TS 23.283, Third Generation Partnership Project, 2019.

[3GP19l] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mission Critical Services Common Requirements (MCCoRe); Stage 1. Technical Report TS 22.280, Third Generation Partnership Project, 2019.

[3GP19m] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Policy and charging control architecture. Technical Report TS 23.203, Third Generation Partnership Project, 2019.

[3GP19n] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Proximity-based services (ProSe); Stage 2. Technical Report TS 23.303, Third Generation Partnership Project, 2019.

[3GP19o] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Security of the mission critical service. Technical Report TS 33.180, Third Generation Partnership Project, 2019.

[3GP19p] 3GPP. Group Communication System Enablers (GCSE). Technical Report TS 22.468, 3GPP, 2019.

[3GP19q] 3GPP. IP Multimedia Subsystem (IMS) emergency sessions. Technical Specification (TS) 23.167, 3rd Generation Partnership Project (3GPP), 09 2019. Version 16.0.0.

[3GP19r] 3GPP. LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception. Technical Report TS 36.101, Third Generation Partnership Project, 2019.

[3GP19s] 3GPP. Mission Critical Push To Talk (MCPTT) media plane control; Protocol specification). Technical Report TS 24.380, 3GPP, 2019.

[3GP19t] 3GPP. Mission Critical Video (MCVideo) media plane control; Protocol specification. Technical Report TS 24.581, 3GPP, 2019.

[3GP19u] 3GPP. NR; Physical channels and modulation. Technical Specification (TS) 38.211, 3rd Generation Partnership Project (3GPP), 9 2019. Version 15.7.0.

[3GP19v] 3GPP. NR; Physical layer procedures for control. Technical Specification (TS) 38.213, 3rd Generation Partnership Project (3GPP), 9 2019. Version 15.7.0.

[3GP19w] 3GPP. Procedures for the 5G System (5GS). Technical Specification (TS) 23.502, 3rd Generation Partnership Project (3GPP), 12 2019. Version 16.3.0.

[3GP19x] 3GPP. Rail Telecommunications (RT); Next Generation Communication System; Radio performance simulations and evaluations in rail environment; Part 1: Long Term Evolution (LTE. Technical Report ETSI TR 103 554-1, ETSI, 2019.

[3GP19y] 3GPP. Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC). Technical Specification (TS) 38.824, 3rd Generation Partnership Project (3GPP), 3 2019. Version 16.0.0.

[3GP19z] 3GPP. System architecture for the 5G System (5GS). Technical Specification (TS) 23.501, 3rd Generation Partnership Project (3GPP), 12 2019. Version 16.3.0.

[3GP20a] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol A (LPPa) . Technical Report TS 36455, 3GPP, 2020.

[3GP20b] 3GPP. 3G security; Network Domain Security (NDS); IP network layer security. Technical Report TS 33.210, 3GPP, 2020.

[3GP20c] 3GPP. 3GPP System Architecture Evolution (SAE); Security architecture. Technical Report TS 36.823, 3GPP, 2020.

[3GP20d] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mission critical services support in the isolated operation for public safety (IOPS) mode of operation; Functional architecture and information flows; . Technical Report TS 23.180, Third Generation Partnership Project, 2020.

[3GP20e] 3GPP. 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Study on mission critical services over 5G multicast broadcast system. Technical Report TR 23.774, Third Generation Partnership Project, 2020.

[3GP20f] 3GPP. AT command set for User Equipment (UE). Technical Report TS 27.007, 3GPP, 2020.

[3GP20g] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA); LTE Positioning Protocol (LPP). Technical Report TS 36.355, 3GPP, 2020.

[3GP20h] 3GPP. Mission Critical Data (MCData) media plane control; Protocol specification). Technical Report TS 24.582, 3GPP, 2020.

[3GP20i] 3GPP. Network Domain Security (NDS); Authentication Framework (AF). Technical Report TS 36.823, 3GPP, 2020.

[3GP20j] 3GPP. Service accessibility. Technical Report TS 22.011, 3GPP, 2020.

[3GP20k] 3GPP. Support for Internet Protocol (IP) based IP Multimedia Subsystem (IMS) Emergency calls over General Packet Radio Service (GPRS) and Evolved Packet Service (EPS)). Technical Report TS 23.869, 3GPP, 2020.

[ABC$^+$14] J.G. Andrews, S. Buzzi, Wan Choi, S.V. Hanly, A. Lozano, A.C.K. Soong, and J.C. Zhang. What will 5g be? *Selected Areas in Communications, IEEE Journal on*, 32(6):1065–1082, 06 2014.

[ABF$^+$19] G. Avino, P. Bande, P. A. Frangoudis, C. Vitale, C. Casetti, C. F. Chiasserini, K. Gebru, A. Ksentini, and G. Zennaro. A mec-based extended virtual sensing for automotive services. *IEEE Transactions on Network and Service Management*, 16(4):1450–1463, 2019.

[ACC17] I. Ahmad, W. Chen, and K. Chang. Lte-railway user priority-based cooperative resource allocation schemes for coexisting public safety and railway networks. *IEEE Access*, 5:7985–8000, 2017.

[ÁlvarezDMR12a] A. Álvarez, A. Díaz, P. Merino, and F. J. Rivas. Field measurements of mobile services with android smartphones. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 105–109, 01 2012.

[ÁlvarezDMR12b] A. Álvarez, A. Díaz, P. Merino, and F. J. Rivas. Mobile application profiling with testeldroid. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 36–37, 01 2012.

[ÁlvarezDMR12c] A. Álvarez, A. Díaz, P. Merino, and F. J. Rivas. Mobile application profiling with testeldroid. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 36–37, 01 2012.

[AR16] Manzalini A. and Verdone R. 5g experimental facilities in europe. Technical report, Networld2020 ETP, 02 2016.

[ARK$^+$17] A. Aldalbahi, M. Rahaim, A. Khreishah, M. Ayyash, and T. D. C. Little. Visible light communication module: An open source extension to the ns3 network simulator with real system validation. *IEEE Access*, 5:22144–22158, 2017.

[ASS17] A. Al-Shuwaili and O. Simeone. Energy-efficient resource allocation for mobile edge computing-based augmented reality applications. *IEEE Wireless Communications Letters*, 6(3):398–401, 06 2017.

[AT19] FRMCS Functional Working Group & Architecture and Technology. Future Railway Mobile Communication System User Requirements Specification. Specification FU-7100, International Union of Railways (UIC), 2019.

[Aut16] Various Authors. 5G Experimental Facilities in Europe. Technical report, NetWorld 2020 ETP, 2016.

[BCD$^+$11] K. C. Budka, T. Chu, T. L. Doumi, W. Brouwer, P. Lamoureux, and M. E. Palamara. Public safety mission critical voice services over lte. *Bell Labs Technical Journal*, 16(3):133–149, 12 2011.

[BDLP17] R. Bruschi, F. Davoli, P. Lago, and J. F. Pajo. A scalable sdn slicing scheme for multi-domain fog/cloud services. In *2017 IEEE Conference on Network Softwarization (NetSoft)*, pages 1–6, 04 2017.

[BMZA12] Flavio Bonomi, Rodolfo Milito, Jiang Zhu, and Sateesh Addepalli. Fog computing and its role in the internet of things. In *Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing*, MCC '12, pages 13–16. ACM, 2012.

[BRT$^+$14] V. Buenestado, J. M. Ruiz-Avilés, M. Toril, S. Luna-Ramírez, and A. Mendo. Analysis of throughput performance statistics for benchmarking lte networks. *IEEE Communications Letters*, 18(9):1607–1610, 2014.

[CCB$^+$16] E. Cau, M. Corici, P. Bellavista, L. Foschini, G. Carella, A. Edmonds, and T. M. Bohnert. Efficient exploitation of mobile edge computing for virtualized 5G in EPC architectures. In *2016 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud)*, pages 100–109, 03 2016.

[CFH14] Woon Hau Chin, Zhong Fan, and R. Haines. Emerging technologies and research challenges for 5g wireless networks. *Wireless Communications, IEEE*, 21(2):106–112, 04 2014.

[CKC$^+$17] M. Charitos, D. Kong, J. Cao, D. Berkovskyy, A. A. Goulianos, T. Mizutani, F. Tila, G. Hilton, A. Doufexi, and A. Nix. Lte-a virtual drive testing for vehicular environments. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–5, 06 2017.

[CL16] A. Cornejo and S. Landeros. Simulation model and comparison for satellite links in ku and ka bands for standards-based on dvb-s2 and ns3. *IEEE Latin America Transactions*, 14(4):1689–1697, 04 2016.

[CLML18] R. Chen, W. Long, G. Mao, and C. Li. Development trends of mobile communication systems for railways. *IEEE Communications Surveys Tutorials*, 20(4):3131–3141, 9 2018.

[CLS$^+$20] M. Centenaro, D. Laselva, J. Steiner, K. Pedersen, and P. Mogensen. System-level study of data duplication enhancements for 5g downlink urllc. *IEEE Access*, 8:565–578, 2020.

[CMD$^+$16] A. F. Cattoni, G. C. Madueño, M. Dieudonne, P. Merino, A. D. Zayas, A. Salmeron, F. Carlier, B. S. Germain, D. Morris, R. Figueiredo, J. Caffrey, J. Baños, C. Cardenas, N. Roche, and A. Moore. An end-to-end testing ecosystem for 5g. In *2016 European Conference on Networks and Communications (EuCNC)*, pages 307–312, 06 2016.

[CML$^+$14] J. Cao, M. Ma, H. Li, Y. Zhang, and Z. Luo. A survey on security aspects for lte and lte-a networks. *IEEE Communications Surveys Tutorials*, 16(1):283–302, 01 2014.

[CQC$^+$18] H. Chang, B. Qiu, C. Chiu, J. Chen, F. J. Lin, D. de la Bastida, and B. P. Lin. Performance evaluation of open5gcore over kvm and docker by using open5gmtc. In *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6, 04 2018.

[DDT$^+$13] T. Doumi, M.F. Dolan, S. Tatesh, A. Casati, G. Tsirtsis, K. Anchan, and D. Flore. LTE for public safety networks. *Communications Magazine, IEEE*, 51(2):106–112, 02 2013.

[DMPR07] A. Díaz, P. Merino, L. Panizo, and A. M. Recio. Evaluating video streaming over gprs/umts networks: A practical case. In *2007 IEEE 65th Vehicular Technology Conference - VTC2007-Spring*, pages 624–628, 04 2007.

[DSM$^+$17] Almudena Díaz, Alberto Salmeron, Pedro Merino, Andrea F. Cattoni, German Corrales Madueno, Michael Diedonne, Frederik Carlier, Bart Saint Germain, Donald Morris, Ricardo Figueiredo, Jeanne Caffrey, Janie Banos, Carlos Cardenas, Niall Roche, and Alastair Moore. Triangle: 5G applications and devices benchmarking. In *Building the Future Internet through FIRE*, pages 561–572. River Publishers, 03 2017.

[DZGPMG14a] A. Díaz Zayas, C.A. García Pérez, and P. Merino Gomez. Third-generation partnership project standards for delivery of critical communications for railways. *Vehicular Technology Magazine, IEEE*, 9(2):58–68, 06 2014.

[DZGPMG14b] A. Díaz Zayas, C.A. García Pérez, and P. Merino Gomez. Third-generation partnership project standards for delivery of critical communications for railways. *Vehicular Technology Magazine, IEEE*, 9(2):58–68, 06 2014.

[DZGPMG15] A. Díaz Zayas, C.A. García Pérez, and P. Merino Gómez. PerformLTE: A testbed for LTE testing in the future internet. In *Wired/Wireless Internet Communications: 13th International Conference, WWIC 2015, Malaga, Spain, May 25-27, 2015, Revised Selected Papers*, volume 9071, page 46. Springer, 2015.

[DZGPRPM16] A. Díaz-Zayas, C. A. García-Pérez, Á. M. Recio-Pérez, and P. Merino. 3gpp standards to deliver lte connectivity for iot. In *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 283–288, 04 2016.

[DZGPRPM18] Almudena Díaz Zayas, César A. García Pérez, Álvaro M. Recio Pérez, and Pedro Merino. 3gpp evolution on lte connectivity for iot. In Raffaele Gravina, Carlos E. Palau, Marco Manso, Antonio Liotta, and Giancarlo Fortino, editors, *Integration, Interconnection, and Interoperability of IoT Systems*, pages 1–20. Springer International Publishing, Cham, 2018.

[DZPB+18] Almudena Díaz Zayas, Laura Panizo, Janie Baños, Carlos Cárdenas, and Michael Dieudonne. Qoe evaluation: The triangle testbed approach. *Wireless Communications and Mobile Computing*, 2018:1–12, 12 2018.

[DZRPGPM16] A. Díaz-Zayas, Á. M. Recio-Pérez, C. A. García-Pérez, and P. Merino. Remote control and instrumentation of android devices. In *2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, pages 190–195, 02 2016.

[ETS14] ETSI. GS NFV-MAN 001, Group Specification; Network Functions Virtualisation (NFV); Management and Orchestration. Technical report, European Telecommunications Standards Institute, 12 2014.

[ETS20a] ETSI. Rail Telecommunications (RT); Future Rail Mobile Communication System (FRMCS); Study on system architecture. Technical Report ETSI TS 103.389, ETSI, 2020.

[ETS20b] ETSI. Rail Telecommunications (RT); Next Generation Communication System; Part 2: New Radio (NR) radio performance simulations and evaluations in rail environment. Technical Report ETSI TS 103.554-2, ETSI, 2020.

[FGMM16] R. Fantacci, F. Gei, D. Marabissi, and L. Micciullo. Public safety networks evolution toward broadband: sharing infrastructures and spectrum with commercial systems. *IEEE Communications Magazine*, 54(4):24–30, 04 2016.

[FPSB13] R. Ferrus, R. Pisz, O. Sallent, and G. Baldini. Public safety mobile broadband: A techno-economic perspective. *IEEE Vehicular Technology Magazine*, 8(2):28–36, 06 2013.

[GDM$^+$18] César A. García, Almudena Díaz, Álvaro Martín, Pedro Merino, Álvaro Ríos, Donal Morris, Ricardo Figueiredo, Terry O'Callaghan, and Pilar Rodríguez. EPC Experiment C report. Deliverable D2.2, Q4Health Consortium, 02 2018. Version 1,1.

[GFSH18] B. Göktepe, T. Fehrenbach, T. Schierl, and C. Hellge. Reduced cbg harq feedback for efficient multimedia transmissions in 5g for coexistence with urllc traffic. In *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5, 6 2018.

[GP15] César A. García-Pérez. Video for Mission Critical Applications over LTE. Master's thesis, Universidad de Málaga, Málaga, Spain, 2015.

[GP17] César A. García-Pérez. Low Latency Communications for LTE networks. Master's thesis, Universidad de Málaga, Málaga, Spain, 2017.

[GPDZR$^+$17a] Cesar García-Pérez, Almudena Díaz-Zayas, Álvaro Rios, Pedro Merino, Kostas Katsalis, Chia-Yu Chang, Shahab Shariat, Navid Nikaein, Pilar Rodríguez, and Donal Morris. Improving the efficiency and reliability of wearable based mobile ehealth applications. *Pervasive and Mobile Computing*, 40:674 – 691, 2017.

[GPDZR$^+$17b] César Augusto García-Pérez, Almudena Díaz-Zayas, Álvaro Ríos, Pedro Merino, Kostas Katsalis, and Donal Morris. Supporting new applications and services over {LTE} public safety networks. In Daniel Camara and Navid Nikaein, editors, *Wireless Public Safety Networks 3*, pages 113 – 132. Elsevier, 2017.

[GPM16] C. A. García-Pérez and P. Merino. Enabling low latency services on lte networks. In *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, pages 248–255, 09 2016.

[GPM17] César A. García Pérez and P. Merino. Experimental evaluation of fog computing techniques to reduce latency in lte networks. *Transactions on Emerging Telecommunications Technologies*, 0(0):e3201, 2017.

[GPMMR17] Almudena Díaz; César A. García-Pérez, Álvaro Martín, Pedro Merino, and Álvaro Ríos. Performnetworks: a testbed for exhaustive interoperability and performance analysis for mobile networks. In *Building the Future Internet through FIRE*, pages 189–209. River Publishers, 03 2017.

[GPRM$^+$16] César A. García-Pérez, Álvaro Ríos, Pedro Merino, Kostas Katsalis, Navid Nikaein, Ricardo Figueiredo, Donal Morris, and Terry O'Callaghan. Q4health: Quality of service and prioritisation for emergency services in the lte ran stack. In *Networks and Communications (EuCNC), 2016 European Conference on*, 06 2016.

[GPRM$^+$17] César A. García-Pérez, Álvaro Ríos, Pedro Merino, Kostas Katsalis, Navid Nikaein, Ricardo Figueiredo, Donal Morris, Terry O'Callaghan, and Pilar Rodriguez. Q4Health: Mission Critical Communications Over LTE and Future 5G Technologies. In *Building the Future Internet through FIRE*, pages 613–624. River Publishers, 03 2017.

[GPRPRG$^+$16] César A. García-Pérez, Álvaro M. Recio-Pérez, Álvaro Ríos-Gómez, A. Díaz-Zayas, and Pedro Merino. Extensive and repeatable experimentation in mobile communications with programmable instruments. In *2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, pages 30–36, 02 2016.

[Gro15] GSM-R Functional Group. EIRENE functional requirements specification. Technical report, International Union of Railways, 2015.

[Gro19] FRMCS Functional Working Group. Future Railway Mobile Communication System Use Cases. Specification MG-7900, International Union of Railways (UIC), 2019.

[GSA19] GSA. GSA Networks, Status of the LTE Ecosystem. Technical report, Global mobile Suppliers Association, 09 2019.

[GSM14a] GSMA. IMS Profile for Voice and SMS. Technical Report IR.92, Global Supplier Mobile Association (GSMA), 03 2014. Version 12.0.0.

[GSM14b] GSMA. IMS Profile for Voice and SMS. Technical Report IR.92, Global Supplier Mobile Association (GSMA), 03 2014. Version 12.0.0.

[GSN$^+$17] I. Gomez, P. Sutton, A. Nag, A. Selim, L. Doyle, V. Ramachandran, and A. Kokaram. A software radio lte network testbed for video quality of experience experimentation. In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–3, 04 2017.

[GWL$^+$17] Z. Geng, X. Wei, H. Liu, R. Xu, and K. Zheng. Performance analysis and comparison of gpp-based sdr systems. In *2017 7th IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE)*, pages 124–129, 10 2017.

[GZW18] P. Gawlowicz, A. Zubow, and A. Wolisz. Enabling cross-technology communication between lte unlicensed and wifi. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pages 144–152, 04 2018.

[Hil18] Kelly Hill. Meeting the public safety communications imperative. Report, RCRWireless, 05 2018.

[HLT$^+$18] K. Han, S. Li, S. Tang, H. Huang, S. Zhao, G. Fu, and Z. Zhu. Application-driven end-to-end slicing: When wireless network virtualization orchestrates with nfv-based mobile edge computing. *IEEE Access*, 6:26567–26577, 2018.

[HMP$^+$16] T. Heyn, J. Morgade, S. Petersen, K. Pfaffinger, E. Lang, M. Hertlein, and G. Fischer. Integration of broadcast and broadband in lte/5g (imb5) - experimental results from the embms testbeds. In *2016 European Conference on Networks and Communications (EuCNC)*, pages 319–324, 06 2016.

[HNS$^+$17a] A. Huang, N. Nikaein, T. Stenbock, A. Ksentini, and C. Bonnet. Low latency mec framework for sdn-based lte/lte-a networks. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6, 05 2017.

[HNS$^+$17b] A. Huang, N. Nikaein, T. Stenbock, A. Ksentini, and C. Bonnet. Low latency mec framework for sdn-based lte/lte-a networks. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6, 05 2017.

[HPK$^+$14] Johanna Heinonen, Tapio Partti, Marko Kallio, Kari Lappalainen, Hannu Flinck, and Jarmo Hillo. Dynamic tunnel switching for sdn-based cellular core networks. In *Proceedings of the 4th Workshop on All Things Cellular: Operations, Applications, & Challenges*, AllThingsCellular '14, pages 27–32. ACM, 2014.

[HPS$^+$15] Yun Chao Hu, Milan Patel, Dario Sabella, Nurit Sprecher, and

Valerie Young. ETSI White Paper No. 11. Mobile Edge Computing A key technology towards 5G - White paper. Technical report, ETSI, 09 2015.

[HYA18] L. He, Z. Yan, and M. Atiquzzaman. Lte/lte-a network security data collection and analysis for security measurement: A survey. *IEEE Access*, 6:4220–4242, 2018.

[IDA15] IDATE Consulting. ERA/2014/04/ERTMS/OP Evolution of GSM-R. Technical report, European Railway Agency, 04 2015.

[Int19] GSMA Intelligence. The mobile economy 2019. Technical report, GSMA, 2019.

[IRB15] E. A. Ibrahim, M. R. M. Rizk, and E. F. Badran. Study of lte-r x2 handover based on a3 event algorithm using matlab. In *2015 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1155–1159, 10 2015.

[(IT01] International Telecommunications Union (ITU). Perceptual evaluation of speech quality (PESQ) : An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. Recommendation P.862, International Telecommunications Union (ITU), 2001.

[JKY^{+}17] B. Jin, S. Kim, D. Yun, H. Lee, W. Kim, and Y. Yi. Aggregating lte and wi-fi: Toward intra-cell fairness and high tcp performance. *IEEE Transactions on Wireless Communications*, 16(10):6295–6308, 2017.

[JLHW11] A. Jurgelionis, J. Laulajainen, M. Hirvonen, and A. I. Wang. An empirical study of netem network emulation functionalities. In *2011 Proceedings of 20th International Conference on Computer Communications and Networks (ICCCN)*, pages 1–6, 07 2011.

[KDR^{+}17] M. R. Kousri, V. Deniau, J. Rioult, G. Romero, and E. P. Simon. Comparative study of transient disturbances impact on 2g and 4g telecommunication systems in a railway context. In *2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, pages 1–4, 08 2017.

[KFF^{+}18] Sami Kekki, Walter Featherstone, Yonggang Fang, Pekka Kuure, Alice Li, Anurag Ranjan, Debashish Purkayastha, Feng Jiangping, Danny Frydman, Gianluca Verin, KuoWei Wen, Kwihoon Kim, Rohit Arora, Andy Odgers, Luis M. Contreras, and Salvatore Scarpina. MEC in 5G networks. Technical report, ETSI, 2018.

[KKGM17] A. Kumbhar, F. Koohifar, İ. Güvenç, and B. Mueller. A survey on legacy and emerging technologies for public safety communications. *IEEE Communications Surveys Tutorials*, 19(1):97–124, 2017.

[KNH18] K. Katsalis, N. Nikaein, and A. Huang. Jox: An event-driven orchestrator for 5g network slicing. In *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, pages 1–9, 04 2018.

[Kru17] Lennard G. Kruger. The First Responder Network (FirstNet) and Next-Generation Communications for Public Safety: Issues for Congres . Technical report, US Government, 01 2017.

[KS18] Udit Narayana Kar and Debarshi Kumar Sanyal. An overview of device-to-device communication in cellular networks. *ICT Express*, 4(4):203 – 208, 2018.

[KTG+18] H. Koumaras, D. Tsolkas, G. Gardikis, P. M. Gomez, V. Frascolla, D. Triantafyllopoulou, M. Emmelmann, V. Koumaras, M. L. G. Osma, D. Munaretto, E. Atxutegi, J. S. d. Puga, O. Alay, A. Brunstrom, and A. M. C. Bosneag. 5genesis: The genesis of a flexible 5g facility. In *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1–6, 2018.

[LBP+14] F. Lobillo, Z. Becvar, M. A. Puente, P. Mach, F. Lo Presti, F. Gambetti, M. Goldhamer, J. Vidal, A. K. Widiawan, and E. Calvanesse. An architecture for mobile computation offloading on cloud-enabled lte small cells. In *Wireless Communications and Networking Conference Workshops (WCNCW), 2014 IEEE*, pages 1–6, 04 2014.

[LCC17] Signals Analytics LCC. Mission Critical Push-To-Talk (MCPTT) Implementation for Colorado. Technical report, Signals Analytics LCC, 2017.

[LLL+20] C. Li, Y. Lin, Y. Lai, H. Chien, Y. Huang, P. Huang, and H. Liu. Transparent aaa security design for low-latency mec-integrated cellular networks. *IEEE Transactions on Vehicular Technology*, 69(3):3231–3243, 2020.

[LLM+17] M. A. Lema, A. Laya, T. Mahmoodi, M. Cuevas, J. Sachs, J. Markendahl, and M. Dohler. Business case and technology analysis for 5g low latency applications. *IEEE Access*, 5:5917–5935, 2017.

[LM17] Yannick Lair and Georg Mayer. Mission critical services in 3gpp. Technical report, 3GPP, 07 2017.

[LMRZ16] M. Labib, V. Marojevic, J. H. Reed, and A. I. Zaghloul. How to enhance the immunity of lte systems against rf spoofing. In *2016 International Conference on Computing, Networking and Communications (ICNC)*, pages 1–5, 02 2016.

[LNPW14] Q. Li, H. Niu, A. Papathanassiou, and G. Wu. Edge cloud and underlay networks: Empowering 5g cell-less wireless architecture. In *European Wireless 2014; 20th European Wireless Conference; Proceedings of*, pages 1–6, 05 2014.

[LQNL17] C. Li, Z. Qin, E. Novak, and Q. Li. Securing sdn infrastructure of iot-fog network from mitm attacks. *IEEE Internet of Things Journal*, PP(99):1–1, 2017.

[LSRH15] Riku Luostari, Jari Salo, Jussi Reunanen, and Harri Holma. *VoLTE Optimization*, chapter 14, pages 293–331. John Wiley & Sons, Ltd, 2015.

[LSRM$^+$12] M. Laner, P. Svoboda, P. Romirer-Maierhofer, N. Nikaein, F. Ricciato, and M. Rupp. A comparison between one-way delays in operating hspa and lte networks. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012 10th International Symposium on*, pages 286–292, 05 2012.

[LYS16] J. Liu, Q. Yang, and G. Simon. Optimal and practical algorithms for implementing wireless cdn based on base stations. In *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, pages 1–5, 04 2016.

[Mak15] O. Makinen. Streaming at the edge: Local service concepts utilizing mobile edge computing. In *Next Generation Mobile Applications, Services and Technologies, 2015 9th International Conference on*, pages 1–6, 09 2015.

[MBR$^+$18] M. Miozzo, N. Bartzoudis, M. Requena, O. Font-Bach, P. Harbanau, D. Lopez-Bueno, M. Payara, and J. Mangues. Sdr and nfv extensions in the ns-3 lte module for 5g rapid prototyping. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, 04 2018.

[MGM$^+$18] Germán Corrales Madueno, Cesar Augusto García, Alvaro Martín, Iván González, Pedro Merino, Donal Morris, Ricardo Figueiredo, Janie Baños, Manuel García, Carlos Cárdenas, and

Oscar Castañeda. Report on the networking capabilities provided by the testbed. Deliverable D4.1, TRIANGLE Consortium, 10 2018. Version 1,1.

[MLZ+20] H. Ma, S. Li, E. Zhang, Z. Lv, J. Hu, and X. Wei. Cooperative autonomous driving oriented mec-aided 5g-v2x: Prototype system design, field tests and ai-based optimization tools. *IEEE Access*, 8:54288–54302, 2020.

[MMA19] N. A. Mohammed, A. M. Mansoor, and R. B. Ahmad. Mission-critical machine-type communication: An overview and perspectives towards 5g. *IEEE Access*, 7:127198–127216, 2019.

[MMP16] C. Maton, S. Marcus, and L. Pushparatnam. Study on migration of Railway radio communication system from GSM-R to other solutions. Study ERA_RS1_DLV_023, European Railway Agency (ERA), 2016.

[MRHR17] V. Marojevic, R. M. Rao, S. Ha, and J. H. Reed. Performance analysis of a mission-critical portable lte system in targeted rf interference. In *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, pages 1–6, 09. 2017.

[MYZ+17a] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief. A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys Tutorials*, PP(99):1–1, 2017.

[MYZ+17b] Yuyi Mao, Changsheng You, Jun Zhang, Kaibin Huang, and Khaled Ben Letaief. Mobile edge computing: Survey and research outlook. *CoRR*, abs/1701.01090, 2017.

[MZK+15] N. Makris, C. Zarafetas, S. Kechagias, T. Korakis, I. Seskar, and L. Tassiulas. Enabling open access to lte network components; the nitos testbed paradigm. In *Proceedings of the 2015 1st IEEE Conference on Network Softwarization (NetSoft)*, pages 1–6, 04 2015.

[Net09] Nokia Siemens Networks. The impact of latency on application performance. Technical report, Nokia Siemens Networks, 2009.

[OAL+17] J. Ordonez-Lucena, P. Ameigeiras, D. Lopez, J. J. Ramos-Munoz, J. Lorca, and J. Folgueira. Network slicing for 5g with sdn/nfv: Concepts, architectures, and challenges. *IEEE Communications Magazine*, 55(5):80–87, 5 2017.

[OMA11] OMA. Push to Talk over Cellular Requirements. Technical Report OMA-RD-PoC-v2.1, Open Mobile Alliance (OMA), 03 2011.

[PC19] T. Pierre and G. Christophe. Railway network evolution why it is urgent to wait for 5g? In *2019 IEEE 2nd 5G World Forum (5GWF)*, pages 1–6, 9 2019.

[PH16] J. Pacheco and S. Hariri. Iot security framework for smart cyber infrastructures. In *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, pages 242–247, 09 2016.

[PJR^{+}20] M. N. Patwary, S. Junaid Nawaz, M. A. Rahman, S. K. Sharma, M. M. Rashid, and S. J. Barnes. The potential short- and long-term disruptions and transformative impacts of 5g and beyond wireless networks: Lessons learnt from the development of a 5g testbed environment. *IEEE Access*, 8:11352–11379, 2020.

[PLG^{+}18a] V. Petrov, M. A. Lema, M. Gapeyenko, K. Antonakoglou, D. Moltchanov, F. Sardis, A. Samuylov, S. Andreev, Y. Koucheryavy, and M. Dohler. Achieving end-to-end reliability of mission-critical traffic in softwarized 5g networks. *IEEE Journal on Selected Areas in Communications*, 36(3):485–501, 03 2018.

[PLG^{+}18b] V. Petrov, M. A. Lema, M. Gapeyenko, K. Antonakoglou, D. Moltchanov, F. Sardis, A. Samuylov, S. Andreev, Y. Koucheryavy, and M. Dohler. Achieving end-to-end reliability of mission-critical traffic in softwarized 5g networks. *IEEE Journal on Selected Areas in Communications*, 36(3):485–501, 03 2018.

[PLGB18] N. Patriciello, S. Lagen, L. Giupponi, and B. Bojovic. 5g new radio numerologies and their impact on the end-to-end latency. In *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pages 1–6, 9 2018.

[PTMK09] L. Pushparatnam, T. Taylor, D. Mandoc, and K. Konrad. GSM-R Implementation and Procurement Guide . Technical report, UIC, 2009.

[RDM16] F. J. Rivas-Tocado, A. Díaz-Zayas, and P. Merino. Improving mobile communications research and education with uxm wireless test set. In *2016 13th International Conference on Remote Engineering and Virtual Instrumentation (REV)*, pages 186–189, 02 2016.

[ROJS10] Thierry Rakotoarivelo, Maximilian Ott, Guillaume Jourjon, and Ivan Seskar. OMF: A control and management framework for networking testbeds. *SIGOPS Oper. Syst. Rev.*, 43(4):54–59, January 2010.

[RV18] J. Rao and S. Vrzic. Packet duplication for urllc in 5g: Architectural enhancements and performance analysis. *IEEE Network*, 32(2):32–40, 03 2018.

[SAHI18] I. Shayea, T. Abd. Rahman, M. Hadri Azmi, and M. R. Islam. Real measurement study for rain rate and rain attenuation conducted over 26 ghz microwave 5g link system in malaysia. *IEEE Access*, 6:19044–19064, 2018.

[SDM14] Stephan Sand, Armin Dammann, and Christian Mensing. *Scenarios and Models*, chapter 6, pages 119–133. John Wiley & Sons, Ltd, 2014.

[SFB14] Robert Horvitz Simon Forge and Colin Blackman. Is Commercial Cellular Suitable for Mission Critical Broadband? . Technical report, SCF Associates Ltd., 01 2014.

[SGK+17] Junaid Shuja, Abdullah Gani, Kwangman Ko, Kyoungyoung So, Saad Mustafa, Sajjad A. Madani, and Muhammad Khurram Khan. Simdom: A framework for simd instruction translation and offloading in heterogeneous mobile architectures. *Transactions on Emerging Telecommunications Technologies*, pages e3174–n/a, 2017. e3174 ett.3174.

[SJK+18] R. Selvaraju, M. H. Jamaluddin, M. R. Kamarudin, J. Nasir, and M. H. Dahri. Mutual coupling reduction and pattern error correction in a 5g beamforming linear array using csrr. *IEEE Access*, 6:65922–65934, 2018.

[SKCK16] Y. Song, J. Kim, S. W. Choi, and Y. Kim. Long term evolution for wireless railway communications:testbed deployment and performance evaluation. *IEEE Communications Magazine*, 54(2):138–145, 02 2016.

[SSA+18] R. Solozabal, A. Sanchoyerto, E. Atxutegi, B. Blanco, J. O. Fajardo, and F. Liberal. Exploitation of mobile edge computing in 5g distributed mission-critical push-to-talk service deployment. *IEEE Access*, 6:37665–37675, 01 2018.

[SSBL19] A. Sanchoyerto, R. Solozabal, B. Blanco, and F. Liberal. Analysis of the impact of the evolution toward 5g architectures on mission critical push-to-talk services. *IEEE Access*, 7:115052–115061, 2019.

[TNW13] C.-L. Taylor, D. Nolan, and S. Wainberg. Priority capabilities in lte supporting national security and emergency preparedness next generation network priority services. In *Technologies for Homeland*

Security (HST), 2013 IEEE International Conference on, pages 584–588, 11 2013.

[TZ13] Mike Thorpe and Ewald Zelmer. LTE Location Based Services Technology Introduction. White Paper 1SP05, Rohde & Schwarz, 2013.

[VVDM17] T. Van Haute, B. Verbeke, E. De Poorter, and I. Moerman. Optimizing time-of-arrival localization solutions for challenging industrial environments. *IEEE Transactions on Industrial Informatics*, 13(3):1430–1439, 06 2017.

[WN10] Karl Wale and Sridharan Natarajan. Implementing ATCA Serving Gateways for LTE Networks. Technical report, Radisys, 2010.

[WZ17] W. Wang and W. Zhou. Computational offloading with delay and capacity constraints in mobile edge. In *2017 IEEE International Conference on Communications (ICC)*, pages 1–6, 05 2017.

[YCPK16] B. Yang, W. K. Chai, G. Pavlou, and K. V. Katsaros. Seamless support of low latency mobile applications with nfv-enabled mobile edge-cloud. In *2016 5th IEEE International Conference on Cloud Networking (Cloudnet)*, pages 136–141, 10 2016.

[Zha16] M. Zhanikeev. Fog cloud caching at network edge via local hardware awareness spaces. In *2016 IEEE 36th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pages 184–188, 06 2016.

[ZQK+16] G. Zhang, T. Q. S. Quek, M. Kountouris, A. Huang, and H. Shan. Fundamentals of heterogeneous backhaul design, analysis and optimization. *IEEE Transactions on Communications*, 64(2):876–889, 02 2016.

[ZSTM17] A. D. Zayas, A. Salmerón, F. J. R. Tocado, and P. Merino. Advanced testing of mobile applications and devices. In *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, pages 801–804, 06 2017.