Universidade de Aveiro Departamento de Eletrónica, Telecomunicações e Informática
2013

Bruno de Oliveira Cruz

Impacto das Comunicações M2M em Redes Celulares de Telecomunicações

Impact of M2M Communications on Cellular Telecommunications Networks

Universidade de Aveiro Departamento de Eletrónica, Telecomunicações e Informática
2013

Bruno de Oliveira Cruz

Impacto das Comunicações M2M em Redes Celulares de Telecomunicações

Impact of M2M Communications on Cellular Telecommunications Networks

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Eletrónica e de Telecomunicações, realizada sob a orientação científica do Doutor Manuel de Oliveira Duarte, Professor Catedrático do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro e coorientação do Mestre Ricardo Jorge Moreira Ferreira, Gestor de Negócio da PT Inovação.

O Júri

Presidente	Professor Doutor José Rodrigues Ferreira da Rocha Professor Catedrático do Departamento de Engenharia Eletrónica, Telecomunicações e Informática da Universidade de Aveiro
Vogal - Arguente Principal	Professora Doutora Ana Cristina Costa Aguiar Professora Auxiliar do Departamento de Engenharia Eletrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto
Vogal - Orientador	Professor Doutor Aníbal Manuel de Oliveira Duarte Professor Catedrático do Departamento de Engenharia Eletrónica, Telecomunicações e Informática da Universidade de Aveiro

Agradecimentos

Em primeiro lugar quero agradecer ao Professor Manuel de Oliveira Duarte por todo o apoio, motivação, e orientação, e por me ter dado todas as condições para desenvolver um trabalho do qual gosto bastante, numa das minhas áreas de maior interesse científico. De igual forma agradeço também ao Mestre Ricardo Ferreira por toda a sua colaboração ao longo deste trabalho, e pelas suas cuidadas observações que foram fundamentais para a consistência do mesmo. A ambos tenho a agradecer todas as aprendizagens que me transmitiram.

Gostaria de agradecer também o apoio prestado por todos os colegas do grupo de investigação que de alguma forma estiveram envolvidos neste trabalho. Igualmente a todos os colegas e amigos que fizeram parte desta caminhada, com uma palavra especial para as amizades feitas em Trento, e para os colegas da Associação Académica, o meu obrigado por terem tornado a minha passagem pela Universidade um momento de grande crescimento pessoal, sem vocês tudo seria mais difícil.

Por último, agradeço à minha família por sempre me ter dado todas as condições e apoio para desenvolver os meus estudos.

Palavras-chave

Análise de Tráfego, Internet das Coisas, M2M, Orquestração de Serviços, Planeamento de Redes, Redes de Acesso, Redes Core, Redes Móveis, UMTS.

Resumo

As comunicações Máquina-Máquina (M2M) apresentam um crescimento muito significativo e algumas projeções apontam para que esta tendência se acentue drasticamente ao longo dos próximos anos.

O tráfego gerado por este tipo de comunicações tem caraterísticas muito diferentes do tráfego de dados, ou voz, que atualmente circula nas redes celulares de telecomunicações. Assim, é fundamental estudar as caraterísticas dos tipos de tráfego associados com comunicações M2M, por forma a compreender os efeitos que tais caraterísticas podem provocar nas redes celulares de telecomunicações.

Esta dissertação procura identificar e estudar algumas das caraterísticas do tráfego M2M, com especial enfoque na sinalização gerada por serviços M2M. Como resultado principal deste trabalho surge o desenvolvimento de modelos que permitem a construção de uma ferramenta analítica de orquestração de serviços e análise de rede. Esta ferramenta permite orquestrar serviços e modelar padrões de tráfego numa rede UMTS, possibilitando uma análise simultânea aos efeitos produzidos no segmento *core* da mesma rede.

Ao longo deste trabalho procura-se que a abordagem aos problemas apresentados permita que os resultados obtidos sejam válidos, ou adaptáveis, num âmbito mais abrangente do que apenas as comunicações M2M.

Keywords

Access Networks, Business Analytics, Core Networks, Internet of Things, M2M, Mobile Networks, Network Planning, Service Orchestration, Traffic Analysis, UMTS.

Abstract

Machine to Machine (M2M) communications present significant growth and some projections indicate that this trend is going to increase dramatically over the coming years.

The traffic generated by this type of communication has very different characteristics when compared to data or voice traffic currently going through cellular telecommunications networks. Thus, it is essential to study the characteristics of traffic associated with M2M communications in order to understand the effects that its features can imply to cellular telecommunications networks.

This dissertation tries to identify and study some of the characteristics of M2M traffic, with particular focus on signaling generated by M2M services. A number of models, that enable the development of an analytic tool for service orchestration and network analysis, are presented. This tool enables service orchestration and traffic modeling on a UMTS network, with simultaneous visualization of the impacts on the core of such network.

The work presented in this document seeks to approach the problems at study in ways ensuring that its outcomes are valid for a wider scope than just M2M communications.

INDEX

List of Abbreviations

1G	First Generation of Mobile Telecommunications Technology
2G	Second Generation of Mobile Telecommunications Technology
3G	Third Generation of Mobile Telecommunications Technology
3GPP	3rd Generation Partnership Project
3GPP2	3rd Generation Partnership Project 2
4G	Fourth Generation of Mobile Telecommunications Technology
6LoWPAN	IPv6 over Low power Wireless Personal Area Networks
AAL5	ATM Adaptation Layer type 5
AAA	Authentication, Authorization, and Accounting
AC	Alternating Current
ADSL	Asymmetric Digital Subscriber Line
ADSL2+	Asymmetric Digital Subscriber Line Two Plus
ADSL2++	Asymmetric Digital Subscriber Line Two Plus Plus
AMI	Advanced Metering Infrastructure
AMPS	Advanced Mobile Phone System
ANSI	American National Standards Institute
AP	Access Point
APN	Access Point Name
App	Application
ARPU	Average Revenue per User
AS	Access Stratum
ATM	Asynchronous Transfer Mode
AuC	Authentication Centre
BH	Busy Hour
BLE	Bluetooth Low Energy
BSC	Base Station Controller
BSS	Business Support System or Base Station System
BT	Bluetooth
BTS	Base Transceiver Station
BYOD	Bring Your Own Device
CAN	Controller Area Network
CAPEX	Capital expenditure
CATV	Community Access Television
CDMA	Code Division Multiple Access
CDR	Call Detail Record
CEN	European Committee for Standardization
CENELEC	European Committee for Electrotechnical Standardization
CG	Charging Gateway
CMTS	Cable Modem Termination System
CN	Core Network
COAX	Coaxial cable
CP	Customer Premises
CPU	Central Processing Unit
CS	Circuit Switched domain
DC	Dual Carrier
DHCP	Dynamic Host Configuration Protocol
DL	Downlink
DLMS	Device Language Message specification
DMA	Direct Memory Access
DNS	Domain Name System
DOCSIS	Data over Cable Service Interface Specification
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DSP	Digital Signal Processor
DTV	Digital Television

DVB	Digital Video Broadcasting
DVB-H	Digital Video Broadcasting - Handheld
DVB-RCS	Digital Video Broadcasting - Return Channel via Satellite
E-UTRAN	Evolved Universal Terrestrial Radio Access
EC	Echo Canceler
ECHONET	Energy Conservation and Homecare Network
EDGE	Enhanced Data rates for GSM Evolution
EDR	Enhanced Data Rate
EE	Energy Eficiency
EIR	Equipment Identity Register
eNB	E-UTRAN NodeB
EPC	Evolved Packet Core
EQAM	Edge Quadrature Amplitude Modulation
ERM	Edge Resource Manager
ETSI	European Telecommunications Standards Institute
EU	European standards
FE	Fast Ethernet
FDD	Frequency-Division Duplexing
FITH	Fiber-in-the-Home
FSO	Free Space Optics
FSK	Frequency-Shift Keying
FTTC	Fiber to the Curb/Closet/Cabinet
FTTH	Fiber to the Home
FTTN	Fiber to the Node/Neighborhood
FTTx	Fiber to the x
GE	Gigabit Ethernet
GERAN	GSM EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
GMM/SM	GPRS Mobility Management and Session Management
GMSK	Gaussian Minimum Shift Keying
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
GTP	GPRS Tunneling Protocol
GTP-C	GTP Control
GTP-U	GTP User
GW	Gateway
HAN	Home Area Network
HART	Highway Addressable Remote Transducer protocol
HDMI	High-Definition Multimedia Interface
HetNet	Heterogeneous Network
HFC	Hybrid Fiber Coax
HG	Home Gateway
HGI	Home Gateway Initiative
HiMSS	Healthcare Information and Management Systems Society
HL7	Health Level Seven International
HLR	Home Location Register
HPNA	Home Phone line Networking
HR	Human Resources
HS	High Speed
HSDPA	High-Speed Downlink Packet Access
HSPA+	Evolved High-Speed Packet Access
HSS	Home Subscriber Server
HSUPA	High-Speed Uplink Packet Access
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force

IHD	In-House Display
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
IoE	Internet of Everything
IoT	Internet of Things
IP	Internet Protocol
IPSEC	Internet Protocol Security
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISA	International Society of Automation
ISDN	Integrated Services Digital Network
ISM	Industrial, Scientific and Medical radio bands
ISO	International Organization for Standardization
ISP	Internet Service Provider
ITU	International Telecommunication Union
IWF	Interworking Function
KPI	Key Performance Indicator
L1	Layer 1
L2	Layer 2
LAI	Location Area Identity
LAN	Local Area Network
LCS	Location Services
LTE	Long Term Evolution
LTE-A	Long Term Evolution – Advanced
M-BUS	Meter-Bus
M2M	Machine-to-Machine
MAC	Media Access Control
MAN	Metropolitan Area Network
MAP	Mobile Application Part
MD	Mobile Device
MGW	Media Gateway
MIM	M2M Identification Module
MIMO	Multiple-Input and Multiple-Output
MM	Mobility Management
MME	Mobility Management Entity
MMS	Multimedia Messaging Service
MNO	Mobile Network Operator
MO	Mobile Originated
MPLS	Multiprotocol Label Switching
MQTT	Messaging Queue Telemetry Transport
MS	Mobile Station
MSC	Mobile services Switching Center
MSRN	Mobile Station Roaming Number
MT	Mobile Terminate
MTPx	Message Transfer Part – Level x
NAS	Non Access Stratum
NAT	Network Address Translation
NE	Network Element
NFC	Near field Communication
NGTP	Next Generation Telematics Pattern
NIST	National Institute of Standards and Technology
NSS	Network Switching Subsystem
OA&M	Operation, Administrative, and Maintenance
OCS	Online Charging System
OLT	Optical Line Termination
OMS	Operations and Maintenance System

ONT	Optical Network Terminal
ONU	Optical Network Unit
OSI	Open Systems Interconnection model
OSS	Operations Support System
PAN	Personal Area Network
PC	Personal Computer
PCEF	Policy Control and Charging Enforcement Function
PCRF	Policy Control and Charging Rules Function
PDCP	Packet Data Convergence Protocol
PDN	Public Data Network
PDP	Packet Data Protocol
PDU	Protocol Data Unit
PGW	Public Data Network Gateway
PLC	Power Line Communication
PLMN	Public Land Mobile Network
PMM	Packet Mobility Management
POF	Plastic Optical Fiber
POTS	Plain Old Telephone Service
PPP	Point-to-Point Protocol
PPS	Packets per Second
PS	Packet Switched
PSK	Phase-Shift Keying
PSTN	Public Switched Telephone Network
QAM	Quadrature Amplitude Modulation
QoE	Quality of Experience
QoS	Quality of Service
RA	Routing Area
RADIUS	Remote Authentication Dial in User Service
RAM	Random Access Memory
RAN	Radio Access Network
RANAP	Radio Access Network Application Part
RAU	Routing Area Update
RF	Radio Frequency
RG	Road Gateway
RLC	Radio Link Control
RNC	Radio Network Controller
RNS	Radio Network Subsystem
ROI	Return on investment
RRC	Radio Resource Control
S-GW	Serving Gateway
S-UMTS	Scalable UMTS
SAE	System Architecture Evolution
SCCP	Signaling Connection Control Part
SCS	Simplified Wiring System
SCTP	Stream Control Transmission Protocol
SDO	Standard-Developing Organization
SEP	Smart-Energy Profile
SFH	Slow Frequency Hopping
SGSN	Serving GPRS Support Node
SGW	Serving Gateway
SIG	Special Interests Group
SIM	Subscriber Identity Module
SLA	Service Level Agreement
SLIC	Serial Line Interface Card
SME	Short Message Entity
SMS	Short Message Service
SMS-GMSC	Gateway MSC for Short Message Service

SMS-IWMSC	Interworking MSC for Short Message Service
SMS-SC	SMS Service Centre
SON	Self-Organizing Networks
SP	Service Provider
SRD	Short Range Devices
SRNS	Serving RNS
SSCOP	Service-Specific Connection-Oriented Protocol
STB	Set-Top Box
SWF	Small Web Format
SWOT	Strengths, Weaknesses, Opportunities, and Threats
TC	Technical Committee
TCAP	Transaction Capabilities Application Part
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDM	Time-Division Multiplexing
TDMA	Time Division Multiple Access
TIA	Telecommunications Industry Association
TMSI	Temporary Mobile Subscriber Identity
TOU	Time of Use
TV	Television
UDP	User Datagram Protocol
UL	Uplink
UMTS	Universal Mobile Telecommunications System
URA	UTRAN Registrations Area
USA	United States of America
USB	Universal Serial Bus
USD	United States Dollar
USIM	Universal/UMTS Subscriber Identity Module
UTRA	Universal Terrestrial Radio Access
UTRAN	Universal Terrestrial Radio Access Network
VDSL	Very-high-bit-rate Digital Subscriber Line
VDSL2	Very-high-bit-rate Digital Subscriber Line Two
VHDSL	Very-High-bit-rate Digital Subscriber Line
VLR	Visitor Location Register
VMSC	Visited Mobile Switching Centre
VOD	Video on Demand
VoIP	Voice over IP
VPN	Virtual Private Network
W-CDMA	Wideband Code Division Multiple Access
W M-BUS	Wireless Meter-Bus
WAN	Wide Area Network
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network

List of Symbols

Ack	Acknowledgment
B	Interface between an MSC and a VLR
bps	Bits per second
Bps	Bytes per second
C	Interface between an HLR and a MSC
cm	Centimeters
D	Interface between an HLR and a VLR
E	Interface between an MSC and a MSC, or between an MSC and a MGW
F	Interface between an MSC and an EIR
Gb	Interface between an SGSN and a Base Station System (BSS)
Gbps	Gigabit per second
Gd	Interface between an SGSN and an SMS-GMSC/IWMSC
Gf	Interface between an SGSN and an EIR
GHz	Gigahertz
Gi	Reference point between GPRS and an external packet data network
Gn	Interface between two GSNs within the same PLMN
Gp	Interface between two GSNs in different PLMNs. The Gp interface allows support of GPRS network services across areas served by the co-operating GPRS PLMNs
Gr	Interface between an SGSN and an HLR
Gs	Interface between an SGSN and an MSC
Gx	Interface between an S-GW and an UTRAN
Iu	Interface between an PCRF and a PDN-GW
Iub	Interface between an RNC and a NodeB
Iur	Interface between RNCs
KB	Kilobyte
kbps	Kilobit per second
kHz	Kilohertz
km	Kilometer
LTE-Uu	Interface between an eNodeB and an MS
m	Meters
min	Minutes
ms	Milliseconds
Mbps	Megabit per second
MHz	Megahertz
n/a	Not Available
PHY	Physical layer of the OSI model
Pkt	Packet
Rx	Interface between a PCRF and operators IP services
s	Seconds
sec	Seconds
S1-MME	Interface between an eNodeB and an MME
S1-U	Interface between an eNodeB and an S-GW
S10	Control interface between MMEs
S11	Interface between an S-GW and an MME
S12	Interface between an S-GW and an UTRAN
S3	Interface between an MME and an SGSN
S4	Interface between an S-GW and a GERAN
S5	Interface between an S-GW and a PDN-GW
S6a	Interface between an MME and an HSS
SGi	Interface between a PDN-GW and operators IP services
TH	Throughput
tx	Transmissions
Um	Interface between a BTS and an MS
Uu	Interface between a NodeB and an MS

List of Figures

List of Tables

List of Charts

List of Equations

1. Introduction

1.1 Motivation

Telecommunications services are on the verge of major changes with the rising of Machine to Machine (M2M) communications and the Internet of Things (IoT). Such changes will profoundly affect life as it is known. Tasks, objects, health, security, houses, almost everything will be affected by the revolutionary effects of the connected world, the IoT, empowering people to live more comfortably, safer, and ultimately even longer. Cities will become more efficient and welcoming with the rise of the smart city, where everything and everyone will be connected, resolving many problems experienced today such as traffic jams, lack of parking spots [1], or pollution, while public transportation systems and emergency response will have the tools to achieve levels of optimization never seen before [2], [3]. Also new climate challenges will be tackled by M2M communications with the provisioning of faster and more adequate preventive action and emergency response [4], [5].

This reality, if it comes to be, will imply a dramatic increase on the number of M2M devices and consequently on the amount of traffic being transported by telecommunications networks. There are growing evidence that this increase will indeed come to be, Ericsson estimates that by 2020 there will be 50 billion things connected to telecommunications networks [6], Cisco estimates are a little more conservative, but still point to 37 billion connections by the same year [3].

Such expressive numbers may still be several years away, but many M2M services are starting to roll out. An example of services that are starting the M2M era is the smart metering and smart grid, which have the power to change the way utilities have been work since decades, and might even change the role of such companies in the marketplace [7]-[15]. The study of smart metering services was one of the main drivers for this work, with particular focus on the characteristics of the data and control (signaling) traffic, generated by such services. Also the car is becoming an interest beacon in the M2M deployment [2], [16], [17], among other areas where developments are already visible.

The effects that this foreseeable increase on the number of devices and connections supported by telecommunications networks will imply on the performance and response of such networks are widely unknown and unstudied. Thus it becomes the fundamental objective of this work to provide some of the required knowledge on this subject, and methods to achieve them; essentially, ways to model the effects of M2M services, with specifiable characteristics, on already deployed telecommunications networks.

As referred, smart metering was a fundamental driver to this study, but a high level of generality was preserved, since it is expected that many, if not most, of its characteristics (that are relevant to the problem of network performance and response) will be similar to other M2M services. Additionally it is expected that the traffic generated by M2M services will present similar characteristics to the traffic generated by modern smartphone applications, mainly social networks, especially in which refers to the ratio of signaling to payload [18]. Therefore it is believed that this work is valid and relevant to services other than M2M.

1.2 Framework

M2M services and applications are expected to generate traffic presenting some different features when compared to conventional data or voice traffic. Such features may include radically different patterns of use among different services, with some services presenting high predictability, and others high unpredictability; downlink to uplink ratios are usually much bigger than those on conventional traffic, most of the traffic is expected to be diurnal, and many services will generate synchronized data transmissions very few times a day, which will result in bursts of aggregated payload [19], [20] (burstiness is a statistics concept that refers to the intermittent increases and decreases in the activity or frequency of an event [21]). Other M2M services can generate a high number of daily transmissions with very small payloads, which may cause problems when dealing with very large amounts of active sessions on a network [19]. Another fundamental characteristic of M2M traffic is the ratio of payload to overhead, which will be much lower than in conventional traffic [20].

Considering that cellular networks have not been designed to deal with this kind of traffic, instead they were designed to carry human-to-human, and human-to-machine, communications, it becomes urgent to understand how massive adoption of M2M services will affect cellular networks [18], [19]. In particular it is important to understand what can occur in terms of number of sessions, number of subscriptions, amount of signaling generated by very small quantities of payload and burstiness transmissions [18]. Such understanding would provide Mobile Network Operators (MNOs) valuable information to understand if redesign and rescale of their cellular networks is necessary, and which are the best ways to implement such changes. Also, knowing the effects that the traffic can pose to the networks, MNOs and other players have the resources to take informed and optimized design options for M2M services, applications and platforms.

Such M2M traffic characteristics imply, and are a consequence of, new business models for M2M services that are considerably different from the conventional business models adopted by telecommunications operators. The ARPU of M2M services is much lower than what these operators are accustomed to deal with (it can be of 10% when compared to handset subscribers) [22], [23]. This poses a big challenge in terms of business models: M2M services will have lower ARPU, with a higher number of active subscriptions (or connections, since it is expectable that a single subscription will aggregate multiple connections), and will present lower CAPEX expenses since these services will, for the most cases, be supported by already deployed networks [22]. It is not expected that the MNOs will be very receptive to implement significant changes on their access networks driven by M2M services, at least in the first waves of M2M deployments. Such comes to reinforce the previously referred importance of optimizing and adequate the existent networks, and optimizing the design of M2M services in order to allow for provisioning of M2M services with enough scale to profit from these extremely low ARPUs, while assuring that the networks will respond within the required performance parameters. As a positive business aspect of M2M services, lower levels of churn are to expect when compared to conventional subscribers [22].

An important business decision or constraint is on the ownership of the M2M devices and, or, collectors (gateways). This could be the difference from a business model where the CAPEX would be close to zero, or a business model where CAPEX could not be disregarded and possible would imply considerable truck-roll costs[1].

Taking in consideration the current needs for the Portuguese market, this study shall be focused on 3G networks, although large amount of it can be applied to other 3GPP[2], or non 3GPP, networks and scenarios.

[1] Truck-roll: dispatch of a truck or other vehicle to an off-site location.

[2] 3GPP: 3rd Generation Partnership Project, unites telecommunications standard development organizations, providing their members with a stable environment to produce reports and specifications that define 3GPP technologies (1G to 4G) [24].

1.3 Objectives and Methodology

The fundamental goals for this work shall be:

- Study factors that might impact a 3G network;
- Identify critical elements, interfaces, and (or) functions of 3G core networks;
- Design models to orchestrate M2M, and other data services;
- Design models to compute throughput values (control, data, and combined) on 3G core networks;
- Design capacity models for 3G core networks;
- Develop an analytic tool based on the previously referred models.

In order to achieve such goals the following methodology shall be applied:

1. Define a reference architecture;
2. Identify performance and capacity metrics and indicators;
3. Identify critical elements, interfaces, and functions.
4. Design processes to orchestrate M2M and other data services;
5. Design models to compute throughput values (control, data, and both) on 3G core networks;
6. Modeling the performance of a 3G core network;
7. Develop an analytic tool based on the previously referred models;
8. Present a case-study based on the use of this tool;
9. Draw conclusions from the case-study;
10. Present conclusions on this work;
11. Present suggestions for future work.

It is important to stress out that the fundamental scope of this dissertation is to develop analytic models and computational tools describing the performance of 3G networks in the presence of M2M traffic. In order to achieve this, a large number of assumptions concerning parameter values have to be done, this assumptions will obviously define the accuracy of the values and results obtained. Presenting a high level of accuracy is out of the scope of this work and it is proposed as future work.

Succeeding on achieving such goals, this work is believed to present novel knowledge on the subject of the effects caused by M2M communications on cellular networks, and on overall telecommunications networks, as well as is believed to provide relevant information, and more important, models and tools to extract valuable information for telecommunication network operators. Also, the models that shall be designed or adapted over the course of this work are believed to provide an interesting framework for future academic studies and developments on this, or related, subjects.

1.4 Document Structure

This Section provides a brief description of the contents and purposes of each one of the nine Chapters that comprise this dissertation:

- **Chapter 1, “Introduction”:** A brief exposition on the motivation, framework, objectives and methodology of this work, intended to provide the reader with the necessary framework to the reading of this document.

- **Chapter 2, “M2M Communications”:** The state of the art on M2M Communications. Its history, current developments, main areas of application, concerns, and opportunities are presented.

- **Chapter 3, “Telecommunication Networks: Technologies and Architectures”:** A state of the art on technologies and architectures of telecommunication networks. Additionally, specific architectures and solutions related to M2M communications are presented.

- **Chapter 4, “Impact of M2M Communications into 3G Wireless Systems”:** A study of the implications of M2M traffic into cellular telecommunications networks, with focus on 3G networks.

- **Chapter 5, “Traffic and Throughput Analysis for UMTS Packet Core Networks”:** In this Chapter models for the analysis of traffic and throughput for UMTS Packet Core are presented, with a certain degree of focus on M2M traffic, but preserving broader application.

- **Chapter 6, “SOTA – Scenario Orchestration and Traffic Analysis Tool”:** This Chapter presents an analytic tool, developed with the objective of providing orchestration of services while simultaneous reporting on their impact on a 3G Core Network, with particular focus on throughput and subscriber capacity.

- **Chapter 7, “Case Study”:** An example of the applicability of the models presented on the previous Chapters, demonstrating the usability and usefulness of the tool presented on Chapter 6.

- **Chapter 8, “Conclusions”:** Conclusions on the legitimacy of the hypothesis, models and tools developed over the course of this work, and on the results achieved by the case sudy.

- **Chapter 9, “Future Work”:** Future developments related to this subject and improvements on proposed models and tool are suggested.

This work is unusually long for the purpose of a master dissertation, the author and the advisors hope that the reader can understand that such is justifiable by the novelty and complexity of the subject at study, and it was considered that the relegation of parts of this work to appendices would compromise the coherence and comprehension of this study. Furthermore, and considering the novelty of the subject, it should be referred that this work has the ambition of serving as a framework for future dissertations, or other academic works, on this subject.

2. M2M Communications

2.1 Definition

M2M is not a technology but a concept. Accordingly there is no unique definition for it. The two definitions presented are believed to be both clear and concise:

- ““Machines” using network resources to communicate with remote application infrastructure for the purposes of monitoring and control, either of the “machine” itself, or of the surrounding environment.” - ABI Research [25].

- “M2M communications is the networking of intelligent, communications-enabled remote assets. It allows key information to be exchanged automatically without human intervention, and covers a broad range of technologies and applications which connect the physical world – whether machines or monitored physical conditions – to a back-end IT infrastructure.” – Vodafone [26].

In the view of the author, M2M communications is the ecosystem comprised by devices capable of, without human interference, communicate in some kind of network, by the network itself, and by the applications and services that can be enabled by such devices and networks.

2.2 Origins and Evolution

The concept of M2M can be seen as an evolution of Telemetry, which goes back to 1845, when the first data-transmission circuits were developed between the Russian Tsar's Winter Palace and army headquarters [27]. Since then Telemetry has dramatically increased its presence and importance on everyday life.

In the beginning of the nineteen-nineties the advancements on remote communications systems for Formula One and the first Smart Metering systems can be seen as the step-up from Telemetry to today’s concept of M2M Communications.

The advent of mobile data communications in the two thousands marks the explosion on the number of M2M applications, services and devices. The evolution on the number of connected things and the development of mobile communications is presented in Figure 1.

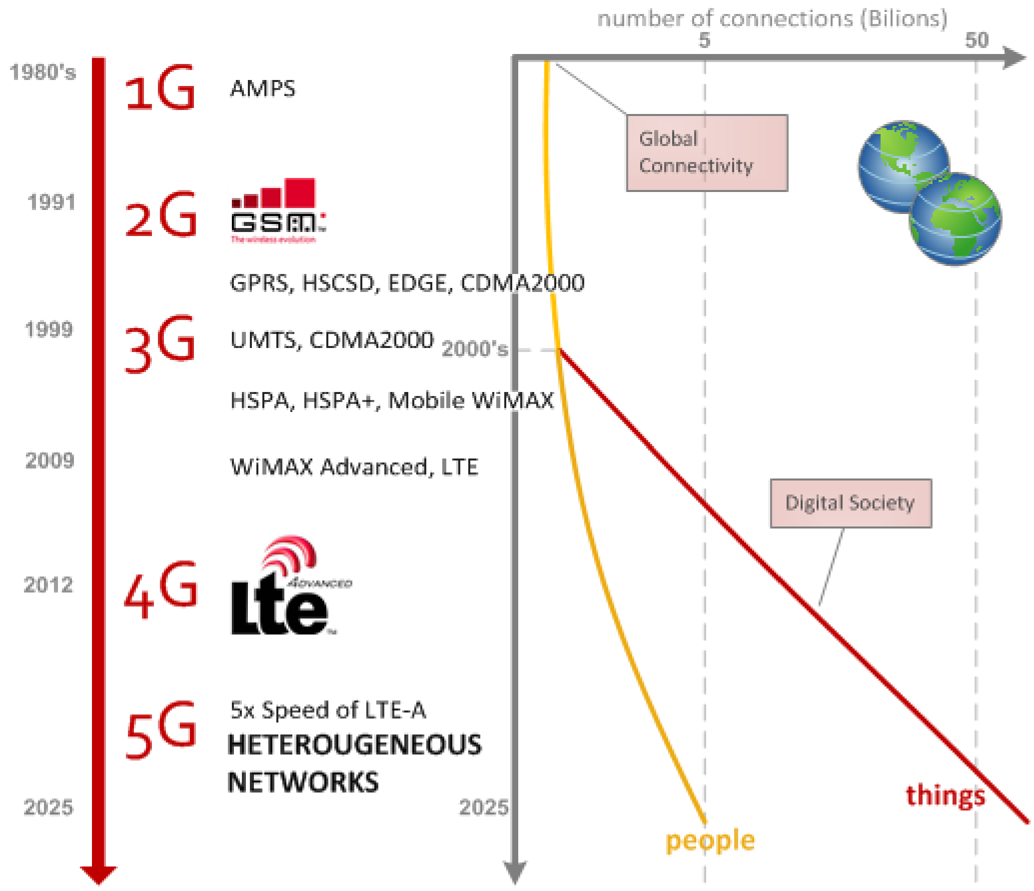


Figure 1 – Evolution of mobile communications and connections (based on [6], [28])

This explosion on the number of M2M applications, services and devices is part of a broader growth in the number of connected subscriptions and dramatic increase on the mobile data traffic, as seen on Chart 1.

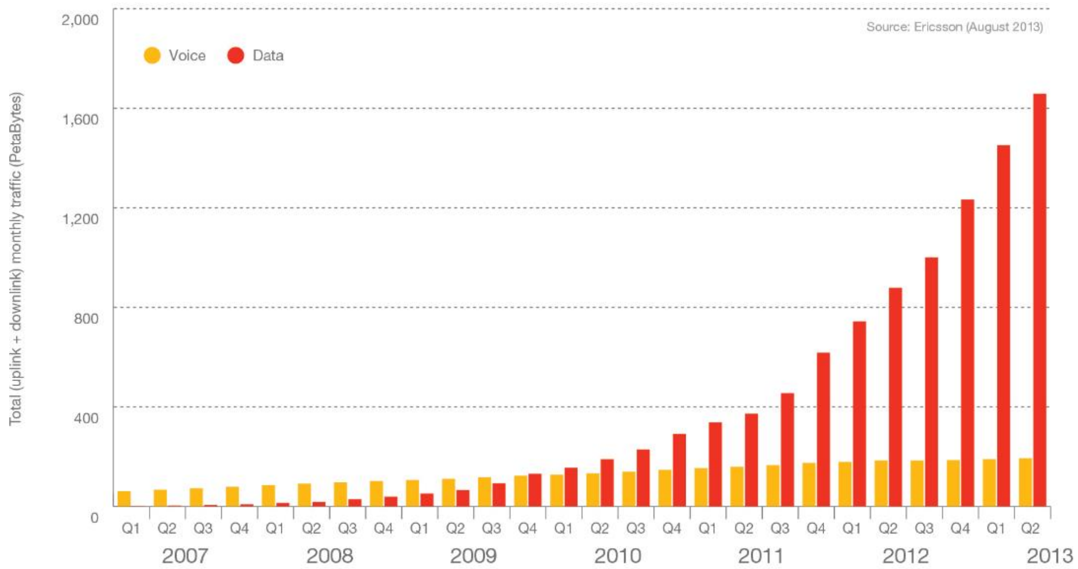


Chart 1 – Global monthly data and voice mobile traffic[3] [29]

[3] Does not include DVB-H, Wi-Fi, or Mobile WiMax. Voice does not include VoIP.

According to Vodafone [30], some of the drivers behind the industry and corporate implementation of M2M solutions are:

- Cost savings by automating manual processes;
- Process and productivity improvements;
- Greater competitive advantage;
- Need for faster decision-making;
- Need for consistent global offering;
- Opportunity to generate additional revenue;
- Demand from end users;
- Greater transparency/predictability of costs;
- Declining M2M costs;
- Demand for greater environmental focus;
- New legislation;
- Regulatory pressures.

Equally, Mobile Network Operators (MNOs) are facing rates of penetration of about 100% for many of their current services, which means that the added economic value based on providing mobility to subscribers has been considerably exploited. Consequently, supply M2M connectivity and related services is a fundamental revenue source in order to enable business growth. It can be said that "the Blue Ocean Strategy of creating new economic value is to shift the paradigm to connecting devices instead of connecting consumers" [31].

But connecting people is not the same as connecting things, typical monthly ARPU per handset subscriber is more than 50 USD [32]. But the typical monthly ARPU per device connected is 5-15 USD, and could be as low as 4 USD a year [32]. Connectivity costs must be kept down in order to ensure operator revenues with such low ARPU values.

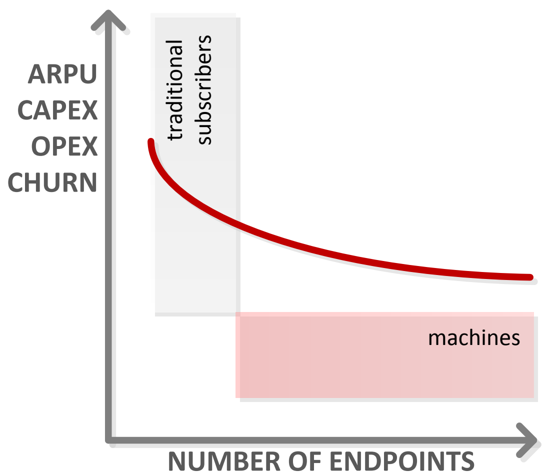


Figure 2 – Paradigm shift from connecting people to connecting devices (based on [33])

Considering the high hopes for M2M Communication (50 to 80 billion connected devices in 2020 [6], [31]) it is interesting to analyze its current state of penetration. In terms of the Hype Cycle [34], it can be said that in the current moment M2M Communication is "On the Rise" (Figure 3) or "At the Peak" depending on the service/technologies.

Therefore, it should be expected that it would slide "Intro the Trough", but at this point is fundamental to remember that "M2M is not a technology but a concept" (Section 2.1). M2M Communication makes use, and is enabled by, a large group of technological solutions that are, in most cases, in the "Slope of Enlightenment" or in the "Plateau of Productivity". Therefore it is possible to look at such hopes and predictions with the security of knowing that M2M solutions will most likely follow a growing path in the next years.

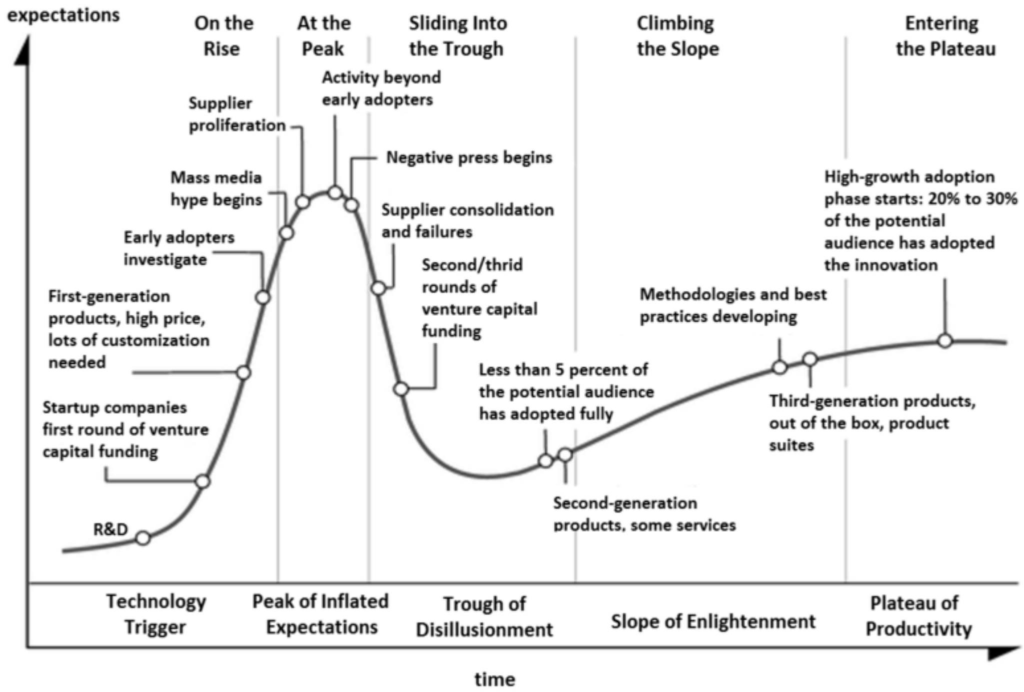


Figure 3 – General Hype Cycle [35]

It is also important to look at how the companies that implemented M2M solutions are evaluating their ROI. According to Vodafone, 94% of inquired businesses see some return from M2M, with 36% seeing significant return [30]. Transports and logistics together with energy and utilities register the best results; automotive by far the least [30].
As mentioned before, M2M Communications are enabled by a wide range of technologies, regarding both devices and connectivity. In the context of this document the main focus is on the connectivity choices.

Machina Research predicts that the evolution of global M2M connections will behave as presented on Chart 2.

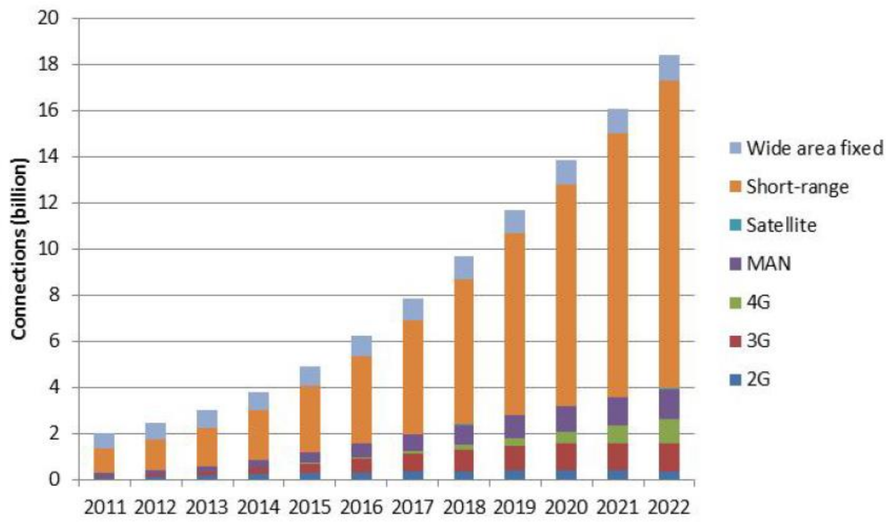


Chart 2 – Global M2M connections 2011-22 by technology [36]

It can be seen that most of the connectivity will predictably be provided by short range technologies, which makes sense considering that most of the M2M devices and applications will be in the context of HANs or LANs. Many of such HANs and LANs need to be connected to the operator (or operators) network, and therefore the impact of short-range M2M connections in the various access technologies should be considered as having the potential to cause major repercussions in the operator's network capacity.

Machina Research predictions point to a scenario, by 2022, where only 40% of the revenues will be associated with provisioning of connectivity, considering that the rest of the revenues will come from services, installation, and devices. From this 40%, as can be seen on Chart 3, also the big majority of revenues will come from service enablement platforms, and only a residual part will come from actual traffic.

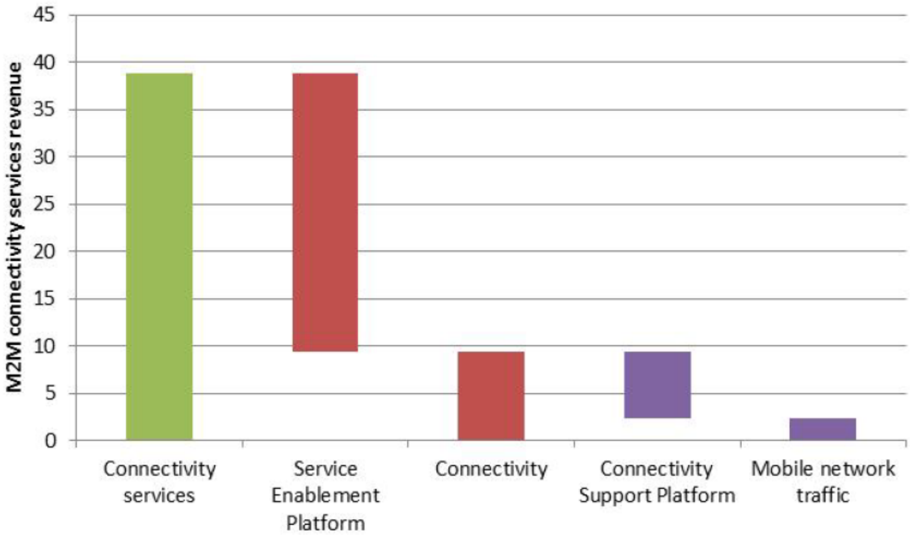


Chart 3 – Global M2M connectivity revenue, 2022 (in billion USD) [36]

2.3 Application Areas

M2M services and devices can be deployed to infinite tasks and application areas. Nevertheless there are some verticals that are leading, or are believed to lead soon, the implementation of M2M solutions.

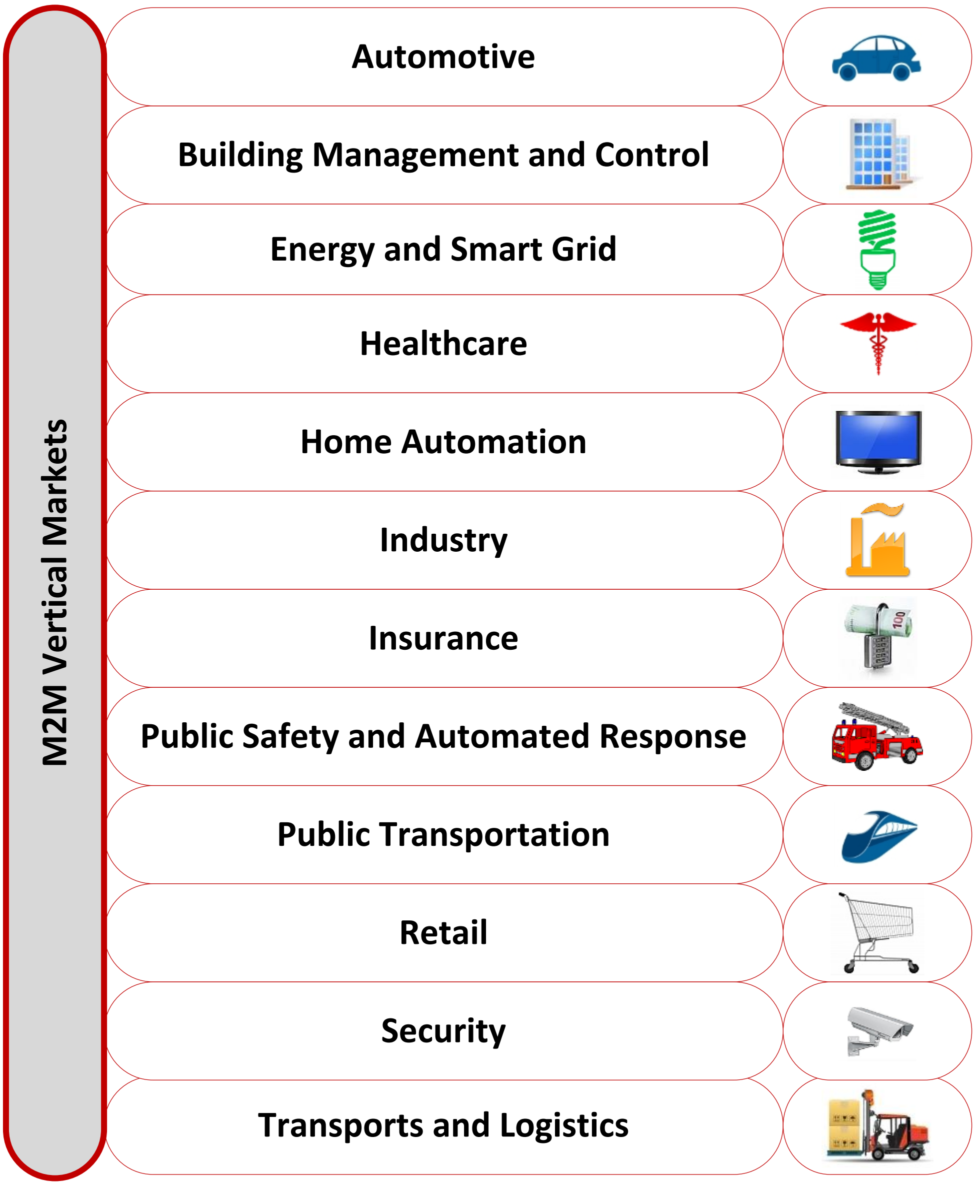


Figure 4 – M2M vertical markets

2.3.1 Automotive

Cars are one of the symbols of our developed world and a lot of people spend a considerable amount of time in their car every day. Therefore, it comes as no surprise that the car is getting connected to an increasingly number of "things", increasing the comfort and security of its occupants. Figure 5 depicts some of the main features and technologies enabling the connected car.

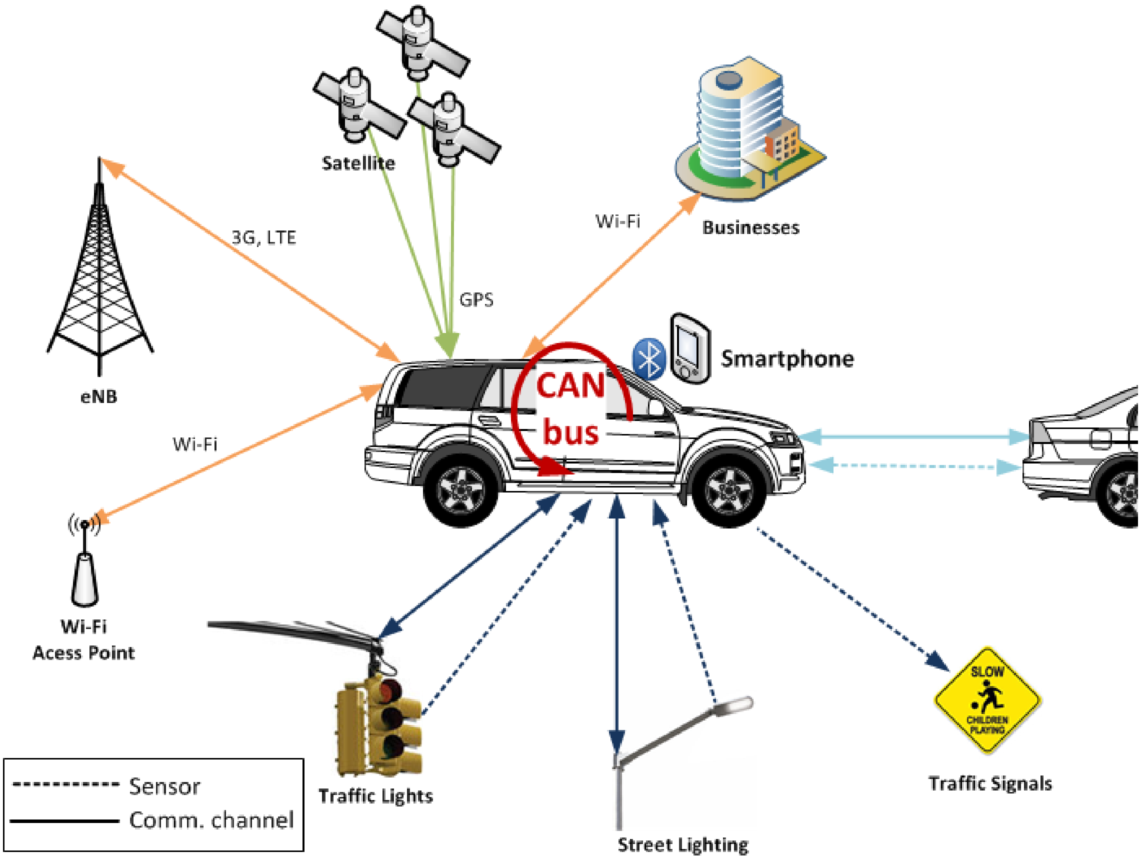


Figure 5 – The connected car

Five main relationships that are/will be enabled by the connected car can be distinguished:

- **The car connected with the road:** Traffic lights, street lights and electronic road panels are some of the road elements that by being connected to the cars would increase road safety and energy efficiency, as well as reduce traffic congestion. The entire traffic system could be interconnected in a cloud service capable of providing real time route information such as dangerous situations, traffic congestion and weather conditions to drivers.

- **The car connected with other cars:** Cars communicating with each other would present enormous safety advantages, for example in case of an accident at close proximity, the car could brake even before the driver react to the situation. Overtaking and braking maneuvers could be dramatically safer than they are today.

- **The car connected with the outside world:** A connected car poses several new communication possibilities, some of them safety related and others business related:

    - Regarding safety, a good example is the European Commission's e-Call initiative, which targets the installation of a communication system in each new car sold in Europe that, in case of accident, automatically calls the nearest emergency center. Even if no occupant is able to communicate, a "Minimum Set of Data" is sent, including the location of the crash site. The initiative's objective is to have the system seamlessly functioning throughout Europe by 2015, with the coordination of Member States, car manufacturers, telecom operators and emergency centers [2].

  - Regarding business opportunities, short range connections such as Wi-Fi would allow new types of marketing, a stopped car at a traffic jam or at a traffic light could be receiving information from nearby businesses that could be directly targeted to the occupants (e.g. offering a discount in cold beverages to occupants stopped at a traffic jam on a summer day).

- **The driver connected to the outside world:** An important feature of the connected car will be to ensure driver safety while maintaining the driver connected to the outside world. On this subject manufactures and app designers will play a major role in providing solutions that ensure these principles. One interesting possibility is the concept of Bring Your Own Device (BYOD) applied to vehicles. Basically this concept intends to provide to drivers a safe and seamless experience regardless of switch of vehicle, without having to personalize its functioning, based on the driver smartphone. The smartphone will provide communication and processing capability and the car head unit will provide an easy and safe interface for driver use [16].

- **The passengers connected to the outside world:** Passengers present good opportunities for the telecommunications operators to explore the Wi-Fi/3G/LTE connectivity: personalized marketing, entertainment services, video streaming and cloud services could be exploited both for short and long journeys. LTE will be the main connectivity solution, especially when the vehicles are in motion, although it will face coverage difficulties in rural areas, but Wi-Fi will be a more cost-effective solution when the vehicles are stopped and in urban context [17].

2.3.2 Building Management and Control

Either for economical or environmental motivations all kind of enterprises and individuals are looking for new ways to improve energetic efficiency and improve environmental sustainability. Buildings are responsible for a tremendous amount of energetic waste and therefore many applications based on M2M solutions are being implemented in private and public buildings all over the world to help monitoring, manage and control energy related applications. More information on this subject is available in Subsection 2.3.5.

2.3.3 Energy and Smart Grid

The same environmental concerns, scarce resources and carbon targets to meet, cost savings, faster decision-making demands, demand from end users for faster information, fear of churn, regulatory pressures and pursue of new revenues are steering Utilities to adopt M2M solutions [7].

The old centralized energy network is being replaced by a new distributed, interconnected, and intelligent network, currently named Smart Grid.

The interconnection of the energy network with the telecommunications network enabled by M2M solutions like remote monitoring and management, integrated communications, and even smart power generation and distribution has the potential to make the energy grid, for the first time in history, effectively self-healing and load balanced. Such goals, if accomplished, will mean a significantly increase on the quality of service and consequently on consumer satisfaction and utilities revenues.

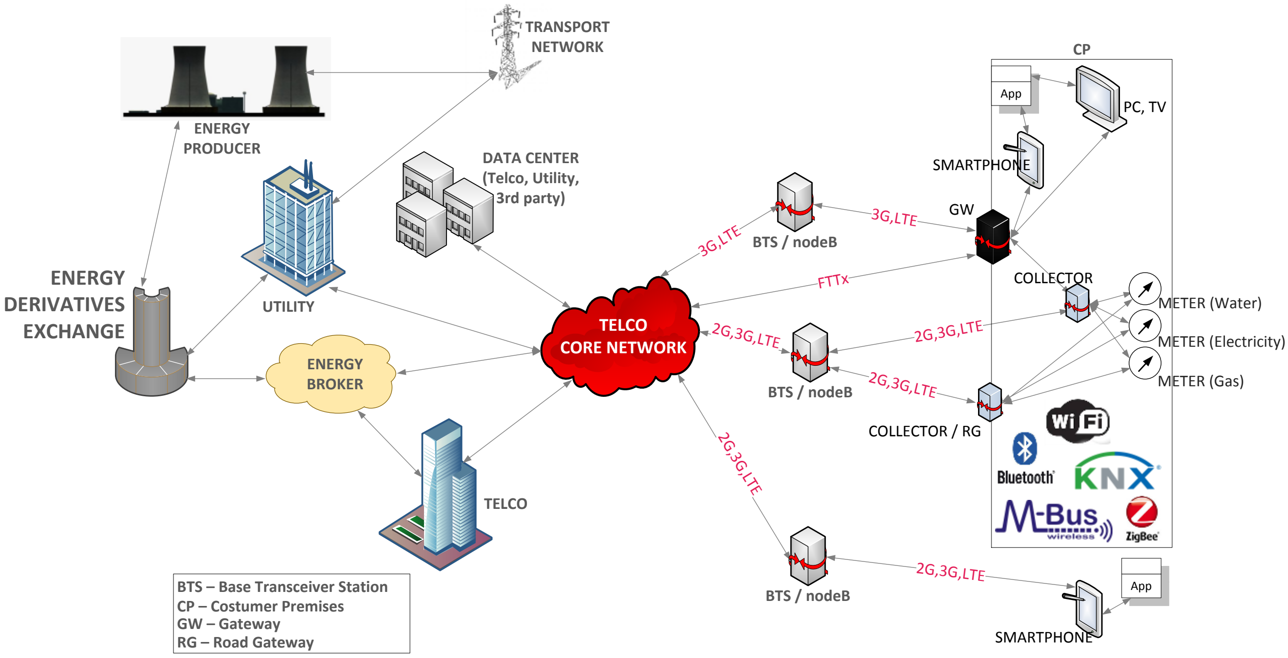


Figure 6 – The Energy-Telco ecosystem

As presented in Figure 6, many utilities are choosing to use commercial cellular networks instead of self-owned infrastructure. Accordingly to J. Berst [8], the main reasons backing up this decision are:

- **Focusing on core business:** turning over communications issues to telecom operators;
- **No last mile connection:** Cellular technologies cover about 99% of the customers in most areas;
- **Proven Technology:** Cellular technologies are mature and very well established, enjoying the robust security, reliability and performance that come from successfully connecting billions of people worldwide;
- **Pricing:** All the major carriers have a M2M division, sometimes with special focus on smart grid. Prices are very competitive;
- **Long-term stability:** Cellular networks are constantly evolving. Most carriers will contractually guarantee support and backwards compatibility for years to come. Many of the latest chipsets support multiple 3G technologies, plus backwards compatibility with 2G and extensibility with 4G;
- **QoS:** Carriers offer quality-of-service guarantees and improved abilities to prioritize traffic.

But the MNOs may grab more of the Smart Grid opportunity and expand their business other than providing connectivity. Analyzing Figure 6 it can be seen that MNOs are in a good position to perform a number of roles in this ecosystem. The following SWOT analyses were designed as way to understand which could, such roles, be, and what viability could they present. These analyses are to be seen as a purely hypothetical exercise from the author.

- **Connectivity Enabler;**

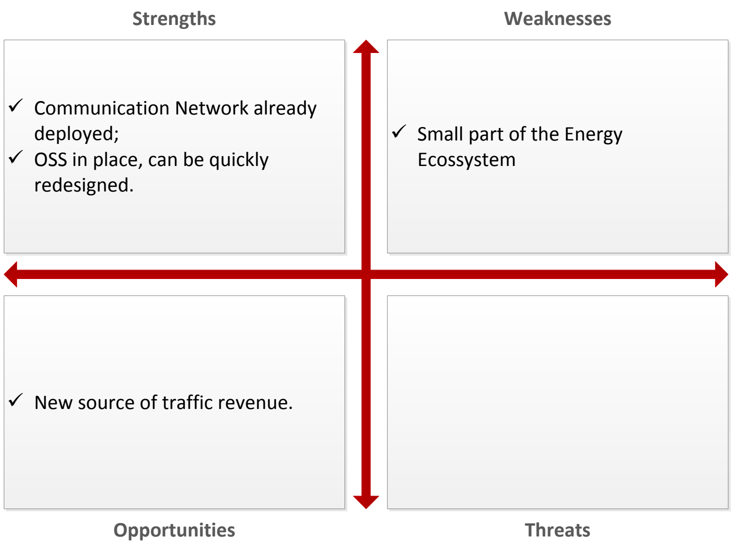


Figure 7 – SWOT analysis for Connectivity Enabler scenario

- **Virtual Energy Utility;**

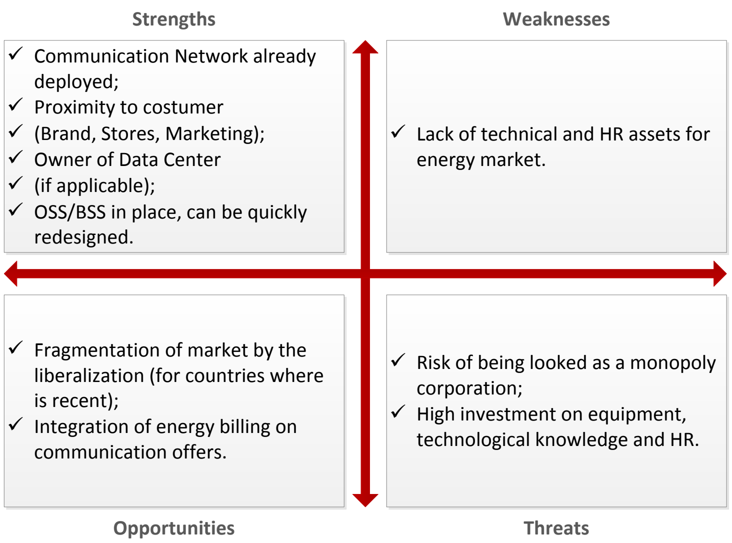


Figure 8 – SWOT analysis for Virtual Energy Utility scenario

- **Data/Billing collector and/or processor;**

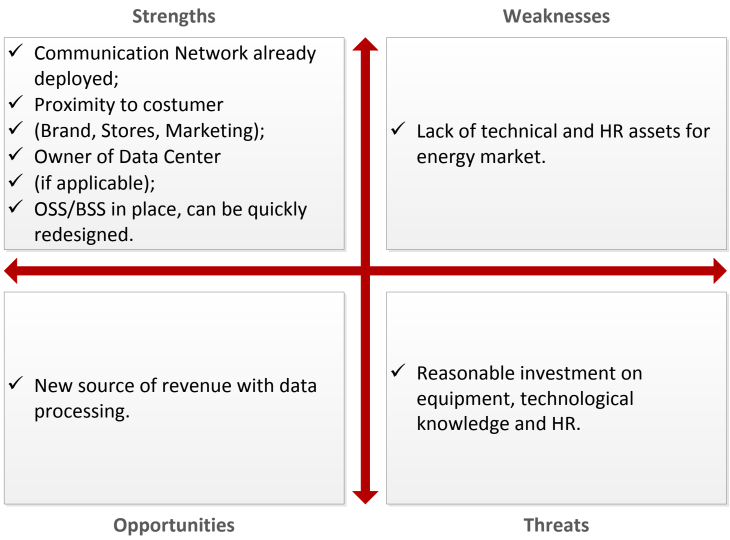


Figure 9 – SWOT analysis for Data/Billing collector and processor scenario

Energy Brokers increase their relevance in this ecosystem; such role can be played by the utility, the MNO or by other companies specialized on the task. The Smart Grid can, in theory, allow for a direct relationship between the consumer and the broker, enabling consumers to buy energy almost in a stock exchange fashion.

The Smart Grid will be constituted by many business segments, from the already referred smart power generation and distribution, to intelligent public lighting, electric cars and smart metering, among countless others.

Smart metering is the strongest one at this moment, having many benefits, from which two are fundamental:

- On one hand there are the cost savings associated with not having to send personnel to read a meter in order to produce an invoice for a subscriber. The data from the smart meter simplify the billing process that can now be fully automatic. Instead of estimated bills, the invoices are based on real data, possibly on real-time metering. This is a clear benefit to the utility company because it reduces direct costs;

- The second benefit is less obvious but of bigger importance. The best way to achieve lower energy consumptions is educating consumers themselves to rationalize energy consumption. The most effectively way to do so is presenting real time consumer data, highlighting wasteful behaviors. This will allow new utility business models, based on real time data.

Such new business models can be:

- **Improved peak and off peak tariffs:** The pricing could be established on real-time accordingly to the demand and offer of energy. Consumers could be warned when turning on devices at peak hours, motivating them to wait for lower rate hours.

- **Prepaid:** Smart metering allows for pre-paid energy plans, in fact DEFG, a management consulting company focused on energy, says participation in prepaid energy services resulted in average reductions in energy use of 11% [9] which means remarkable economic savings for both consumers and utilities.

- **Time-of-use (TOU) rates:** TOU rates are being well received by consumers [10], and are strengthening the appeal of energy management. The first deployments resulted in average reductions in energy use of 7% [11].

The Smart Grid and the new business models enabled by smart metering require standardization efforts, which are not yet well developed. Some of the most relevant initiatives in this area are:

- **Open Meter (Open Public Extended Network Metering):** An standardization initiative promoted by the European Commission under the Seventh Framework Programme with the objective of “specify a comprehensive set of open and public standards for AMI, supporting electricity, gas, water and heat metering, based on the agreement of all the relevant stakeholders in this area, and taking into account the real conditions of the utility networks so as to allow for full implementation”. The OPEN meter project finalized its work during the second half of 2011. The result of the project was a set of draft standards, based on already existing and accepted standards wherever possible. These standards include the IEC 61334 series PLC standards, the IEC 62056 DLMS/COSEM standards for electricity metering, the EN 13757 series of standards for utility metering other than electricity using M-Bus and other media [12].

- **SEP 2.0 (Smart-Energy Profile 2.0):** “Being developed to create a standard and interoperable protocol that connects smart energy devices in the home to the Smart Grid. While the original work for SEP 2.0 was done via a joint liaison agreement between the ZigBee Alliance and the HomePlug Alliance, the standard itself is designed to run over Transmission Control Protocol / Internet Protocol (TCP / IP) and is therefore media access control (MAC) and physical layer agnostic” [13]. SEP 2.0 will “help transform smart meters into smart sensors in the connected home, where multiple communication pathways will coexist and different players will interact with energy data in different ways. Big data and analytics firms, which specialize in disaggregation and energy modeling, highlighted the tremendous potential for data to transform products, services, and business relationships” [11].

Along with standardization, governmental regulation for the Smart Grid and smart metering is necessary. The European Parliament and the Council has laid ground for the implementation and regulation of this market, presenting the directives 2009/73/EC (natural gas) and 2009/72/EC (electricity).

Considering the case of electricity metering, this directive (2009/72/EC) stipulates that “Member States shall ensure the implementation of intelligent metering systems that shall assist the active participation of consumers in the electricity supply market”. It also states that “the implementation of those metering systems may be subject to an economic assessment of all the long-term costs and benefits to the market and the individual consumer or which form of intelligent metering is economically reasonable and cost-effective and which timeframe is feasible for their distribution”. In the cases where such economic assessment is positive the Member States “shall prepare a timetable with a target of up to 10 years for the implementation of intelligent metering systems. Where roll-out of smart meters is assessed positively, at least 80 % of consumers shall be equipped with intelligent metering systems by 2020” [14].

Member States transposed these directives to their own law, presenting results of the economic assessments and setting technical requirements for the meters. As an example, Portuguese Government published the ordinance 231/2013 stating that the economical assessment was positive for the electrical market and negative for the natural gas market, therefore establishing design rules only for electricity meters [15].

2.3.4 Healthcare

Mobile technology and M2M solutions will be the next big revolution of the heath sector, patients will enjoy better health outcomes, medical professionals will be more effective and efficient, governments will see their expenses tumbling and society will be more health conscious [37]. M2M solutions will allow many health issues to be monitored remotely and regularly, easing crowding of hospitals, which will eventually fade away as the center of health service delivery [37]. The smartphone will also be at the heart of this revolution, enabling connectivity and data processing for many PAN equipments.

Some of the most relevant benefits from the implementation of M2M health solutions are:

- Elderly will be able to age more independently and with closer monitoring;
- Chronic diseases will be better controlled, almost in real-time, increasing patient quality of life and reducing crowding and costs for the health system;
- The medical data from patients will be stored on cloud, providing fast and ubiquitous access, plus increased capability for quicker and more accurate diagnosis;
- Fitness equipments will feed from and to the health cloud;
- Patients in remote areas will benefit from telemedicine with increased functionalities and quality;
- Specialists from around the world will have easy access to data of patients, regardless of their location;
- Prescription and consumption of medication will be more adequate;
- Urgent patients will be diagnosed and treated more effectively even before reaching the hospital;
- Emergency services will reduce time of response, with more detailed information on location and health issues (Subsection 2.3.8).

Being one of the most profitable industries in the world, healthcare represents a huge percentage of government budgets in countries where the health sector is dominated by State. This motivates governments to implement and regulate a “mHealth” system, while time suppliers and providers could try to slow down the massification of such a system [37].

Once again the standardization and interoperability are key factors to assure the success of M2M services and devices in the health sector. One relevant initiative towards this goal is Continua Health Alliance, “a non-profit, open industry organization of healthcare and technology companies (…) dedicated to establishing a system of interoperable personal connected health solutions with the knowledge that extending those solutions into the home fosters independence, empowers individuals and provides the opportunity for truly personalized health and wellness management” [38].

2.3.5 Home Automation

The Home Automation context can be seen as an extension of the context presented on Subsection 2.3.2, “Building Management and Control”, but it goes much further than just building control.

There are M2M solutions for an extended number of home related applications and services, many times crossing over with other vertical markets:

- Energy monitoring and control (Subsection 2.3.3);
- Safety and security (Subsection 2.3.11);
- Multimedia and entertainment;
- Home appliances;
- Healthcare (Subsection 2.3.4);
- Automotive (Subsection 2.3.1).

The Home Automation paradigm is still dubious; at this point many alternative communication standards and technologies are present on a modern home. Consuming electronic manufacturers must pursue common connectivity choices in order to make a reality out of “Smart Home” concept.

Wireless short range technologies will be fundamental to enable the Home Automation in its plenitude; Wi-Fi, Zig-Bee, 6LoWPAN and W M-Bus seem to be on the front of this race. Most likely the existence of an Home Gateway - capable of manage interoperability between a wide number of wired and wireless communication technologies – will be fundamental to allow successful interoperability between all the things and services present at modern homes. An overview on what a Gateway enabling M2M services should be like is presented on Subsection 3.4.1.

Obviously the center of any home should be its residents, therefore the human will also be connected to this environment and it should be expected to see smartphones and tablets as the control center of many home features and appliances.

2.3.6 Industry

Industry has long used M2M solutions, even before the concept was named. Many M2M solutions are being used and many more are being designed to improve processes, increase sustainability and generate economical savings.

General Electrics (GE) introduced the concept of Industrial Internet, which can be seen as “M2M on steroids”. The concept was originated by GE associations with helping airlines operate and maintain its aircraft fleets. Doing so, GE realized that if captured and analyzed, data from sensors attached to jet engines, airlines could save expressive maintenance costs [39].

Accordingly with GE [40], the Industrial Internet is enabling a global convergence of the industrial system thanks to:

- Advanced computing analytics;
- Low-cost sensing;
- New levels of connectivity enabled by the Internet.

By meshing the digital world with the world of machines, GE predicts that there is potential for “profound transformation to global industry, and in turn, to many aspects of daily life” [40].

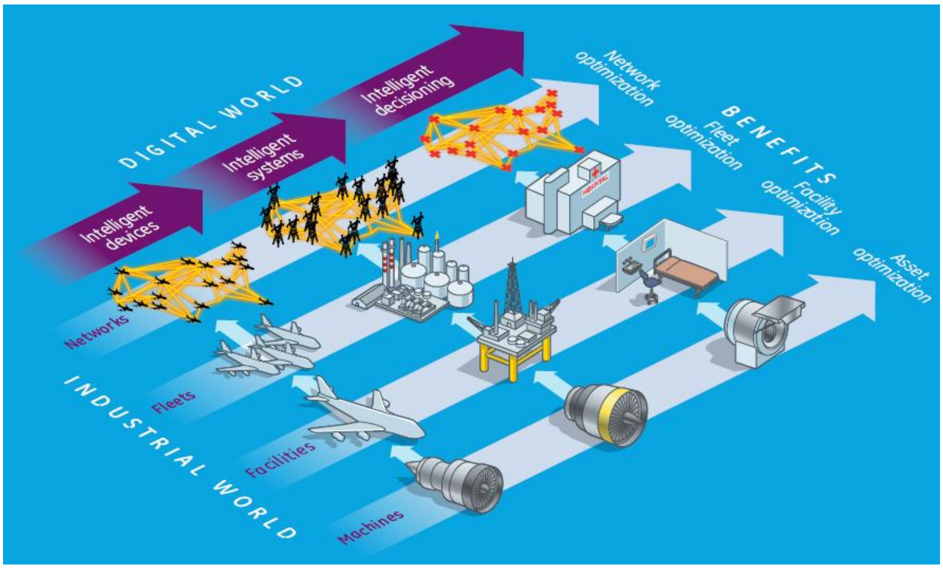


Figure 10 – Applications of the Industrial Internet [40]

Industrial Internet promises to increase speed and efficiency to industries as diverse as aviation, rail transportation, power generation, oil and gas development, and healthcare [40]. Some estimates on performance gains for these key sectors are presented on Table 1.

Industry	*Type of Savings*	*Estimated Value Over 15 years (Billion nominal USD)*
Commercial Aviation	1% Fuel Savings	30
Power: Gas-fired Generation	1% Fuel Savings	66
Healthcare	1% Reduction in System Inefficiency	63
Rail Freight	1% Reduction in System Inefficiency	27
Exploration and Development of Oil & Gas	1% Reduction in CAPEX	90

Table 1 – Industrial Internet performance gains for some key sectors [40]

A fundamental aspect of Industrial Internet is the generation, transport, storage and processing of enormous amounts of data. Consequently, telecommunication networks need to be reconfigured to deal with this paradigm. Big data is further analyzed on Section 2.4.

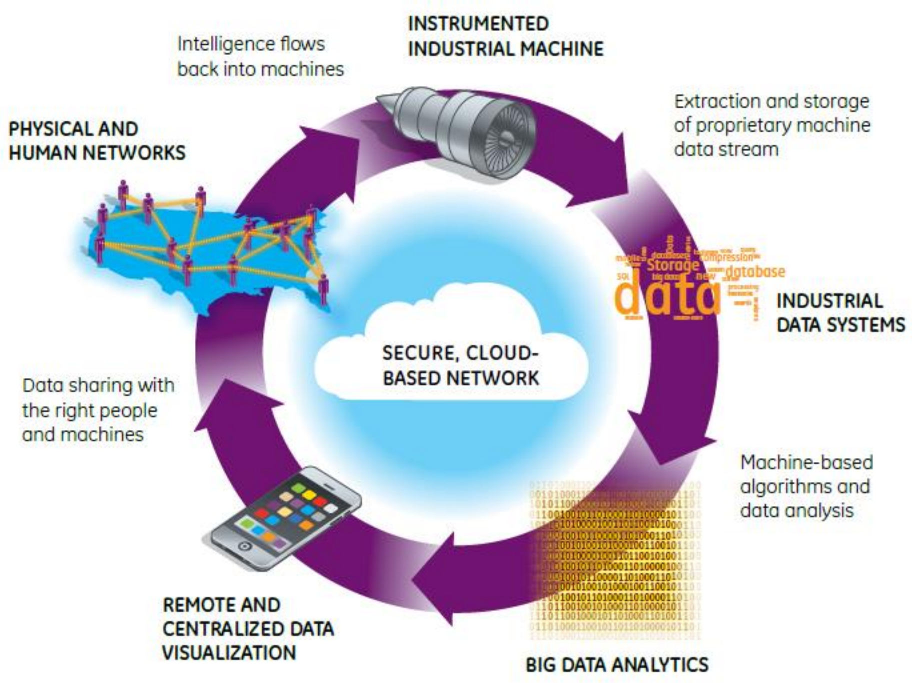


Figure 11 – Industrial Internet data loop [40]

2.3.7 Insurance

Inclusion of this vertical on the list of M2M application areas might be a surprise, but this is in fact one of the areas where M2M solutions might be more disruptive.

Car insurance in particular may change dramatically over the course of the next years. O2 UK and a joint venture of Telefónica and Generali Insurances are developing similar systems to exploit the M2M opportunity in this vertical. After the European Union ban gender discrimination on insurance policy and as response for the high insurance premiums of young drivers, these companies decided to use the omnipresent cellular coverage and advanced telematics of modern vehicles to create a system where the driver pays accordingly to how she or he drives [41], [42]. Vehicle telematics return records of driver behavior against which his or her accident risk (and consequently higher or lower premiums) is calculated. This information is stored on a secure cloud and additionally used to feed customer applications. These applications are an important part of this system.

As a response to eventual "big-brother" fears, these companies developed mobile applications that can help the customers to control the information they are generating, their scores, and also improve their driving quality. This way the gamification aspect of the system improves the road safety, reduces good drivers premiums and provides insurance companies an exact method to determine driver premiums and increase interaction and consumer loyalty [42].

2.3.8 Public Safety and Automated Response

M2M communications are making an impact also in public safety and automated response. A wide range of devices and sensors are being used by vulnerable or other at-risk personnel (e.g. the elderly or lone workers) to automatically call help, either family, co-workers or emergency services. As presented on Subsection 2.3.1 also the vehicles will be able to automatically connect to emergency services with initiatives like e-Call [2].

2.3.9 Public Transportation

M2M solutions are also being used by many public transportation systems all over the world. Energy efficiency, scheduling compliance, increased security and management improvements represent cost savings and enhanced user satisfaction.

2.3.10 Retail and Consumer Goods

Retail companies have not yet explored many of M2M solutions, but accordingly to Vodafone, 83% of interviewed business executives believe that M2M will be fundamental for future customer interaction [43].

Improved supply chain effectiveness (getting products to stores faster), optimization of stores and warehouses operation, energy efficiency gains, and customer engagement will be some of the benefits from M2M solutions in this vertical [43].

2.3.11 Security

M2M solutions won't change the security industry core business; instead they are making security devices and networks more reliable, more effective and less expensive, which can account for massive growth of this vertical.

Conventional CCTV cameral and alarms rely in wired connectivity, which can be easily disconnected by criminals; instead M2M solutions rely on cellular or satellite communication, being more reliable and enabling tracking functionalities. Such tracking ability improves security of assets on transit as well as increases stolen goods recovery [44].

2.3.12 Transports and Logistics

Since the dawn of history transports and logistics played a fundamental role for globalization and economic development. Thus is no surprise that this vertical is at the front line of M2M adoption, with clear applications and visible results (According to Vodafone 67% of transports and logistic companies using M2M say that returns are “significant”) [45].
The main drivers behind M2M adoption in this vertical are:

- Business agility;
- Consistency of delivery across markets;
- Cost reduction;
- Environmental sustainability;
- Fleet safety;
- Improved customer satisfaction;
- Process and productivity improvements [45], [46].

Such benefits are result of four key M2M services:

- Asset tracking;
- Driver management;
- Fleet management;
- Vehicle monitoring [46].

2.4 IoT Ecosystem

M2M communications are a part of a broader ecosystem being developed for the years to come, the Internet of Things (IoT). Different companies have alternative terminology; Ericsson calls this ecosystem the “Network Society” [47] while Cisco calls it the “Internet of Everything (IoE)” [3], although there may be some differences on their definitions, they are fundamentally the same thing.

Cisco defines IoE as “bringing together people, process, data, and things to make networked connections more relevant and valuable than ever before—turning information into actions that create new capabilities, richer experiences, and unprecedented economic opportunity for businesses, individuals, and countries” [3].
The kind of holistic thinking behind this ecosystem is that with decreasing costs of connectivity, everything that can benefit from being connected, will be connected [32]; therefore enabling horizontal services between vertical markets, many of them presented on Section 2.3. This will be the ultimate expression of Metcalfe's law, “the value of a telecommunications network is proportional to the square of the number of connected users of the system” [48].

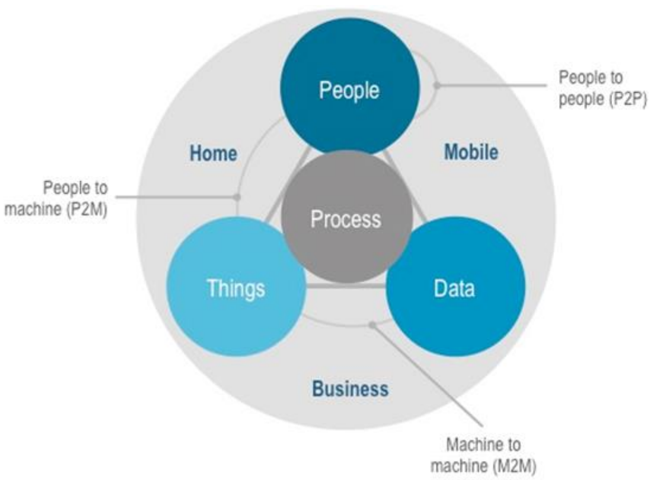


Figure 12 – The Internet of Everything according to Cisco [3]

The Internet of Things will strongly impact quotidian life, especially in urban centers, where many changes will happen, sparking the much-hyped concept of “Smart City”.

New ways of marketing will be introduced thanks to Near Field Communications (NFC) – and other short range communication technologies – allowing marketers to more effectively target and engage with users. NFC can be implemented in almost everything, from buildings to magazines, and newspapers. However, the use of this technology won’t be limited to marketing; public transportation, security and health will strongly benefit from it too. Many modern smartphones are already embedded with NFC technology.
Other field where smartphones, smart watches and another NFC enabled devices will make a big impact is finance and retail. Credit and debit cards will be something of the past when services like “Visa payWave” [49] are widely widespread.

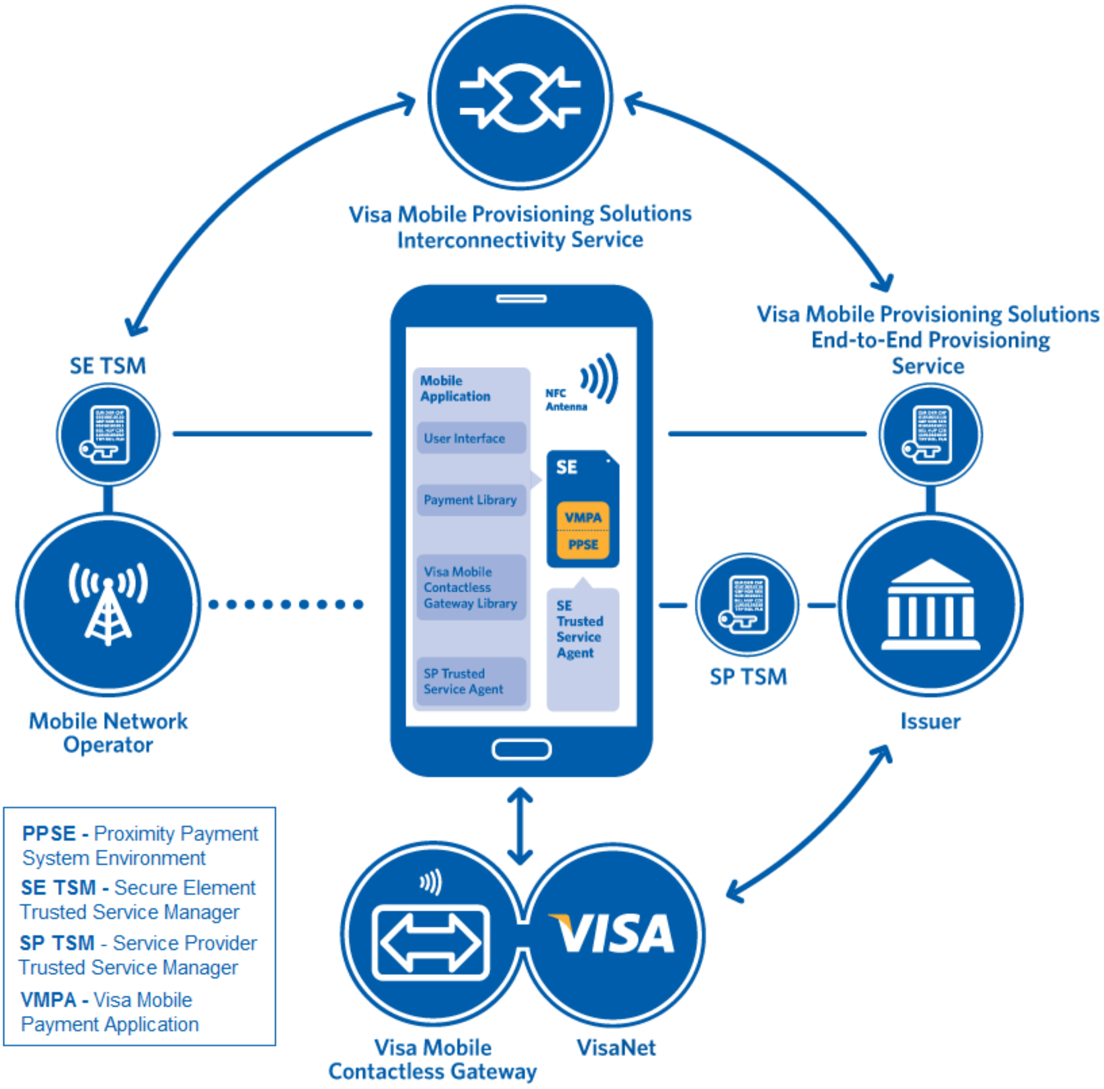


Figure 13 – Visa payWave for mobile ecosystem [49]

It becomes clear that personal devices such as smartphone, tablets or smart watches will be at the heart of this “Network Society”, serving as gateways for all sorts of connected things.

A fundamental aspect of modern urban lifestyle is commute, therefore transportation systems will be profoundly affected by this IoT ecosystem. Other than the benefits for public transportation (described on Subsection 2.3.9), highways are also being turned into “smart roads” with installation of sensors to monitor bridges, tunnels and even road surfaces [50]. Also parking is being addressed with innovative mobile applications to help find parking spaces and to allow for new models of pay for those spaces [1], [51].

A discrete but important result of the Smart City environment is that by interconnecting all things and communication networks is being deployed a massive redundant network for emergency response, as well as the generated data are being used for relief efforts. The first grand example was the devastating Hurricane Sandy that affected North America on October 2012 [4], [5].

2.4.1 Big Data

Relief efforts are in fact an example of an application of Big Data, but every single M2M and IoT service and application will generate enormous amounts of data that if adequately analyzed can generate major economic and social benefits.

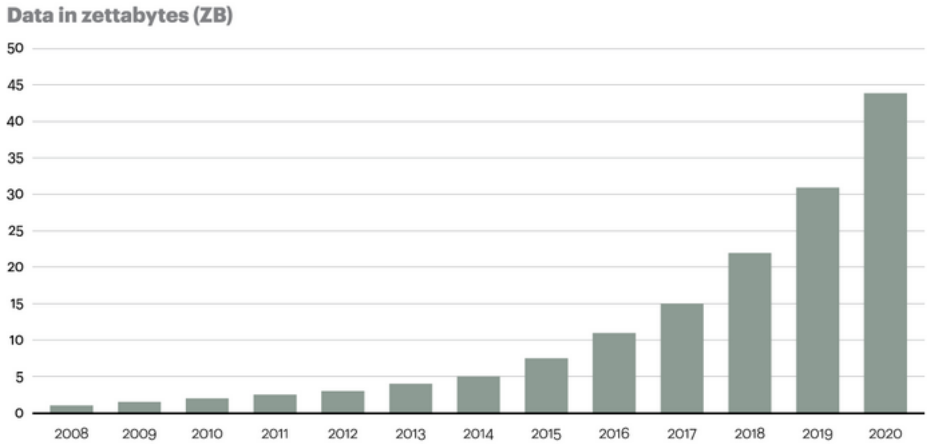


Chart 4 – Oracle estimations on annual growth of business data [52]

As Chart 4 illustrates the amount of data that will be collected over the next years will be astronomical and therefore it will be fundamental to understand how to deal with it. Quoting G. Hollingworth and J. Hoffman, this abundance of data "is only valuable if interpretations can be gathered from it in a timely fashion and if the generated interpretations are possible to use in the ongoing business process" [53].

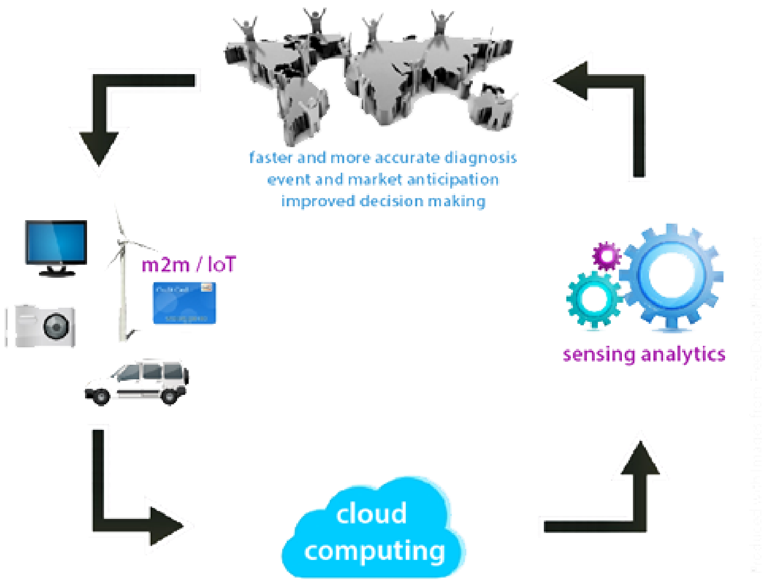


Figure 14 – The Big Data ecosystem

This poses an analytic problem, detecting relevant patterns within all of this data generated by M2M devices. And there is also the question of where they will be stored. The solution will most likely be recourse to cloud computing [54]. But the resource to cloud computing raises several security concerns: profiling, tracking, discrimination, exclusion, government surveillance and loss of control over privacy, to name a few. And there is also the question of who 'owns' data related to people, devices or processes [55].

An interesting topic to reflect upon is raised by M. Jauchius at McKinsey & Company Chief Marketing & Sales Officer Forum [56]: "The concept I would urge most marketers to think about is the notion of what you're really doing with that big data. You are trying to change the way decisions get made in marketing. And academically, we can say, 'Well, of course I am. I apply big data. I have statistically more valid techniques. And therefore, I have better decisions.' But ultimately, every organization is about people. In marketing, the people who have had the control and the power for generations are often the creatives. In truth, the creative process does start with data and so forth. But at some point, you have to have inspiration. You have to have creativity. Forgetting the people part of this art-to-science journey could mean a disaster". Although this quote is related to marketing in particular it is an important message to everyone involved in the Big Data business.

Considering all of the information presented to this point, the telecommunication operator can enter the Big Data market as a provider of connectivity, enabler of cloud storage and processing and as vendor of related services.

2.5 Standardization

There are many standards and standard organizations than will play a role in the M2M/IoT ecosystem. Figure 15 depicts some of the most important standard groups and organizations in this context.

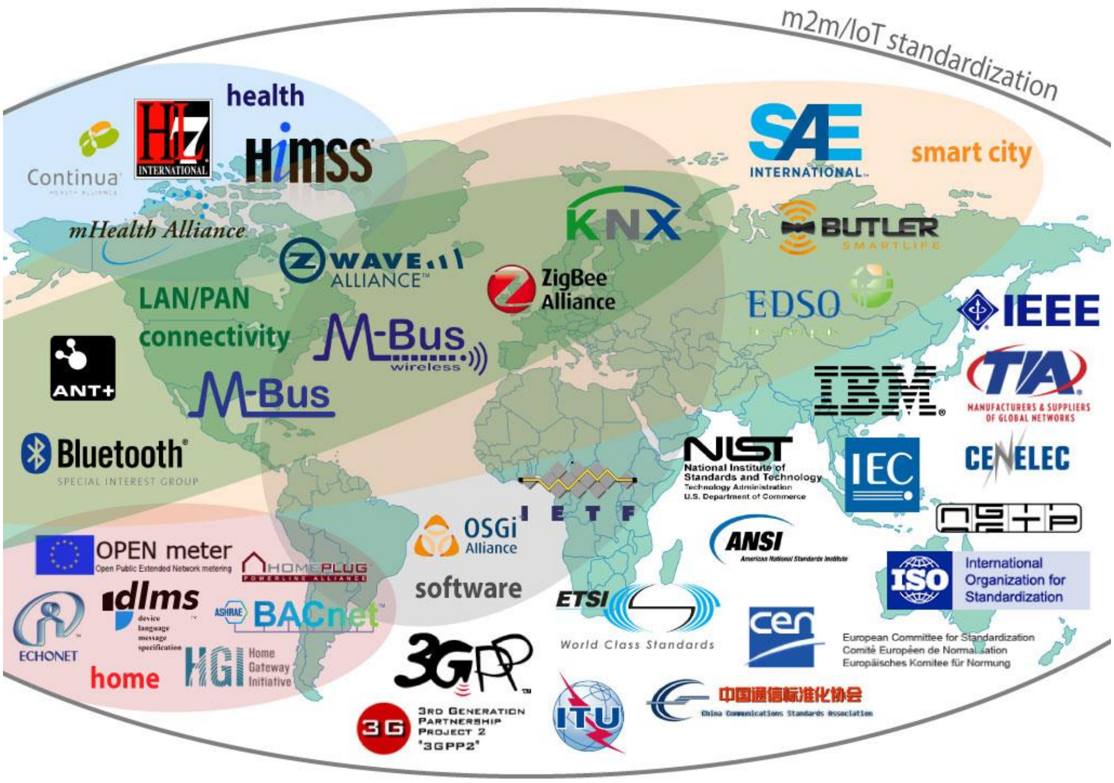


Figure 15 – M2M/IoT standardization groups and organizations (based on [57]-[60])

Figure 15 shows that there is a big number of organizations and groups dealing with M2M standardization. Therefore it may seem that there is no need for new standards (excepting specific needs), and that the important step at this point is to understand how the various standards and technologies can coexist and interoperate. But there are still several standardization tasks to complete, as it will be made clear over this Section. Considering connectivity choices, most applications will be agnostic to which technology is enabling their connectivity [61]. This is something common already, with many devices in the market than can switch seamlessly between technologies. Further developments in cognitive radio technologies will also contribute to this.

Figure 16 depicts the M2M communications high level architecture specified by ETSI.

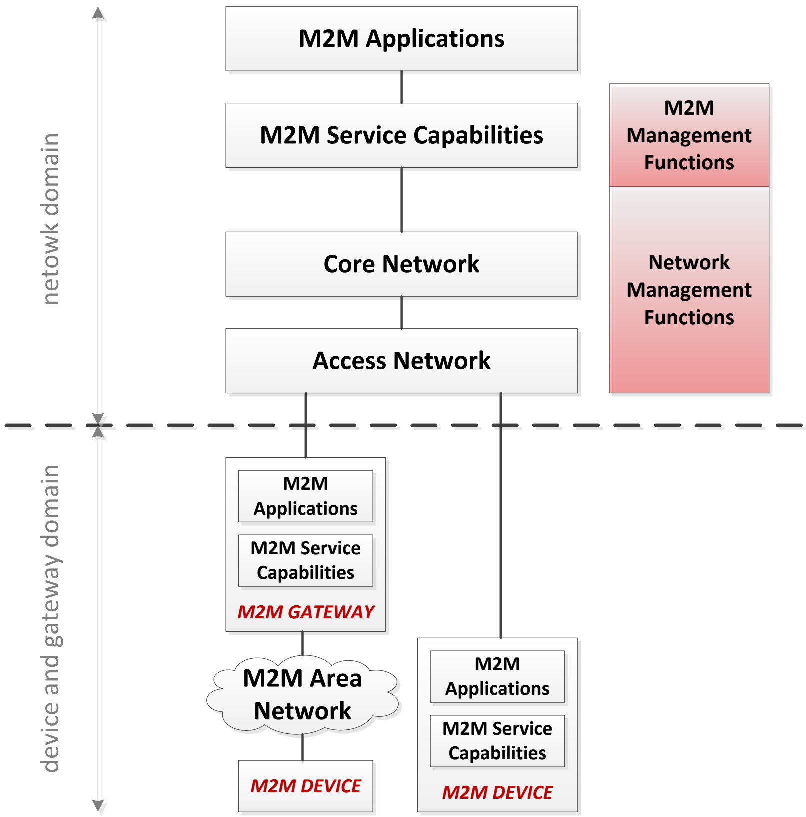


Figure 16 – ETSI M2M comm. high level architecture specification (based on [62])

Considering Figure 16 as a reference, and based on the organizations presented on Figure 15, a description of the main standardization efforts taking place is presented for each Section of the architecture:

M2M Applications and Service Capabilities

The fundamental goal for standardization efforts in this area is to enable interoperable with cost-efficient solutions [57].

- Continua Health Alliance, Health Level Seven International (HL7), Healthcare Information and Management Systems Society (HIMSS) and mHealth Alliance are developing standards and enabling interoperable health systems and devices.

- Butler is a consortium co-financed under the 7th framework program of the European Commission working with the purpose of “enabling the development of secure and smart life assistant applications thanks to a context and location aware, pervasive information system" [59]. Its main study domains are "smart home/office", "smart shopping", "smart mobility/transport", "smart healthcare/wellness" and "smart cities" [63].

- Open Meter and SEP 2.0 are smart metering standardization efforts, already presented on Subsection 2.3.3.

- DLMS User Association has produced a number of specifications for CEN and IEC regarding smart metering for electricity, gas, heat and water [60].

- IBM has made its Messaging Queue Telemetry Transport (MQTT) protocol open source, in order to allow this protocol to become an M2M standard [64]. It was "designed for constrained devices and low-bandwidth, high-latency or unreliable networks. The design principles are to minimize network bandwidth and device resource requirements whilst also attempting to ensure reliability and some degree of assurance of delivery" [65].

Core and Access Network

The fundamental goal for standardization efforts in core network is to protect networks against negative effects of M2M traffic [57]. 3GPP, 3GPP2, IEEE, IETF and ITU are working on M2M related standards. 3GPP Rel. 11 and Rel. 12 include a number of standards for M2M communications [66].

M2M Area Network and M2M Gateway

The fundamental goal for standardization efforts in the M2M Area Network is to support services of multiple verticals [57], ensuring interoperability between such services and between different technologies. Groups like HGI, Home Plug Alliance, Echonet, BACnet, ZigBee Alliance, M-BUS, Z-Wave Alliance, Bluetooth Special Interests Group, Ant+, KNX and the Internet of Things Consortium are working on M2M standardization for HANs, LANs and PANs.

The M2M Gateway is further studied on Subsection 3.4.1.

In order to avoid the definition of competing standards, seven standard-developing organizations (SDOs) have integrated their efforts on a single initiative, called "oneM2M" [57]. The scope of oneM2M is to "prepare, approve and maintain the necessary set of Technical Specifications and Technical Reports for:

- Use cases and requirements for a common set of Service Layer capabilities;
- Service Layer aspects with high level and detailed service architecture, in light of an access independent view of end-to-end services;
- Protocols/APIs/standard objects based on this architecture (open interfaces & protocols);
- Security and privacy aspects (authentication, encryption, integrity verification);
- Reachability and discovery of applications;
- Interoperability, including test and conformance specifications;
- Collection of data for charging records (to be used for billing and statistical purposes);
- Identification and naming of devices and applications;
- Information models and data management (including store and subscribe/notify functionality);
- Management aspects (including remote management of entities); and
- Common use cases, terminal/module aspects, including Service Layer interfaces/APIs between:
- Application and Service Layers;
- Service Layer and communication functions" [57], [67].

As such oneM2M has inherited the work done by the ETSI TC M2M that had been focused on achieving a standardized horizontal M2M service layer [57].

Fundamental issues to consider in the standardization process are the security concerns posed by M2M communications. ETSI recommends securing every link in the chain:

- "Physical device security (e.g. tamper-resistance);
- Communication security on application level (e.g. IP encryption end-to-end);
- Modem security;
- SIM / MIM / embedded Secure Element security;
- Network security;
- Application backend server security" [68].

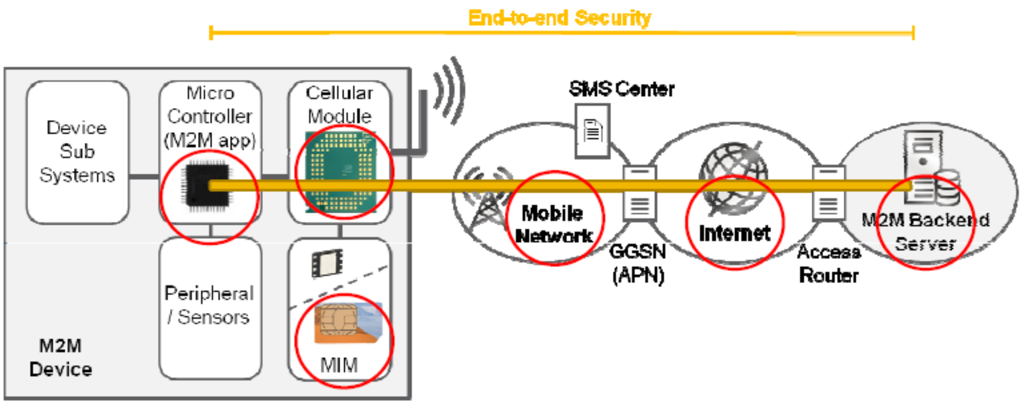


Figure 17 – ETSI recommended chain security [68]

Further information on this subject can be found in ETSI security work from 3GPP, TC TISPAN, TC SCP and TC M2M.

2.6 Security

The security measures presented on the previous Section are fundamental to avoid issues like hacking of crucial devices (e.g. heart monitors [69] and insulin pumps [70]), theft (e.g. car theft using on-board diagnostics ports [71]) and privacy breaches (e.g. hacking of baby monitors [72], and disclosing of personal data [73]). But M2M communications pose many other security concerns, some of them already addressed in this document. First of all, concerning M2M devices and connections a few security threats can be enumerated:

- More and more things are becoming "connected", meaning they will be designed for online working, thus the potential to problems when connection is lost are huge. Devices, systems, and the people who operate them, may not be prepared for standalone functioning.

- Connected devices usually imply a greater dependence on electric supply, thus power outages can become a nightmare.

- Data collectors and aggregators are usually installed at insecure locations. Furthermore installers usually prioritize cost over physical security and therefore the data privacy and integrity can be compromised [74]. Orchestrated attacks on communications networks are a growing concern.

- As data is moved from a remote site through an ISP, through various third-party cloud services and through servers, the reliability of the chain of custody of the data might be at risk [75].

Such problems call for increased concerns while design networks and devices and recommends the use of redundancy when possible (e.g. phone lines are less susceptible to orchestrated attacks on communications networks; likewise, remote monitoring networks based on phone lines are able to bypass the chain of custody previously referred and collect data directly from the remote monitoring system [75]).

The IoT ecosystem poses unprecedented security challenges: data privacy, safety, governance and trust to name a few [74]. Risks like “big-brother” governments and excessive dependence on machines and autonomous systems should be considered. Although automation has showed over time that most of time does a much better job than humans in assuring safety, the truth is that the human element is still fundamental in extreme circumstances and as a backup plan when technology fails (e.g. nowadays commercial airplanes can fly without (direct) human intervention, but in emergency situations the control is still given to pilots). IoT and M2M solutions designers must understand and consider this.

2.7 Social Implications

The age of IoT, supported by M2M communications, implies a number of social consequences. Other than the already referred fears of “big-brother” societies, the new paradigm of big data will improve consumer profiling, giving marketers and retailers a huge advantage on targeted marketing and advertising. Consumers should be aware – and prepared – for this. Governments and regulators must ensure that such profiling does not pose a risk for privacy.

Another direct result of an M2M connected world is the replacement of jobs by machines. Probably the most obvious example is smart-metering, which will replace the job of thousands of people that nowadays go door-to-door reading meters (commonly people with lower qualifications and middle aged, which might face difficulties on finding new jobs). Evidently this is not a new situation (it happens since the industrial revolution) and can always be argued that new business possibilities will provide economic growth for corporations and states, and consequently for its citizens. Anyway this kind of implication should be considered by companies entering the M2M world.

2.8 Future Evolution

Many previously presented features of IoT may never come to be, as the hype for this ecosystem falls into reality, but even so there are a few (even more) futuristic features worth mentioning.

Google is preparing commercial deployment of “Google Glass”, that promises to revolutionize (or at least open a new era for) personal devices. It might not be the death of smartphones for now, but it has the potential to shift the center of personal connectivity and increase the volume of data in the networks (especially video upload and streaming).

Figure 18 – Google Glass [76]

IBM is working with the goal of enabling computers with the sense of smell. If succeeded this will be a historical landmark in the history of technology and will allow unprecedented advancements on context sensing and cognitive computing, with enormous benefits for business, safety and health [77].

Figure 19 – IBM 5 in 5 2012: Smell [78]

Future advancements in nanotechnology will also enable concepts like Smart Dust: “conceived as a system of millimeter scale autonomous devices that form the basis for massively distributed wireless sensor networks. Smart dust motes have been demonstrated that cram sensors, interfaces, power sources, digital control communications and processing circuitry into a few cubic millimeters volume. Various self-assembly methods have been proposed. Commercial systems are already available” [79]. Among others, military applications are being studied [80].

Going even further, IoT will proceed to the interior of the human body with “nano” robots diagnosing and fixing health problems.

2.9 Summary

Over this Section M2M Communications were defined, their history and foreseeable evolution was presented, showing that this kind of communications are based on old principles but that are being used for new types of services which will present explosive growth over the course of the next years. The fundamental areas of application (vertical markets) were presented, with special focus on verticals that are believed to comprise the first significant wave of M2M deployment (e.g. energy, automotive, and healthcare). These verticals make no sense without a horizontal overview and connectivity, so the concept of Internet of Things (IoT) was presented with all of the arising implications to future societies.

An overview of current state on standardization and standardization efforts was presented, along with some considerations on standardization concerns that are still to be fully addressed. M2M Communications and the IoT pose unprecedented security challenges and risks, such were addressed on Section 2.6. The disruptive nature of M2M Communications and of the IoT Ecosystem have the potential to profoundly change the society, thus it will inevitably involve social implications as presented on Section 2.7.

3. Telecommunication Networks: Technologies and Architectures

Telecommunication networks are commonly designed with an architecture comprised by a customer network, an access network, and a core/aggregation network. Such networks are comprised of wired and/or wireless technologies.

The following Sections present the main architectures and technologies available. The concept of heterogeneous networks (HetNets) is also presented, as well as specific features for M2M communications.

3.1 Customer Network

Subscriber networks can vary in size and characteristics, depending on the physical media(s), the technology used, and on the network structure:

- **Local Area Network (LAN):** Interconnects devices in a limited area, than can be as small as an office or house, or as big as a University Campus. It can be private, semi-private (limited to authorized users, e.g. staff of a business office) or open.

- **Home Area Network (HAN):** Private network that interconnects devices inside the home perimeter.

- **Personal Area Network (PAN):** Private network that interconnects devices of a single user, e.g. Bluetooth devices associated with a specific smartphone.

Wired solutions typically provide high data-rates and are less affected by interference, but their implementation is more expensive and the flexibility and reconfiguration of the network is reduced (also, the need for opening holes on buildings is an unmotivating factor for many consumers).

Wireless solutions have lower data-rates and are more affected by interference, but they can be easily and quickly reconfigured and their implementation is much cheaper. Although they can present higher OPEX since many times imply the use of batteries, or other electrical supply source, on remote devices. Wireless networks are also a source of health concerns, which have not been fully understood up to this date.

Some of the most common physical mediums and technologies are presented in the following Subsections.

3.1.1 Wired Technologies

COAX

COAX, or coaxial cable, is the most commonly found wired media in houses. It is usually used to distribute digital or analogical TV and Internet signal.
Figure 20 and Figure 21 present common home networks based on coaxial media.

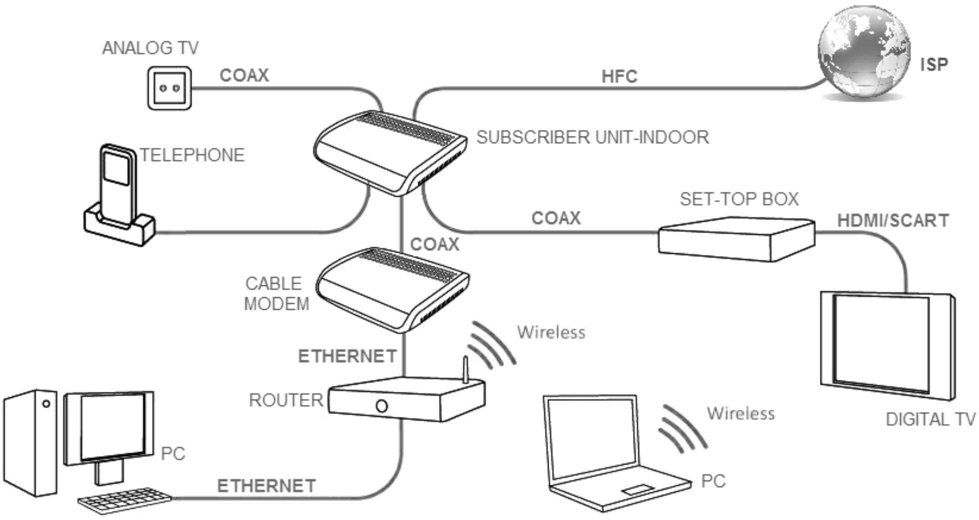


Figure 20 – Coaxial home distribution network for HFC service (based on [81])

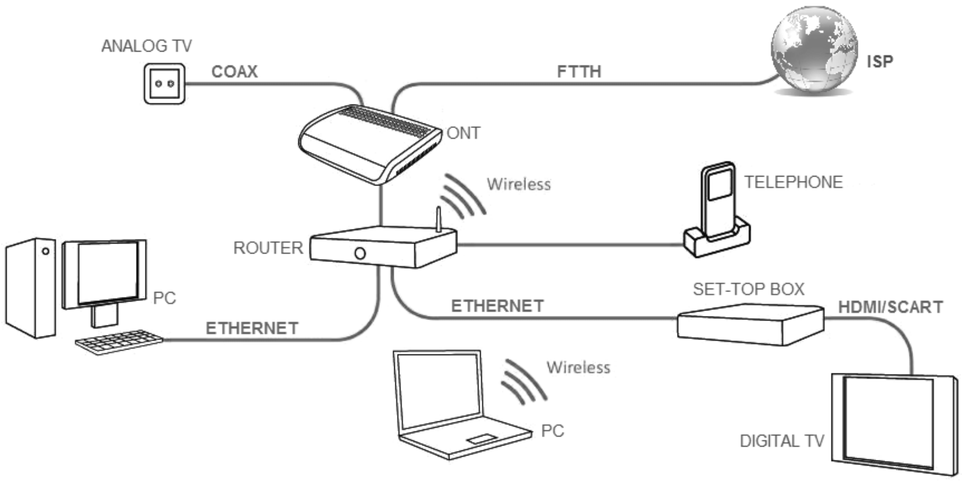


Figure 21 – Coaxial home distribution network for FFTH service (based on [81])

Ethernet

Ethernet (IEEE 802.3 [82]) based networks are not common in domestic contexts (except for the router–terminal connection, as seen in Figure 20 and Figure 21) but this it is the most used technology for office and industrial LANs.
Traditionally it was enabled by copper wiring, but in recent years is also supported by optical fiber. It is believed that Ethernet over fiber can provide data rates up to 400 Gbps [83].

M-Bus

M-Bus is a European standard (EN 13757-2 physical and link layer, EN 13757-3 application layer) for the remote reading of gas, heat and electricity meters [84]. It is a twisted-pair communication system with a high level of security and highly scalable [85].

Figure 22 – M-BUS standard official logo [84]

KNX

KNX is a worldwide standard (ISO/IEC 14543-3) "for applications in home and building control, ranging from lighting and shutter control to various security systems, heating, ventilation, air conditioning, monitoring, alarming, water control, energy management, metering as well as household appliances, audio and lots more. The technology can be used in new and existing buildings" [86].
KNX can be implemented over twisted-pair, PLC, RF and Ethernet [86].

Figure 23 – KNX standard official logo [86]

PLC

Power Line Communication technology uses the electric wiring as communication media, overlaying a modulated carrier signal. Although it is an unregulated and low cost transmission media, it has many problems such as high noise levels, high attenuation, impedance variations and most of all, low privacy and security. Thus there is a tradeoff between range and data rates, which in top of the low security explains the reason why PLC was never as widespread as other technologies.

FITH

FITH, or Fiber-in-the-Home, are domestic networks based on Fiber Optics. Although common glass fiber FITH implementations are scarce, advancements in Plastic Optical Fiber (POF) may change that. POFs are cheaper and present a major feature for building installation: POFs are resistant to physical stress and can be bended to higher radius.
POFs polymers have been recognized as a cost-effective solution and good competitors to copper lines in the broadband transmission of short range, especially as indoor infrastructure for broadband Gigabit Ethernet [87]. POF has already showed to be feasible for a transmission rate of 40Gbps over a distance of 200m [88].

Other Wired Technologies

Other wired technologies commonly present in customer networks include:

- **HDMI:** High-Definition Multimedia Interface [89].
- **IEEE 1394:** also known as "FireWire" (Apple), "i.LINK" (Sony), and "Lynx" (Texas Instruments) [90].
- **SCS:** "Sistema Cablaggio Semplificato" (Simplified Wiring System) [91].
- **USB:** Universal Serial Bus [92].

3.1.2 Wireless Technologies

It should be noted that in this Subsection only short-range technologies are considered, although wide-range technologies are also used to provide customer area connectivity. For the purpose of this document such fact is overlooked and wide-range wireless technologies are presented on Subsection 3.2.2.

Wi-Fi

Wi-Fi is a Wi-Fi Alliance trademark. Wi-Fi networks use radio technologies specified by IEEE 802.11, providing secure, reliable, fast wireless connectivity. A Wi-Fi network can be used to connect electronic devices to each other, to the Internet, and to wire networks which use Ethernet technology. Wi-Fi operates in the 2.4 or 5 GHz radio bands, or both, in case of dual-band products [93].

Is one of the most successful technology families worldwide, and with more than 500 Wi-Fi Alliance member companies, interoperability is one of its main characteristics [93]. There are countless devices with Wi-Fi certification (Wi-Fi certified products bear the official logo - Figure 24).

Figure 24 – Official Wi-Fi certified product logo [93]

There are several Wi-Fi Generations, as presented on Table 2.

Wi-Fi Technology	*Frequency Band*	*Maximum Bandwidth*	*Approximated Indoor Range*	*Approximated Outdoor Range*
802.11a	5 GHz	54 Mbps	40 m	100 m
802.11b	2.4 GHz	11 Mbps	70 m	150 m
802.11g	2.4 GHz	54 Mbps	80 m	200 m
802.11n	2.4 GHz, 5 GHz, 2.4 or 5 GHz (selectable), 2.4 and 5 GHz (concurrent)	450 Mbps	100 m	250 m
802.11ac	5 GHz	1.3 Gbps	n/a	n/a

Table 2 – Wi-Fi generations [93], [94]

ZigBee

ZigBee is a specification for high level communication protocols, used to create customer area networks built from small, low-power digital radios. ZigBee is based on the IEEE 802.15 standard [95]. Though low-powered, ZigBee enables data transmission over long distances by relaying of data through intermediate devices, creating a mesh network.
ZigBee operates globally in the 2.4GHz frequency band according to IEEE 802.15.4 and regionally in the 915 MHz (Americas), 868 MHz (Europe) and 920 MHz (Japan) [95].

In response to growing concerns of industries over energy efficiency, in the end of 2012 ZigBee launched a Green Power Feature. This feature allows switches and dimmers, among other devices, resort to energy harvesting sources instead of batteries or AC power, creating what ZigBee Alliance calls a "no-maintenance, environmentally friendly solution" [96].

With ZigBee PRO[4] Green Power products, "consumers and businesses will appreciate the install it and forget it simplicity" [97]. ZigBee devices can be easily installed, including locations where power is unavailable, not allowed for safety reasons or for historical preservation purposes [97].

Figure 25 – ZigBee certified product official logo [98]

Z-Wave

Z-Wave is a wireless communications protocol designed for building automation, specialized on remote control applications for residential and commercial environments.
It is a low powered RF communications technology that supports full mesh networks without the need for a coordinator node. It operates in the sub-1GHz band, being impervious to interference from Wi-Fi and other wireless technologies in the 2.4-GHz range (Bluetooth, ZigBee, etc.). It supports data rates up to 100 kbps, IPV6, and multi-channel operation. MAC and PHY layers are described by ITU-T G.9959 specification [99].

W M-BUS

W M-Bus, or Wireless M-Bus, is the Wireless European standard (EN 13757-4 [100]) correspondent to the M-Bus standard (Subsection 3.1.1). It defines several operating modes for the remote reading of gas, heat and electricity meters, operating at 169MHz [100], [101], 433MHz [100], [102], and in the 868 MHz to 870 MHz SRD band [101]-[103].

Figure 26 – W M-BUS official logo [84]

The W M-Bus specification EN 13757-4:2010 defines six fundamental modes to exchange data with remote meters that can be further divided on different modes [103], [104]:

- **Mode S:** stationary;
- **Mode T:** frequent transmit;
- **Mode R2**: frequent receive;
- **Mode C:** compact;
- **Mode N:** narrowband VHF;
- **Mode F:** frequent receive and transmit.

6LowPan

IPv6 over Low power Wireless Personal Area Networks (6LoWPAN) is an IETF working group [105]. The 6LoWPAN concept originated from the idea that "Internet Protocol could and should be applied even to the smallest devices," and that low-power devices with limited processing capabilities should be part of the IoT ecosystem [106]. The 6LoWPAN group has defined encapsulation and header compression mechanisms that allow IPv6 packets to be transferred over IEEE 802.15.4 (e.g. ZigBee) based networks [105].

[4] ZigBee PRO: enhanced ZigBee specification.

Bluetooth

Bluetooth is a wireless technology for short range data exchange, created by Ericsson in 1994. In 1998, the original group of promoter companies – Ericsson, Intel, Nokia, Toshiba and IBM – came together to form the Bluetooth Special Interest Group (SIG) [107]. Since that date billions of devices, ranging from mobile phones and computers to medical devices and home entertainment products, have been designed to rely on Bluetooth connectivity. The key features of Bluetooth technology are robustness, low power, and low cost [108].

Bluetooth is one of the most used technologies to enable PAN networks. "It operates in the unlicensed Industrial, Scientific and Medical (ISM) band at 2.4 to 2.485 GHz, using a spread spectrum, frequency hopping, full-duplex signal at a nominal rate of 1600 hops/sec" [108].

Class	***Approx. Range (m)***
1	1
2	10
3	100

Table 3 – Bluetooth radio classes [108]

Over the course of the years several Bluetooth versions were released, Table 4 resumes some of the key features of every version.

Core Version	***Theoretical Max. Data Rate***	***Key features***
1.0 (1999)	n/a	Discontinued version.
1.2 (2003)	1 Mbps	-
2.0 + EDR (2004)	3 Mbps	Introduction of Enhanced Data Rate (EDR) for faster data transfer; Introduction of radios up to 100 meter range; Introduction of NFC support.
3.0 + HS (2004)	24 Mbps	Introduction of 802.11 as a High Speed (HS) transport.
4.0 + BLE (2010)	-	Introduction of Bluetooth Low Energy (BLE), aimed at ultra-low power applications.

Table 4 – Bluetooth core versions [109], [110]

Bluetooth Low Energy (BLE) enables a simple communication link between devices that must operate with extreme low power consumption (able to operate for months or years on button-cell batteries [111]), with very low cost. BLE it is particularly suited for devices that transfer small data amounts within close range [110].

Consequently BLE provoked an explosion on the number of Bluetooth devices, which led Bluetooth SIG to create two Bluetooth Smart Marks:

- **Bluetooth Smart Ready** devices can efficiently receive data from Bluetooth Smart devices;

- **Bluetooth Smart** devices are designed to gather a specific piece of information and send it to a Bluetooth Smart Ready device [111].

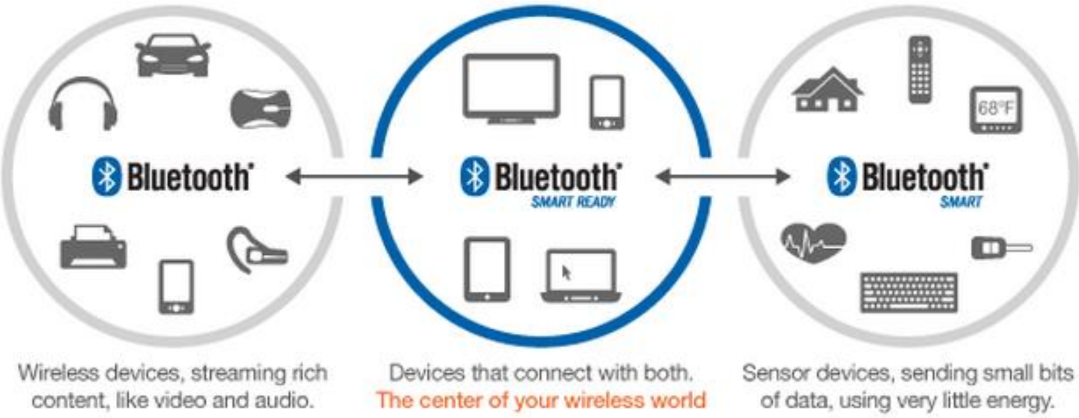


Figure 27 – Bluetooth ecosystem [111]

NFC

Near Field Communication (NFC) is a set of short-range wireless standards, typically requiring a distance of 4 cm or less to initiate a connection. NFC enables sharing of small payloads of data between an NFC tag and an NFC device, or between NFC devices [112].

NFC contactless technologies enable solutions in areas such as:

- Access control;
- Consumer electronics;
- Healthcare;
- Information collection and exchange;
- Loyalty and coupons;
- Payments;
- Transports [113].

Figure 28 – NFC certified products official logo [114]

Other Wireless Technologies

Other wireless technologies commonly present in customer networks include:

- **ANT:** Wireless sensor network [115].
- **DASH7:** Open-source wireless sensor networking standard [116].
- **ISA100.11a (IEEE 802.15.4):** Wireless networking technology standard developed by the International Society of Automation (ISA) [117].
- **MiWi (IEEE 802.15.4):** wireless protocol [118].
- **ONE-NET**: Open-source standard designed for wireless low-cost, low-power (battery-operated) control networks [119].
- **Wireless Heart (IEEE 802.15.4):** Wireless technology based on the Highway Addressable Remote Transducer (HEART) protocol [120].

3.1.3 Summary

	Technology	Applications	Max. Data Rate	Range	Security	Interference Immunity	Strong Interoperability	Flexible and Scalable	Energy Saving Functions	Low-Cost
Wired	**COAX**	Broadband	▲120 Mbps ▼160 Mbps	kms	high	good	no	no	no	no
	Ethernet (Copper/ Fiber)	Broadband	400 Gbps	up to 70 km (fiber)	high	excellent	no	environment dependent	no	no
	M-Bus	Residential Automation	57,6 kbps	up to 1 km	high	good	no	yes	no	yes
	KNX	Residential Automation	n/a	up to 1 km	high	good	yes	yes	yes	yes
	PLC	Automation	200 Mbps	data rate dependent	very low	very low	no	yes	no	yes
	FITH (POF)	Broadband	40 Gbps	kms	high	excellent	no	no	no	no
Wireless	**Wi-Fi**	Broadband	1,3 Gbps	up to 250 m	medium	low	yes	yes	yes	yes
	ZigBee	Automation	250 kbps	up to 100 m	low	low	yes	yes	yes	yes
	Z-Wave	Residential Automation	100 kbps	up to 100 m	low	n/a	yes	yes	yes	yes
	W M-Bus	Residential Automation	225 kbps	up to 5 km	low	low	no	yes	no	yes
	6LowPan	Automation	250 kbps	up to 100 m	low	low	n/a	yes	no	yes
	Bluetooth	Automation	24 Mbps	up to 100 m	low	good	yes	yes	yes	yes
	NFC	Automation	424 kbps	up to 4 cm	high	good	yes	yes	n/a	yes

Table 5 – Summary of customer area technologies [83], [85], [88], [121]-[124]

3.2 Access Network

The Access Network is the middle segment of a telecommunication network; it interconnects the customer network (or device) to the core network of the telecommunications operator. In most countries the oldest access network is the telephone network (copper wiring on poles or ducts), that goes back to 1870s and Bell Telephone Company [125]. Around 1950s two new technologies started to be developed – CATV and DSL – and by the end of the century they become widespread in several developed countries. Technological evolutions, respectively HFC and ADSL would come to be the most successful versions of these technologies to this day.

Also in the end of the XX century the first wireless digital cellular access networks were implemented, it was the beginning of GSM, the second generation of mobile telecommunications technology (2G). In the turn of the millennium the third generation of mobile telecommunications technology (3G) provided data rates of at least 200 kbps. Advancements on 3G technologies would eventually provide mobile broadband access of several Mbps to smartphones and mobile modems in laptop computers: the dawn of mobile broadband access.

The beginning of the XIX century also saw the widespread commercially deployment of a new wired access network: Fiber to the x (FTTx), a generic term for any broadband network architecture using optical fiber to provide all or part of the access network used for last mile telecommunications. 2012 marks the deployment of the first commercial LTE-A network, the first technology complying with the requirements of 4G, mainly, peak data rates up to approximately 100 Mbps for high mobility (such as mobile access) and up to approximately 1 Gbps for low mobility (such as nomadic/local wireless access) [126].

Another mobile access technology is WiMAX. It was created by the WiMAX Forum, which was formed in 2001 and that describes WiMAX as "a standards-based technology enabling the delivery of last mile wireless broadband access as an alternative to cable and DSL" [127]. This technology never reached considerable commercial success, eventually becoming part of 3G and 4G families [128].

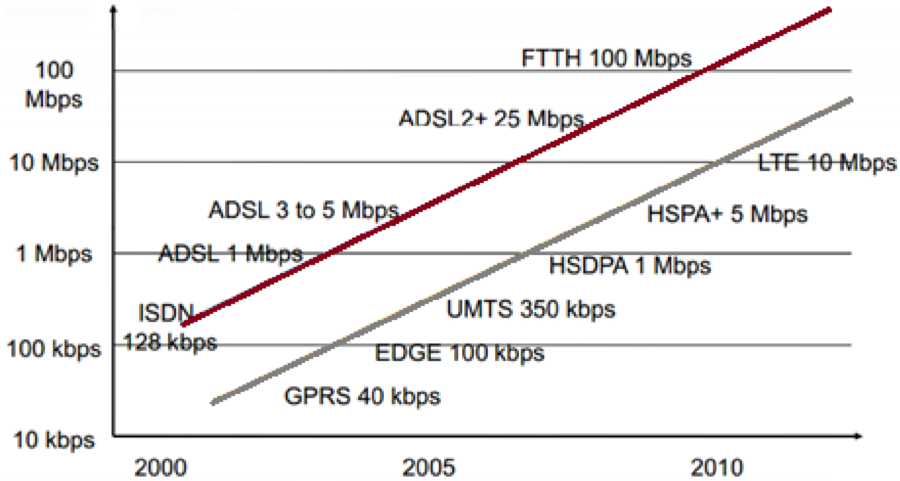


Figure 29 – Wired and wireless advances on access technologies (based on [128])

Satellite is also a successful access technology, especially in remote areas where wired broadband access is not available and mobile broadband also has poor coverage or is more expensive. Another interesting technology, mainly for urban context, is Free Space Optics (FSO).

This Section aims to provide an overview of these access network alternatives.

3.2.1 Wired Technologies

xDSL

Digital Subscriber Line (DSL) is a family of technologies that provide Broadband Internet access, that use the copper telephone network to transmit digital data. DSL family is comprised of several technologies:

- **Asymmetric Digital Subscriber Line (ADSL):** The most commercially adopted version of DSL, can be distributed over short distances from the telephone exchange typically providing a range of up to 5.5 km, depending on the diameter of the cable and its quality [129]. The first version of ADSL provided data rates up to 10 Mbps (downstream) and 1 Mbps (upstream), further advancements increased downstream rates (20 Mbps – ADSL2+ and 52 Mbps – ADSL2++) [130].

- **Very-high-bit-rate Digital Subscriber Line (VDSL or VHDSL):** Provides higher data rates than ADSL, up to 52 Mbps downstream and 10 Mbps upstream [130], and even higher rates (around 85 Mbps – up and downstream) for coaxial cable implementations [131]. Second-generation systems (VDSL2) can provide copper/fiber hybrid networks with up to 100 Mbps data rates (up and downstream) [132]. The maximum data rates are achieved at a range of about 300 meters; speeds of around 10 Mbps are possible for connections up to 1.3 km [130].

- **Other DSL technologies.**

Typical xDSL access networks interconnect the customer premises to the Digital Subscriber Line Access Multiplexer (DSLAM) located at the telephone exchange. xDSL architecture is presented on Figure 40**.**

HFC

Hybrid Fiber-Coaxial (HFC) is the name given to telecommunication networks that combine optical fiber and coaxial cable. HFC networks are enabled by Data Over Cable Service Interface Specification (DOCSIS), an international telecommunications standard that provides high-speed data transfer to existing cable TV (CATV) systems.

There are several DOCSIS versions, their maximum bandwidths are presented on Table 6.

Version	***Downstream***	***Upstream***
1.0 / 1.1	50 Mbps	10 Mbps
2.0	50 Mbps	30 Mbps
3.0	160 Mbps	120 Mbps
3.1 (in development)	10 Gbps	1 Gbps

Table 6 – DOCSIS bandwidths [121], [133]

The architecture of an HFC network is presented in Figure 41.

FTTx

Fiber to the x (FTTX) is a generic term for any broadband network architecture using optical fiber to provide all or part of the local loop used for last mile connectivity.

FTTx implementations can range from FFTN to FFTH[5]:

- **Fiber to the Node/Neighborhood (FTTN):** Optical Fiber is terminated in a street cabinet (where the ONU is located), with copper as the last mile connection. Typically the distance from the cabinet to Customer Premises (CP) is reasonable (up to several km);

- **Fiber to the Curb/Closet/Cabinet (FTTC):** Similar to FTTN, but the street cabinet or pole (ONU) is closer to the CP, typically within 300 m, the necessary range for high-bandwidth copper technologies;

- Fiber to the premises (FTTP):
  - **Fiber to the Building (FTTB):** The optical fiber reaches the boundary of the building (e.g. basement in a multi-dwelling unit), or the corridor, with the final connection from the ONT to the independent dwelling being established by other wired or wireless technology;

  - **Fiber to the Home/Office (FTTH/O):** The optical fiber (one subscriber gets one exclusive fiber, from the OLT to the ONT) reaches the boundary of the dwelling. The ONT is installed within the CP and the connectivity to LAN/HAN devices is enabled by other wired or wireless technology [134], [135].

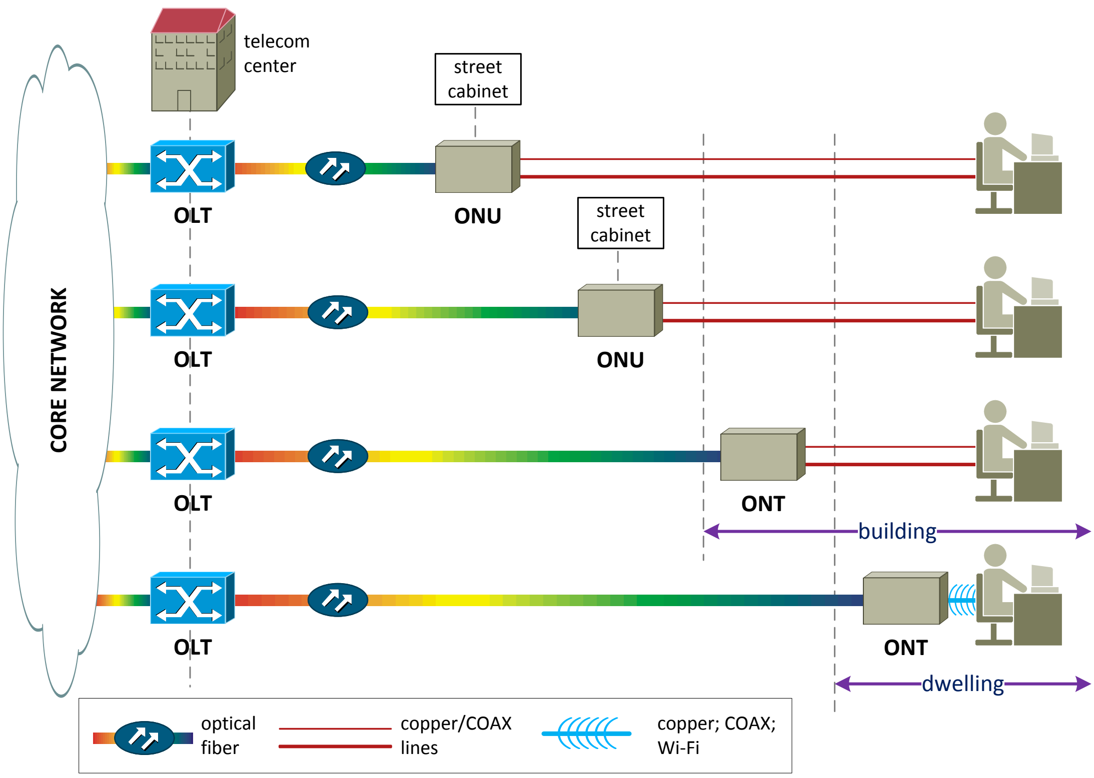


Figure 30 – FTTx architectures

[5] There are other FTTx terms applicable to specific implementations. Their study is not relevant to this document.

The choice between these FTTx architectures defines the available bandwidth per user, as seen on Table 7.

Architecture	***Distance OLT ↔ ONU***	***Distance ONU ↔ CP***	***Bandwidth per Customer***	***ONU Capability [6]***	***ONU Interfaces***
FTTN	< 100 km	< 3 km	2 to 25 Mbps	hundreds	POTS; xDSL
FTTC	< 100 km	< 300 m	50 to 100 Mbps	n/a	FE; POTS; VDSL
FTTB	< 20 km	< 50 m	50 to 100 Mbps	tens	FE; POTS; VDSL
FTTH/O	< 20 km	< 20 m	>= 100 Mbps	single home/office	FE, GE; POTS; RF; W-Fi

Table 7 – FTTx features [134]-[136]

3.2.2 Wireless Technologies

GSM

Global System for Mobile Communications (GSM) is a 1992 standard set developed by ETSI to describe protocols for 2G digital cellular networks. It was improved over time with new releases, especially with packet data transport via GPRS and EDGE [137], [138]. GSM operates in the 800 MHz band in Europe and in the 1900 MHz band in the USA, supporting FDD for uplink and downlink. Each radio channel has 200 kHz. GSM900 supports 124 two-way channels (assigned to a cluster of seven cells), while GSM1800 supports 374 two-way channels [138].

TDMA is the multi-access technique used on GSM, with eight users sharing each 200 kHz channel. Consequently each of those users will have 25 kHz of available bandwidth [138]. To avoid channel errors GSM has error correction and interleaving. The modulation scheme used is Gaussian Minimum Shift Keying (GMSK), a type of Frequency Shift Keying (FSK) [138]. Slow Frequency Hopping (SFH) is used at a slow frame rate. Each frame is sent in a repetitive pattern, hopping through frequencies on all the available channels. Frequency Hopping reduces the effects of fading, improving link performance [138].

As stated before, the GSM system was enhanced by the implementation of the General Packet Radio Service (GPRS). Its main enhancement is support of packet data. GPRS uses packet switching to transmit at high-speed data and improve signaling efficiency. Both radio and network usage resources are optimized with GPRS, which preserves severe separation of radio and network subsystems, enabling the network subsystem to be used with other radio access technologies [138]. GPRS defined new radio channels, allowing dynamic channel allocation for each user. An active user can take up to eight slots per TDMA frame. Several channel coding schemes are defined, enabling bit rates from 9 kbps to more than 150 kbps, for user [138].

GPRS was designed to support both bursty and intermittent data transmissions. Four different QoS classes are defined. User data is transferred transparently between the mobile station and the external data networks with the use of encapsulation and tunneling. Such data can be compressed and protected with retransmission to ensure efficiency and reliability [138].

Figure 31 – GSM official logo, used to identify compatible devices [139]

[6] The capacity of a single fiber on the ONU.

Further enhancements on GSM/GPRS systems were introduced by Enhanced Data Rates for Global Evolution (EDGE), a new radio interface technology capable of boost network capacity and user data rates [138]. EDGE increases GSM/GPRS data rates by up to three times, enabling services such as email, multimedia, web browsing and video calls [138]. Like GSM, EDGE uses TDMA frame structure, logic channel, and 200 kHz channel bandwidth. It introduces 8-PSK modulation, providing up to 400 kbps per carrier, and peak rates up to 473 kbps per user. The EDGE system is also characterized by adaptive modulation and coding scheme, which increases system efficiency [138].

A fundamental EDGE feature is the link quality control, through link adaptation and incremental redundancy. The link adaptation technique regularly estimates the channel quality and successively selects the most appropriated modulation and coding scheme for the transmission, maximizing the user bit rate. In the incremental redundancy scheme, information is first sent with limited codding, yielding a higher bit rate if decoding is well succeeded. If not succeeded, more coding bites are sent, generating a low bit rate [138]. EDGE devices are fully backwards compatible with GPRS [138].

GSM/GPRS architecture is presented on Figure 42.

UMTS

3GPP defines UMTS as “an umbrella term for the third generation radio technologies (3G) developed within 3GPP” [140]. UMTS radio access specifications support Frequency Division Duplex (FDD) and Time Division Duplex (TDD) variants, enabling UTRA technology to operate on a wide range of bands and co-exist with other radio access technologies [140].

UMTS includes the original W-CDMA scheme (Release 99 and Release 4). High Speed Packet Access (HSPA) was introduced in Release 5 (Downlink) and 6 (Uplink) increasing substantially the bit rates and improving packet-switched applications [140]. Networks with Evolved HSPA (HSPA+) – Release 7 to 10 – can achieve up to 168 Mbps on the downlink (HSPDA) and 23 Mbps on the uplink (HSUPA) [141]. UMTS radio access network connects to the core network which is an evolution from the GSM core. “3GPP has expanded its capabilities, in principle allowing most services to be delivered over either 2G GERAN (GSM/EDGE) or 3G UTRAN” [140].

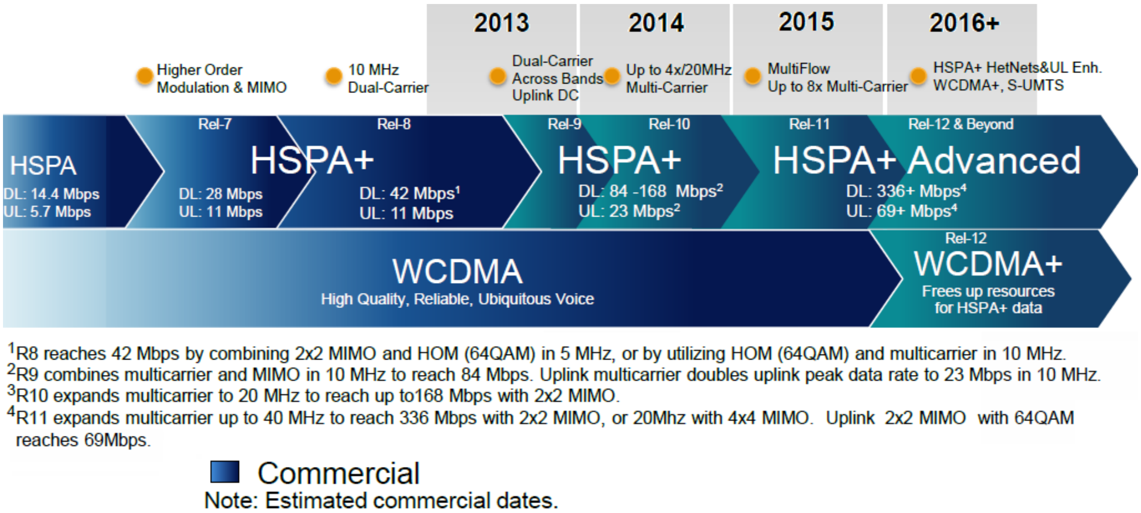


Figure 32 – HSPA evolution [141]

The UMTS architecture is presented on Figure 45 and Figure 46.

LTE

The explosive increase of data generated by the modern society, and the need to collect and share that data with mobility led 3GPP to design a new standard for wireless communication of high-speed data, it was born the Long Term Evolution (LTE).

LTE is based on GSM/EDGE and UMTS/HSPA technologies, presenting increased capacity and speed thanks to a different radio interface and core network improvements.
The fundamental features of LTE are:

- OFDMA based;
- Frequency-Division Duplexing (FDD) and Time-Division Duplexing TDD support;
- High data rates;
- Low Latency;
- Flexible spectrum support;
- FDD and TDD interworking;
- Seamless 3G interworking [142].

3GPP release 10 defined LTE-Advance (LTE-A) as the first standard that met the 4G requirements, and that will increase the capacities of mobile broadband for the next years.

Figure 33 – LTE and LTE-A official logos [143]

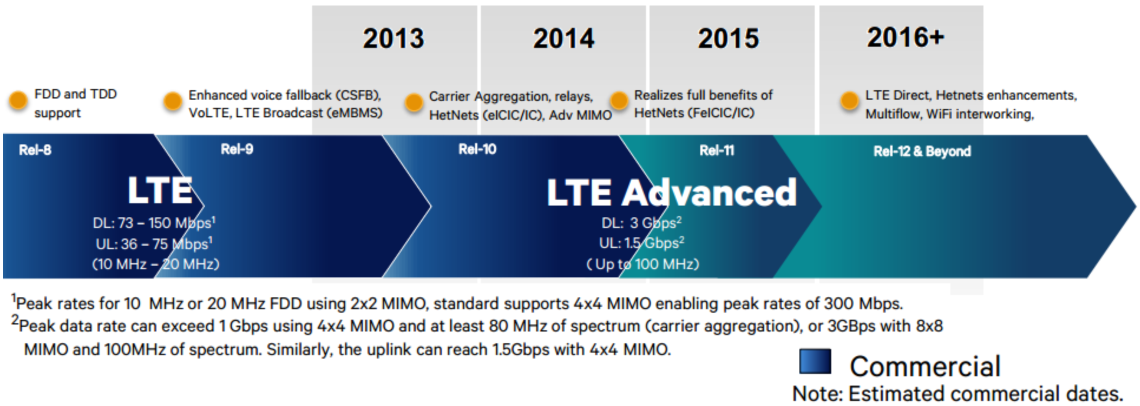


Figure 34 – LTE evolution [142]

The LTE architecture is presented on Figure 47.

WiMAX

WiMAX, or World Wide Interoperability for Microwave Access, is a last-mile mobile access technology, which is a part of 3G and 4G technology families.

The latest developments on the technology (IEEE 802.16m Air Interface standard) provided improvements on latency, power efficiency, interoperability with other technologies (HSPA/LTE) and transmission rates [144]. Using MIMO techniques (similar to LTE) it is possible to achieve up to 100Mbps or 1Gbps for low mobility users [144].

Figure 35 – WiMAX certified product logo [145]

IEEE 802.16m also defined several QoS classes, as seen on Table 8.

Class	Data Rate
Low Multimedia	up to 144 kbps
Medium Multimedia	up to 2 Mbps
High Multimedia	up to 30 Mbps
Super High Multimedia	up to 100 Mbps, or up to 1 Gbps for low mobility users

Table 8 – IEEE 802.16m QoS classes [144]

Satellite

Satellite powered data transmission are supported by the Digital Video Broadcasting - Return Channel via Satellite (DVB-RCS) specification. Worldwide use of DVB specification has enabled broadband access to regions away from other broadband accesses [136]. The data rates provided by the DVB-RCS system are dependent on the link budget, and can go up to 20 Mbps (downlink) and 5 Mbps (uplink) [136].

In 2011 the Second Generation DVB Interactive Satellite System (DVB-RCS 2) was submitted to standardization (subsequently approved). This evolution "is natively developed for IP services and is specified for interactive satellite services in several market segments, ranging from professional to consumer" [146].

Figure 36 – DVB-RCS2 official logo [146]

FSO

Free Space Optics (FSO) is a long developed technology by NASA and the United States Department of Defense [147]. It consists on a point-to-point optical link, a laser beam transmitted directly in the atmosphere. Can be implemented as earth-to-earth or earth-to-space communication. More recent implementations based on high powered LEDs are also available [87]. It is currently in mass deploying through the world for military and Mobile Network Operators (MNOs) applications, and it is becoming a viable solution for last mile access. It current transfer rates from 100Mbps up to 2.5Gbps, and it is believed that this capacity can be extended as far as 160Gbps [148].

Some of the fundament features of FSO are:

- Protocol independent;
- Can be implemented as part of networks with multiple architectures (cellular, mesh or point-to-point);
- Can be installed in roof-tops, through office windows or in the interior of the buildings;
- Fast deployment: Once a suitable line of sight is identified, a point-to-point link can, typically, be operational in one hour or less;
- Easily re-deployable;
- Outdoor availability can be as good as 99.9% or more if implemented with a backup system (typically RF link or xDSL);
- FSO's links don't need to be constantly aligned, and building motion is not sufficient to take it down;
- High Security;
- Eye Safe;
- Outdoor range can go as far as 7.7 km in light rain, and 3.2 km in thin fog [148], [149].

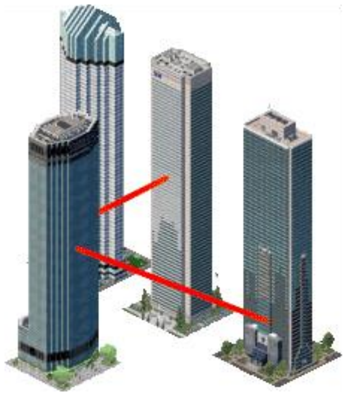


Figure 37 – Example of FSO implementation on high-rise buildings [150]

3.2.3 Summary

	Technology	Applications	Max. Bandwidth	Range	Security	Interference Immunity
Wired	ADSL	Broadband	▲1 Mbps ▼52 Mbps	up to 5.5 km	high	good
Wired	VDSL	Broadband	10 Mbps	up to 1.3 km	high	good
Wired	HFC	Broadband	▲120 Mbps ▼160 Mbps	up to dozens of km	high	good
Wired	FTTx	Broadband	Gbps	architecture dependent	very high	good
Wireless	GSM	Voice/Data	40 kbps	WAN	high	good
Wireless	HSPA+	Broadband	▲23 Mbps ▼168 Mbps	WAN	high	good
Wireless	HSPA+ A [7]	Broadband	▲69 Mbps ▼336 Mbps	WAN	high	good
Wireless	LTE	Broadband	▲75 Mbps ▼150 Mbps	WAN	high	good
Wireless	LTE-A [8]	Broadband	▲1.5 Gbps ▼3 Gbps	WAN	high	good
Wireless	WiMAX	Broadband	> 100 Mbps [9]	WAN	high	good
Wireless	Satellite	Broadband	▲5 Mbps ▼20 Mbps	worldwide	high	good
Wireless	FSO	Broadband	2.5 Gbps	up to 7.7 km	very high	low

Table 9 – Summary of access technologies [129]-[132], [136], [139]-[142], [144], [148]

[7] In development; it should not be commercially deployed before 2016 [141].
[8] Current LTE-A deployments do not yet support such bandwidths values.
[9] Up to 1 Gbps for low mobility users [144].

Mobile technologies such as UMTS, LTE and WiMAX are nowadays used to connect multiple devices, without the existence of a customer network. Thus it is useful to make the comparison between these technologies and the wireless technologies presented in Subsection 3.1.2 (Table 10).

Technology	*Max. Bandwidth*	*Range*	*Security*	*Interference Immunity*	*Global Availability*	*Cost[10] of Deployment*	*Cost[11] of Connectivity*
Wi-Fi	1,3 Gbps	up to 250 m	medium	low	yes	low	inexistent
ZigBee	250 kbps	up to 100 m	low	low	no	low	inexistent
W M-Bus	225 kbps	up to 5 km	low	low	no	low	inexistent
GSM	40 kbps	WAN	high	high	yes	inexistent	low
HSPA	▲23 Mbps ▼168 Mbps	WAN	high	high	yes	inexistent	high
HSPA+ A[12]	▲69 Mbps ▼336 Mbps	WAN	high	high	yes	inexistent	high
LTE	▲75 Mbps ▼150 Mbps	WAN	high	high	in development	inexistent	high
LTE-A[13]	▲1.5 Gbps ▼3 Gbps	WAN	high	high	in development	inexistent	high
WiMAX	> 100 Mbps[14]	WAN	high	high	no	inexistent	high

Table 10 – Comparison of mobile technologies [85], [123], [124], [139]-[142], [144]

3.2.4 Heterogeneous Networks

The increase on the number of broadband subscribers (10 fold by 2016 – Ericsson [151]), on the number of connected devices (50 billion by 2020 – Ericsson [151]), and on the customer demand for seamless connectivity everywhere and anytime is driving the development of a new type of access networks: Heterogeneous Networks (HetNets). HetNets are no more than the interconnectivity and interoperability of several different technologies in order to offer a seamless user experience. The development of HetNets is fundamental to MNOs success, enabling them to provide high QoE, especially in terms of coverage, data throughput, and low latency [152].

HetNets are mainly created by the deployment of small-cells (micro-cells, indoor/femto-cells and Wi-Fi hotspots) as a complement to the already existing serving macro cells. Their main characteristic is flexibility, thus there are multiple technology, size, and architecture alternatives. This also implies that the several accesses comprising the HetNets don't need to be owned by the same entity, in fact they can be owned by different MNOs, utilities, and even individual or enterprise customers that can benefit from some kind of compensation in exchange for providing public offloading capacity. Figure 38 depicts this situation.

[10] The cost for the subscriber.
[11] The cost for the subscriber.
[12] In development; it should not be commercially deployed before 2016 [141].
[13] Current LTE-A deployments do not yet support such bandwidths values.
[14] Up to 1 Gbps for low mobility users [144].

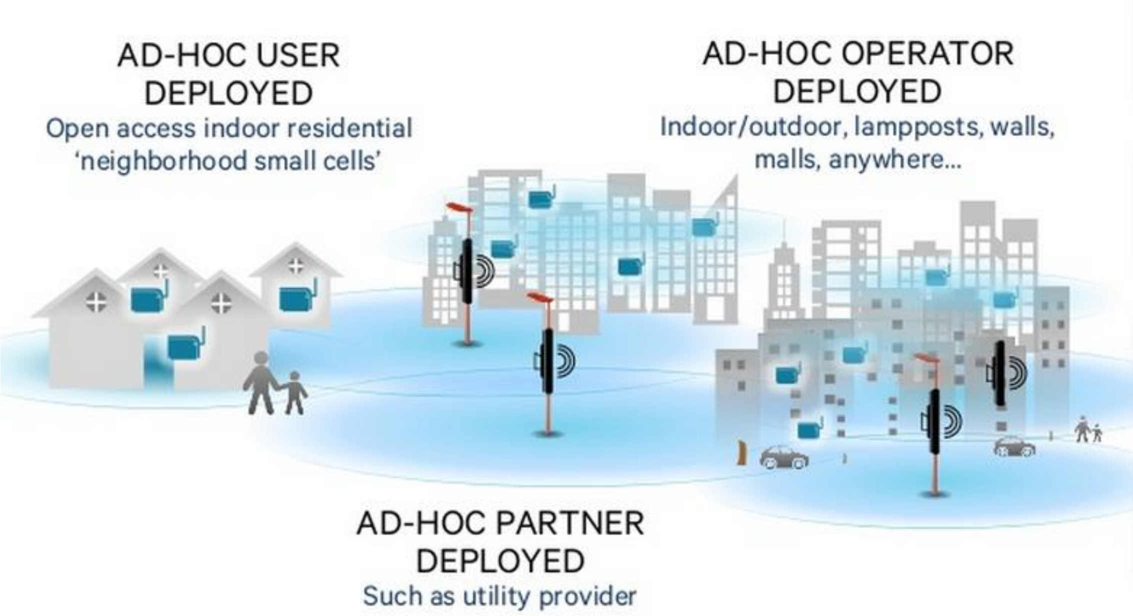


Figure 38 – Ad-Hoc small cells deployments [153]

There is a technology that is clearly on the front line of HetNets. This technology is Wi-Fi and it is being deployed by MNOs as a way to offload their 3GPP technologies based networks [154], [155].

Network performance is the principal driver of loyalty to mobile operators [156], thus it is fundamental ensuring that Wi-Fi does not threaten their customers' perceptions of network performance. While using Wi-Fi access, users can experience QoE degradation [155]. Figure 39 depicts the main reasons for QoE degradation.

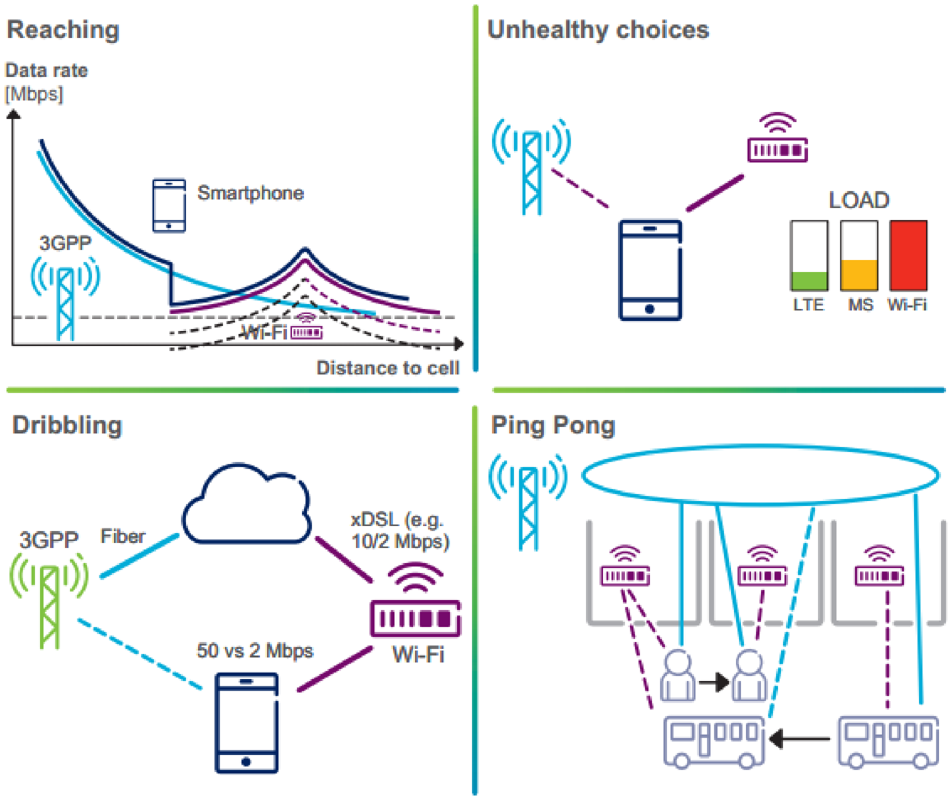


Figure 39 – QoE degradation from Wi-Fi's bad habits [155]

According to Ericsson [155], such reasons are:
- "**"Reaching"** refers to the Wi-Fi enabled device's habit of attaching to Wi-Fi when it's within range, even if that connection won't deliver a good user experience. Reaching not only affects the devices that connect from too far away, it also saps resources from the Wi-Fi AP, resulting in a bad user experience even for user devices closer to the Access Point (AP);
- **"Unhealthy choices"** refers to the Wi-Fi-enabled device's common practice of attaching to a loaded Wi-Fi network, even while 3G/4G remain more lightly loaded;
- **"Dribbling"** refers to the fact that, in many cases, Wi-Fi employs a lower speed backhaul to enable cost-efficient network deployment. in such scenarios, the user experience – coming from a higher capacity 3GPP network – will be degraded in the move to Wi-Fi.
- **"Ping-pong"** refers to the Wi-Fi-enabled device's habit of connecting to Wi-Fi, even when only temporary access is available. This can apply to pedestrians on the move or commuters on a bus".

In order to solve such problems a new level of Wi-Fi and 3GPP integration is required. Such integration is outside the scope of this work thus it won't be studied any further. Relevant documentation on the subject is available on the following references: [155], [157], and [158].

3.3 Core Network and Architectures

3.3.1 Wired Technologies

xDSL

The xDSL architecture is presented on Figure 40.

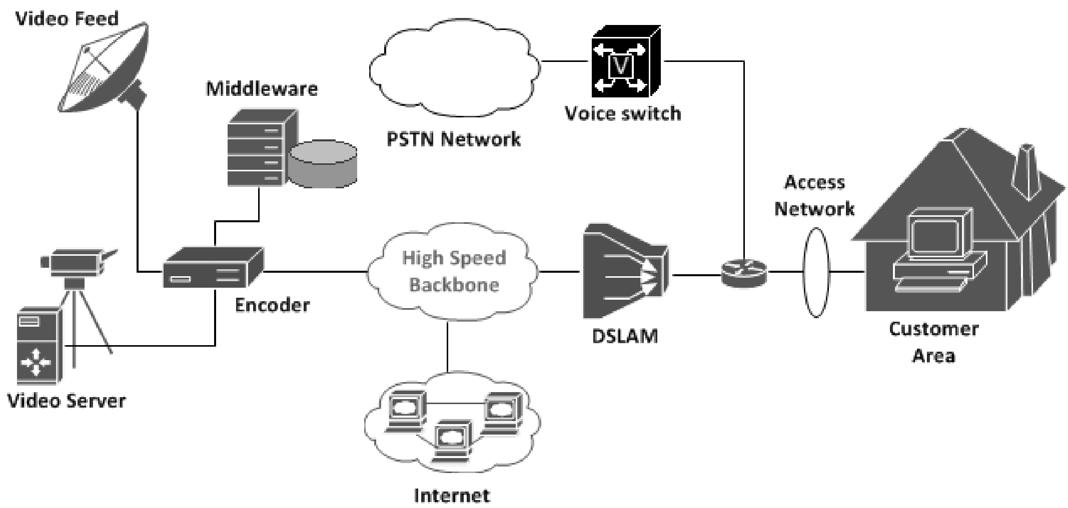


Figure 40 – xDSL architecture (based on [136])

Digital Subscriber Line Access Multiplexer (DSLAM): A network device, commonly located at telecommunication operators telephone exchanges. It separates the voice-frequency signals from the high-speed data traffic, and controls and routes xDSL traffic between the customer equipment and the network service provider network [152].

HFC

The HFC architecture is presented on Figure 41.

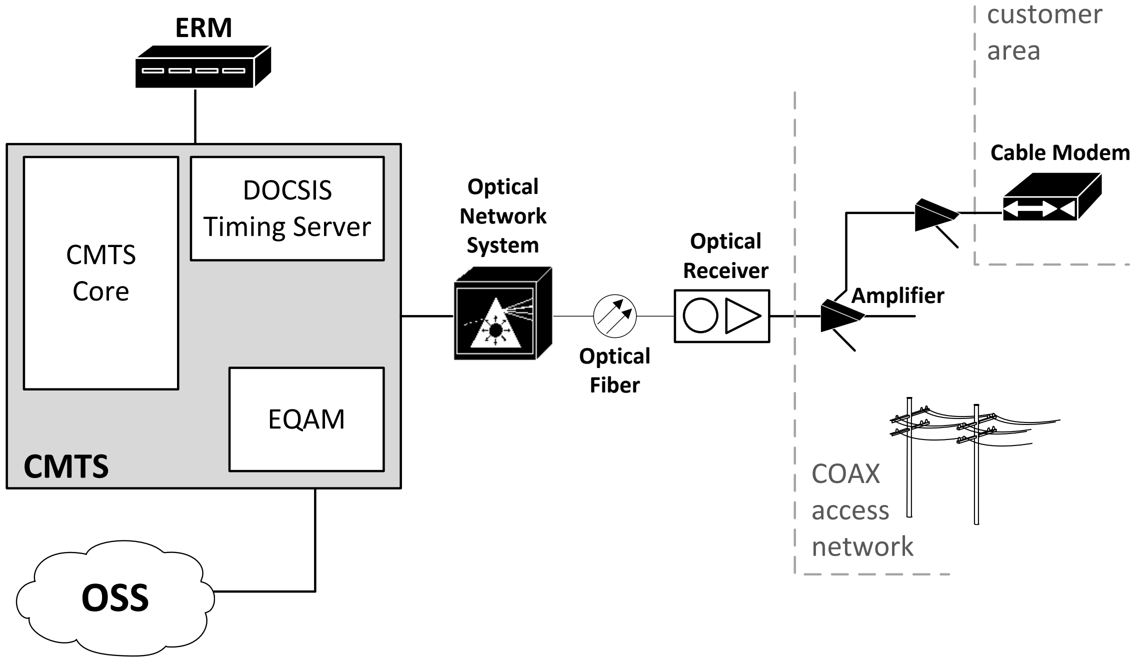


Figure 41 – HFC architecture (based on [134], [160])

Cable Modem Termination System (CMTS): Typically located at cable companies head ends, it is used to provide high speed data services. Many of its functions are similar to the ones provided by the DSLAM in an xDSL system.

EQAM: Edge Quadrature Amplitude Modulation.

Edge Resource Manager (ERM): Allows advanced services to share network resources. Reallocates bandwidth without rewiring and reconfiguring. Provides operators with access to processing statistics, configuration information, complete session lists and diagnostics [161].

Operations Support System (OSS): Hardware and (mostly) software applications that "support back-office activities which operate a telecommunications operator network, provision and maintain customer services" [162]. For further detail on OSS, please refer to www.tmforum.org.

3.3.2 Wireless Technologies

GSM

A GSM network is comprised by four main elements:

- **Mobile Station (MS):** The part of network that connects the subscriber to the network;
- **Base Station System (BSS):** Deals with the RF connection between the MS and the communication system;
- **Network Switching System (NSS):** Interconnects the GSM network with the public network (PSTN), manages the data bases and processes the information through the interfaces;
- **Operations and Maintenance System (OMS):** Responsible for the operation, administration and maintenance of the network [163].

A group of such components – the GSM network of one operator – is called Public Land Mobile Network (PLMN).

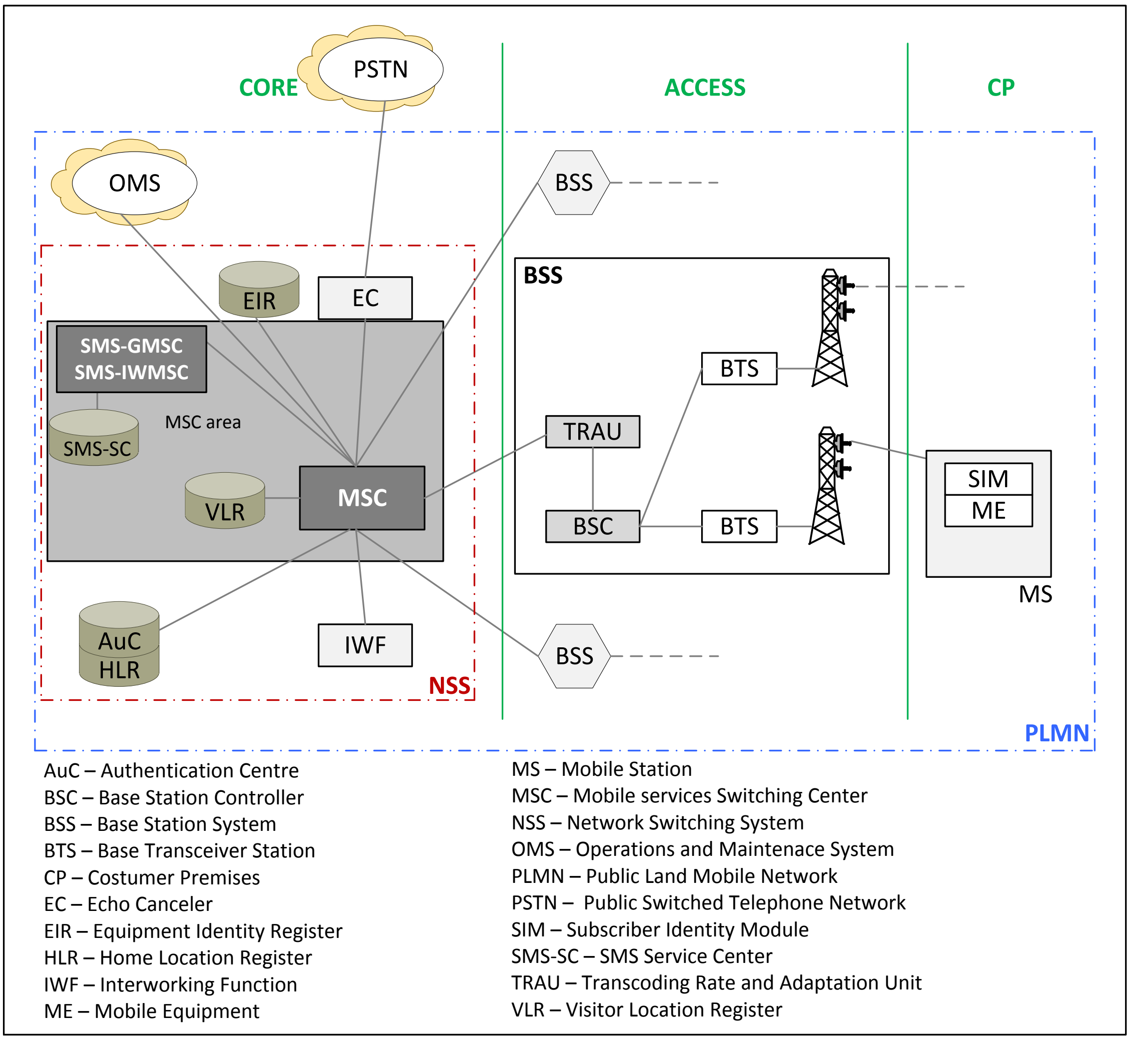


Figure 42 – GSM architecture (based on [163], [164])

The **Mobile Station (MS)** is comprised by the Mobile Equipment (ME) and by the Subscriber Identity Module (SIM). The ME interconnects the subscriber to the GSM network. Each ME has a unique International Mobile Equipment Identity (IMEI) that allows the system to identify the ME. The IMEIs are registered on the Equipment Identity Register (EIR) [163]. In GSM systems the identity of the device and of the subscriber is separated, the SIM determines the directory numbers and the calls billed to a subscriber. The SIM is a database on the user end that communicates directly with the VLR and indirectly with the HLR [164]. A ME can only successfully connect to a PLMN if the subscriber has a valid SIM card. The only exception is for emergency calls, for which the MS can connect to the network even without a SIM card.

The Access Network is comprised by the **Base Station System (BSS)**, which contains several elements and interfaces. BSS main elements are:

- **Base Station Controller (BSC):** Controls a group of BTSs [163], [164].

- **Base Transceiver Station (BTS):** Is made of RF hardware and antennas, which establish the radio link with the MS [163], [164].

- **Transcoding Rate and Adaptation Unit (TRAU):** Converts the voice signals that are received from the MSC (64 Kbits/s) to the GSM specifications, allowing for an optimal use of the frequency resources, maximizing the use of the available bandwidth, and consequently maximizing the number of calls that can be established with that same bandwidth [163], [164].

The Core Network (CN) consists of many elements and interfaces, the main elements are:

- **Authentication Centre (AuC):** is the entity responsibly for authentication and encryption functions. Such functions are performed, simultaneous, in the AuC and in the MS. Their objective is to provide security to the system (e.g. avoiding cloning of MSs). The AuC communicates solely with its associated HLR (H-interface), and is typically installed in the HLR hardware [163], [165].

- **Echo Canceler (EC):** suppresses the echo present in the MSC – PSTN connections [163].

- **Equipment Identity Register (EIR):** a centralized data base containing the International Mobile Equipment Identity (IMEI). This data base is comprised by three lists: White list – authorized IMEIs; Black list – non-authorized IMEIs (e.g. stolen or cloned MS); and Grey list – IMEIs correspondent to MSs with some kind of problem, but still authorized to access the network (e.g. hardware or software failure or MS on authorized maintenance). The EIR is remotely accessed by the MSC, which translates the IMEI on the EIR logic address [163].

- **Home Location Register (HLR):** this register is responsible for administrating and control the data base of local subscribers. It controls and maintains any change on a subscriber profile. The HLR data profiles are remotely accessed by the MSC and VLR. The fundamental profile data are: International Mobile Subscriber Identity (IMSI); current subscriber location in the VLR; supplementary services associated to subscriber (and related information); subscriber status (attached/non-attached); and authentication key [163].

- **Interworking Function (IWF):** responsible for providing the GSM interface with other public and private networks. Its fundamental tasks are data rate adaptation and protocol translation [163], [165].

- **Mobile services Switching Centre (MSC):** is the central element of the Network Switching System (NSS), performing switching and signaling functions (CS) for mobile stations [165]. Its fundamental tasks are: call processing (establishment and termination of calls, handover between BSSs and between MSCs); operation, maintenance and supervision (data base management and traffic data measurement); inter-functioning (management of interfaces between GSM and other networks, e.g. PSTN or ISDN); and charging [163].

- **Operations and Maintenance System (OMS):** responsible for the administration, operation, maintenance and centralized remote supervision of the elements of the GSM network [163].

- **Service Centre (SC):** Function responsible for the relaying and store-and-forwarding of a short message between a Short Message Entity[15] (SME) and an MS. The SC is not a part of the GSM PLMN, however MSC and SC may be integrated [166].

- **Gateway MSC for Short Message Service (SMS-GMSC):** A function of an MSC capable of receiving a short message from an Service Centre (SC), interrogating an HLR for routing information and SMS info, and delivering the short message to the Visited Mobile Switching Centre (VMSC) of the recipient MS [165], [166].

- **Interworking MSC for Short Message Service (SMS-IWMSC):** A function of an MSC capable of receiving a short message from within the PLMN and submitting it to the recipient SC [165], [166].

- **Visitor Location Register (VLR):** in order to minimize load for HLR, it is responsible for the maintenance of a copy of the subscribers main profile data stored on the HLR, such as: MS state (e.g. available, busy, non-responsive); Location Area Identity (LAI); Temporary Mobile Subscriber Identity (TMSI); and Mobile Station Roaming Number (MSRN) [163], [165].

The GSM Interfaces provide interconnection of the different network elements, as well as the implementation of mobile services and applications between those same elements. The interconnection scheme of GSM interfaces is presented on Appendix A – GSM Interfaces.

The GSM functional structure is based on the OSI model [163], depicted on Figure 43.

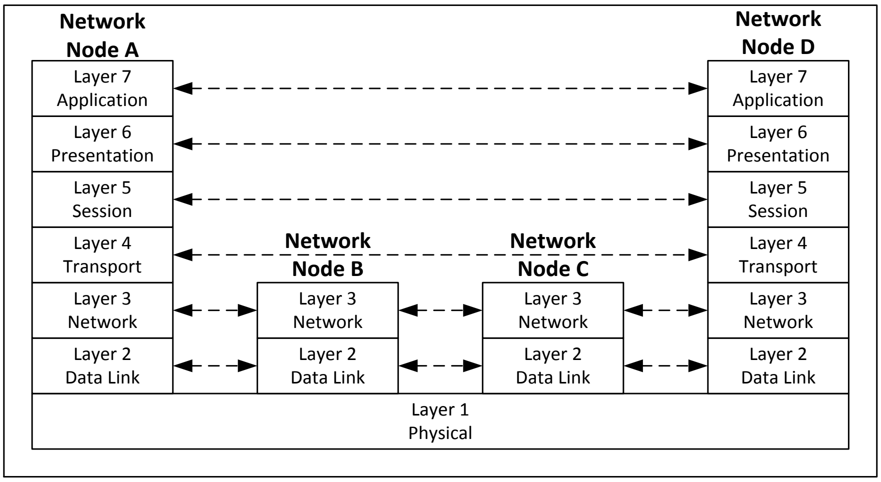


Figure 43 – OSI model architecture, with intermediary nodes (based on [164])

[15] SME: an entity which may send or receive Short Messages. The SME may be located in a fixed network, an MS, or an SC [157].

Each layer is comprised by protocols that use services offered by inferior layers, and that offer services to superior layer's protocols. Layers 1 to 3 comprise the information transfer functions from one network node to another. Layers 4 to 7 define functions of end-to-end communication.

Layer 1 Physical	Defines the functional, mechanical and electrical characteristics of the physical media of transmission and of the interface responsible for the data transport.
Layer 2 Data Link	Is in charge of transmission errors correction, improving the quality of service offered by the physical layer.
Layer 3 Network	Routes the data across the network.
Layer 4 Transport	Enables end-to-end communication, allowing the user to be abstracted of any knowledge about the transmission characteristics.
Layer 5 Session	Enables the data transmission between users by organizing and synchronizing the data exchange. Offers to layer 6 services of session establishing and session ending, special data exchange, connection synchronization, etc.
Layer 6 Presentation	Transforms the data formatting between the layer 7 and the rest of the system.
Layer 7 Application	Specifies the communication nature as well as the meaning of the exchanged information.

Table 11 – OSI layers description (based on [163])

Messages are exchanged between superior and inferior layers, and the opposite. However from a logical point of view each layer communicates directly with the same layer in the other node/end. While transmitting each protocol of each layer will add a header on the message. On the reception the opposite process will happen. This process will repeat itself for all layers down to layer 2, where it will be added an error detection and correction field to the message [163]. This process is exemplified on Figure 44.

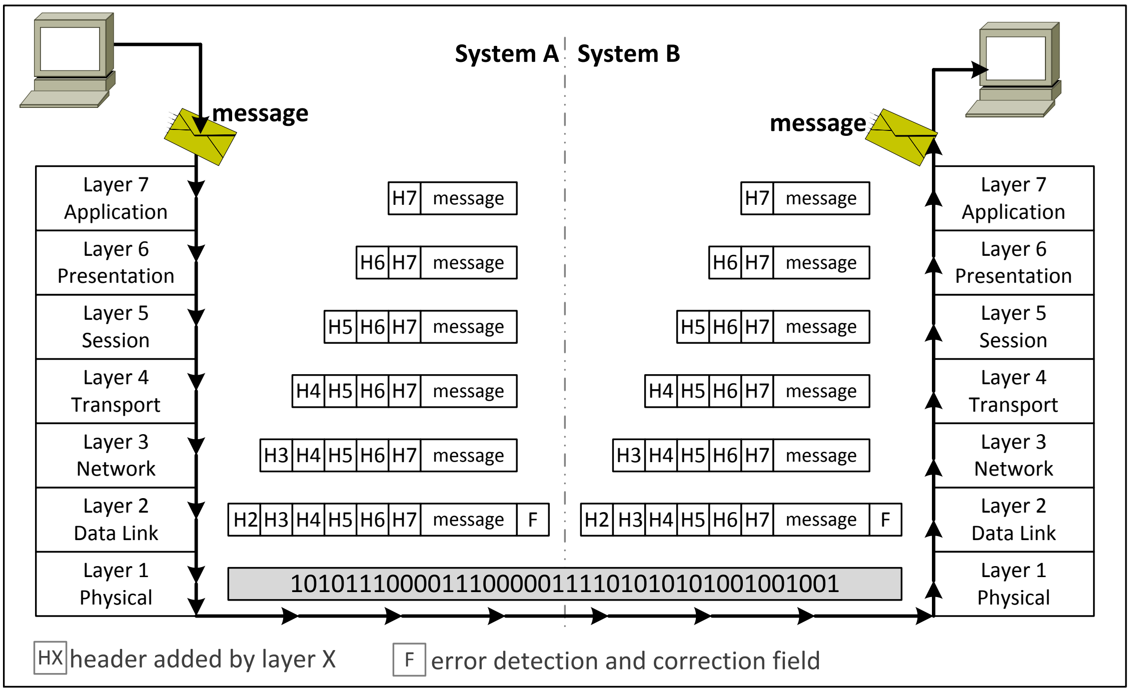


Figure 44 – Relationship between layers of OSI model (based on [163])

UMTS

Figure 45 and Figure 46 depict UMTS core and access architecture, whit distinction of Circuit and Packet Switched domains.

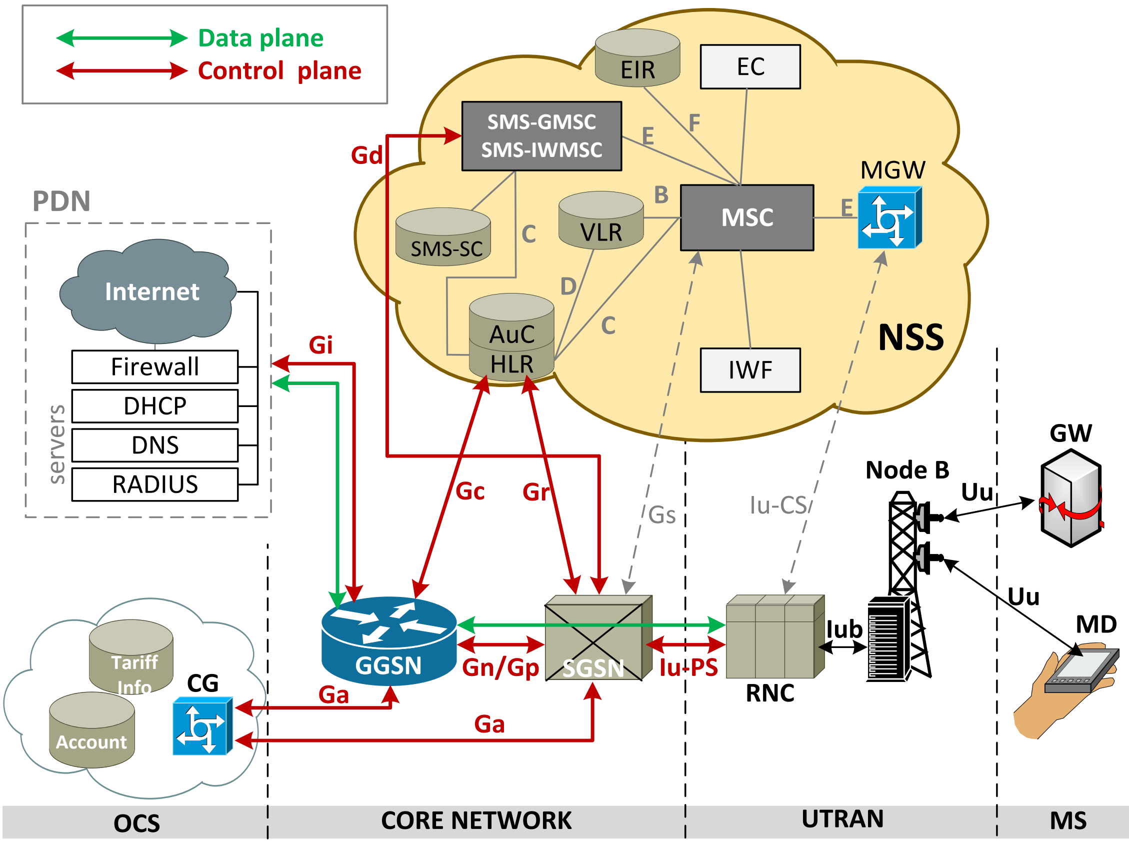


Figure 45 – UMTS architecture with emphasis on data plane (based on [163], [166]-[169])

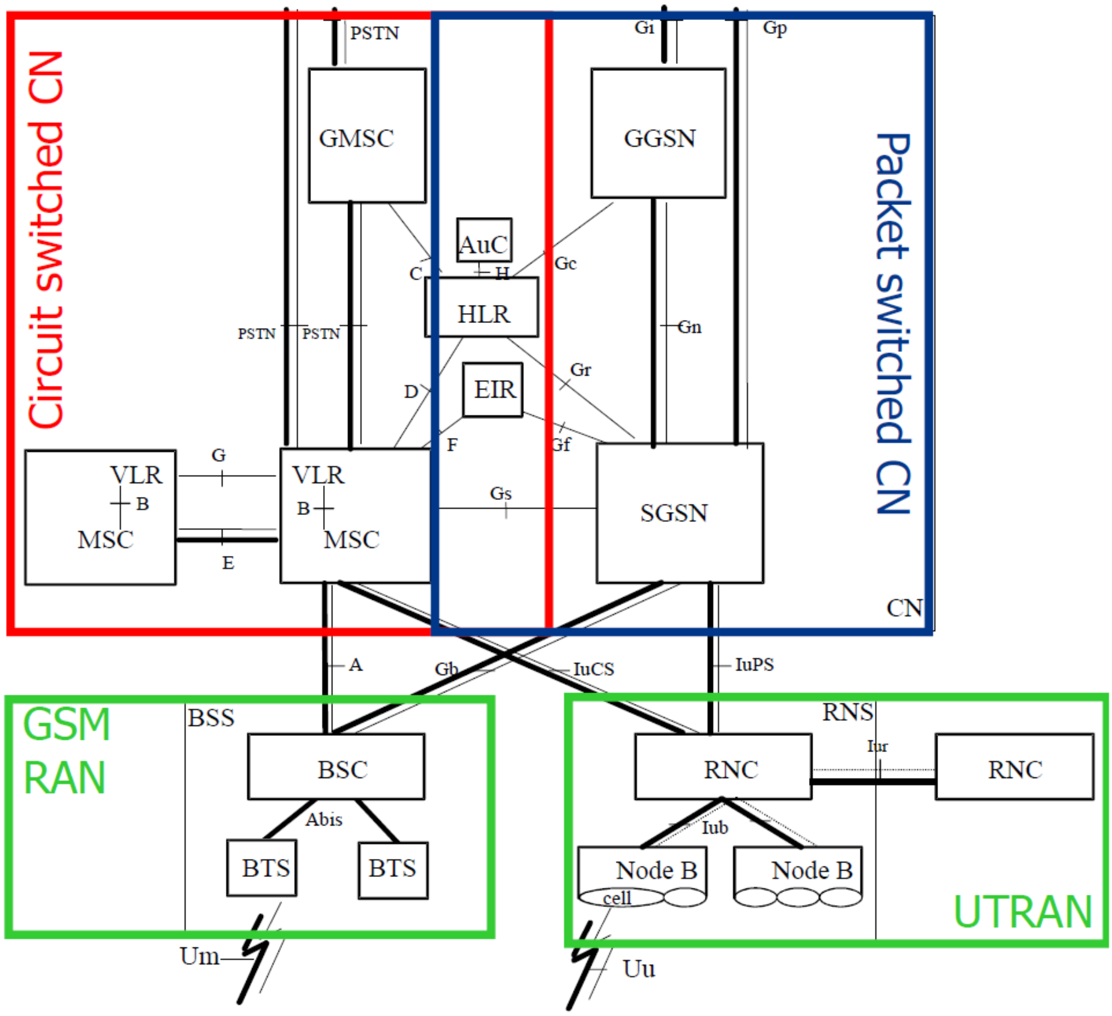


Figure 46 – Basic UMTS network architecture [165]

The Mobile Station (MS) is comprised by the Mobile Device (MD) or Gateway (GW) and the UMTS Subscriber Identity Module (USIM), a smartcard that stores the subscriber identity, authentication and encryption keys, and some subscription information that is necessary at the terminal. The USIM also implements authentication algorithms [138].

The UTRAN is a set of Radio Network Subsystems (RNS) that comprise the UMTS access network. A RNS comprises one Radio Network Controller (RNC) and several NodeBs:

- **NodeB:** controls the data flow between Uu and Iub interfaces. The NodeB terminates the physical layer, by extracting the MAC protocol data units and transporting them to the RNC. It is also a part of the radio resource management [138], [165].

- **RNC:** controls the radio resources of its own domain. Its basic functions are:
  - Manage radio resources: the RNC contains an admission control module that decides whether a voice call/data packet session is admitted and the radio resources that need to be allocated;
  - Process radio signaling: the RNC processes all the radio signaling, originated both from the UE or the Core Network;
  - Perform call setup and termination: the RNC exchanges signaling with the UE and the NodeB in order to set up bearer service during a communication session establishment process. It also deals with the reverse process;

- Perform soft, hard and intersystem handoffs: the RNC decides (using radio signal measurements) when to command the UE to establish handovers;
- Provide Operation, Administrative, and Maintenance (OA&M) capabilities: the RNC sends keep alive messages across the various interfaces that connects it to other RNCs, NodeBs, and other access points for OA&M purposes [138], [165].

The Core Network (CN) consists of many elements and interfaces, the main elements are (elements inherited from GSM are presented on 3.3.2 – GSM):

- **Charging Gateway (CG):** is the billing unit for Packet Switched (PS) domain. It collects, merges, filters and stores the Call Detail Records (CDR) – and other charging information from the GGSN and the SGSN. It communicates with the billing center, and subsequently transfers the sorted CDR to the billing center. The CG can be coupled with the SGSN [163], [165], [169].

- **Gateway GPRS Support Node (GGSN):** is a gateway between UMTS PS/GPRS network and external data networks (e.g. Internet). The GGSN executes a number of functions such as, routing and data encapsulation between a MS and external data networks, security control, network access control and network management. Considering the UMTS PS/GPRS context, a MS selects a GGSN as its routing device between itself and the external network in the activation process of a PDP[16] context in which Access Point Name (APN) defines the access point to destination data network. Considering the external network, the GGSN is a router capable of address all MS IPs in UMTS PS/GPRS network [138], [165], [169].
  The GGSN supports functions such as:
    - GPRS Tunneling Protocol (GTP);
    - Lawful Interception[17];
    - Route resolve through HLR interaction;
    - Static/Dynamic IP address assignment;
    - Packet data routing;
    - Authentication, Authorization, and Accounting (AAA);
    - Traffic management and control (e.g. reject of irrelevant packets and congestion control);
    - Service performance measurements [138], [165].

- **Home Location Register (HLR):** As previously described on 3.3.2 – GSM, this register is responsible for administrating and control the data base of local subscribers. It controls and maintains any change on a subscriber profile. The HLR data profiles are remotely accessed by the MSC and VLR. The fundamental profile data are: International Mobile Subscriber Identity (IMSI); current subscriber location in the VLR; supplementary services associated to subscriber (and related information); subscriber status (attached/non-attached); and authentication key [163]. In 3G networks, the HLR is also responsible for storage of zero or more Packet Data Protocol (PDP) addresses and for providing permission for the GGSN to dynamically allocate PDP addresses for a subscriber [165].

- **Serving GPRS Support Node (SGSN):** the SGSN is responsible for the data packet delivery from and to MSs within its serving area. Some of its fundamental tasks are: packet routing and transfer; mobility management (attach/detach and location

[16] PDP: network protocol used by an external packet data network interfacing to GPRS.
[17] Lawful Interception: the process by which law enforcement agencies conduct electronic surveillance of circuit and packet-mode communications as authorized by judicial or administrative order [170].

management); logical link management; authentication; and charging functions [165], [169]. Thus, it supports functions of:

- MS registration with its HLR;
- MS authentication;
- Authorization and admission control of packet services for a subscriber;
- Packet session establishment, maintenance, and termination;
- PDP context activations;
- Packet switching/routing for UMTS/GPRS packet services;
- Maintenance of a database of subscribers currently being served;
- GPRS attach processes;
- GPRS Tunneling Protocol (GTP);
- Collection of charging data [138], [165].

Clusters of SGSNs can be connected to regional GGSNs. A trade-off must be reached between the cost of equipments and the cost of transport in order to define how many clusters should be implemented and how they should be configured [138]. A paper published by G. Atkinson, S. Strickland, and M. Chuah [171] presents a study on this subject, where the costs of transport over long distances and the cost of owning several GGSNs was compared. The paper conclusion is that is more expensive to have a centralized network due to the high cost related with the transport of information.

For a more detailed description on the architecture and elements of 3G 3GPP networks, Appendix B and 3GPP TS 23.002 [172] should be consulted. For a brief description of UMTS protocol architecture please refer to Appendix C.

LTE

LTE is an evolution of the UMTS system, therefore its radio access network is known as Evolved UTRAN (E-UTRAN) and the non-radio part of the network is known as "System Architecture Evolution", which includes the Evolved Packet Core (EPC) [172].

Figure 47 depicts the LTE functional architecture.

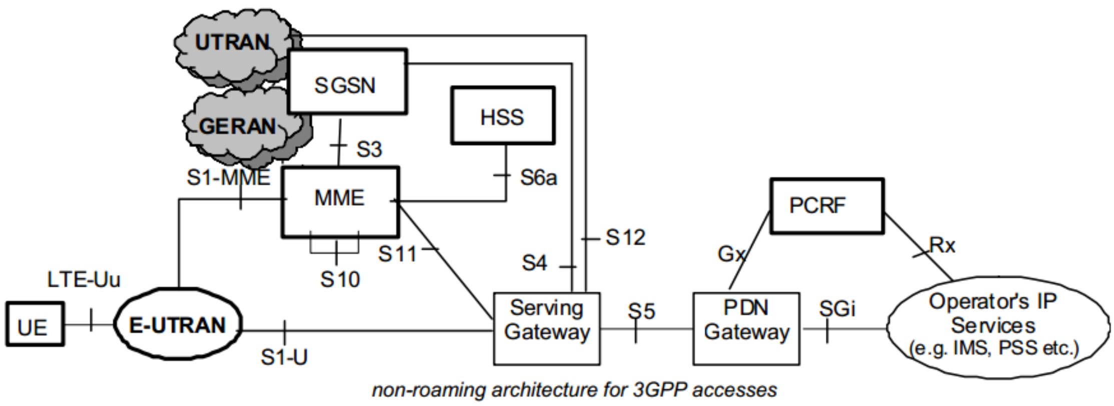


Figure 47 – LTE architecture [173]

The LTE access network (E-UTRAN) is comprised by a single type of node, the enhanced NodeB (eNodeB) that connects the User Equipment (UE) to the network. eNodeBs may be interconnected with each other by X2 interfaces, that support enhanced mobility, inter-cell interference management, and SON functionalities [173].

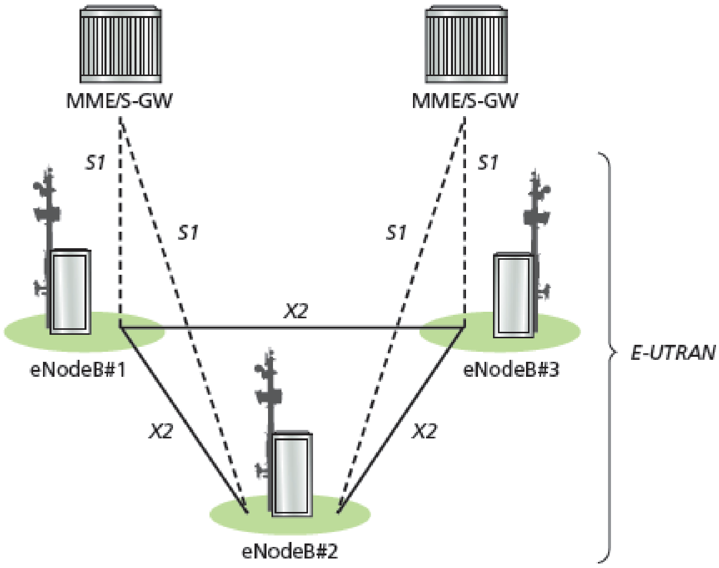


Figure 48 – Overall E-UTRAN architecture [174]

Regarding the LTE functional split, the **eNB** hosts the following functions:

- Radio Resource Management functions;
- Measurement and measurement reporting configuration for mobility and scheduling;
- AS security;
- IP header compression and encryption of user data stream;
- Selection of an MME at UE attachment when no routing to an MME can be determined from the information provided by the UE;
- Routing of User Plane data towards the Serving Gateway;
- Scheduling and transmission of paging messages (originated from the MME);
- Scheduling and transmission of broadcast information (originated from the MME or OA&M) [173].

The **(Mobility Management Entity) MME** hosts the following functions:

- NAS signaling;
- NAS signaling security;
- AS security control;
- Inter CN node signaling for mobility between 3GPP access networks;
- Tracking Area list management;
- PDN GW and Serving GW selection;
- MME selection for handovers with MME change;
- SGSN selection for handovers to 2G or 3G 3GPP access networks;
- Roaming;
- Authentication;
- Bearer management functions including dedicated bearer establishment;
- UE reachability in idle state (including control and execution of paging retransmission) [173].

The **Serving Gateway (S-GW)** hosts the following functions:
- The local Mobility Anchor point for inter-eNB handover;
- Mobility anchoring forinter-3GPP mobility;
- E-UTRAN idle mode downlink packet buffering and initiation of network triggered service request procedure;
- Lawful Interception;
- Packet routing and forwarding;
- Transport level packet marking in the uplink and the downlink;
- Accounting on user and QCI granularity for inter-operator charging;
- UL and DL charging per UE, PDN, and QCI [173].

The **PDN Gateway (PDN-GW)** hosts the following functions:
- Per-user based packet filtering (by e.g. deep packet inspection);
- Lawful Interception;
- UE IP address allocation;
- Transport level packet marking in the downlink;
- UL and DL service level charging, gating and rate enforcement;
- Credit control for online charging [173].

The **Home Subscriber Server (HSS)** holds the dynamic information that keeps record of the MME identities whose subscribers are connected. It includes data for subscriber System Architecture Evolution (SAE) such as QoS profiles and roaming restrictions. It also holds the PDN information that allows subscribers to connect with PDNs. Additionally the HSS is fundamental to authentication and security since it is able to integrate the Authentication Center (AuC) [172].

The **Policy Control and Charging Rules Function (PCRF)** fundamental task is to control the Policy Control and Charging Enforcement Function (PCEF) functionalities [172].

3.4 Enabling M2M Communications

The inclusion of M2M devices and applications on everyday life will require new technological solutions and approaches to provide connectivity and services. This Section presents some of those solutions.

3.4.1 M2M Gateway

Many households are already connected to the serving network (or networks) through a Home Gateway (HG), which can be defined as “an intelligent network interface device located at the consumer premises [175]”.

The most common example of a HG is the set-top box (STB). Either connected via cable, fiber or xDSL this device is typically used to provide connectivity for triple play services. On the customer side, these devices can include multiple interfaces such as Ethernet, USB, HDMI, SCART, Bluetooth, Wi-Fi, Memory Cards, HPNA (phone line), among others. Another example of HGs are the DTV receivers, 3G/WiMAX/LTE receivers, and even a mobile phone when working as a relay node. There are also Road Gateways (RG), which provide connectivity and/or concentration of data. One example are the utility data concentrators, installed on public grounds, serving a street or a neighborhood. This kind of device can use technologies such as Zig-Bee [176] or Wi-Fi.

Extensive implementation of M2M communications may be a driver to new designs for HGs, supporting a broader number of technologies and interfaces, since M2M communications can be implemented through a wide range of technologies, as exemplified on Figure 49.

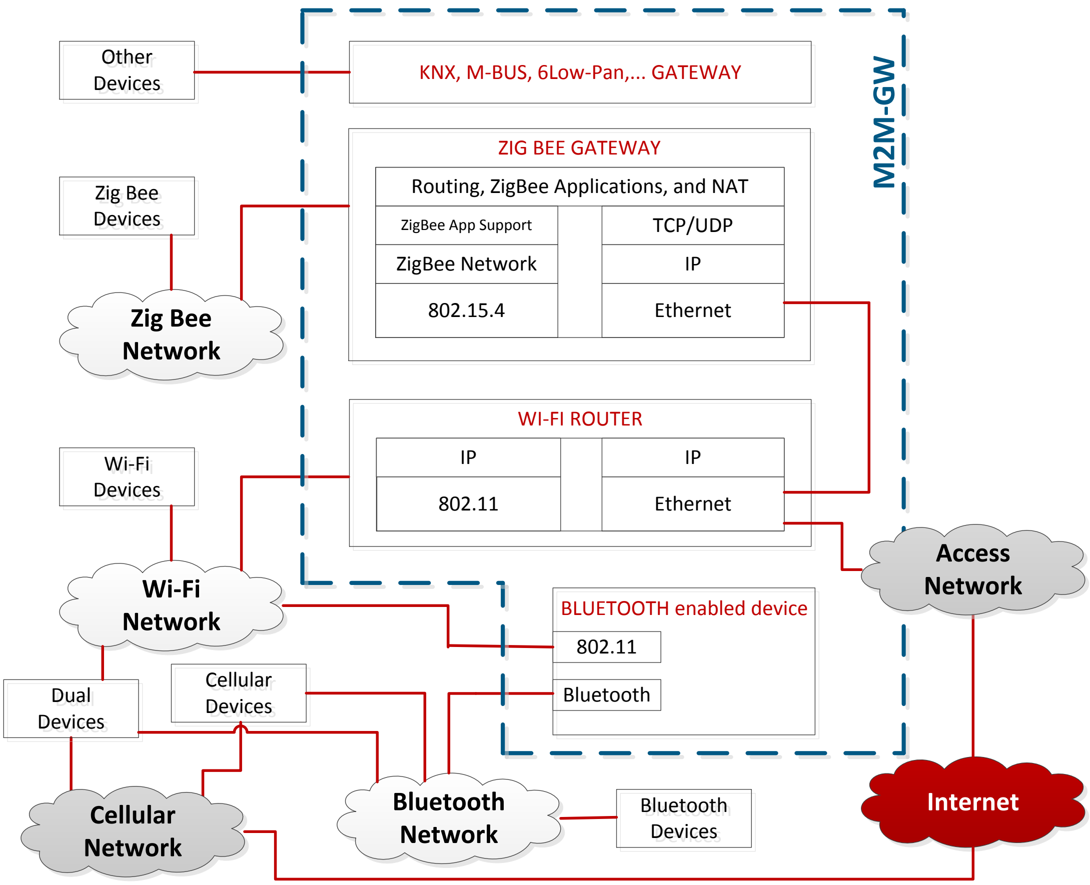


Figure 49 – Interconnection between different technologies on a domestic environment

The HG can perform bridging, routing, collection of information, and protocol and address translation between the external broadband network and the internal domestic network(s) [175].

HGs present key benefits both to users as to Service Providers (SPs), many of them derived from M2M solutions. Users benefit from:

- Internet sharing;
- In-home file/print sharing;
- VPN connectivity;
- Firewall and parental protection;
- VOIP;
- IP VOD;
- Remote heath-monitoring;
- Security services;
- Home automation;
- Remote meter reading [175].

Service Providers benefit from:

- New revenue opportunities due to new and differentiating services and applications;
- Capacity to deliver multiple services through a single equipment on the CP;
- Reduced “truck-roll” costs [175].

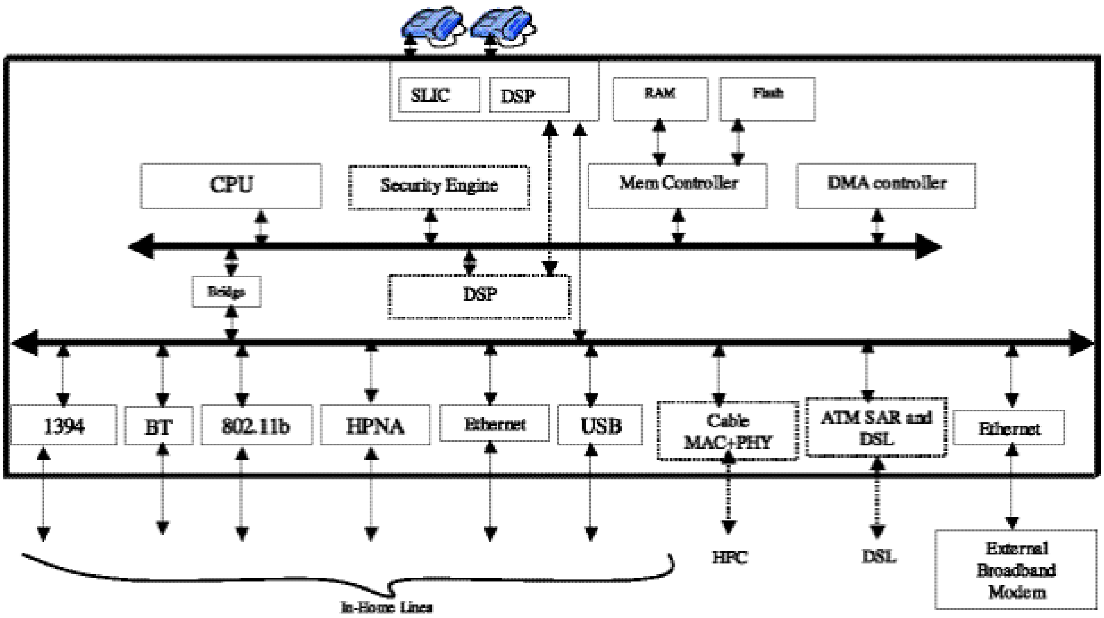


Figure 50 – HG hardware block diagram [175]

3.4.2 M2M Connectivity

It is important to understand if current connectivity solutions are a good fit for delivering M2M services, or if they need to be complemented, or even replaced.

Considering short-range connectivity such as LANs, HANs, or PANs, there can be multiple technologies at work, as presented on Section 3.1: 6LowPan, ANT, Bluetooth, COAX, DASH 7, Ethernet, FITH, ISA100.11a, KNX, M-BUS, MiWi, One-Net, PLC, W M-BUS, Wi-Fi, WirelessHEART, Z-Wave, Zig-Bee, among others. Obviously 2G, 3G, 4G, Satellite, and White Spaces can be used as an alternative for such technologies.

In terms of long-range connectivity the options are much more limited, connectivity can be provided by four fundamental types of technologies:

- Wired technologies;
- Satellite;
- Cellular networks (2G, 3G, 4G);
- White Spaces.

Wired technologies

Wired technologies such as COAX, FFTx or HFC are mature technologies that can provide connectivity for many M2M services, but will not be able to do so for many others since it fails to provide mobility and ubiquitous coverage.

Satellite

Satellite has ubiquitous coverage but is too expensive for provide M2M services in a cost-effective way, and capacity problems would probably rise.

Cellular Networks

Cellular networks have important characteristics to support M2M services, such as ubiquitous coverage (a reality for most of the habited territory in developed countries), already deployed infrastructure, and excellent support for mobility.

The deployment of M2M services supported by 2G technologies might be a risky move, since several MNOs through the world are planning to shut down their 2G networks in order to reallocate spectrum for other services [177], [178]. 3G, and especially 4G technologies are believed to offer longevity for support of M2M services [179], although they present higher power consumption when compared to 2G [180].

White Spaces

Another technological option is to provide M2M service connectivity with radio technology based on “White Spaces”, that is nothing more than the use of unregulated spectrum, or the use of spectrum freed from the analog TV shut down.
There are several companies already developing and deploying White Spaces based solutions, e.g. SIGFOX [181] and NEUL [182].

Summary

A brief comparative analysis for the previously presented options is depicted on Chart 5.

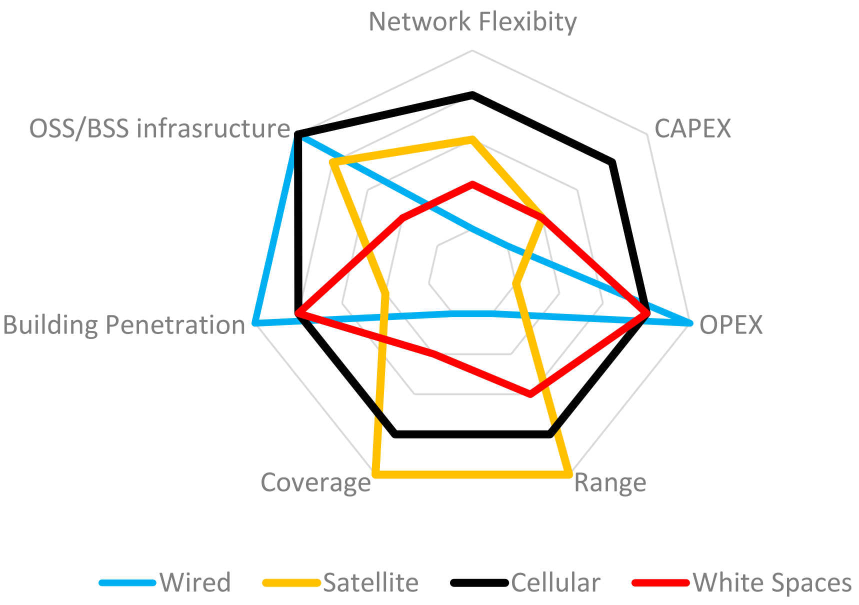


Chart 5 – Comparative of M2M connectivity options (based on assumptions and [183])

This analysis shows that wired solutions face many constraints in order to provide M2M connectivity, Satellite would not be a cost-effective solution, and Cellular technologies present better results than White Spaces when analyzing fundamental characteristics such as Coverage, Range, and an already deployed OSS/BSS infrastructure.
Additionally, MNOs want to take full advantage of their already developed mobile infrastructure, fundamentally the 3G network, which is widely available and has the ability to provide a broad range of M2M services. This considerations support the decision to study cellular network, and particularly the UMTS (3G) technologies.

3.5 Summary

Through this Chapter the state of the art on telecommunications networks was presented with distinction on customer network, access network and core network. Wired and wireless solutions were presented with special focus on the architectures of cellular networks.

Section 3.4 provides an analysis of the specific characteristics that gateways should present in order to provide M2M services, and compares several options for M2M connectivity, providing a basic analysis of which are the strengths and weaknesses of the several presented options. Such analysis resulted on the conclusion that cellular access networks present, overall, good characteristics to provide connectivity to M2M services, and MNOs are eager to fully explore their capabilities. Thus, justifying the option to study cellular networks, with particular focus on 3G technologies.

4. Impact of M2M Communications on 3G Wireless Systems

As seen previously it is expected a massive deployment of new M2M devices and services, able to communicate remotely, using the available networks but, possibly, with a higher pressure on the conventional mobile networks. Current core network architectures are designed mainly for Human-to-Human communication, and are not prepared to deal with the foreseeable increase on signaling traffic. Such increase is driven by growing M2M and smartphone signaling traffic (growing 50% faster than data traffic [18]). Furthermore the overall impacts on network capacity and performance, caused by adding a large number of M2M subscriptions to current networks are generally unknown, and may require new levels of scalability, in terms of subscription handling and in terms of mobility and resource management [18].

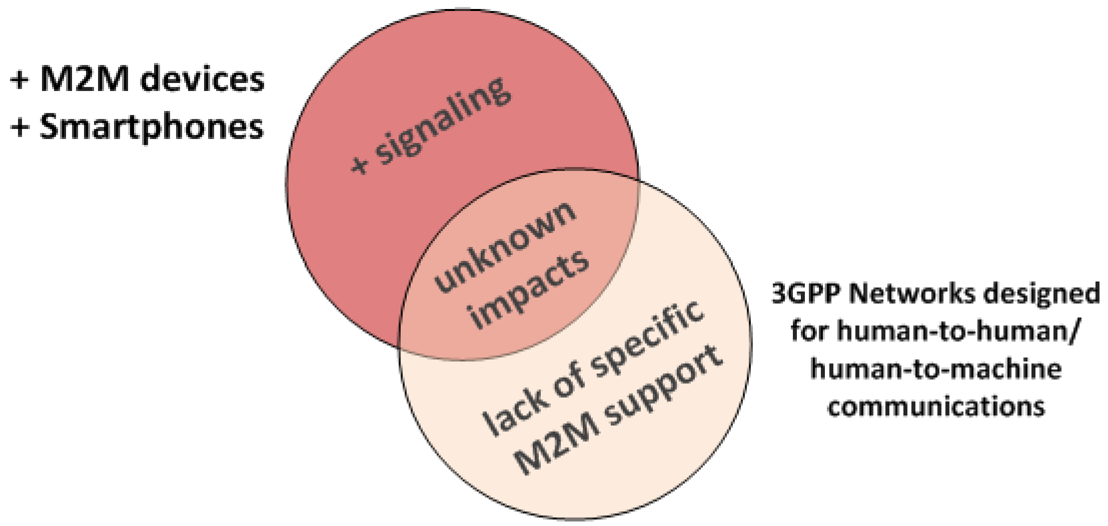


Figure 51 – 3GPP networks current paradigm

The capacity analysis of packet switched domain is challenging. Dynamic resource allocation for data users makes this analysis much more complicated than for the circuit switched domain [138]. Additionally, traffic processes and characteristics of packet data applications (packet traffic generates bursty traffic) are much more diverse and intricate than conventional voice data traffic [138]. On top of that, circuit switched systems have been studied and modeled for a long time, on the contrary, models for analyzing and designing packet switched systems are not so developed. Considering the explosion of M2M communications, and given that designing a system based on the worst case scenario is very costly [138], it is important to design models to describe the networks on large M2M deployment scenarios, allowing dimensioning based on average values with some sort of overload control strategies.

The following Sections present some possible factors, metrics and elements that can be of interest to a better understanding of the impact of M2M communications into 3G wireless systems. The work presented over these Sections, and subsequent Chapters (that present models intended to estimate such impact) is resultant of research work that was conducted over the course of this study, which it is believed to be of groundbreaking relevance for this research field, since, to the knowledge of the author and advisors, similar research applied to M2M communications are non-existent to date.

4.1 M2M related factors that might have impact on Core Network performance

4.1.1 M2M Traffic Temporal Regime

As presented over this document M2M communications present a broad number of applications and services. Thus is no surprise that the M2M traffic temporal regime can be very heterogeneous among different applications and usage scenarios, and consequently its characteristics are diverse and hard to predict. For example, smart-metering operations might take place a few times per hour, or per day, while health devices and cars might have to communicate in time windows of seconds or less.
This is a fundamental difference from conventional circuit switched traffic, adding up to the complexity of analysis of this issue.

4.1.2 Burstiness

Burstiness is a statistics concept that refers to the intermittent increases and decreases in the activity or frequency of an event, which is characteristic of most of M2M related traffic. Understanding the burstiness behavior of data traffic is fundamental, since burstiness introduces sudden peak loads to the network, and is relevant for design and QoS purposes [21]. For example, several meters (or other M2M devices) can be constantly reporting to a collector, which will only transmit the collected data to the network on a specific moment, generating a sudden peak load on the network. This situation will be even more relevant if there are multiple collectors, configured to transmit in synchronism.
One of the fundamental design issues related to burstiness is to determine which solution is better: to have devices always in PDP active state or to have them constantly activating and deactivating PDP sessions.

4.1.3 Relation between information payloads and signaling overheads

One of the implications of such decision is on the payload to signaling ratio. Sessions being constantly activated and deactivated will increase dramatically this ratio, with consequences that are not yet known. Some of them might be congestion on devices/functions such as AAA, GGSN, OCS and SGSN.
For example, a health M2M device that transmits the glycemic value every half an hour, will generate more signaling traffic than actual payload. Another good example of this comes from social networks usage on smartphones, where instant messaging applications generate more signaling than payload per each message exchanged.

4.1.4 Addressability

On the other hand if every single device is always with an active PDP session there will be a need for more IP addresses, and probability an expansion on several databases such as HLR and AuC. Such a scenario if implemented would require the deployment of IPv6, since IPv4 would not sustain such a network for a long period of time.
It is known that the IPv4 addresses are running out [184], so it is easy to predict that such would be a major issue if every person would have several devices, each with its own IP address.

4.2 Indicators and Performance Metrics

There is a number of indicators and metrics that need to be measured and/or calculated by the models presented on the next Chapter. Such metrics and indicators are presented on the following Subsections.

4.2.1 Latency

Low latency will be of fundamental importance for some M2M applications, while for others it won't be a significant constraint. For example some automotive services related with safety and accident avoidance will require extremely low latencies. Also health systems might require almost real-time latency. Applications that require user interaction must present low values of latency in order to avoid user withdrawal (The old eight-second rule[18]). Other services, such as metering will not be so sensitive to latency.

Thus it is fundamental to model latency requirements for a wide range of M2M services in order to define adequate QoS classes based on this metric. According to Robert D. Woolley [185], in order to study and/or measure latency of web transactions such metric must be decomposed on the following components:

- Browser processing time;
- Wide Area Network time;
- Web server processing time;
- Local Area Network time;
- Application and transaction server processing time;
- Database server processing time.

The sum of all of these components represents the latency of the studied/measured connection. The study of the response time for each of these components will allow to identify network bottlenecks and to define the application latency constraints: "any application is only as good as its weakest (slowest) link" [185]. Although this process is designed to web transactions, M2M communications can be studied with a similar approach, just implementing some changes. For example, while studying M2M communications it is important to study the latency imposed by AAA processes.

4.2.2 Throughput (including Signaling)

Subscriber and network throughput are fundamental performances metrics to correctly design the network. It is also important to understand how much throughput feeding the network is payload (data plane) and how much is signaling (control plane). An adequate model for the throughput generated by a group of subscribers of circuit and data services (including M2M) is a key tool to help MNOs designing their network in terms of number (or configuration) of sites, number of Core Network (CN) elements, and CN architecture.

The design of models for throughput dimensioning for UMTS Packet Core Networks (considering M2M services and conventional data traffic) is one of the main goals of this work. Such model is presented on Chapter 5.

[18] The "eight-second rule" states that no user will wait more than eight seconds for a web-page to load [185]. However this rule is out of date since nowadays, with the generalization of broadband, most users would quit much before the eight seconds.

4.2.3 Bandwidth

Although most M2M services and applications are not expected to generate large amounts of data, and that the control plane will be the most affected by the increase of signaling throughput, it is still important to consider that the network must be capable to offer the minimum bandwidth required by the subscribers. Failing to do so might result in low Quality of Experience (QoE), and consequently on churn and poor client satisfaction.

4.2.4 Addressability / Number of Subscribers

As referred on the previous Section, addressability is an important feature to consider while adding M2M services to mobile networks. It is important to model the market scenario in order to understand how many IP addresses are needed and how to dimensioning several network elements such as AuCs, CGs, HLRs, SMS-CSs, and VLRs.

4.2.5 Bearers

It is important to understand the number of available bearers, so that their number is enough to serve the subscribers connected to the network. This work is not focused on the radio access and therefore this metric will not be considered.

4.2.6 Energy Efficiency (EE)

Another important performance metric is the Energy Efficiency (EE). Energy has already a huge weight on the MNOs OPEX, therefore improvements on EE are fundamental. In respect to M2M applications and devices the focus on improving EE should be on the device/subscriber side, since M2M related EE improvements on the network are unlikely.

With broad implementation of M2M services and devices this situation presents a new importance, since it must be clear who will pay the energy bill of M2M devices, especially on scenarios that present high truck roll costs. EE constraints will need to be considered when designing the processes for M2M services. For example, Table 12 shows how the communication interval can dramatically change the lifetime of batteries installed on water meters.

Communication Interval	***Battery Lifetime***
Real-time	5 years
15 minutes	8 years
2 times per day	10 years

Table 12 – Lifetime of batteries for water metering devices [186]

4.3 Critical Network Elements

Now that the relevant metrics and indicators are identified it is convenient to understand which are the metrics and indicators relevant to each network element (NE). And what is their impact on each NE.

4.3.1 Radio Network Controller (RNC)

According to M. C. Chuah and Q. Zhang [138], the capacity of an RNC can be described in terms of:

- Maximum number of STM-1 interfaces;
- Maximum raw throughput on interface Iu (C_{Iu});
- Maximum raw throughput on interface Iur, typically $\frac{C_{Iu}}{3}$;
- Maximum raw throughput on interface Iub, typically $\frac{3\,C_{Iu}}{2}$;
- Maximum number of cell-connected data users;
- Maximum number of UTRAN Registrations Area (URA) connected MS that can handle simultaneously.

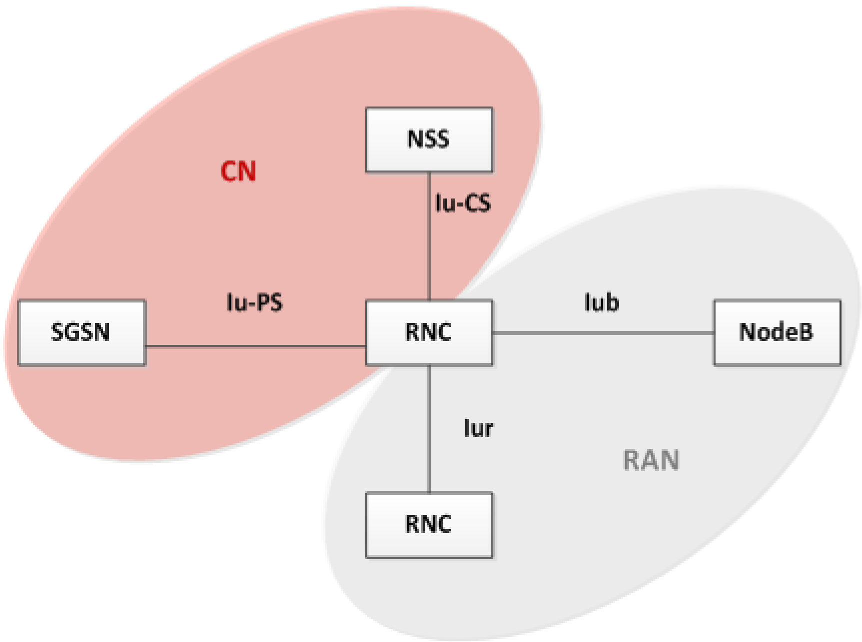


Figure 52 – RNC interfaces

4.3.2 Serving/Gateway GPRS Support Node (SGSN/GGSN)

The SGSN/GGSN must be designed in order to fulfill the MNO desired performance requirements. This means that a traffic model must be defined to determine how much throughput, and how much active PDP contexts, need to be sustained by a SGSN. According to M. C. Chuah and Q. Zhang [138], the capacity of a SGSN/GGSN can be described in terms of:

- Maximum number of active PDP contexts without IPSEC[19];
- Maximum number of active PDP contexts with IPSEC;
- Maximum throughput, considering an average value of *P* bytes per data packet;
- Maximum number of RNCs that can be connected to the SGSN (for SGSNs);
- Maximum number of SGSNs that can be connected to the GGSN (for GGSNs).

When dimensioning a GGSN or a SGSN, QoS/performance parameters and service parameters (KPIs) must be considered [169]. SGSNs and GGSNs can perform overload control in a number of ways:

- **Drop of GTP-C request messages:** in case of signaling overload the SGSN/GGSN can drop GTP-C messages according to pre-defined algorithms when either the network resource consumption or the CPU consumption exceeds a predefined threshold;
- **Shedding traffic:** the SGSN/GGSN can control the overload by shedding traffic from PDP contexts with lower QOS requirements first. The GGSN shall first deny service to incoming requests for establishing new PDP contexts [138].

4.3.3 Home Location Register (HLR)

As presented on Subsection 3.3.2 the HLR is responsible for administrating and control the data base of local subscribers. It controls and maintains any change on a subscriber profile. Consequently the performance of an HLR can be affected by the number of sessions (addressability) and the signaling load (control plane throughput).

4.3.4 DHCP server

The Dynamic Host Configurator Protocol (DHCP) server is a central unit that manages all IP addresses. It allocates an IP address to the subscriber via GGSN in a PDP activation request [168]. As such the DHCP server performance can be affected by the number of subscribers and the number of PDP requests. Again, the shortage of IPv4 addresses might be a major constraint for the DHCP performance.

4.3.5 RADIUS server

The Remote Authentication Dial In User Service (RADIUS) server implements Authentication, Authorization and Accounting (AAA) functions. As such the RADIUS server performance can be affected by the number of subscribers and the number of PDP requests.

[19] Internet Protocol Security (IPSEC) is a technology protocol suite for securing IP communications by authenticating and/or encrypting each IP packet of a communication session.

4.3.6 Online Charging System (OCS)

According to the specified on ETSI TS 132 296 [167], the OCS supports mechanisms for:

- "Online bearer charging towards access/core network entities (e.g. SGSN, PCEF, WLAN);
- Online charging of applications/services that are provided to subscribers via service nodes (outside the core network) e.g. MMS and LCS;
- Account balance management towards external account management servers e.g. recharge server, hot billing server;
- Generation of Charging Data Records and their transfer to the operator's post-processing system;
- Spending limit and balance monitoring and reporting based on subscription or configuration within OCS, towards Policy and Charging Rule Function".

Thus, OCS may be affected by M2M subscriptions requiring a high (and differentiated) number of charging operations.

4.3.7 Authentication Centre (AuC)

As stated on Subsection 3.3.2 the AuC is the entity responsibly for authentication and encryption functions. Therefore an increase on the number of subscriptions will require a reinforcement on the AuC storage and processing capacity.

4.3.8 Summary

Table 13 presents a summary of the metrics that might significantly affect the previously presented Network Elements.

Network Element / System	***Throughput (data plane)***	***Signaling Load (control plane)***	***Addressability / Number of Subscribers***
AuC	-	affected by	affected by
DHCP server	-	affected by	affected by
HLR	-	affected by	affected by
OCS	-	affected by	affected by
RADIUS server	-	affected by	affected by
SGSN / GGSN	-	-	affected by

Table 13 – Summary of critical Network Elements

This Chapter presented the M2M related factors that might seriously impact the UMTS Packet Core Network, also explaining the novelty of the work presented. In order to study such factors, a group of indicators and metrics were defined, and it was developed an analysis of some of the Network Elements that are believed to be the UMTS Packet CN critical nodes. The following Chapters present models to help study such features and NEs.

5. Traffic and Throughput Analysis and Modeling for UMTS Packet Core Networks

This Chapter presents the SGSN Control Plane fundamental functions, providing the framework for understanding the messages and processes involved on Mobility Management and Session Management.
Based on such knowledge, a number of throughput models are defined in order to compute data throughput and signaling throughput values for each of the UMTS interfaces at study (Iu-PS, Gn, Gp, Gr, Gi).
The combined use of the models defined for each interface allows for the design of a capacity model for the UMTS Packet CN.

5.1 SGSN Control Plane Fundamental Functions

The SGSN control plane provides the following fundamental functions:

- Mobility Management;
- Session Management;
- Path Management;
- Short Message Service (SMS) [168].

In the context of this document it is relevant to study in detail some of the messages involved in Mobility Management and Session Management. The steps presented in the figures of this Section are fully detailed on ETSI TS 123 060 [187].

5.1.1 Mobility Management

Attach Process

When a Mobile Station (MS) is activated it scans the vicinity for the strongest control channel. Then the mobile station tunes into it and decodes the signal in the channel [138]. Subsequently the MS will register itself on the network, it begins the Attach process.

A MS is considered to be PMM-Attached when there is communication between the MS and the SGSN [138]. The Attach process is a fundamental step for a mobile station to obtain access to a packet switched network. This process is depicted in Figure 53.

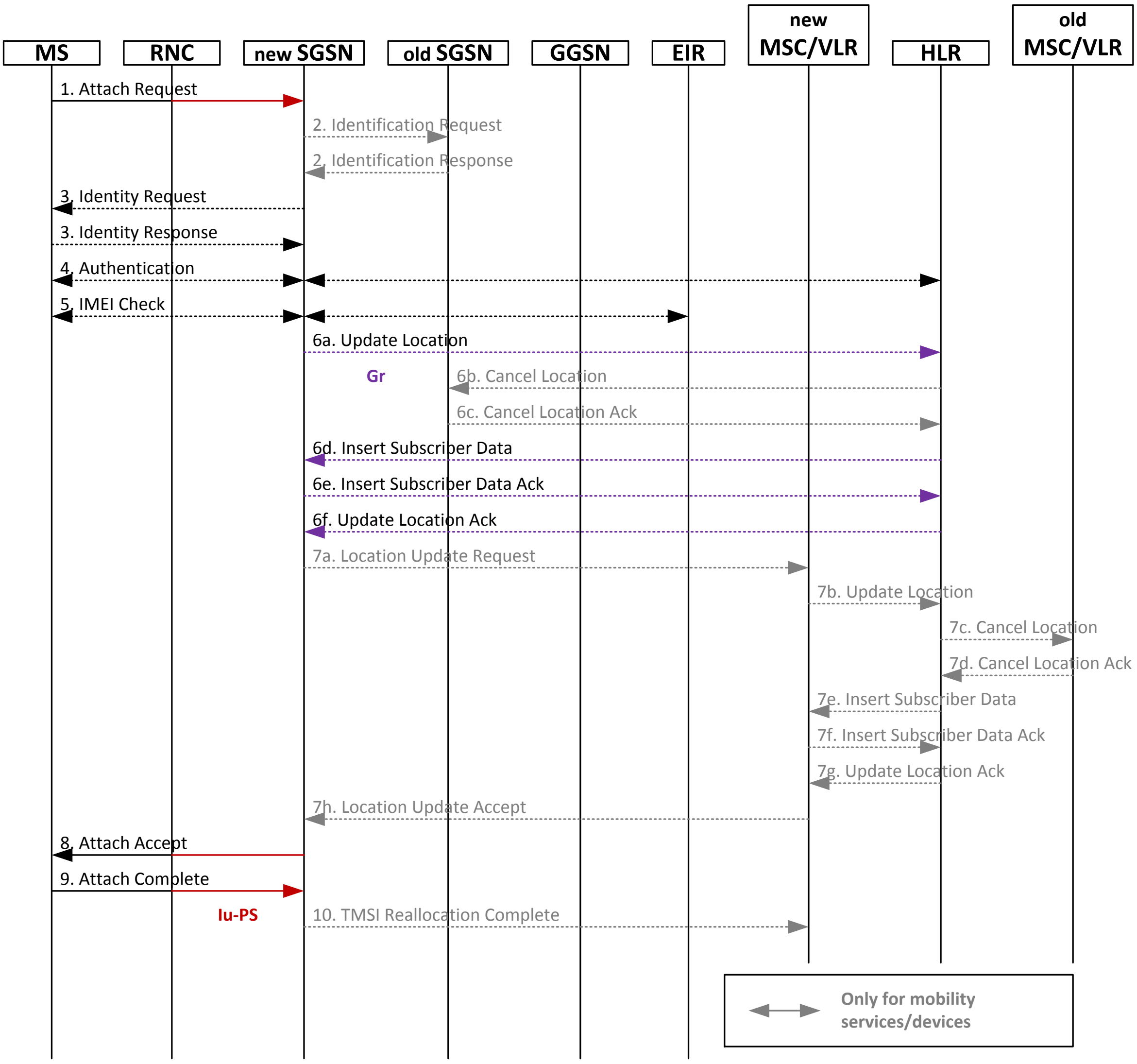


Figure 53 – Combined GPRS / IMSI attach procedure (based on [187])

Steps 1, 8 and 9 contribute to the throughput in the control plane of Iu-PS interface. Step 6 contributes to the throughput of Gr interface.

Detach Process

The GPRS Detach procedure allows an MS to inform the network that it does not intends to access SGSN-based services any longer; and the network to inform an MS that it does not have access to the SGSN-based services any more [187]. The MS is detached either explicitly or implicitly:

- **Explicit detach:** The network or MS explicitly requests detach. A Detach Request (Cause) is sent by the SGSN to the MS, or by the MS to the SGSN;

- **Implicit detach:** The network detaches the MS, without notifying it, a configuration-dependent time after the mobile reachable timer expired, or after an irrecoverable radio error causes disconnection of the logical link [187].

In the MS Detach Request message there is an indication to tell if the detach is due to switch off or not. The indication is needed to know whether a Detach Accept message should be returned or not. On network-originated Detach Request message there may be an indication to tell the MS

that it is requested to initiate GPRS Attach and PDP Context Activation procedures for the previously activated PDP contexts [187].
The MS-Initiated Detach Procedure is depicted on Figure 54.

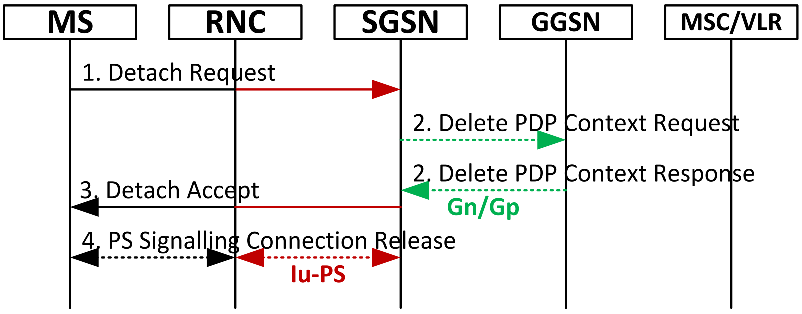


Figure 54 – MS-initiated combined GPRS / IMSI detach procedure (based on [187])

Steps 1, 3 and 4 contribute to the throughput in the control plane of Iu-PS interface.

The Network-Initiated Detach Procedure can be SGSN-Initiated or HLR-Initiated. The SGSN-Initiated Detach Procedure is depicted on Figure 55.

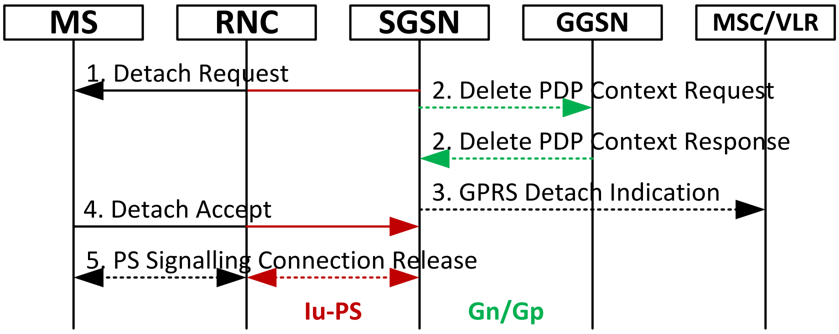


Figure 55 – SGSN-Initiated GPRS detach procedure (based on [187])

Steps 1, 4 and 5 contribute to the throughput in the control plane of Iu-PS interface.

The HLR-Initiated Detach Procedure is depicted on Figure 56.

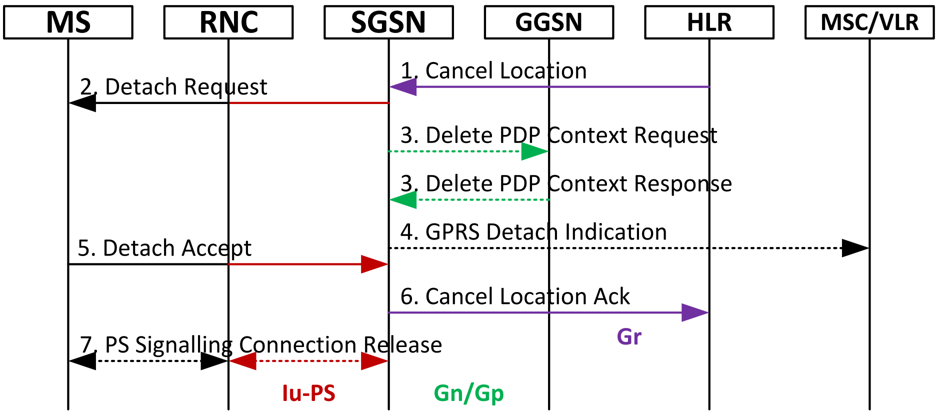


Figure 56 – HLR-initiated GPRS detach procedure (based on [187])

Step 1 and 6 contributes to throughput of Gr interface. Step 2, 5 and 7 contribute to the throughput in the control plane of Iu-PS interface.

Authentication Process

The authentication process is executed by the Mobile Station (MS), SGSN, HLR and AuC [187]. The authentication process is depicted on Figure 57.

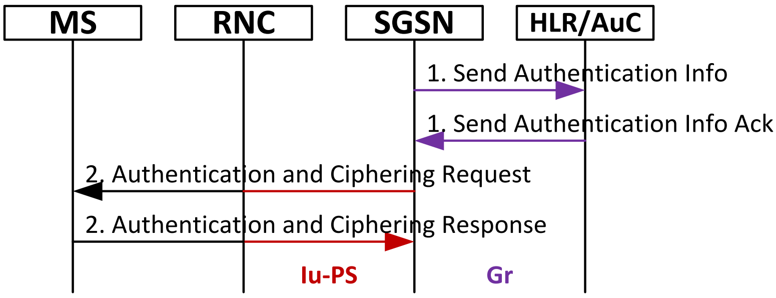


Figure 57 – UMTS authentication procedure (based on [187])

Messages of authentication request and authentication response (step 2) are part of the throughput in control plane of Iu-PS interface. Step 1 messages contribute to Gr interface throughput.

Routing Area Update Process (RAU) - Intra SGSN

Cellular networks are mainly designed for mobility scenarios. When some entity wants to communicate with a Mobile Station (MS), the network consults the HLR to determine in which MSC is the MS registered at that moment [138]. Thus it is important that a MS informs the network of its location so that calls and data can be routed to the correct destiny.

A routing area update occurs when an attached MS detects that it has entered a new Routing Area (RA), when the periodic RA update timer has expired, or when the MS has to indicate new access capabilities to the network. The SGSN detects that it is an intra-SGSN routing area update by noticing that it also handles the old RA. In this case, the SGSN has the necessary information (mobility management/PDP contexts) about the MS and there is no need to inform the GGSNs or the HLR about the new MS location. A periodic RA update is always an intra SGSN routing area update [138], [187]. In the context of this work it shall be assumed the worst case: that each RNC spans a different area. Therefore, any change of RNC will trigger either an intra-SGSN RAU or an inter-SGSN RAU [138]. The Intra SGSN Routing Area Update is depicted on Figure 58.

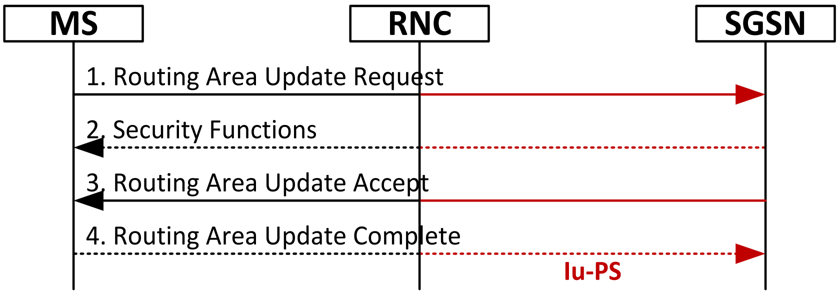


Figure 58 – Intra SGSN Routing Area Update procedure (based on [187])

All four steps contribute to throughput of user plane of Iu-PS interface.

Routing Area Update Process - Inter SGSN

If the SGSN detects that it is being requested an inter-SGSN routing area update, it will request mobility management and PDP contexts to the old GGSN. After subscriber authentication the SGSN will send an update PDP context request to the GGSN as well as an update location to the HLR [138]. This process is depicted on Figure 59.

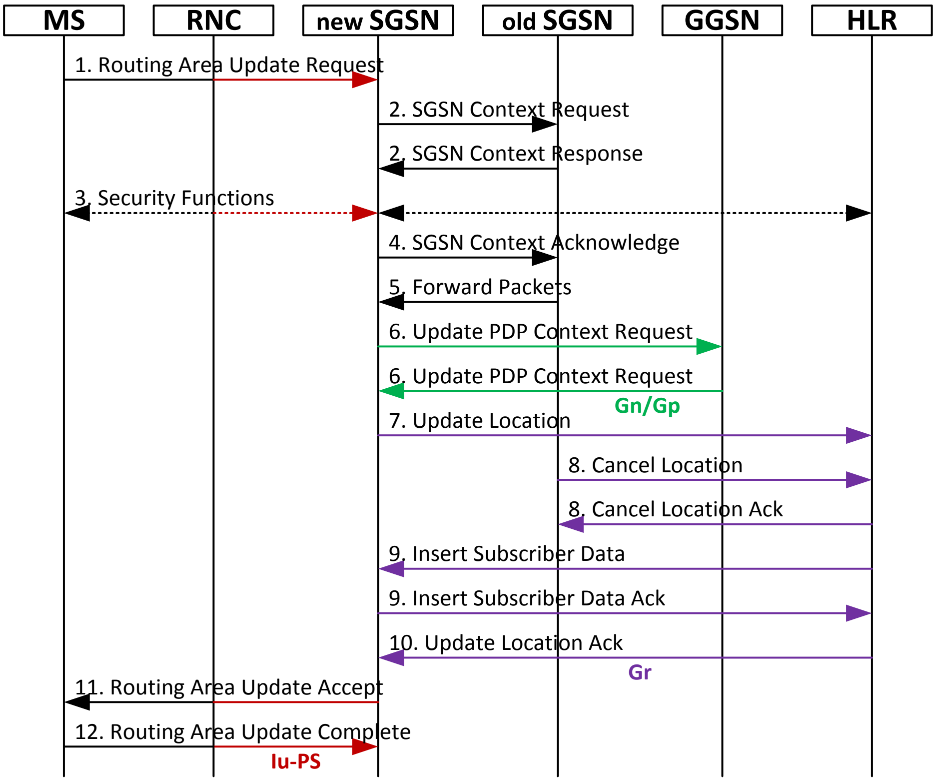


Figure 59 – Inter SGSN Routing Area Update procedure (based on [187])

Step 1, 11 and 12 contribute to the control plane throughput of Iu-PS interface. Step 6 contributes to the control plane throughput of Gn/Gp interface. Steps 7 to 10 contribute to the throughput of Gr interface.

Routing Area Update Process – Service RNC Route Update

Service Radio Network Controller (RNC) route update, or Serving Radio Network Subsystem (SRNS) relocation, is a procedure used in mobility scenarios when the control of the SRNS is changed to another Radio Network Subsystem.

There are three statuses of Mobility Management (MM) in Packet Switched (PS) domain:

- Packet Mobility Management connected (PMM-connected);
- PMM-detached;
- PMM-idle [168].

These statuses are stored in the SGSN MM context. Service RNC relocation is only implemented in MM-connected status [168].

The Service RNC relocation is depicted on Figure 60.

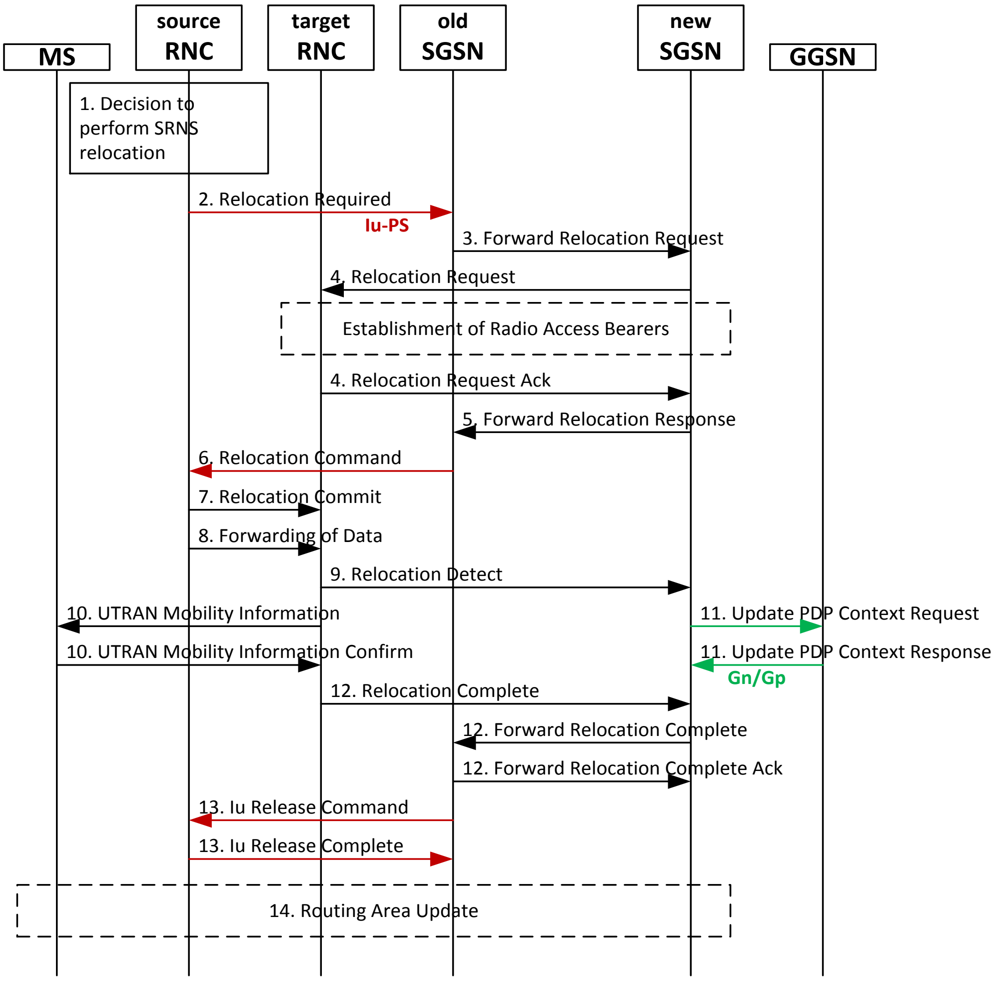


Figure 60 – Service RNC relocation procedure (based on [187])

Steps 2, 6 and 13 contribute to the throughput for control plane of Iu-PS interface.

5.1.2 *Session Management*

PDP Activation Process

After establishing the SGSN service request, the MS can request the activation of a PDP context [138]. Figure 61 depicts the PDP activation process.

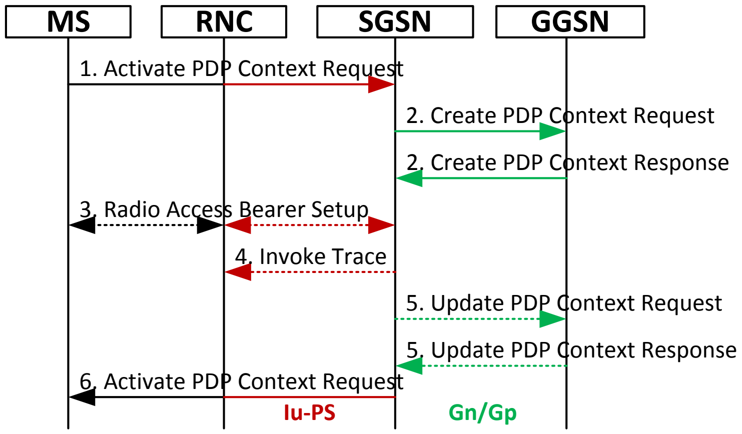


Figure 61 – PDP context activation procedure for UMTS (based on [187])

For each PDP Address a different quality of service (QoS) profile may be requested. For example, some PDP addresses may be associated with E-mail that can tolerate lengthy response times. Other applications cannot tolerate delay and demand a very high level of throughput, interactive applications being one example. These different requirements are reflected in the QoS profile. If a QoS requirement is beyond the capabilities of a PLMN, the PLMN negotiates the QoS profile as close as possible to the requested QoS profile. The MS either accepts the negotiated QoS profile, or deactivates the PDP context [187]. If the PDP Context Activation Procedure fails or if the SGSN returns an Activate PDP Context Reject (Cause, PDP Configuration Options) message, the MS may attempt new activation to the same APN up to a maximum number of attempts [187].

Step 1, 3, 4 and 6 contribute to the throughput in the control plane of Iu-PS interface. Step 2 and 5 contribute to the throughput in the control plane of Gn/Gp.

The PDP context activation can also be Network-Requested, as depicted in Figure 62.

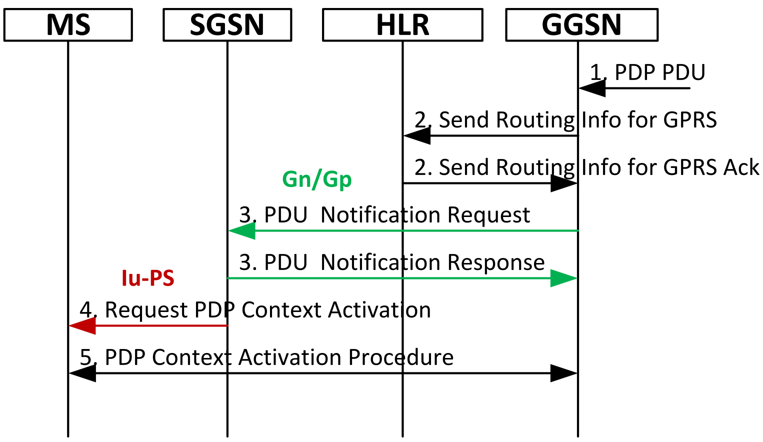


Figure 62 – Successful network-requested PDP context act. procedure (based on [187])

Step 3 contributes to the throughput in the control plane of Gn/Gp interface. Step 4 contributes to the throughput in the control plane of Iu-PS interface.

PDP Deactivation process

As in the case of the PDP activation, PDP deactivation process also has two trigger models: MS initiated and Network initiated [187]. In terms of the contribution to the throughput in control plane of Iu-PS interface both models are similar [168], and therefore only MS initiated deactivation is depicted on Figure 63.

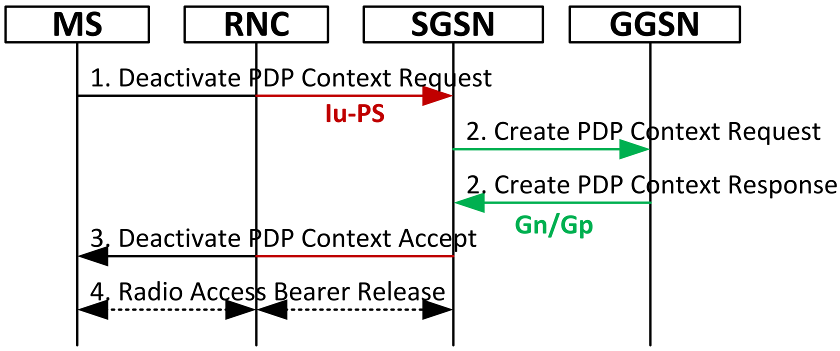


Figure 63 – MS initiated PDP context deactivation procedure for UMTS (based on [187])

Steps 1 and 3 contribute to the throughput in the control plane of Iu-PS interface. Step 2 contributes to the throughput in the control plane of Gn/Gp interface.

5.2 Definition of Packet Switched Aggregation Node

Following Sections will make use of the concept of "Packet Switched Aggregation Node", therefore it shall be defined as a node that is responsible to connect the access network to the customer network, Local Area or single device. Figure 64 depicts examples of aggregation nodes, which can be.

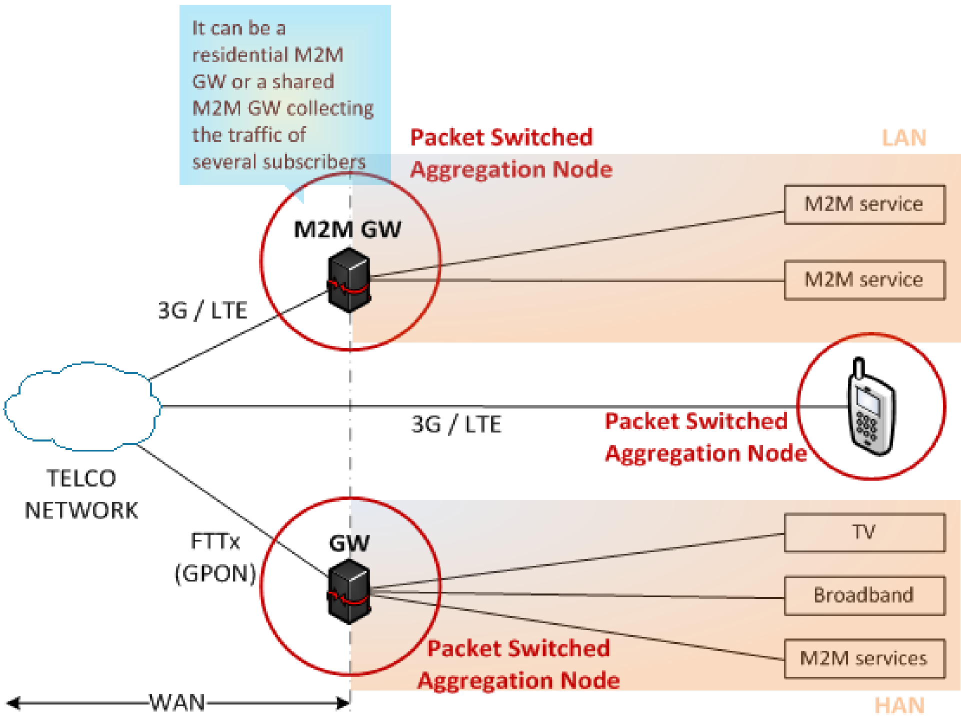


Figure 64 – Packet Switched Aggregation Node

5.3 Iu-PS Interface

The Iu interface is comprised of two connections, the Iu-PS interface that interconnects the Radio Network Controller (RNC) and the Serving GPRS Support Node (SGSN) and the Iu-CS interface that connects the RNC and the Media Gateway (MGW). The MGW is part of the circuit switched domain (CS) [168] and therefore the Iu-CS interface will not be considered in the following models. The Iu-PS interface protocol stack is depicted on Table 14.

<table>
<tr><th colspan="2">Radio Network Control Plane</th><td rowspan="7"></td><th>PS Data Plane</th><th>Header Size</th><th>OSI reference model</th></tr>
<tr><td colspan="2">RANAP</td><td rowspan="2">Iu-UP</td><td rowspan="2">HIu-UP</td><td rowspan="4">layer 4, Transport</td></tr>
<tr><td colspan="2">SCCP</td></tr>
<tr><td>MTP3-B</td><td>SCTP</td><td>GTP-U</td><td>HGTP</td></tr>
<tr><td>SSCF-INI</td><td rowspan="2">IP</td><td>UDP / TCP</td><td>HUDP</td></tr>
<tr><td>SSCOP</td><td>IP</td><td>HIP</td><td>layer 3, Network</td></tr>
<tr><td colspan="2">MPLS</td><td>MPLS</td><td>HMPLS</td><td>layer 2, Data Link</td></tr>
<tr><td colspan="6">layer 1, Physical</td></tr>
</table>

Table 14 – Protocol stack of Iu-PS interface (based on [168], [186], [187])

The GTP protocol allows multi-protocol packets to be tunneled between GPRS Support Nodes (GSNs) in the UMTS/GPRS backbone network. It includes both the GTP control plane (GTP-C) and data transfer (GTP-U) [188]. In the user plane, GTP-U is used to provide a service for carrying user data packets. The GTP-U protocol is implemented by SGSNs and GGSNs in the UMTS/GPRS Backbone and by Radio Network Controllers (RNCs) in the UTRAN [188].

The header of a GTP-U protocol has variable size [168]; for the purpose of this document it will be considered its maximum size, 12 bytes [168]. The considered header size for Iu-PS interface is presented on Table 15.

User Plane	*Header Size [bytes]*
Iu-UP	4
GTP-U	12
UDP	8
IP	20
MPLS	8
Total	**52**

Table 15 – Header size for Iu-PS interface (based on [168])

5.3.1 Throughput of data plane in Iu-PS interface

$$RO_{Iu-PS} = \frac{S_{Packet} + H_{IuUP} + H_{GTP} + H_{UDP} + H_{IP} + H_{MPLS}}{S_{Packet}}$$

Equation 1 – Overhead ratio in Iu-PS interface

Where:
- S_{Packet} is the average IP packet size [bytes];
- H_{IuUP} is the header size of Iu-UP packet [bytes];
- H_{GTP} is the header size of GTP-U packet [bytes];
- H_{UDP} is the header size of UDP packet [bytes];
- H_{IP} is the header size of IP packet [bytes];

- H_{MPLS} is the header size of MPLS packet [bytes].

Consequently the throughput of data plane in Iu-PS interface is defined by:

$$D_{IuPS} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Iu-PS} * f_d$$

Equation 2 – Throughput of data plane in Iu-PS interface (based on [138], [168])

Where:

- D_{IuPS} is the data plane throughput in Iu-PS interface [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;
- $R_{\frac{Active}{Attach}}$ is the rate of attached aggregation nodes that activate PDP;
- Th_{Node} is the average throughput per aggregation node per second [bps];
- RO_{Iu-PS} is the overhead ratio in Iu-PS interface;
- f_d is the data throughput redundancy factor (prevents the network from traffic overflow).

5.3.2 Signaling Load of Iu-PS interface

Table 16 lists eleven basic types of messages that can be estimated by the mobile operators and that comprise most throughput of Iu-PS control plane. The first eight messages are introduced in Section 5.1 and the other three are optional messages that can also be applied to Equation 3. The signaling load of Iu-PS interface is given by:

$$S_{IuPS} = N_N * R_{Attach} * \sum_{i=1}^{11} \left(N_{IuPS_i} * L_{IuPS_i}\right) * 8 * \frac{1}{3600} * f_s$$

Equation 3 – Signaling load of Iu-PS interface (based on [138], [168])

Where:

- S_{IuPS} is the Iu-PS interface signaling load [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;
- N_{IuPSi} refers to Table 16;
- L_{IuPSi} refers to Table 16;
- 8 is used to convert bytes to bits;
- $\frac{1}{3600}$ is used to convert hour to seconds;
- f_s is the signaling redundancy factor.

i	N_{IuPS_i}	L_{IuPS_i} $[bytes]$	***Relevant for M2M comm.***
1	Authentication times per hour	Length of messages per authentication	yes
2	Attachment times per hour	Length of messages per attachment in Iu-PS	
3	Detachment times per hour	Length of messages per detachment in Iu-PS	
4	Inter SGSN route update times per hour	Length of messages per inter SGSN route update	only if M2M service requires mobility
5	Intra SGSN route update times per hour	Length of messages per intra SGSN route update	
6	Intra SGSN SRNC route update times per hour	Length of messages per intra SGSN SRNC	
7	PDP activation times per hour	Length of messages per PDP activation	yes
8	PDP deactivation times per hour	Length of messages per PDP deactivation	
9	Periodic SGSN route area update (RAU) times per hour	Length of messages per periodical SGSN route update	
10	SMS Mobile Originated (SMS MO) times per hour	Length of messages per SMS service	only if M2M service requires SMS service
11	SMS MT times per hour		

Table 16 – Footnotes for Equation 3 (based on [138], [168])

Table 16 messages refers to signaling on Iu-PS interface. Messages such as P-Temporary Mobile Subscriber Identity (TMSI) re-allocation message, identification check message, and service request message are not considered in Equation 3 due to their small size and reduced usage. If proven necessary to integrate them into the Equation, a redundancy factor can be imposed (f_s).

The number of periodic Route Area Updates (RAUs) is determined by applying Equation 4.

$$N_{Route_{periodic}} = \frac{N_N\ (1 - R_{Active})}{P_{Refresh} * 3600}$$

Equation 4 – Periodic RAUs (based on [138])

Where:

- N_N is the number of 3G packet switched aggregation nodes;
- R_{Active} is the ratio of aggregation nodes with activate PDP session;
- $P_{Refresh}$ is the periodic RAU interval [s];
- 3600 is used to convert seconds to hour;
- $N_N\ (1 - R_{Active})$ represents the number of idle aggregation nodes.

5.3.3 Total throughput in Iu-PS interface

Total throughput in Iu-PS interface is obtained by the sum of Equation 2 and Equation 3:

$$T_{IuPS} = D_{IuPS} + S_{IuPS}$$

Equation 5 – Total throughput in Iu-PS interface

5.4 Gn/Gp Interface

Gn interface interconnects SGSNs to GGSNs on the same Public Land Mobile Network (PLMN), or interconnects two SGSNs [163]. Gp interface interconnects SGSNs to GGSNs on different PLMNs, providing security functions required for this inter-PLMN communication [163]. Both interfaces implement the same protocol stack (Table 17), in which GPRS Tunneling Protocol (GTP) is used to transport encapsulated packets between SGSNs and GGSNs.

User Plane	*Header Size [bytes]*
GTP	12
UDP	8
IP	20
LLC	n/a
MAC	n/a

Table 17 – Protocol stack of Gn/Gp interface [168]

GTP messages are classified into three categories, GTP-U, GTP-C and GTP'. Only GTP-U and GTP-C are defined for Gn/Gp interfaces [188]. GTP-C messages contain path management messages, tunnel management messages, mobility management messages, and location management messages. GTP-U will contribute more than GTP-C to the throughput of Gn/Gp interfaces [168].

5.4.1 Throughput of GTP-U (data plane) in Gn/Gp interface

The overhead ratio in Gn/Gp interface is obtained by applying Table 17 values to Equation 6:

$$RO_{Gn/Gp} = \frac{S_{Packet} + H_{GTP} + H_{UDP} + H_{IP}}{S_{Packet}}$$

Equation 6 – Overhead ratio in Gn/Gp interface [168]

Where:

- S_{Packet} is the average IP packet size [bytes];
- H_{GTP} is the header size of GTP-U packet [bytes];
- H_{UDP} is the header size of UDP packet [bytes];
- H_{IP} is the header size of IP packet [bytes].

Therefore, the throughput of GTP-U (data plane) on Gn/Gp interface is given by:

$$D_{Gn/Gp} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Gn/Gp} * f_d$$

Equation 7 – Throughput of GTP-U (data plane) in Gn/Gp interface (based on [138], [168])

Where:

- $D_{Gn/Gp}$ is the throughput of GTP-U (data plane) on Gn/Gp interface [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;
- $R_{\frac{Active}{Attach}}$ is the ratio of attached aggregation nodes that activate PDP;
- Th_{Node} is the average throughput per aggregation node per second [bps];
- $RO_{Gn/Gp}$ is the ratio of overhead in Gn/Gp interface;
- f_d is the data throughput redundancy factor.

5.4.2 Signaling Load of Gn/Gp interface

It is assumed that the major GTP-C messages going through Gn/Gp interface, and that consequently will contribute to the throughput of Gn/Gp interface are PDP context activation and PDP context deactivation [168]. Therefore the throughput of GTP-C in Gn/Gp interface is given by:

$$S_{Gn/Gp} = N_N * R_{Attach} * \sum_{i=1}^{2}\left(N_{Gn_i} * L_{Gn_i}\right) * 8 * \frac{1}{3600} * f_s$$

Equation 8 – Signaling load of Gn/Gp interface (based on [138], [168])

Where:

- $S_{Gn/Gp}$ is the Gn/Gp interface signaling load [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;
- N_{Gni} refers to Table 18;
- L_{Gni} refers to Table 18;
- 8 is used to convert bytes to bits;
- $\frac{1}{3600}$ is used to convert hour to seconds;
- f_s is the signaling redundancy factor.

i	N_{Gn_i}	L_{Gn_i} $[bytes]$	***Relevant for M2M comm.***
1	Create PDP context times per hour	Length of messages per create PDP context	yes
2	Delete PDP context times per busy hour	Length of messages per delete PDP context	

Table 18 – Footnotes for Equation 8 (based on [138], [168])

5.4.3 Total throughput in Gn/Gp interface

Total throughput in Gn/Gp interface is achieved by the sum of Equation 7 and Equation 8:

$$T_{Gn/Gp} = S_{Gn/Gp} + S_{Gn/Gp}$$

Equation 9 – Total throughput in Gn/Gp interface

5.5 Gr Interface

Gr interface interconnects the HLR and the SGSN. Its fundamental task is to transport subscriber related information in PS domain. Data exchange between these two elements usually happen when a subscriber requests PS services, and the SGSN or HLR respond to the request [168]. The protocol stack for this interface is presented on Table 19.

MAP
TCAP
SCCP
MTP3
MTP2
MTP1

Table 19 – Protocol stack of Gr interface [168]

To obtain the signaling load for Gr interface it must be considered four fundamental types of messages: authentication messages (step 1 in Figure 57); attach messages (step 6 in Figure 53); detach messages (step 1 and 6 in Figure 56); and inter SGSN Routing update messages (step 7, 8, 9, and 10 in Figure 59).

5.5.1 Signaling Load of Gr interface

Based on these four messages, and applying the data in Table 20 to Equation 10 it is possible to obtain the throughput on Gr interface.

$$T_{Gr} = S_{Gr} = N_N * R_{Attach} * \sum_{i=1}^{3} \left(R_{Gr_i} * N_{Gr_i} * L_{Gr_i} \right) * 8 * \frac{1}{3600} * f_s$$

Equation 10 – Signaling in Gr interface (based on [138], [168])

Where:

- T_{Gr} is the throughput on Gr interface [bps];
- S_{Gr} is the signaling throughput on Gr interface [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;
- R_{Gri} refers to Table 20;
- N_{Gri} refers to Table 20;
- L_{Gri} refers to Table 20;
- 8 is used to convert bytes to bits;
- 3600 is used to convert hour to seconds;
- f_s is the signaling redundancy factor.

i	R_{Gr_i}	N_{Gr_i}	L_{Gr_i} [bytes]	Relevant for M2M comm.
1	Authentication Rate	Authentication times per hour	Length of messages per authentication	yes
2	Attach Rate	Attach times per hour	Length of messages per attachment	
3	Not applicable	Inter SGSN route update times per hour	Length of messages per inter SGSN route update	only if M2M service requires mobility

Table 20 – Footnotes for Equation 10 (based on [138], [168])

5.6 Gi Interface

Gi interface interconnects the GGSN to external Packet Data Networks (PDNs), e.g. Internet. Gi commonly specifies the end of packet switched domain as the PDN typically belongs to another service provider or carrier [168]. Before subscribers are allowed to access the network, Authentication, Authorization, and Accounting (AAA) information is checked by the RADIUS server, which holds the subscriber related data such as login name, password, profile information and can additionally collect and record accounting statistics [168]. This process is done through Gi interface.
Another network entity connected via Gi interface is the DHCP server, which is a central unit that manages the IP addresses. It is DHCP server task allocation of IP address to subscribers via GGSN in a PDP activation request [168].

5.6.1 Throughput in Gi Interface

Data going through Gi interface are comprised by data payload (such as IP datagram), optional VPN encapsulation, and header (such as Ethernet header) [168]. Assuming that most users are not using VPN encapsulation, the ratio of overhead in Gi interface can be obtained by Equation 11.

$$RO_{Gi} = \frac{S_{Packet} + H_{datagram}}{S_{Packet}}$$

Equation 11 – Overhead ratio in Gi interface [168]

Where:

- S_{Packet} is the average IP packet size [bytes];
- $H_{datagram}$ is the media header on Gi interface [bytes].

Therefore the throughput on Gi interface is given by:

$$T_{Gi} = D_{Gi} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Gi} * f_d$$

Equation 12 – Throughput in Gi Interface (based on [168])

Where:

- T_{Gi} is the throughput on Gi Interface [bps];
- D_{Gi} is the data throughput on Gi Interface [bps];
- N_N is the number of 3G packet switched aggregation nodes;
- R_{Attach} is the ratio of attached aggregation nodes;

- $R_{\frac{Active}{Attach}}$ is the ratio of attached aggregation nodes that activate PDP;
- Th_{Node} is the average throughput per aggregation node per second [bps];
- RO_{Gi} is the rate of overhead on Gi interface;
- f_d is the data throughput redundancy factor.

5.7 UMTS Packet Core Network Modeling

Applying the models presented on the previous Sections it is possible to design an overall model of the capacity and usage of an UMTS Packet Core Network. This model considers five main PS interfaces (Iu-PS, Gn, Gp, Gi, and Gr), four Network Elements (SGSN, GGSN, HLR, and AuC) and the Online Charging System (OCS), as depicted on Figure 65. Using the same methodology this model can be further extended to other interfaces and NEs.

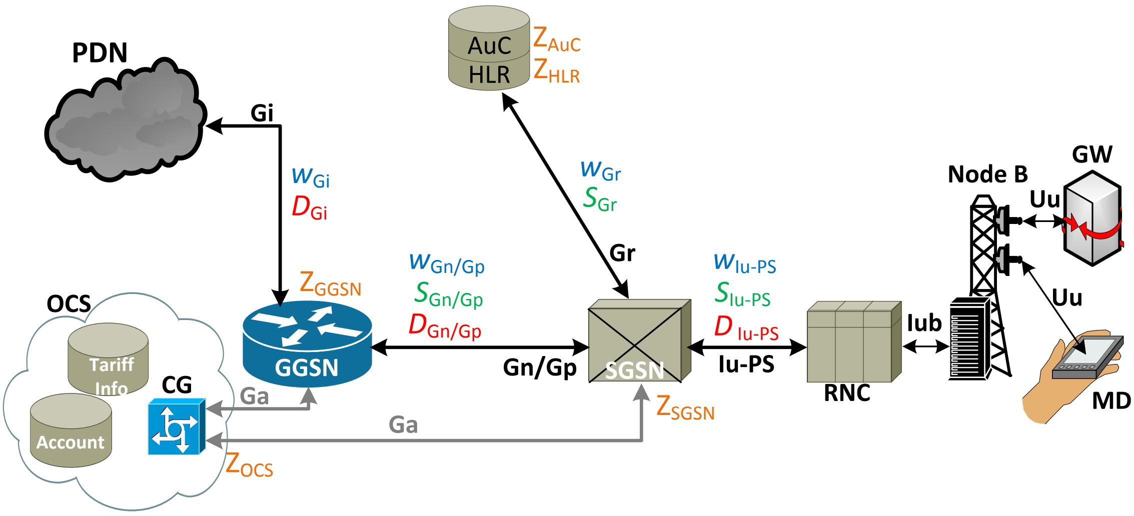


Figure 65 – Mathematical description for the UMTS Packet CN model

As depicted on Figure 65, three types of variables per interface and an array per NE/System were defined, where:

- W_i is the raw throughput capacity of interface i [bps];
- S_i is the signaling throughput on interface i [bps];
- D_i is the data throughput on interface i [bps];
- Z_X is an array of capacity values for the NE/System “x”;

This array is comprised by a set of capacity values, where:

- $AtSUB_X$ is the attached subscribers capacity of NE/System “x”;
- $AcSUB_X$ is the active subscribers capacity of NE/System “x”;
- SIG_{i_X} is the signaling throughput capacity of the interface i of NE/System “x” [bps];
- W_{i_X} is the raw throughput capacity of the interface i of NE/System “x” [bps].

5.8 Profiling and Orchestration of 3G Services

Telecommunications networks should be able to support and provide multiple services to all of its subscribers regardless of the time of day, which is to say that while on Busy Hour (BH) the network should be able to support every request. Therefore it is fundamental to understand the size and characteristics of the traffic generated by the subscribers of this network. Traffic modeling and/or measurement is of vital importance to design a close to optimal network.

In order to conduct a study on the impacts of M2M services on a 3G network, a model to orchestrate service profiles for such network must be designed. This Section presents a model for Packet Switched services. Circuit Switched services will be considered on the network usage baseline (Chapter 0). To define such orchestration model, a number of parameters (many of them correspondent to variables defined on Sections 5.3 to 5.6) shall be defined for each service at study, as presented on Table 21.

Variable	Corresponds to[20]	Description
Number of connections	N_N	Number of devices (Smartphones, Aggregation Nodes, Cars, etc) that are service subscribers.
Throughput	Th_{Node}	Node throughput (aggregated services throughput values) [kbps].
Sessions / Hour	N_{IuPS_7}, N_{IuPS_8}, N_{Gni}	Number of service sessions per hour; Considered to be correspondent to the number of PDP activation/deactivation requests.
Attachments / Hour	N_{IuPS_2}, N_{IuPS_3}, N_{Gr_2}	Number of attachment times per hour; Considered to be correspondent to the number of detachment times per hour.
Authentications / Hour	N_{IuPS_1}, N_{Gr_1}	Number of authentication operations per hour.
Intra SGSN route updates / Hour	N_{IuPS_5}, N_{IuPS_6}	Number of Intra SGSN and Intra SGSN SRNC route update times per hour.
Inter SGSN route updates / Hour	N_{IuPS_4}, N_{Gr_3}	Number of Inter SGSN route update times per hour.
Periodic SGSN RAUs / Hour	N_{IuPS_9}	Number of periodic SGSN route area update (RAU) times per hour.
Authentication Rate	R_{Gr_1}	Ratio of authentication that needs to get parameters from HLR [%].
SMSs MO / Hour	$N_{IuPS_{10}}$	Number of SMSs Mobile Originated per hour.
SMSs MT / Hour	$N_{IuPS_{11}}$	Number of SMSs Mobile Terminated per hour.
Premium subscriber ratio		Ratio of Premium subscribers [%].
Regular subscriber ratio		Ratio of Regular subscribers [%].
Basic subscriber ratio		Ratio of Basic subscribers [%].
Attached subscribers ratio	R_{Gr_2}, R_{Attach}	Ratio of attached subscribers [%].
Active/Attached subscribers ratio	$R_{\frac{Active}{Attach}}$	Ratio of attached subscribers with active session [%].

Table 21 – Parameters for service orchestration

[20] Correspondent variables of Section 5.3, 5.4, 5.5, and 5.6.

To the purpose of this work the throughput of each service shall be determined by a profiling model simple enough that it can be used without exaggerated mathematical complexity. Thus, such model should be considered as an approximation to the reality, good enough to serve as input method for the models presented on previous Sections. It also intends to serve as a framework for future work on this subject (please refer to Chapter 9 for more detail on related future work).

The proposed model will be presented for ten commonly used mobile services (additional Packet Switched services shall be considered on the network usage baseline, along with Circuit Switched services):

- Email;
- Web-Browsing;
- Application update;
- Instagram;
- Facebook;
- Other Social Networks;
- Instant Messaging (Includes Facebook Messenger);
- VoIP;
- Audio Streaming (e.g. Spotify, Google Music);
- Video Streaming (e.g. YouTube).

For each service (except VoIP) three different Service Level Agreements (SLAs) were defined:

- **Premium:** Heavy and Business subscribers;
- **Regular:** Common subscribers;
- **Basic:** Light subscribers.

Each service is characterized by a number of input parameters, as presented on Table 22 and Table 23, where colored cells refer to input values.

This model was inspired by the work of A. M. de Oliveira Duarte and B. R. Jarmo [189], N. K. Shankaranarayanan et all, on “User-perceived performance of Web-browsing and interactive data in HFC cable access networks” [190] and John T. Chapman on “Multimedia Traffic Engineering - The Bursty Data Model” [191].

Service	***SLA***	***Average size of data (δ)[21] [bytes]***	***Average packet size (P_{size})[22] [bytes]***	***Mean time between packets within a packet burst (t_{bP})[23] [sec]***	***Time between file transfer / session (t_{δ})[24] [sec]***
Email (1 email)	premium	75.000,00	1000	0,015	180
	regular	75.000,00	800	0,015	1200
	basic	75.000,00	600	0,015	3600
Web-Browsing (1 web-page)	premium	623.800,00	1000	0,002	2
	regular	623.800,00	800	0,002	4
	basic	623.800,00	600	0,002	7
App Update (1 update = 1 session)	premium	10.000.000,00	1000	0,015	0
	regular	10.000.000,00	800	0,015	0
	basic	10.000.000,00	600	0,015	0
Instagram (15 photos = 1 session)	premium	1.500.000,00	1000	0,015	0
	regular	1.500.000,00	800	0,015	0
	basic	1.500.000,00	600	0,015	0
Facebook (1 session)	premium	4.000.000,00	1000	0,001	1800
	regular	4.000.000,00	800	0,001	1200
	basic	4.000.000,00	600	0,001	3600
Other Social Networks (1 session)	premium	900.000,00	1000	0,005	3600
	regular	900.000,00	800	0,005	2400
	basic	900.000,00	600	0,005	3600
Instant Messaging (1 session)	premium	10.000,00	1000	0,015	300
	regular	10.000,00	800	0,015	300
	basic	10.000,00	600	0,015	1200
Audio Streaming (1 song)	premium	10.000,00	1000	0,001	270
	regular	10.000,00	800	0,001	270
	basic	10.000,00	600	0,001	270
Video Streaming (1 video)	premium	5.000.000,00	1000	0,001	300
	regular	5.000.000,00	800	0,001	600
	basic	5.000.000,00	600	0,001	900

Table 22 – Throughput profiling: downlink service characterization

[21] The average size for a web-page was estimated by measuring the size of various web-pages using a web page analyzer tool (available at: http://websiteoptimization.com/services/analyze/); The average size for an app update was estimated by measuring the size of various app updates available at Google Play store (This updates refer to commonly used Apps such as Dropbox, Google+, Instagram, and Out of Milk); The average size for an Instagram photo was estimated by measuring the size of various photos from the service. The presented value refers to 15 photos (= 15 * average value of a single photo); Average size for Audio and Video streaming was estimated by measuring the size of various songs and videos from Spotify and YouTube, respectively. Other size values were assumed as being reasonable values.
[22] Assumed values.
[23] Assumed values.
[24] Assumed values.

Service	*SLA*	*Average size of data (δ)*[25] *[bytes]*	*Average packet size (P_{size})*[26] *[bytes]*	*Mean time between packets within a packet burst (t_{bP})*[27] *[sec]*	*Time between file transfer / session (t_{δ})*[28] *[sec]*
Email (1 email)	premium	20.000,00	1000	0,015	600
	regular	20.000,00	800	0,015	1200
	basic	20.000,00	600	0,015	3600
Web-Browsing (1 web-page)	premium	0,00	1000	0,002	2
	regular	0,00	800	0,002	4
	basic	0,00	600	0,002	7
App Update (1 update = 1 session)	premium	0,00	1000	0,015	0
	regular	0,00	800	0,015	0
	basic	0,00	600	0,015	0
Instagram (1 photo = 1 session)	premium	100.000,00	1000	0,020	0
	regular	100.000,00	800	0,020	0
	basic	100.000,00	600	0,020	0
Facebook (1 session)	premium	400.000,00	1000	0,020	1800
	regular	400.000,00	800	0,020	1200
	basic	400.000,00	600	0,020	3600
Other Social Networks (1 session)	premium	270.000,00	1000	0,020	3600
	regular	270.000,00	800	0,020	2400
	basic	270.000,00	600	0,020	3600
Instant Messaging (1 session)	premium	10.000,00	1000	0,015	300
	regular	10.000,00	800	0,015	300
	basic	10.000,00	600	0,015	1200
Audio Streaming (1 song)	premium	0,00	1000	0,0010	270
	regular	0,00	800	0,0010	270
	basic	0,00	600	0,0010	270
Video Streaming (1 video)	premium	0,00	1000	0,0010	300
	regular	0,00	800	0,0010	600
	basic	0,00	600	0,0010	900

Table 23 – Throughput profiling: uplink service characterization

The number of packets (referring to one file/session), the time of transfer (of one file/session), and ultimately the throughput generated by each service, for each SLA, can be calculated by applying the following equations:

$$N_{Packets} = \frac{\delta}{P_{size}}$$

Equation 13 – Number of packets per file/session

Where:

[25] The average size for an Instagram photo was estimated by measuring the size of various photos from the service. Other size values were assumed as being reasonable values.
[26] Assumed values.
[27] Assumed values.
[28] Assumed values.

- $N_{Packets}$ is the number of packets needed to transfer the required data;
- δ is the average size of the data message to be transmitted [bytes];
- P_{size} is the average packet size [bytes].

$$\boldsymbol{t_t = N_{Packets} * t_{bP}}$$

Equation 14 – File/Session time of transfer

Where:
- t_t is the time of transfer [s] (Figure 66);
- $N_{Packets}$ is the number of packets needed to transfer the required data;
- t_{bP} is the mean time between Packets within a Packet burst [s] (Figure 66).

$$\boldsymbol{Th_k = \frac{\delta}{t_\delta + t_t} * 8}$$

Equation 15 – Service throughput

Where:
- Th_k is the throughput generated by the service “k” [bps];
- δ is the average size of the data message to be transmitted [bytes];
- t_δ is the time between data messages transmissions [s] (Figure 66);
- t_t is the time of transfer [s] (Figure 66);
- 8 is used to convert bytes to bits.

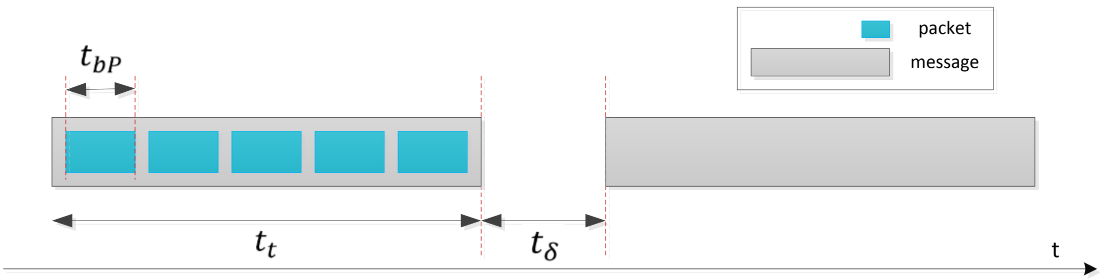


Figure 66 – t_t, t_{bP}, and t_δ time relationships

Based on Cisco information [192], the average throughput for VoIP is considered to be of 6.25 kbps.

Applying this equations and results, the throughput profile is given by Table 24 and Table 25.

Service	***SLA***	***Average size of data (δ) [KB]***	***Average packet size (P_{size}) [bytes]***	***Number of Packets ($N_{Packets}$)***	***t_{bP} [sec]***	***Time between file transfer / session (t_{δ}) [sec]***	***Time of transfer (t_t) [sec]***	***Throughput (Th_k) [kbps]***
Email (1 email)	premium	75,00	1000	75,00	0,015	180	1,13	3,31
	regular	75,00	800	93,75	0,015	1200	1,41	0,50
	basic	75,00	600	125,00	0,015	3600	1,88	0,17
Web-Browsing (1 web-page)	premium	623,80	1000	623,80	0,002	2	1,25	1.536,64
	regular	623,80	800	779,75	0,002	4	1,56	897,63
	basic	623,80	600	1039,67	0,002	7	2,08	549,64
App Update (1 update = 1 session)	premium	10.000,00	1000	10000,00	0,015	0	150,00	533,33
	regular	10.000,00	800	12500,00	0,015	0	187,50	426,67
	basic	10.000,00	600	16666,67	0,015	0	250,00	320,00
Instagram (15 photos = 1 session)	premium	1.500,00	1000	1500,00	0,015	0	22,50	533,33
	regular	1.500,00	800	1875,00	0,015	0	28,13	426,67
	basic	1.500,00	600	2500,00	0,015	0	37,50	320,00
Facebook (1 session)	premium	4.000,00	1000	4000,00	0,001	1800	4,00	17,74
	regular	4.000,00	800	5000,00	0,001	1200	5,00	26,56
	basic	4.000,00	600	6666,67	0,001	3600	6,67	8,87
Other Social Networks (1 session)	premium	900,00	1000	900,00	0,005	3600	4,50	2,00
	regular	900,00	800	1125,00	0,005	2400	5,63	2,99
	basic	900,00	600	1500,00	0,005	3600	7,50	2,00
Instant Messaging (1 session)	premium	10,00	1000	10,00	0,015	300	0,15	0,27
	regular	10,00	800	12,50	0,015	300	0,19	0,27
	basic	10,00	600	16,67	0,015	1200	0,25	0,07
VOIP	-							6,25
Audio Streaming (1 song)	premium	10,00	1000	10,00	0,001	270	0,01	0,30
	regular	10,00	800	12,50	0,001	270	0,01	0,30
	basic	10,00	600	16,67	0,001	270	0,02	0,30
video streaming (1 video)	premium	5.000,00	1000	5000,00	0,001	300	5,00	131,15
	regular	5.000,00	800	6250,00	0,001	600	6,25	65,98
	basic	5.000,00	600	8333,33	0,001	900	8,33	44,04

Table 24 – Downlink throughput profile

Service	***SLA***	***Average size of data (δ) [KB]***	***Average packet size (P_{size}) [bytes]***	***Number of Packets ($N_{Packets}$)***	***t_{bP} [sec]***	***Time between file transfer / session (t_{δ}) [sec]***	***Time of transfer (t_t) [sec]***	***Throughput (Th_k) [kbps]***
Email (1 email)	premium	20,00	1000	20,00	0,015	600	0,30	0,27
	regular	20,00	800	25,00	0,015	1200	0,38	0,13
	basic	20,00	600	33,33	0,015	3600	0,50	0,04
Web-Browsing (1 web-page)	premium	0,00	1000	0,00	0,002	2	0,00	0,00
	regular	0,00	800	0,00	0,002	4	0,00	0,00
	basic	0,00	600	0,00	0,002	7	0,00	0,00
App Update (1 update = 1 session)	premium	0,00	1000	0,00	0,015	0	0,00	0,00
	regular	0,00	800	0,00	0,015	0	0,00	0,00
	basic	0,00	600	0,00	0,015	0	0,00	0,00
Instagram (1 photo = 1 session)	premium	100,00	1000	100,00	0,020	0	2,00	400,00
	regular	100,00	800	125,00	0,020	0	2,50	320,00
	basic	100,00	600	166,67	0,020	0	3,33	240,00
Facebook (1 session)	premium	400,00	1000	400,00	0,020	1800	8,00	1,77
	regular	400,00	800	500,00	0,020	1200	10,00	2,64
	basic	400,00	600	666,67	0,020	3600	13,33	0,89
Other Social Networks (1 session)	premium	270,00	1000	270,00	0,020	3600	5,40	0,60
	regular	270,00	800	337,50	0,020	2400	6,75	0,90
	basic	270,00	600	450,00	0,020	3600	9,00	0,60
Instant Messaging (1 session)	premium	10,00	1000	10,00	0,015	300	0,15	0,27
	regular	10,00	800	12,50	0,015	300	0,19	0,27
	basic	10,00	600	16,67	0,015	1200	0,25	0,07
VOIP	-							6,25
Audio Streaming (1 song)	premium	0,00	1000	0,00	0,001	270	0,00	0,00
	regular	0,00	800	0,00	0,001	270	0,00	0,00
	basic	0,00	600	0,00	0,001	270	0,00	0,00
Video Streaming (1 video)	premium	0,00	1000	0,00	0,001	300	0,00	0,00
	regular	0,00	800	0,00	0,001	600	0,00	0,00
	basic	0,00	600	0,00	0,001	900	0,00	0,00

Table 25 – Uplink throughput profile

When a group of services/tributaries are aggregated on an Aggregation Node, the Node throughput will be given by:

$$T_{node} = \sum Th_k$$

Equation 16 – Aggregation Node throughput

Where:

- T_{node} is the Node throughput [bps];
- Th_k is the throughput generated by the service "k" [bps] (Equation 15).

This model will be applied to the case study of Chapter 0.

5.9 Summary

Combining the models presented over this Chapter it is possible to model the throughputs and capacity usage of the UMTS PC Network at study. And it is possible to do so based on the characteristics of traffic generated by the users. This is of great value in order to estimate the implications on the UMTS PC Network caused by traffic of diverse M2M services. Since this task has a certain degree of mathematical complexity, a prototype for an analytic tool capable of applying the referred models is presented on the next Chapter. The usage of the models presented in this Chapter, as well as of next Chapter's tool, will be further clarified on Chapter 0.

6. SOTA – Scenario Orchestration and Traffic Analysis Tool

This Chapter presents the prototype of an analytic tool capable of modeling throughput and capacity usage of an UMTS PC Network, based on the models presented over Chapter 5. This tool needs to be fed by subscribers' traffic data, which will be the input for the referred models, and will be provided by the Profiling and Orchestration model presented on Section 5.8.
Figure 67 depicts a flowchart of tool structure.

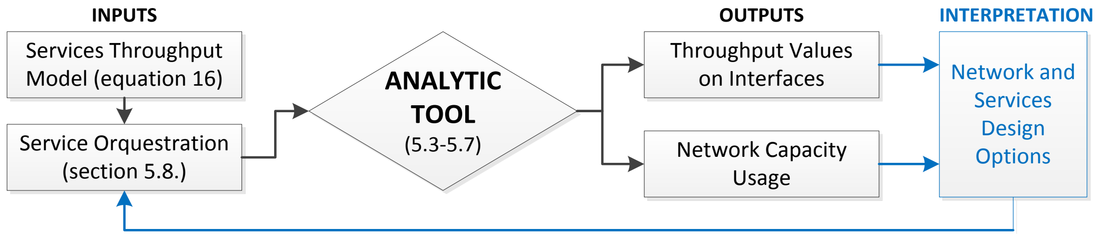


Figure 67 – SOTA flowchart

The tool is presented over the next Sections, with an overall presentation on Section 6.2, a description of the service orchestration tab on Section 6.3, and of the network analysis tab on Section 6.4. Chapter 0 presents a case study that shows in detail how to use the tool and how the mathematical models subjacent to it are applied.

6.2 SOTA Overview

The developed prototype is a Small Web Format (SWF) file which can be executed on devices that have a Flash Player installed (this technical option refers only to the prototype work, an eventual commercial roll-out of this tool could be based on other technologies). The welcome tab for this prototype is depicted on Figure 68.

Figure 68 – SOTA "Welcome" tab

Current version is comprised by six main tabs, as depicted on Figure 69.

Figure 69 – SOTA main tabs

These tabs are the following:

- **Welcome:** Welcome tab, presents a brief description of the tool (Figure 68);
- **Architecture:** Presents the reference architecture for the tool current version (Figure 70). In future releases this tab is expected to become a selection menu for the architecture/technology to consider;
- **Traffic Profile:** Presents the available traffic profiles on the current version of the tool. These profiles are used to characterize the services pattern of usage through the 24 hours of the day (Figure 73);
- **Service Orchestration:** Provides a group of configurable services, that can be selected as to be considered, or not, on the network analysis. This tab will be further explained on Section 6.3;
- **Traffic/Network Analysis:** Presents status reports for network interfaces, elements, and systems, as result of the previous service orchestration. A number of charts and visual indicators are available. Further description of this tab is provided on Section 6.4;
- **Network Configuration:** Enables the user to define a number of network parameters and characteristics. Further detail on this tab is presented on Section 6.5;

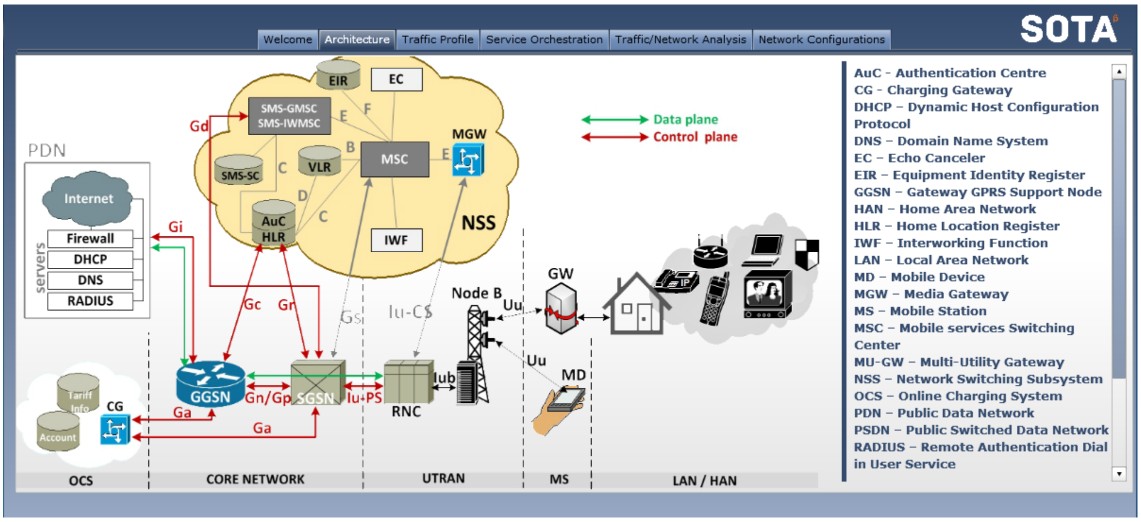


Figure 70 – SOTA "Architecture" tab

6.3 Traffic Modeling: Service Orchestration

The current prototype version allows orchestration of the following M2M services:

- Electricity metering;
- Home automation;
- Healthcare;
- Automotive;
- Security;
- Logistics.

And the following non M2M services:

- Web Browsing;
- Video Streaming.

Different services will have different input variables, which will be applied to the models of Section 5.8.

Figure 71 and Figure 72 depict the graphical interface of “Service Orchestration” tab.

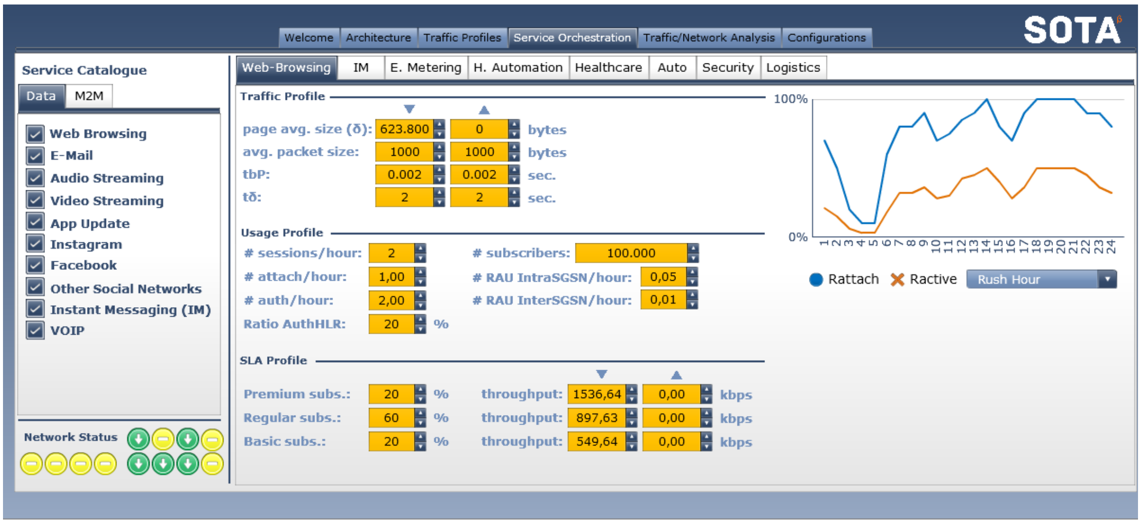


Figure 71 – Web-Browsing orchestration

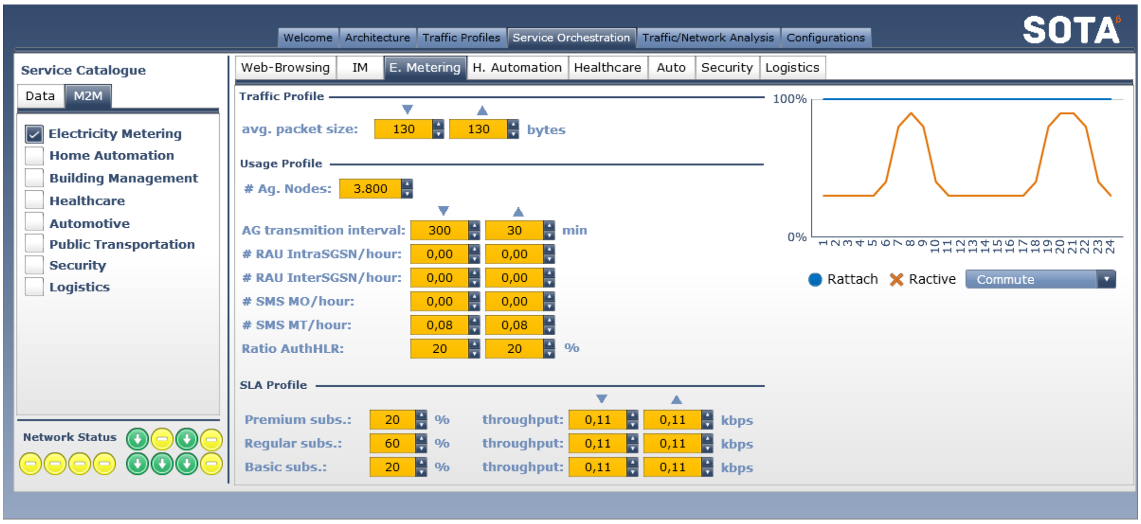


Figure 72 – Electricity Metering orchestration

In its current version the tool permits the choice of one of six traffic profiles for each service, as depicted on Figure 73.

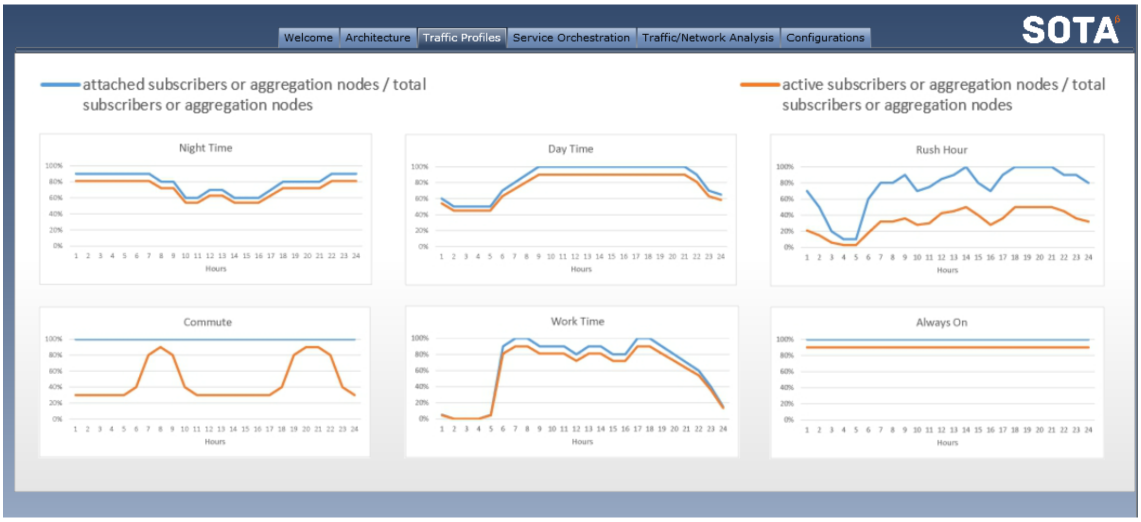


Figure 73 – SOTA traffic profiles

6.4 Network Analysis

This tab presents the network and traffic analysis.
It presents colored indicators through the Core Network that indicate if a specific Interface/NE/System capacity usage is:

- Below 80% (green indicator);
- Above 80% (yellow indicator);
- Above 90% (red indicator).

It also presents Interface/NE/System usage and throughput charts, as depicted on Figure 74 and Figure 75.

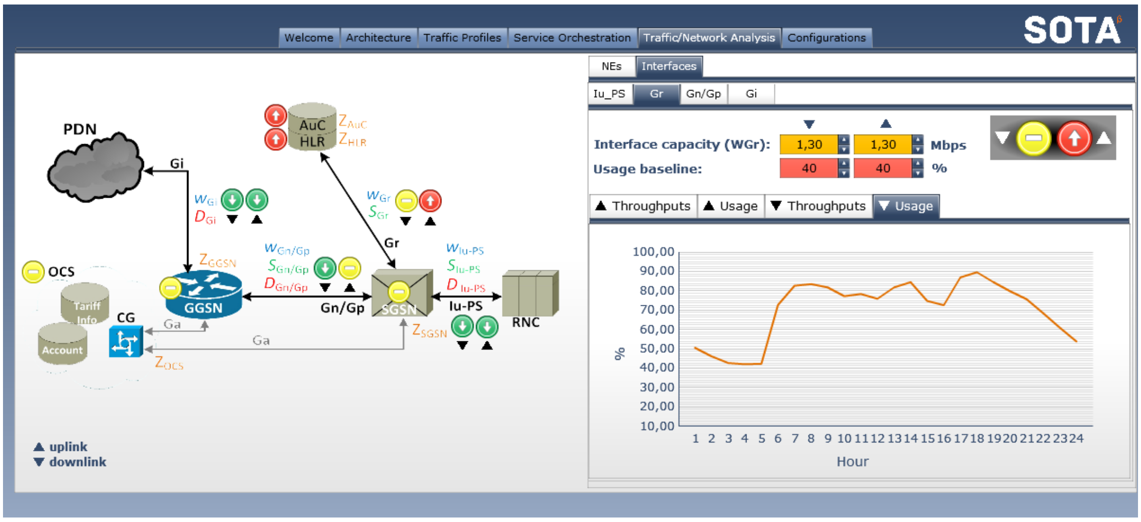


Figure 74 – Gr interface usage report

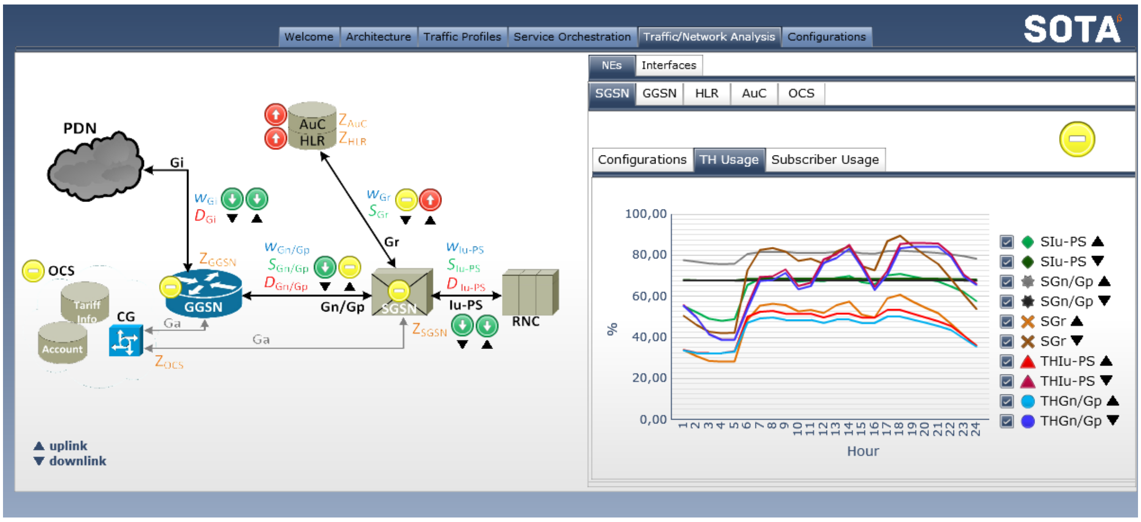


Figure 75 – SGSN usage report

6.5 Network Configurations

Additional configuration inputs are required, being defined in this tab:

- Messages length [bytes];
- Baseline usage signaling to data ratio [%];
- Periodic RAUs interval [s].

Figure 76 depicts this tab.

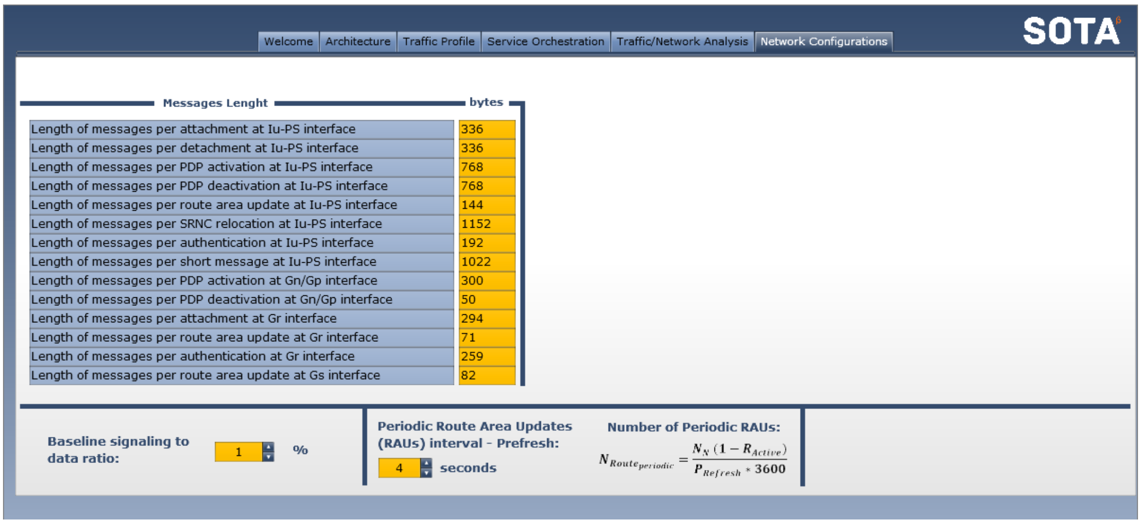


Figure 76 – Network configurations

6.6 Summary

This Chapter presents the prototype for the tool “SOTA”, which applies the previously proposed models, presenting a user friendly interface where the services can be orchestrated and characterized, while simultaneously monitoring their effects on a telecommunications network.

7. Case Study

In order to exemplify the applicability of present work, a case study is presented. This case study is divided over two Sections, the first studies in detail the electricity metering, providing a detailed explanation on how to calculate the values of throughput for this particular service. The second Section provides an example of how to calculate the baseline usage for a cellular network, and which are the effects of adding M2M services in the capacity usage of that network. A sensibility analysis is implemented with the objective of identifying the parameters that present greater potential for severe impact on the network.

The study is done for a period of twenty-four hours, with a granularity of one hour and considering a week day. A Portuguese urban scenario is considered.

The scenarios presented on this chapter do not reflect, or are inspired by, any specific reality of any Portuguese Mobile Network Operator. They are just approximated and more or less informed assumptions, based on the international literature, which were considered to:

- Assess the functionalities of the tool;
- Identify some of the points of possible future constraint on 3G networks arising from the use of M2M services for future detailed research.

7.1 UMTS Packet Core Throughput Generated by Electricity Metering

This Section intends to model the throughput generated by an electricity metering service on a UMTS Packet Core Network.

7.1.1 Subscribers and Aggregation Nodes

It is considered that the electricity metering is executed by a group of meters, inside the subscriber's home, that will report data to a shared Collector/Gateway, which will be the Aggregation Node (AN). This AN will provide the connectivity to the UMTS network. The architecture of the electricity metering service is presented on Figure 77.

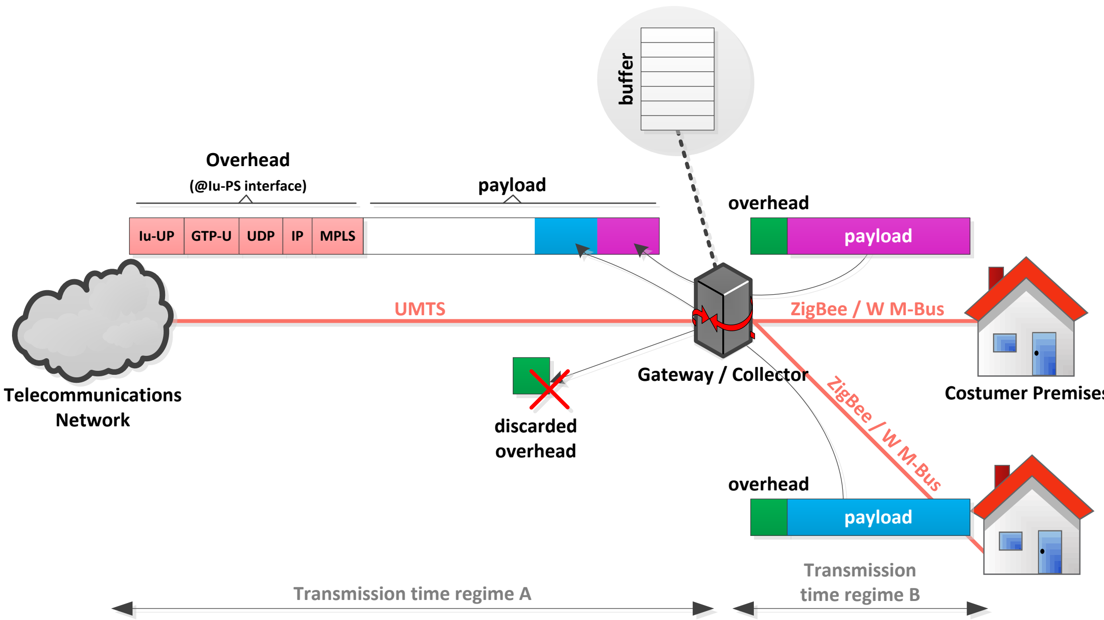


Figure 77 – Gateway/Aggregation Node data collection architecture

As depicted by the figure, the Gateway will collect the payload information from the various tributaries (discarding the overhead), and then it can either send the information to the UMTS network or store it on an internal buffer so that it can be transferred on predefined moments (or when the buffer reaches its maximum capacity).

In the context of this case study the number of connections/ANs it is considered to be:

$$N_N = 3800$$

The average number of subscribers sharing the gateway is defined as:

$$N_{subscribers} = 25$$

7.1.2 Metering

Figure 78 depicts the fundamental technical requirements approved by the Portuguese Government, which are based on European regulation [14], [15].

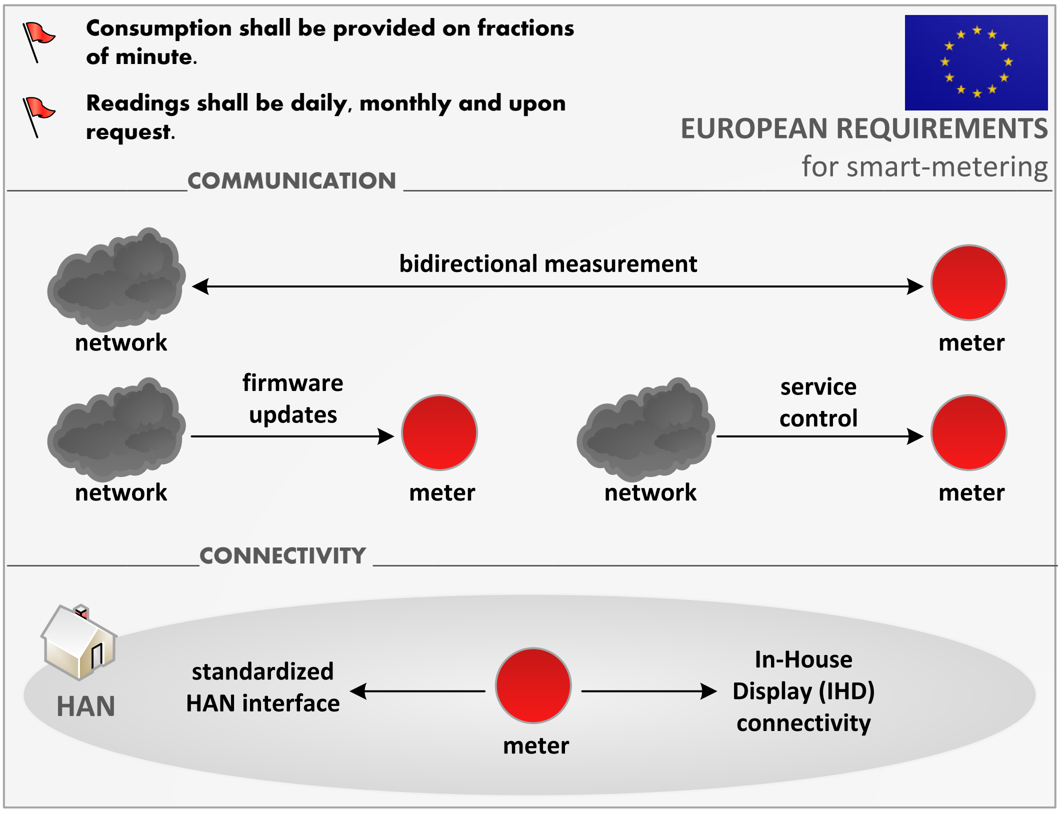


Figure 78 – European approved smart metering technical requirements

Considering the requirements depicted on Figure 78, the following specifications are considered for this case study:

- The meter reads and stores data every minute;
- Every 5 minutes that data is transmitted to the gateway;
- Every 30 minutes the gateway transmits the collected data to the network (AN transmission interval)[29];

[29] Which means that over one hour there will be 2 transmissions. Therefore it shall be assumed that: $N_{IuPS_7} = N_{IuPS_8} = N_{Gni} = N_{IuPS_2} = N_{IuPS_3} = N_{Gr_2} = N_{IuPS_1} = N_{Gr_1} = 2.$

- 2 SMSs are sent to the subscriber every day[30];
- A single downlink operation per day is considered.

Meter-Gateway Communication

Each meter data transmission (to the Gateway, every 5 minutes) is estimated to comprise 160 bytes (1280 bits) of data [186].
As presented on Table 5, some maximum data rates for short-range wireless technologies are:
- **6LowPan:** 250 kbps;
- **ZigBee:** 250 kbps;
- **W M-BUS:** 225 kbps.

Thus, a transmission rate of 250 kbps will be considered between the meter and the gateway, which means that every 5 minutes the meters will send a burst of data to the gateway with the duration of about 5.12 ms.

Gateway-Network Communication

Every hour a single gateway will upload the readings of the meters connected to it:

$$25\ meters * 12\ readings/hour\ * (160 * 8)\ bits\ = 384\ kbits/hour$$

Which can be converted to an average throughput per aggregation node per second (Th_{Node}) of:

$$Th_{Node} = \frac{384\ kbits/hour}{3600\ seconds} = 106.67\ bps$$

It is assumed that there are no differentiated SLAs for the electricity metering service, thus the average throughput of electricity metering per AN will be the same for the three SLAs specified on SOTA:

Figure 79 – Electricity Metering SLA configurations

7.1.3 Network

A fundamental parameter to define is the average size of the IP packet to consider. It is assumed to be:

$$S_{Packet} = 130\ bytes \text{ [193]}$$

Again, there shall be no difference on packet size between SLAs.

[30] Therefore, $N_{IuPS_{10}} = 0$ and $N_{IuPS_{11}} = \frac{2}{24} = 0{,}083$.

In the context of this case study it is considered that there are no intra and inter SGSN route area updates (RAUs), as such, it comes:

$$N_{Route_{intra-SGSN}} = N_{Route_{inter-SGSN}} = N_{IuPS_4} = N_{IuPS_5} = N_{IuPS_6} = N_{Gr_3} = 0$$

Considering the recommended value of $P_{Refresh} = 4$ seconds [194], the number of periodic RAUs is given by Equation 4:

$$N_{Route_{periodic}} = \frac{N_N\,(1 - R_{Active})}{P_{Refresh} * 3600} = \frac{3800\,(1 - R_{Active})}{4 * 3600}$$

Where R_{Active} is defined by the traffic profile correspondent to electricity metering. As depicted on Figure 73, SOTA has six different profiles that can be selected. “Commute” profile is considered to be the one to apply to electricity metering, being depicted on Chart 6.

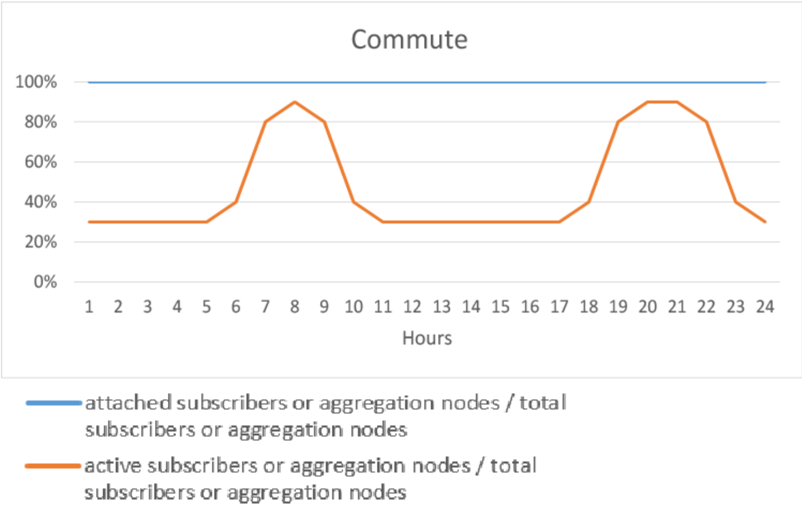


Chart 6 – “Commute” traffic profile

Chart 6 defines R_{Attach}[31] (blue line) and (orange line). In order to apply the equations of Section 5.3, 5.4, 5.5, and 5.6, it is necessary to determine the ratio of active to attached subscriptions ANs:

$$R_{\frac{Active}{Attach}} = \frac{R_{Active}}{R_{Attach}}$$

The ratio of authentication that needs to obtain authentication parameters from HLR it is assumed to be:

$$R_{Auth_{toHLR}} = R_{Gr_1} = 20\%$$

All the necessary configurations on SOTA, correspondent to previously presented values are depicted on Figure 72.

[31] $R_{Gr_2} = R_{Active}$

The length of messages are assumed to be as presented on Table 26.

Parameter	Message Length	Description
$L_{attach_{Iu-PS}}$	336 bytes	Length of messages per attachment at Iu-PS interface
$L_{detach_{Iu-PS}}$	336 bytes	Length of messages per detachment at Iu-PS interface
$L_{PDP_{active_{Iu-PS}}}$	768 bytes	Length of messages per PDP activation at Iu-PS interface
$L_{PDP_{deactive_{Iu-PS}}}$	768 bytes	Length of messages per PDP deactivation at Iu-PS interface
$L_{route_{Iu-PS}}$	144 bytes	Length of messages per route area update at Iu-PS interface
$L_{SRnC_{Iu-PS}}$	1152 bytes	Length of messages per SRNC relocation at Iu-PS interface
$L_{authen_{Iu-PS}}$	192 bytes	Length of messages per authentication at Iu-PS interface
$L_{SMS_{Iu-PS}}$	1022 bytes	Length of messages per short message at Iu-PS interface
$L_{PDP_{active_{Gn}}}$	300 bytes	Length of messages per PDP activation at Gn/Gp interface
$L_{PDP_{deactive_{Gn}}}$	50 bytes	Length of messages per PDP deactivation at Gn/Gp interface
$L_{attach_{Gr}}$	294 bytes	Length of messages per attachment at Gr interface
$L_{route_{Gr}}$	71 bytes	Length of messages per route area update at Gr interface
$L_{authen_{Gr}}$	259 bytes	Length of messages per authentication at Gr interface
$L_{route_{Gs}}$	82 bytes	Length of messages per route area update at Gs interface

Table 26 – Message lengths [168]

A redundancy factor is considered only for data throughput, consequently:

$$f_d = 1{,}1;\ \ f_s = 1{,}0$$

7.1.4 Throughput Calculations

Iu-PS Interface

The first step is to obtain the overhead ratio in Iu-PS interface, given by Equation 1:

$$RO_{Iu-PS} = \frac{S_{Packet} + H_{IuUP} + H_{GTP} + H_{UDP} + H_{IP} + H_{MPLS}}{S_{Packet}}$$

$$RO_{Iu-PS} = \frac{130 + 52}{130} = 1{,}4$$

The Throughput of data plane in Iu-PS interface can be calculated by Equation 2.
Result are presented on Chart 7.

$$D_{IuPS} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Iu-PS} * f_d$$

$$D_{IuPS} = 3800 * 1{,}0 * R_{\frac{Active}{Attach}} * 106{,}67 * 1{,}4 * 1{,}1$$

The signaling load of Iu-PS interface is given by Equation 3.
Result are presented on Chart 7.

$$S_{IuPS} = N_N * R_{Attach} * \sum_{i=1}^{11} \left(N_{IuPS_i} * L_{IuPS_i}\right) * 8 * \frac{1}{3600} * f_s$$

$$S_{IuPS} = 3800 * 1{,}0 * \sum_{i=1}^{11} \left(N_{IuPS_i} * L_{IuPS_i}\right) * 8 * \frac{1}{3600} * 1{,}0$$

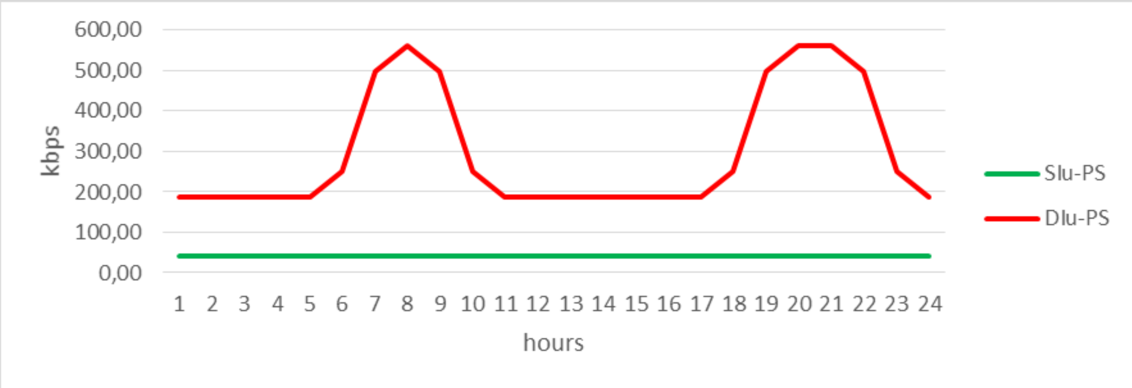


Chart 7 – Electricity Metering case study: results for Iu-PS interface

The same results are achieved by using SOTA, as depicted on Figure 80.

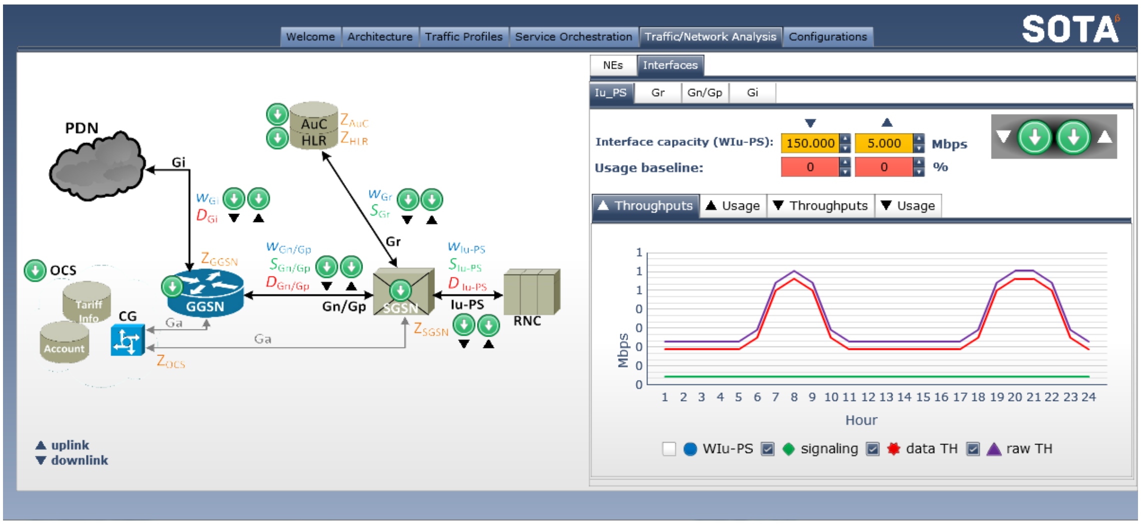


Figure 80 – Iu-PS interface report for Electricity Metering

Gn/Gp Interface

The overhead ratio of Gn/Gp interface is obtained by applying Table 17 values to Equation 6:

$$RO_{Gn/Gp} = \frac{S_{Packet} + H_{GTP} + H_{UDP} + H_{IP}}{S_{Packet}}$$

$$RO_{Gn/Gp} = \frac{130 + 12 + 8 + 20}{130}$$

$$RO_{Gn/Gp} = 1{,}31$$

Therefore, the throughput of GTP-U (data plane) in Gn/Gp interface is given by Equation 7. Results presented on Chart 8.

$$D_{Gn/Gp} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Gn/Gp} * f_d$$

$$D_{Gn/Gp} = 3800 * 1{,}0 * R_{\frac{Active}{Attach}} * 106{,}67 * 1{,}31 * 1{,}1$$

The signaling load of GTP-C in Gn/Gp interface is given by Equation 8. The result is constant for the 24 hours of the day.

$$S_{Gn/Gp} = N_N * R_{Attach} * \sum_{i=1}^{2}(N_{Gn_i} * L_{Gn_i}) * 8 * \frac{1}{3600} * f_s$$

$$S_{Gn/Gp} = 3800 * 1{,}0 * 700 * 8 * \frac{1}{3600} * 1{,}0$$

$$S_{Gn/Gp} = 5{,}91\ kbps$$

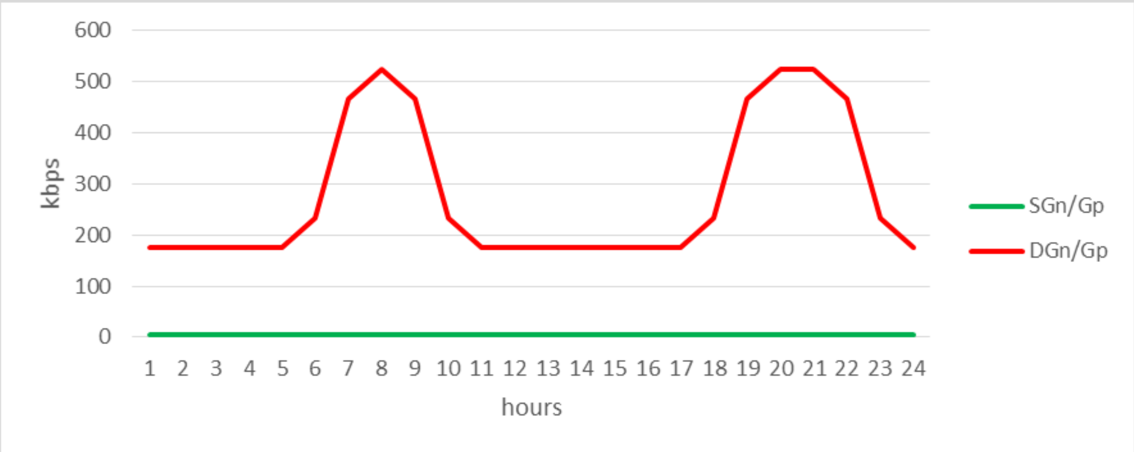


Chart 8 – Electricity Metering Case Study: Results for Gn/Gp interface

Gr Interface

The signaling load of Gr interface is given by Equation 10:

$$S_{Gr} = N_N * R_{Attach} * \sum_{i=1}^{3} \left(R_{Gr_i} * N_{Gr_i} * L_{Gr_i}\right) * 8 * \frac{1}{3600} * f_s$$

$$S_{Gr} = 3800 * 1{,}0 * 691{,}6 * 8 * \frac{1}{3600} * 1{,}0$$

$$S_{Gr} = 5{,}84\ kbps$$

Gi Interface

Considering that IP header carried by MPLS is 14 bytes [168], the ratio of overhead in Gi interface can be obtained by Equation 11:

$$RO_{Gi} = \frac{S_{Packet} + H_{datagram}}{S_{Packet}}$$

$$RO_{Gi} = \frac{130 + 14}{130}$$

$$RO_{Gi} = 1{,}11$$

Therefore the throughput in Gi interface is given by Equation 12.
Results are presented on Chart 9.

$$D_{Gi} = N_N * R_{Attach} * R_{\frac{Active}{Attach}} * Th_{Node} * RO_{Gi} * f_d$$

$$D_{Gi} = 3800 * 1{,}0 * R_{\frac{Active}{Attach}} * 106{,}67 * 1{,}11 * 1{,}1$$

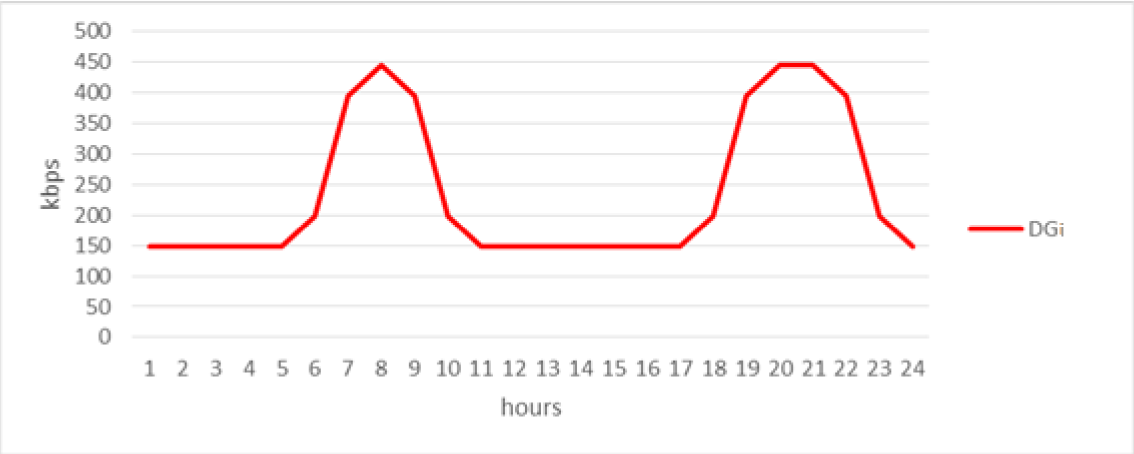


Chart 9 – Electricity Metering case study: results for Gi interface

7.1.5 Sensitivity Analysis

In order to understand which service characteristics are of bigger importance, a sensitivity analysis is presented. This analysis was conducted by varying several parameters of Electricity Metering, for the Uplink connection, having as base parameters:

- 8.500 Aggregation Nodes (ANs);
- Transmission Interval: 30 minutes;
- RAuthHLR: 20%;
- Average Packet size: 130 bytes;
- Gr uplink capacity: 1 Mbps;
- Other interfaces uplink capacity: 10Mbps.

AN transmission interval	***peak Iu-PS usage***	***peak Iu-PS control TH***	***peak Gr usage***	***peak Gn/Gp usage***	***peak Gn/Gp control TH***	***peak Gi usage***
1 min	31,42%	2700 kbps	15,70%	400 kbps	34,93%	9,94%
10 min	15,30%	270 kbps	12,14%	40 kbps	3,50%	9,94%
30 min	13,49%	93 kbps	12,00%	13 kbps	1,16%	9,94%
60 min	13,04%	50 kbps	11,80%	7 kbps	0,58%	9,94%
1440 min	12,60%	< 0,1 kbps	11,74%	< 0,1 kbps	0,02%	9,94%

Table 27 – Varying Aggregation Node transmission interval

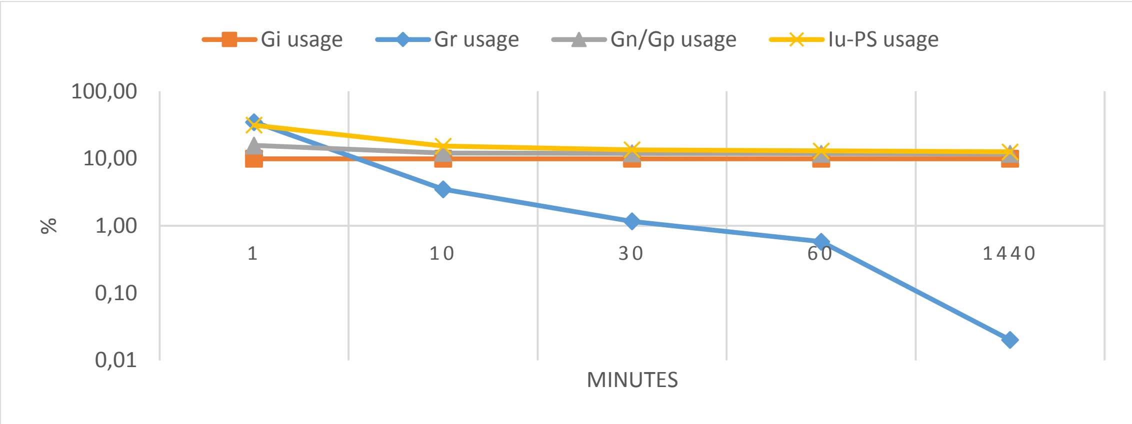


Chart 10 – Varying Aggregation Node transmission interval – Capacity usage

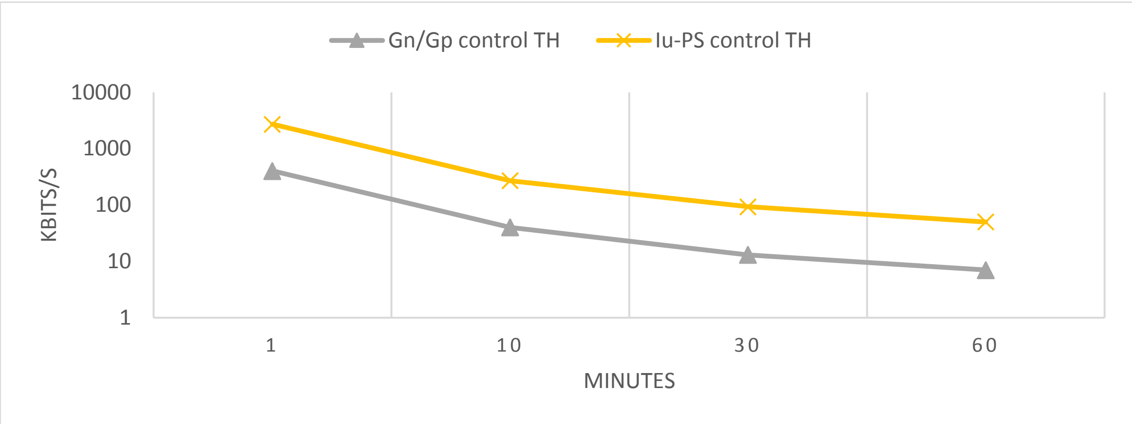


Chart 11 – Varying Aggregation Node transmission interval – Control throughput

R Auth HLR	*peak Iu-PS usage*	*peak Iu-PS control TH*	*peak Gr usage*	*peak Gn/Gp usage*	*peak Gn/Gp control TH*	*peak Gi usage*
10%	13,49%	93 kbps	11,87%	13 kbps	1,14%	9,94%
30%	13,49%	93 kbps	11,87%	13 kbps	1,19%	9,94%
50%	13,49%	93 kbps	11,87%	13 kbps	1,24%	9,94%
70%	13,49%	93 kbps	11,87%	13 kbps	1,30%	9,94%
100%	13,49%	93 kbps	11,87%	13 kbps	1,38%	9,94%

Table 28 – Varying the ratio of authentication that needs to obtain parameters from HLR

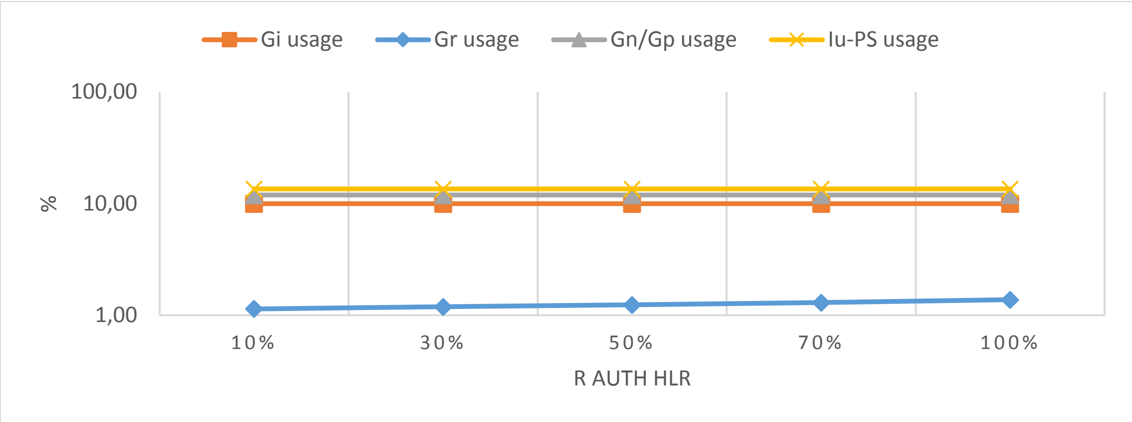


Chart 12 – Varying the ratio of authentication that needs to obtain parameters from HLR – Capacity usage

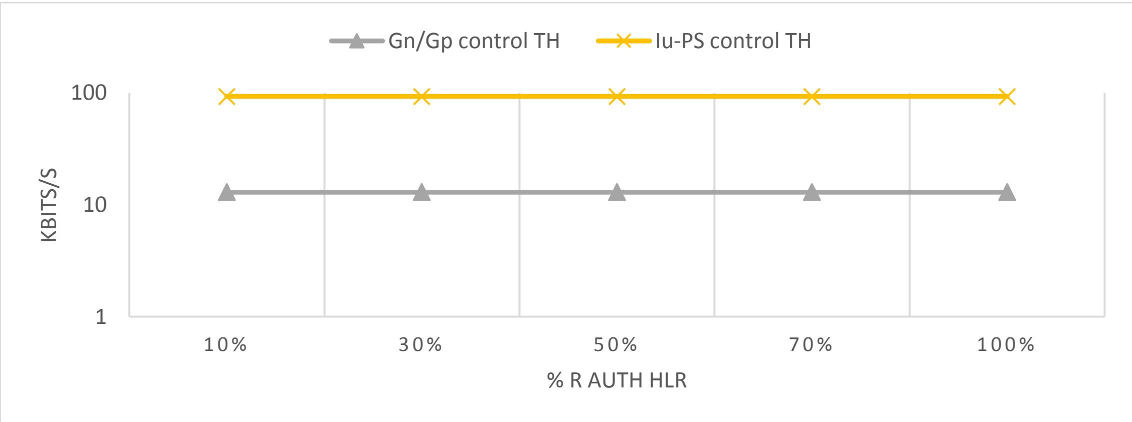


Chart 13 – Varying the ratio of authentication that needs to obtain parameters from HLR – Control throughput

Avg Packet Size	peak Iu-PS usage	peak Iu-PS control TH	peak Gr usage	peak Gn/Gp usage	peak Gn/Gp control TH	peak Gi usage
100 bytes	14,57%	93 kbps	12,70%	13 kbps	1,16%	10,23%
300 bytes	11,46%	93 kbps	10,31%	13 kbps	1,16%	9,40%
700 bytes	10,57%	93 kbps	9,62%	13 kbps	1,16%	9,16%
1000 bytes	10,37%	93 kbps	9,47%	13 kbps	1,16%	9,10%
1500 bytes	10,21%	93 kbps	9,35%	13 kbps	1,16%	9,06%

Table 29 – Varying the average packet size

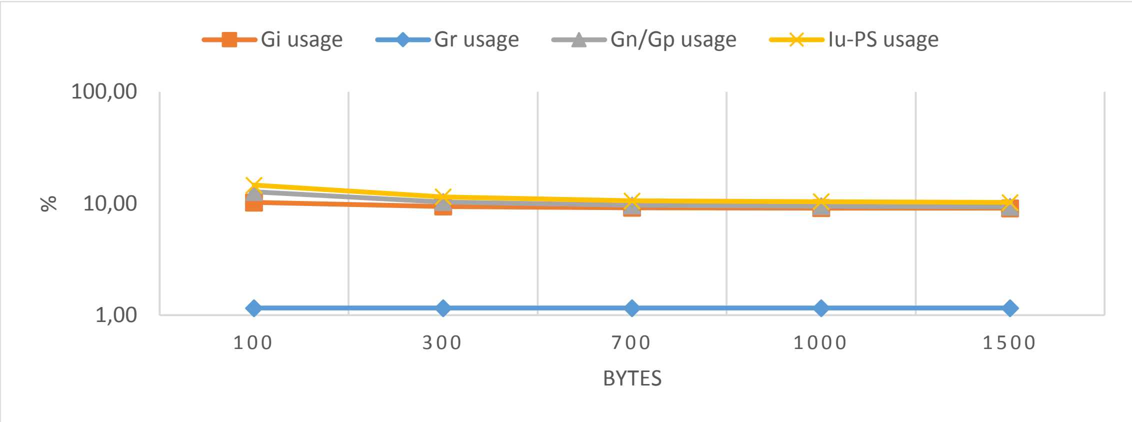


Chart 14 – Varying the average packet size – Capacity usage

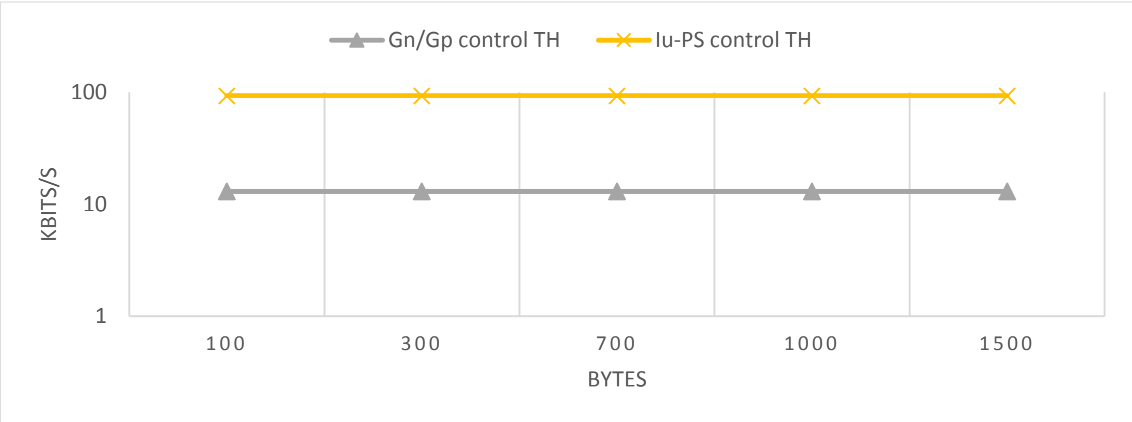


Chart 15 – Varying the average packet size – Control throughput

SMS number (MO/MT)	peak Iu-PS usage	peak Iu-PS control TH	peak Gr usage	peak Gn/Gp usage	peak Gn/Gp control TH	peak Gi usage
0	13,48%	92 kbps	11,87%	13 kbps	1,16%	9,94%
1	13,86%	130 kbps	11,87%	13 kbps	1,16%	9,94%
5	15,41%	285 kbps	11,87%	13 kbps	1,16%	9,94%
10	17,34%	478 kbps	11,87%	13 kbps	1,16%	9,94%
100	52,08%	3953 kbps	11,87%	13 kbps	1,16%	9,94%

Table 30 – Varying the number of SMSs / hour

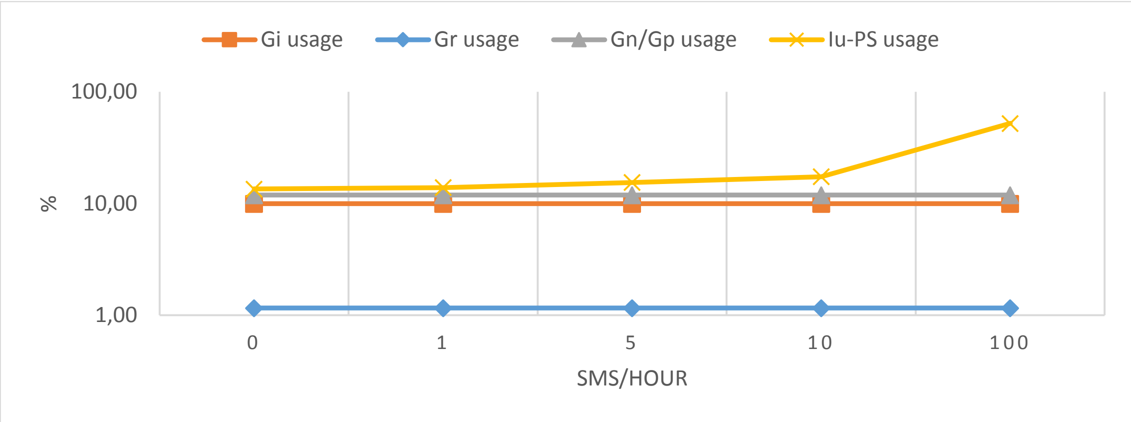


Chart 16 – Varying the number of SMSs / hour – Capacity usage

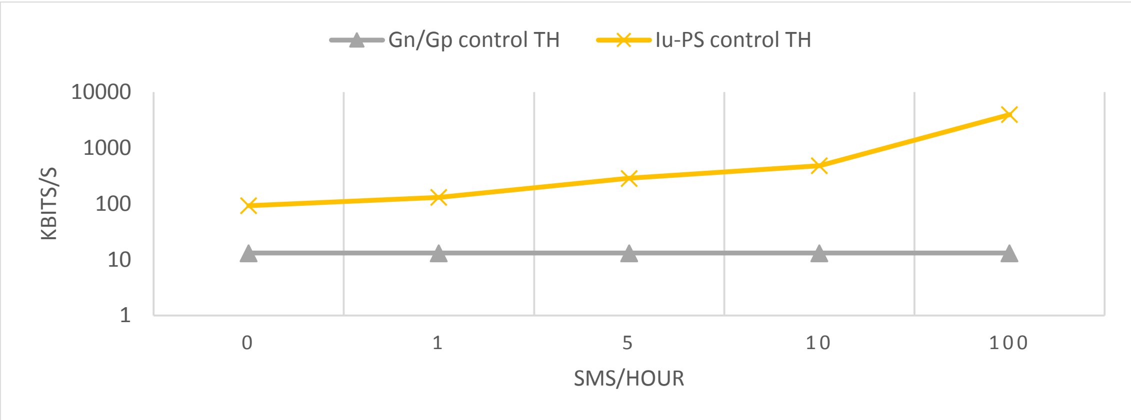


Chart 17 – Varying the number of SMSs / hour – Control throughput

From this analysis some conclusions can be derived:

- **Transmission intervals:** High frequency transmissions (small transmission intervals) imply high Gr interface usage, mainly in control plane. It also imply a not so dramatic increase on Iu-PS and Gn/Gp interface usage;
- **Ratio of authentication that needs to obtain parameters from HLR:** As expected, more authentication requiring parameters from HLR will imply higher Gr usage;
- **Average Packet size:** Bigger Packets mean lower usage of all the three data interfaces (Iu-PS, Gn/Gp, Gi), since the duration of data transmission decreases;
- **Number of SMSs per hour:** A larger number of SMSs per hour implies an increase on Iu-PS usage, mainly on control plane.

7.2 Deploying M2M Services over an existing UMTS Network

This Section provides a study on how the deployment of M2M services could affect an existing cellular network, in this particular case, a UMTS network.

The market scenario to consider is such that there are three telecommunication operators operating in this area, each one owning its wired access infrastructure. The study is elaborated on the perspective of the Incumbent operator, which also owns a wireless access network.

The objective of the incumbent will be to take advantage of the disruptive nature of M2M services and provide such services not only to its customers but also to users who are currently customers of other operators. Thus it can propose to those clients to become incumbent's subscribers in exchange for some sort of discount, or just because the other operators are not providing M2M services. But such users might not want to change operator for a number of reasons such as loyalty contracts, specific TV/VOD content, lack of wired infrastructure by the incumbent, or because they are just too lethargic to switch operator.
For these cases the incumbent could provide M2M connectivity using its cellular network, thus the customer would have separate operators and billing accounts for M2M services and other services such as triple-play. When the incumbent already has a wired access to the customer's premises then it would be advisable to provide M2M connectivity using wired access technologies (It is believed that the impact of M2M communications will be a source of greater concern for wireless access technologies than for wired ones, but that is something that has not been yet studied. Refer to Chapter 9 for suggestions of future work on this subject).

Figure 81 depicts the case study scenario.

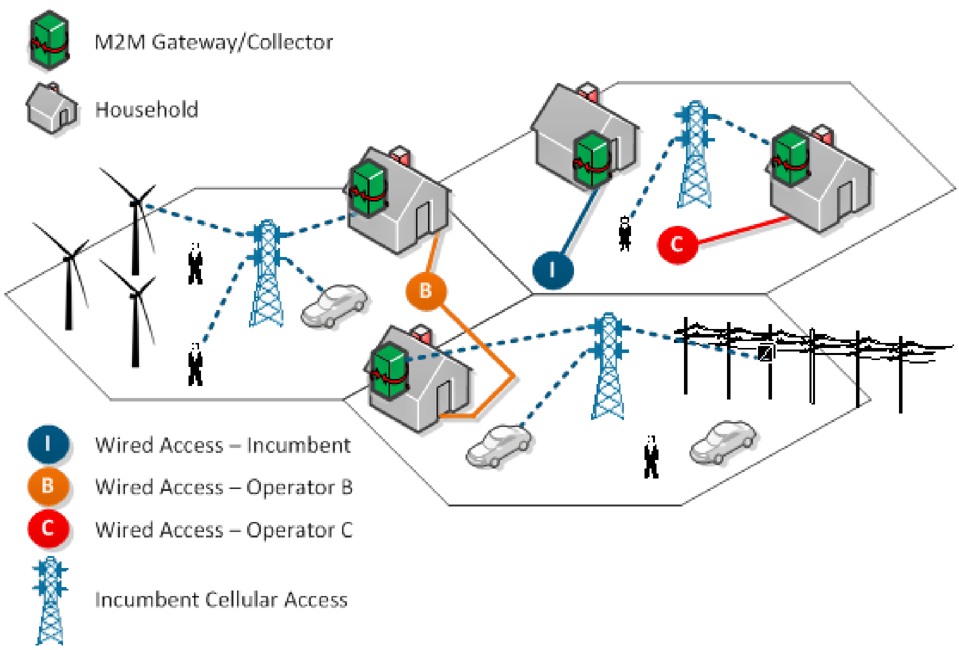


Figure 81 – Case study scenario

7.2.1 Defining the Network Usage Baseline

The first step on this study is to define the capacity and baseline usage of the network at study. In order to do so, let us apply the models presented on Section 5.8, to the same services presented at that Section. Their orchestration parameters are presented on Table 31.

	Web Browsing[32]	*E-Mail*	*Audio Streaming*	*Video Streaming*	*App Update*	*Instagram*	*Facebook*	*Other Social Networks*	*Instant Messaging*	*VOIP*
Number of connections (x 1000)	100	100	100	100	100	100	100	100	100	100
Sessions / Hour	2	2	6	0,5	0,2	2	2	1	10	1,5
Attachments / Hour	1,00	1,00	0,00	0,50	0,00	1,00	1,00	0,00	3,00	1,00
Authentications / Hour	2,00	2,00	6,00	0,50	0,20	2,00	2,00	1,00	10,00	1,50
Intra SGSN route updates / Hour	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Inter SGSN route updates / Hour	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
SMSs MO / Hour	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SMSs MT / Hour	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Authentication rate (HLR)	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
% Premium subscribers	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%
% Regular subscribers	60%	60%	60%	60%	60%	60%	60%	60%	60%	60%
% Basic subscribers	20%	20%	20%	20%	20%	20%	20%	20%	20%	20%

Table 31 – Baseline service orchestration parameters

Throughput values for these services were defined on Section 5.8, where Table 21 explains which are the mathematical variables correspondent to the first column of Table 31.
Other services such as Circuit Switched voice calls, SMSs, and other Packet Switched services are considered on a baseline network usage.

At this point capacity values for the UMTS Packet Core Network model, depicted on Figure 65 of Section 5.7, must be established. Such values are assumed to be as defined to Table 32.

[32] As na exemple, Figure 71 depicts the orchestration of Web Browsing on SOTA.

Variable	Symbol[33]	Capacity	Baseline Usage[34]
Iu-PS interface raw throughput	w_{Iu-PS}	UL: 5.000 Mbps DL: 150.000 Mbps	UL: 30% DL: 30%
Gr interface raw throughput	w_{Gr}	UL: 1,30 Mbps DL: 1,30 Mbps	UL: 40% DL: 40%
Gn/Gp interface raw throughput	$w_{Gn/Gp}$	UL: 5.000 Mbps DL: 150.000 Mbps	UL: 30% DL: 30%
Gi interface raw throughput	w_{Gi}	UL: 5.000 Mbps DL: 120.000 Mbps	UL: 20% DL: 20%
SGSN attached subscribers	$AtSUB_{SGSN}$	180.000	10%
SGSN active subscribers	$AcSUB_{SGSN}$	150.000	10%
SGSN Iu-PS control throughput	$SIG_{Iu-PS_{SGSN}}$	UL: 5,00 Mbps DL: 1000,00 Mbps	-
SGSN Gr control throughput	$SIG_{Gr_{SGSN}}$	UL: 1,30 Mbps DL: 2,00 Mbps	-
SGSN Gn/Gp control throughput	$SIG_{Gn/Gp_{SGSN}}$	UL: 30,00 Mbps DL: 1000,00 Mbps	-
SGSN Iu-PS raw throughput	$W_{Iu-PS_{SGSN}}$	UL: 5.000 Mbps DL: 130.000 Mbps	-
SGSN Gn/Gp raw throughput	$W_{Gn/Gp_{SGSN}}$	UL: 5.000 Mbps DL: 130.000 Mbps	-
GGSN attached subscribers	$AtSUB_{GGSN}$	260.000	30%
GGSN active subscribers	$AcSUB_{GGSN}$	150.000	15%
GGSN Gn/Gp control throughput	$SIG_{Gn/Gp_{GGSN}}$	UL: 30,00 Mbps DL: 1000,00 Mbps	-
GGSN Gn/Gp raw throughput	$W_{Gn/Gp_{GGSN}}$	UL: 5.000 Mbps DL: 150.000 Mbps	-
GGSN Gi raw throughput	$W_{Gi_{GGSN}}$	UL: 5.000 Mbps DL: 120.000 Mbps	-
HLR attached subscribers	$AtSUB_{HLR}$	200.000	20%
HLR active subscribers	$AcSUB_{HLR}$	150.000	5%
HLR control throughput	$SIG_{Gr_{HLR}}$	UL: 1,30 Mbps DL: 1,30 Mbps	-
AuC attached subscribers	$AtSUB_{AuC}$	180.000	10%
AuC active subscribers	$AcSUB_{AuC}$	150.000	5%
AuC control throughput	$SIG_{Gr_{AuC}}$	UL: 1,30 Mbps DL: 1,30 Mbps	-
OCS attached subscribers	$AtSUB_{OCS}$	500.000	60%
OCS active subscribers	$AcSUB_{OCS}$	350.000	50%

Table 32 – UMTS Packet CN model parameters

[33] Symbols referring to the model presented on Section 5.7.
[34] It is considered that 1,5% of each baseline ratio refers to signaling.

7.2.2 Defining the M2M Services to deploy

The network will now support a group of M2M services, which are:
- Electricity Metering;
- Home Automation;
- Healthcare;
- Automotive;
- Security;
- Logistics.

The throughput values for these services are assumed to be as presented on Table 33.

		Electricity Metering	Home Automation	Healthcare	Automotive	Security	Logistics
SLA		**Aggregation Node/Device throughput [kbps]**					
Premium (20%)	Downstream	0,50	0,50	0,50	0,50	0,50	0,50
	Upstream	0,50	0,50	1,00	0,50	10,00	0,50
Regular (60%)	Downstream	0,50	0,50	0,50	0,50	0,50	0,50
	Upstream	0,50	0,50	1,00	0,50	5,00	0,50
Basic (20%)	Downstream	0,50	0,50	0,50	0,50	0,50	0,50
	Upstream	0,50	0,50	0,50	0,50	1,00	0,50

Table 33 – M2M services SLA parameterization

Similarly to baseline services, M2M services must be also orchestrated. Considered orchestration parameters are presented on Table 34. Figure 72 exemplifies the orchestration process using SOTA.

	Electricity Metering	*Home Automation*	*Building Management*	*Healthcare*	*Automotive*	*Public Transportation*	*Security*	*Logistics*
number of connections	3.800			5.000	10.000	500	5.000	700
Transmission interval	UL: 30 min DL: 300 min	UL: 12 hrs DL: 4,8 hrs		-	UL: 2,4 hrs DL: 2,4 hrs	UL: 1 min DL: 1 min	-	UL: 10 min DL: 30 min
Class A device transmission interval	-			UL: 1 min DL: 1440 min	-		UL: 5 min DL: 60 min	-
Class B device transmission interval	-			UL: 5 min DL: 1440 min	-		UL: 15 min DL: 60 min	-
Class C device transmission interval	-			UL: 60 min DL: 1440 min	-			
Class D device transmission interval	-			UL: 1440 min DL: 1440 min	-			
Class A device distribution	-			20%	-		50%	-
Class B device distribution	-			10%			50%	-
Class C device distribution	-			20%	-			
Class D device distribution	-			50%	-			
RAU intra SGSN per hour (UL/DL)	0,00			0,03	0,07	1,00	0,10	1,00
RAU inter SGSN per hour (UL/DL)	0,00			0,01	0,03	0,01	0,01	0,10
SMSs MO per hour (UL/DL)	0,00	0,20	0,00	-	0,00		1,00	0,00
SMSs MT per hour (UL/DL)	0,08	0,10		-	0,80	0,00	1,00	0,00
SMSs MO (class A devices)	-			UL: 0 / DL: 0	-			
SMSs MT (class A devices)	-			UL: 1 / DL: 0	-			
SMSs MO (class B devices)	-			UL: 0 / DL: 0	-			
SMSs MT (class B devices)	-			UL: 1 / DL: 0	-			
SMSs MO (class C devices)	-			UL: 0 / DL: 0	-			
SMSs MT (class C devices)	-			UL: 1 / DL: 0	-			
SMSs MO (class D devices)	-			UL: 0 / DL: 0	-			
SMSs MT (class D devices)	-			UL: 1 / DL: 0	-			
Ratio Auth. (HLR)	20%							

Table 34 – M2M services orchestration parameters

7.2.3 M2M Implications

Now that every input has been established, SOTA can be used to present M2M implications on the network.

Considering the previously defined baseline, the network before supporting M2M services presents capacity usage as depicted on Figure 82.

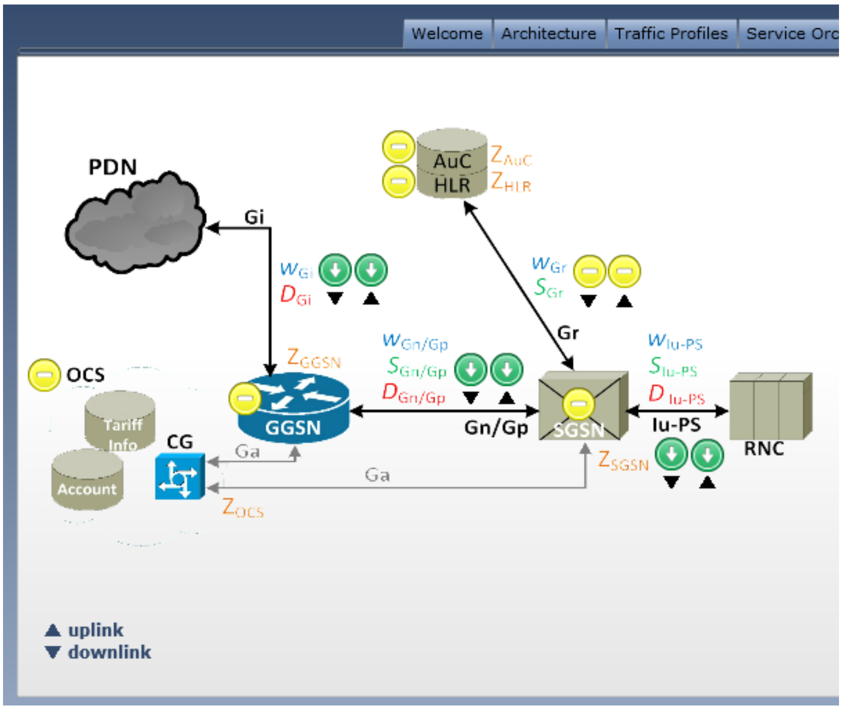


Figure 82 – Network report before adding M2M services

As it can be seen at this point there are already some elements and interfaces that have less than 20% available capacity.

If Electricity Metering is added, there will be no changes in what concerns the network reports since the capacity usage will increase but not enough to exceed the limits that would change indicators status.

But if all the six M2M services at study are added to the network, then the situation changes dramatically, and it becomes visible (Figure 83) that although the network is still able to supply the services (Analyzing the NEs/Interfaces reports it can be seen that every indicator is below 100%), is dangerously close to reach its maximum capacity.

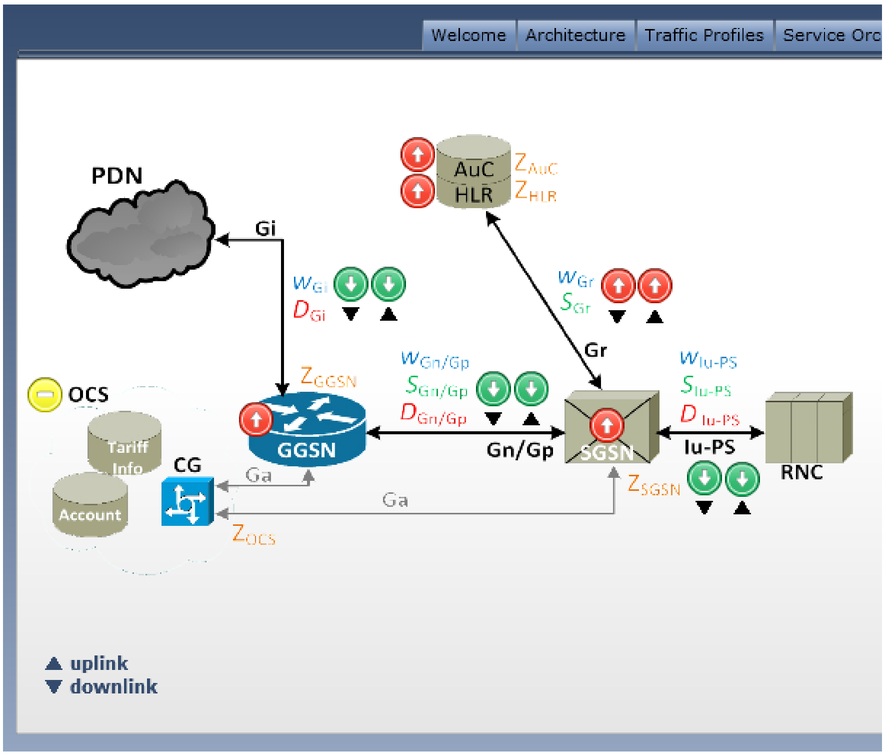


Figure 83 – Network report after adding M2M services

Not surprisingly, control plane and subscriber capacity present the critical capacity usages.

The following charts present some comparisons between the network usage before (left) and after (right) adding M2M services, for the situations where the capacity is close to be exhausted.

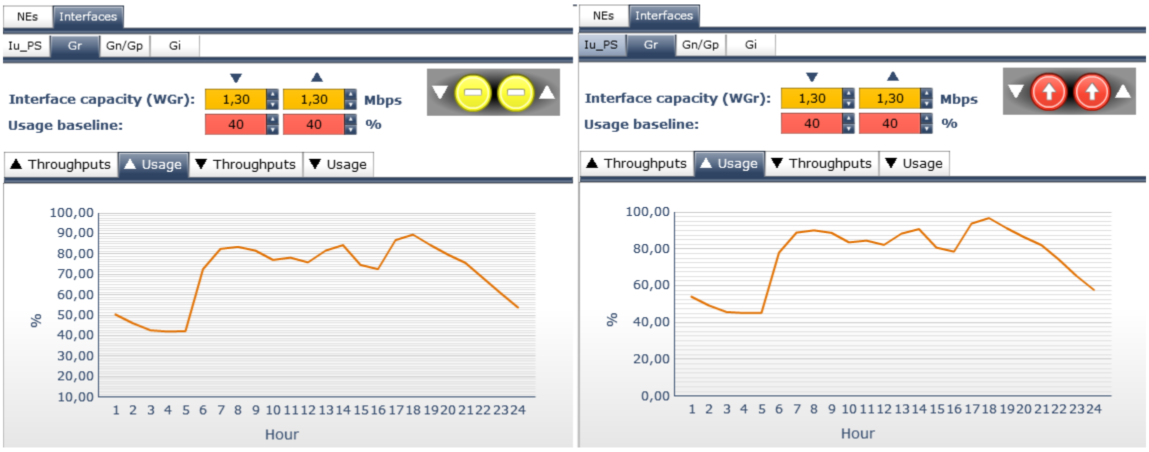


Figure 84 – Gr usage report (uplink)

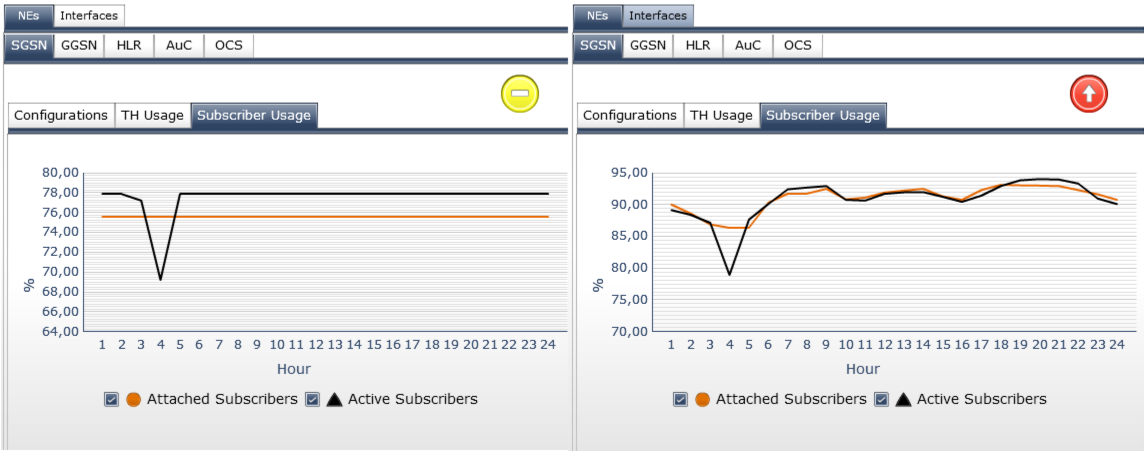


Figure 85 – SGSN subscriber usage report

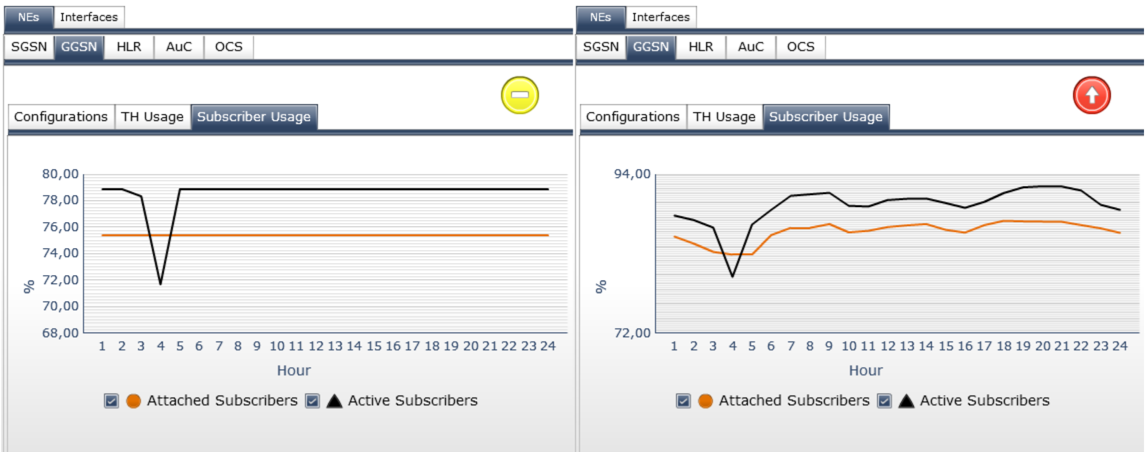


Figure 86 – GGSN subscriber usage report

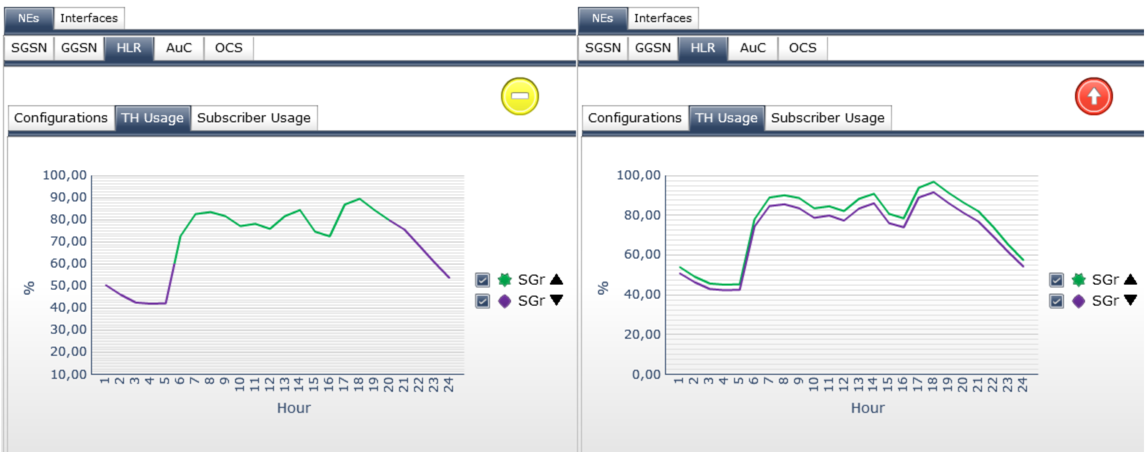


Figure 87 – HLR usage (throughput) report

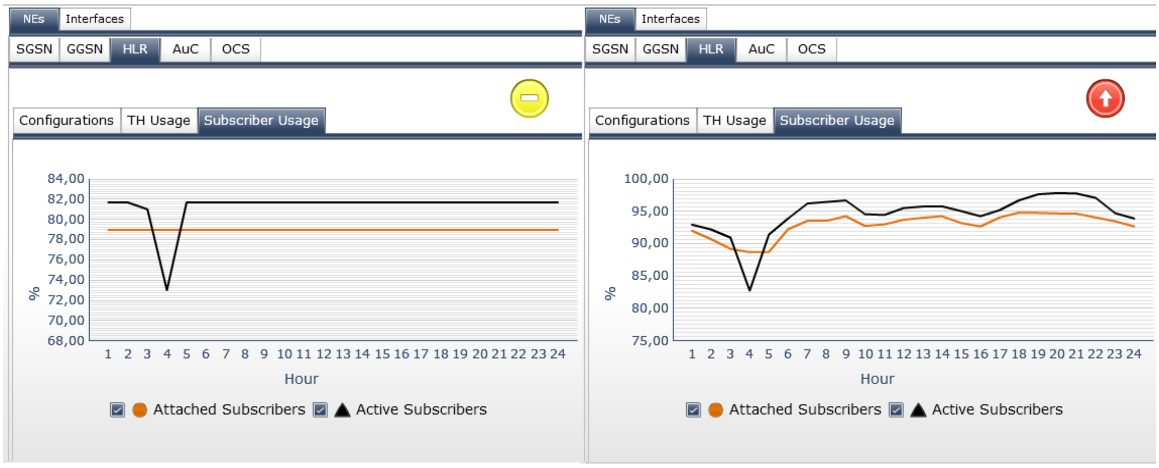


Figure 88 – HLR subscriber usage report

Figure 89 – AuC usage (throughput) report

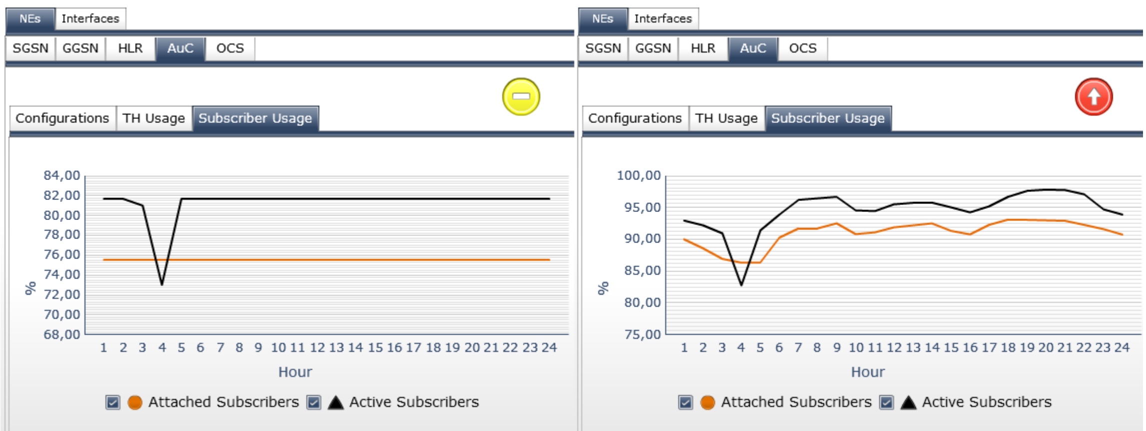


Figure 90 – AuC subscriber user report

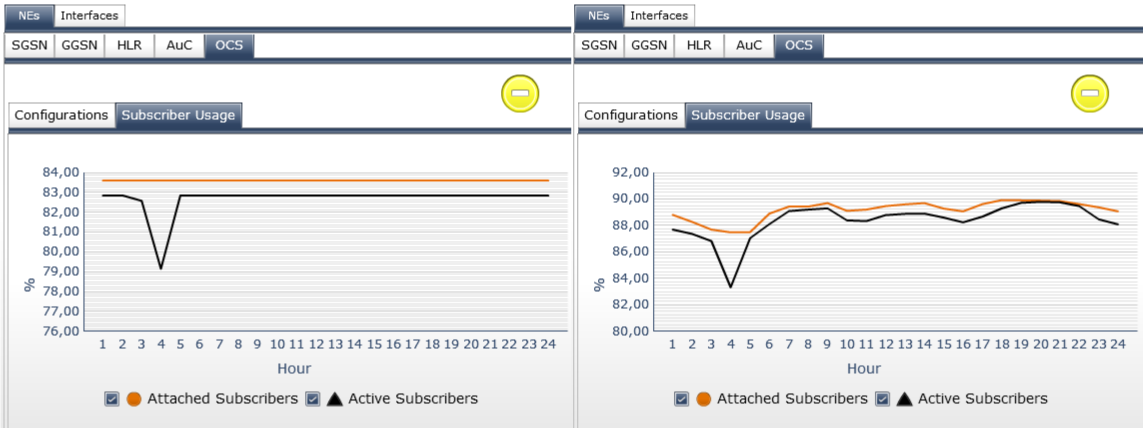


Figure 91 – OCS subscriber capacity report

7.3 Summary

This Chapter provides a case study that illustrates how to apply the models that are presented through this work and also how to apply the tool. The tool functions and utility are depicted, while achieved results, with particular focus on the sensibility analysis presented in 7.1.5, provide the framework for some conclusions over the subject at study: the impacts that M2M services will have on cellular networks, and which are their critical design parameters.

8. Conclusions

This dissertation addresses the influence of M2M communications in cellular networks. Issues like the specific characteristics of this traffic (as compared to conventional data or voice traffic) and the number of connected devices and subscriptions (which will most likely increase intensely over the course of the next years) will pose severe challenges to Mobile Network Operators (MNOs).
In order to study this issues, analytic models and computational tools were developed.

It was demonstrated the need for future cellular networks developments in order to better accommodate this wave of M2M traffic. Fundamentally is necessary that both new and legacy networks present increases on signaling, subscription, session, and throughput capacity.

On Chapter 4 some fundamental indicators and performance metrics to study this problem were presented. Subsequently it were presented and discussed a number of Network Elements (NE) whose performance or capacity can be of critical importance on the overall network performance and the subscribers Quality of Experience (QoE). In the main technological context of this study (3G/UMTS) it can be concluded that NEs that are responsible for functions involving strong loads of signaling (e.g. session establishment, authentication, authorization…) are most likely to cause severe problems of congestion and service degradation.

Models developed to calculate the values of throughput (on both control and data plane) on the UMTS Packet Core proven to be very useful on designing an analytic tool to model the effects of M2M services on a UMTS network. This tool provides a way to orchestrate services (M2M, or of other nature) while simultaneously model the impacts of such services on the UMTS Packet Core network. This is of great usefulness for MNOs since it provides a way to understand which options are more adequate to sustain network performance when designing M2M services and applications. Fundamentally it must always be considered that increasing the number of devices implies exponential demand on sessions, addresses and signaling capacity. At the same time, the number of transmissions generated by those devices per time unit can be a fundamental metric to ensure that M2M services will be successfully accommodated by the network.

The case study provided on Chapter 0 shows how proposed models can be applied to study this problem, and how an analytic tool as SOTA can simplify that work. More than its results, the fundamental value of the work presented on this Chapter is on the processes and models that are depicted. Subsection 7.1.5 provided useful information to understand which could key design decisions be in order to ensure service sustainability, such as balanced transmission intervals, adequate packet size and other relevant parameters.

Section 7.2 proposes a new business model for MNOs wanting to take advantage on early M2M deployments, and depicts the process to – with resort to the analytic tool – study and design M2M services and their implications on cellular networks.

It must be stated that much of the work presented in this document, although focused on 3G/UMTS, is applicable to different technologies and different types of traffic. Being of particular relevance to study the effects of data generated by smartphones, in particular the signaling generated by social network surging usage.

In order to apply the models and tool developed through this work, a large number of assumptions had to be taken, concerning multiple parameters. Obviously the accuracy of those models and corresponding tool depend on these parameters, and their real values should be further investigated.

9. Future Work

The developed analytic models and computational tool must be validated and real values for the parameters considered throughout this work must be further investigated. Such tasks would require the support of telecommunications operators by providing real information for model parameterization, which would allow model and tool calibration, and subsequent validation.

Considering that this work covers a broad range of topics, there are multiple possibilities for future work related to it:

- The analysis presented on Chapter 4 can be detailed for other Network Elements or Systems, and can be extended to LTE technology;
- Models presented over Chapter 5 (Section 5.3 to 5.7) can be extended to other UMTS interfaces, an also to LTE implementations;
- The Radio Access was not considered in this study, therefore it would also be pertinent to study that part of the network in the present context;
- Important aspects such as session number and characteristics, scheduling and buffer strategies (refer to pp. 211-219 [138]), and KPI (refer to [169]) analysis can also be further detailed;
- It could be an interesting approach for techno-economic analysis to consider cost variables for interfaces and Network Elements into the model presented on Section 5.7;
- A detailed study on M2M authentication requirements could also be an interesting development on this work. Authentication configuration options will most likely be of fundamental importance to the performance of telecommunications networks, it would be of interest to study if every M2M device will need authentication, if such authentication can be executed by VPN, and where should the authentications processes be executed (since many authentication functions could be done on the devices/SIMs, which would unload the network, but would implicate energy consequences to the devices;
- Models proposed on Section 5.8 can be further developed, and extended to other types of traffic, such as SMSs and Circuit Switched traffic;
- Statistical processes could also be of use for such models, with particular relevance for the study of overbooking mechanisms for M2M services;
- As proposed on the case study of Section 7.2 it would also be interesting to understand which will be the implications that M2M will pose to wired access networks.

Additionally, Section 9.1 presents some proposals for further developments on SOTA, while Section 9.2 presents some considerations on LTE related developments.

9.1 Developments on SOTA

Future developments on SOTA could be:

- Introduction of Latency metrics;
- Analysis on the number of streams;
- Energy consumption metrics (for CN, access network, and M2M devices);
- Increased detail on baseline usage, with application of usage profile (variable baseline over the day);
- Applicability to other technologies.

9.2 Evolution to 4G/LTE

9.1.1 Signaling

Although LTE implements many enhancements over 3G systems many of the stated problems will be applicable, if not increase on importance. As presented on Chart 6 the average signaling requirement per subscriber is up to 42% higher with LTE than with HSPA, and on Chart 0 it can be seen the forecast of Nokia Siemens Networks for signaling load up to 2015.

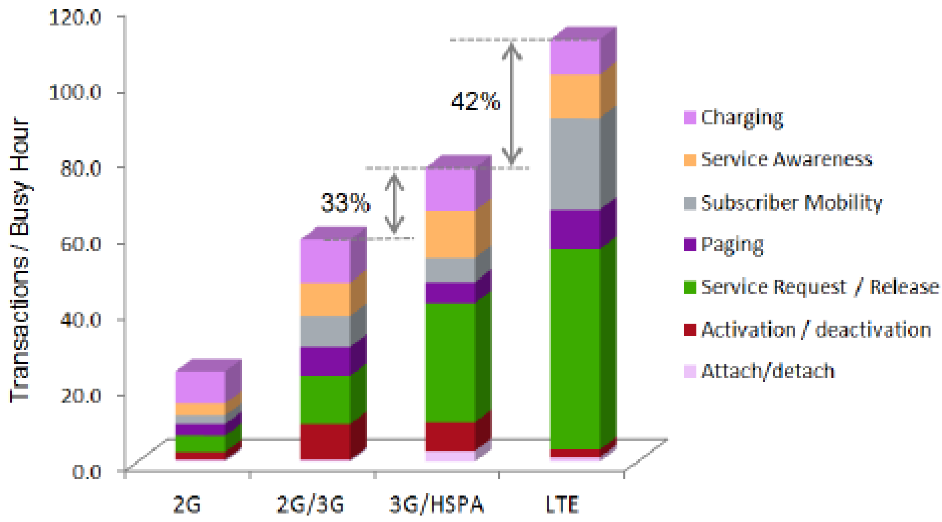


Chart 18 – Average signaling requirement per subscriber [18]

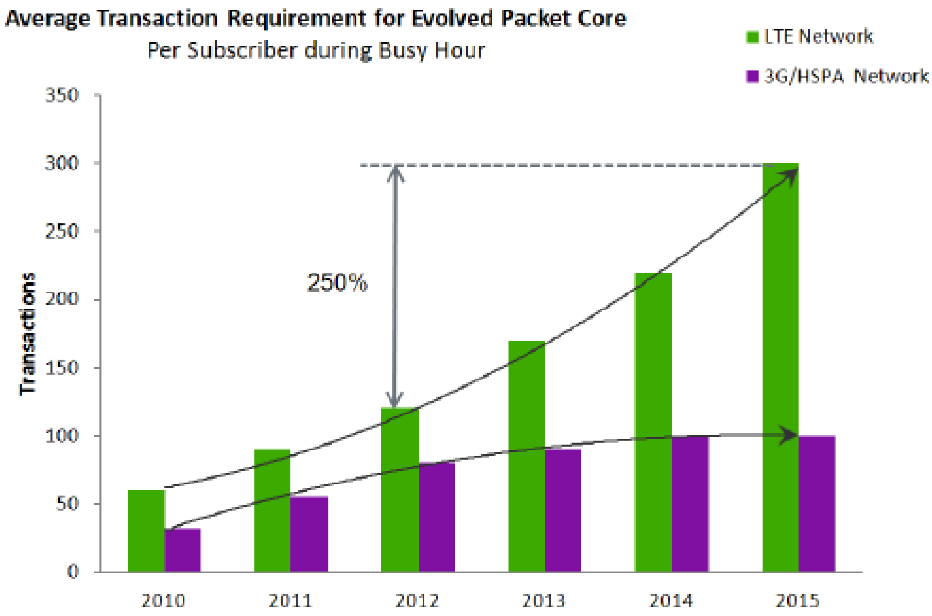


Chart 19 – By 2015 the average number of signaling transactions in the Core Network will increase by 250% [18]

In terms of critical elements on LTE networks is to expect that the most affected element will be the MME, followed by the Serving-GW and PDN-GW, as described on Chart 20, where it can be seen that transactions comprise a different number of messages depending on the network element. For example, one transaction is equivalent to 10 messages in the MME, and 2 messages in the PDN-GW [18].

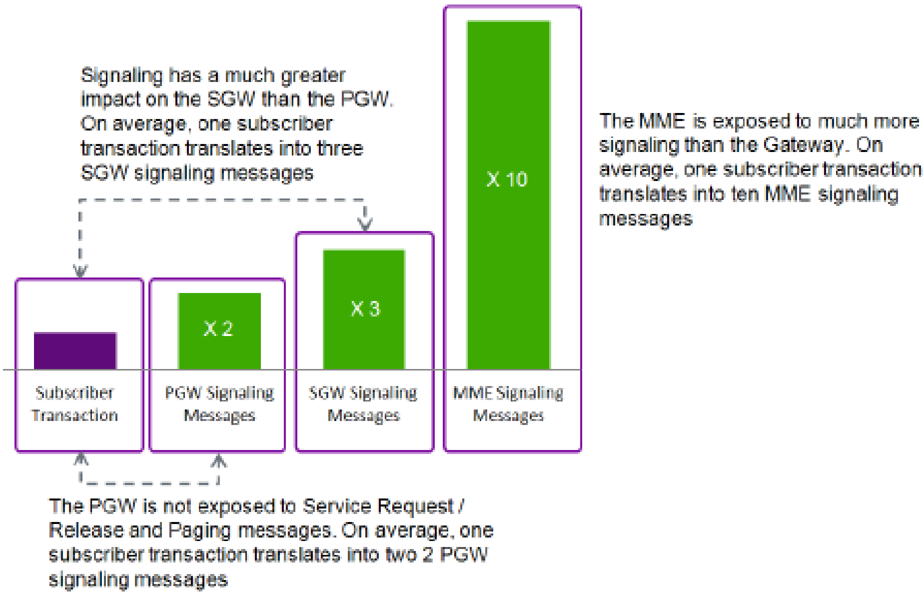


Chart 20 – Signaling on Evolved Packet Core networks [18]

9.1.2 Resource Reservation Mechanisms for M2M Communication

Regarding new levels of scalability to be achieved in the subscription based handling including mobility and resource management, M. Corici, et all [195] present an interesting alternative to the current architecture. They state that “the current network architectures are designed with the human communication in mind”. Which is verifiable since current networks work on a model were every device has a unique subscription (associated with a unique SIM) corresponding to a unique subscription profile stored on the network. Which implies that “the reservation and release of the resources is synchronous with the session establishment and termination, as prior to that moment the network is not aware that the subscriber will initiate a service and on the level of resources required by that service” [195].

With the growth of machine type communication, many M2M devices will be part of the same service e.g. future fire alarms and temperature sensors on a residence could require to be handled by the network as one unique subscription instead of being handled independently, as they will be with current solutions [195]. Moreover M2M communication has, in many cases, predictable resources which are necessary from the network at specifics (and frequently known) moments in time. Something that cannot be assumed for conventional communications. And so M. Corici, et all [195] propose a “new resource reservation concept based on functional grouping of devices and the predictability of the resources required in the mobile core networks”, which could increase the scalability of current or future networks and sustain some of the impacts caused by the M2M explosion.

References

[1] Parking Reinvented, Streetline, accessed Aug 16, 2013.
[2] "eCall: Time saved = lives saved", European Commission, accessed Aug 13, 2013.
[3] D. Evans, "How the Internet of Everything Will Change the World…for the Better", Cisco Blog, Nov 7, 2012.
[4] "How Social Media Is Changing Disaster Response", Scientific American, Jun 7, 2013.
[5] "How Emergency Managers Can Benefit from Big Data", Government Technology, Jul 23, 2013.
[6] M. Alendal, "Operators need an ecosystem to support 50 billion connections", Ericsson Business Review, no. 3, p. 42, 2010.
[7] "The state of M2M in energy and utilities", Vodafone Group, 2013.
[8] J. Berst, "How utilities are using cellular communications today", Smart Grid News, Mar 6, 2013.
[9] "Cutting energy consumption: Is prepay the way to go?", Smart Grid News, Mar 22, 2013.
[10] "UK trial shows customers do want time-of-use rates", Smart Grid News, Mar 27, 2013.
[11] T. Kerber and M. Duchin, "Utilities forming partnerships to enter the connected home", Fierce Energy, Mar 24, 2013.
[12] Open Meter, accessed Aug 13, 2013.
[13] M. Saksena, "Smart Energy Profile (SEP) 2.0 Uncovered", EE Times, Oct 18, 2011.
[14] Directive 2009/72/EC, European Parliament and the Council, Official Journal of The European Union, L211, p. 91, Jul 13, 2009.
[15] Portaria n.º 231/2013, Ministério da Economia e do Emprego, Diário da República, 1st series, no. 139, pp. 4242-4245, Jul 22, 2013.
[16] A. Ewing, "The BYOD connected car", Feb 27, 2013.
[17] "M2M Report: Connected Car Market 2013-2018", M2M Magazine, Mar 24, 2013.
[18] "Signaling is growing 50% faster than data traffic", Nokia Siemens Networks, 2012.
[19] M. Z. Shafiq, L. Ji, A. X. Liu, J. Pang and J. Wang, "A First Look at Cellular Machine-to-Machine Traffic Large Scale Measurement and Characterization", Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA.
[20] A. Orrevad, "M2M Traffic Characteristics", KTH Royal Institute of Technology, Stockholm, Sweden, 2009.
[21] S. Resnick, "Data Network Models of Burstiness", School of Operations Research and Industrial Engineering - Rhodes Hall, Cornell University, Ithaca NY, USA, Aug, 3, 2005.
[22] J. Mohammed, "Realize the promise of M2M – make it work outside the box", Ericsson Business Reviw no. 3, 2010.
[23] "Enabling M2M for Mobile Network Operators", SimService, 2012.
[24] About 3GPP, 3GPP, accessed Nov 12, 2013.
[25] "Maximizing Mobile Operator Opportunities in M2M", ABI Research, p. 4, 2010.
[26] "Global Machine to Machine Communication", Vodafone Group, p. 4, 2010.
[27] "History of M2M Infographic", Aeris Communications, 2013.
[28] J. Rodriguez, "4G: Long Term Evolution (LTE)", class notes for 41591 – Planeamento de Comunicações Móveis, Department of Electronics, Telecommunications and Informatics, University of Aveiro, p. 77 , 2012.
[29] "Interim Ericsson Mobility Report, August 2013", Ericsson, p. 4, Aug, 2013.
[30] "The M2M adoption barometer 2013", Vodafone Group, pp. 11-12, 2013.
[31] M. Menon, "Mega Trends - Asia Pacific Market Insights", Frost & Sullivan, 2011.
[32] "Device Connectivity Unlock Value", Ericsson, Jan, 2011.
[33] S. Mackey, "Cloud Computing for M2M: Simplifying the Experience", Maxis M2M Business Seminar, Kuala Lumpur, Jan 28, 2011.
[34] "Hype-Cycle", Gartner, accessed Aug 12, 2013.
[35] Hype-Cycle-General.png, Wikimedia Commons, accessed Aug 12, 2013.
[36] M. Hatton M, "The Global M2M Market in 2013", Machina Research, p. 3, Jan, 2013.
[37] A. Shehata, "Mobile MD: How mobile devices will revolutionize healthcare", KPMG, accessed Aug 14, 2013.
[38] About the Alliance, Continua Health Alliance, accessed Aug 14, 2013.
[39] T. Dennis, "Industrial internet will blow M2M out of water", Gomo News, Feb 21, 2013.

[40] P. C. Evans and M. Annunziata, "Industrial Internet: Pushing the Boundaries of Minds and Machines", General Electric, pp. 3-12, Nov 26, 2012.
[41] I. D. Scales, "Using M2M to insure the driving more than the car", Telecom TV, Feb 12, 2013.
[42] D. Taylor, Executive Insight: M2M-driven car insurance, Telecom TV, Feb 21, 2013.
[43] "The state of M2M in consumer goods and retail", Vodafone Group, 2013.
[44] "Connect and extend your site and asset security", Vodafone Group, 2012.
[45] "The state of M2M in transport and logistics", Vodafone Group, 2013.
[46] "A new level of efficiency", Vodafone Group, 2013.
[47] Networked Society, Ericsson, accessed Aug 16, 2013.
[48] Metcalfe's law, Wikipedia, accessed Aug 15, 2013.
[49] Visa payWave for Mobile, Visa, accessed Aug 16, 2013.
[50] A. Bielsa, "Smart Roads – Wireless Sensor Networks for Smart Infrastructures: A Billion Dollar Business Opportunity", Libelium, Jan 11, 2013.
[51] "PT lança projecto piloto para pagar parquímetros via telemóvel", Económico, Oct 29, 2012.
[52] C. Hagen et al., "Big Data and the Creative Destruction of Today's Business Models", A.T. Kearney, Jan, 2013.
[53] G. Hollingworth and J. Hoffman, "Changing the Game Before the Game Changes You", Ericsson, p. 6, 2012.
[54] "The Internet of Things is coming. Now what?", Computer World UK, Nov 26, 2012.
[55] "Big Data in M2M: Tipping Points and Subnets of Things", Machina Research, p. 16, 2013.
[56] M. Jauchius, "Advanced analytics for better decisions", McKinsey & Company Chief Marketing & Sales Officer Forum, Mar, 2013.
[57] J. Swetina, "ETSI M2M / oneM2M and the need for semantics", IoT Week 2012, Venice, p. 5-15, Jun, 2012.
[58] N. Akhtar, "Machine Type Communications - 3GPP Standards Perspective", Centre of Excellence in Wireless Technology, p. 4, Dec 3, 2012.
[59] Introducing the Butler Project, Butler, accessed Aug 19, 2013.
[60] What is DLMS/COSEM, DLMS User Association, accessed Aug 19, 2013.
[61] The Big Conversation: "Too many M2M standards or just not the right ones?", Telecom TV, Nov 15, 2012.
[62] ETSI M2M communications functional architecture, ETSI Technical Specification 102 690, Vers. 1.1.1, p. 15, Oct, 2011.
[63] Domains of Applications, Butler, accessed Aug 19, 2013.
[64] "'Publish and subscribe' protocol to become an M2M standard", Telecom TV, Feb 19, 2013.
[65] MQ Telemetry Transport, accessed Aug 20, 2013.
[66] 3GPP Standardization of Machine-type Communications, 3GPP, Vers. 0.2.0, Jun, 2013.
[67] Scope, oneM2M, accessed Aug 19, 2013.
[68] F. Ennesser, "M2M Security Standards: ETSI contributions", ETSI, p. 3, 2011.
[69] "Hacking attacks can turn off heart monitors", The Register, Mar 12, 2008.
[70] "Insulin pump hack delivers fatal dosage over the air", The Register, Oct 27, 2011.
[71] "High tech car theft: 3 minutes to steal keyless BMWs", Network World, Aug 7, 2012.
[72] "It May Have Been Terrifyingly Easy For A Stranger To Hack Into A Baby Monitor And Curse At A 2-Year-Old", Business Insider, Aug 14, 2013.
[73] "Smart meter hacking can disclose which TV shows and movies you watch", Naked Security, Jan 8, 2012.
[74] "The Internet of Things Has Arrived — And So Have Massive Security Issues", Wired, Nov 11, 2013.
[75] S. Sanislo, "Remote Monitoring With Traditional Phone Lines Has Its Advantages", Electronic Design, Nov 9, 2012.
[76] Google Glass, Google, accessed Aug 20, 2013.
[77] H. F. Hamann, "IBM 5 in 5 2012: Smell", IBM Research, Dec 17, 2012.
[78] IBM Portugal, Dec 19, 2012.
[79] J. Barker, Smart dust and smart specks, accessed Aug 20, 2013.
[80] K. S. J. Pister, "Smart Dust", University of California, Berkeley, USA.
[81] ONT, Meo, accessed Aug 21, 2013.

[82] IEEE 802.3 Ethernet Working Group, IEEE, accessed Aug 21, 2013.
[83] 400 Gb/s Ethernet Study Group, IEEE, accessed Aug 21, 2013.
[84] M-Bus, accessed Aug 21, 2013.
[85] A. M. de Oliveira Duarte and J. G. R. de Figueiredo, “Techno-Economic Analysis of Access Networks: Traffic Dimensioning”, Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal, pp. 16-19, 2013.
[86] Introduction, KNX Association, accessed Aug 21, 2013.
[87] A. Neves e Sousa, A. L. Jesus Teixeira and M. J. Neves de Lima, “Thesis Proposal: Transmission Techniques for High Rates in Optical Fiber and Alternative Media”, Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal, 2011.
[88] R. J. Gandhi, A. Polley and S. E. Ralph, "40 Gbps short reach links using plastic optical fiber", in *Lasers and Electro-Optics Society - 19th Annual Meeting of the IEEE*, 2006.
[89] FAQ – “What is HDMI?”, accessed Aug 21, 2013.
[90] IEEE 1394, Wikipedia, accessed Aug 21, 2013.
[91] BUS/SCS Automation system, Legrand.
[92] USB.org, accessed Aug 21, 2013.
[93] Discover and Learn, Wi-Fi Alliance, accessed Aug 21, 2013.
[94] A. Friedman, “Embedding Wi-Fi—What’s Involved?”, Aug, 2012.
[95] Technical Summary, ZigBee Alliance, accessed Aug 21, 2013.
[96] “Zigbee Alliance launches Green Power feature”, Telecom Engine, Dec 26, 2012.
[97] ZigBee PRO Green Power Feature, ZigBee Alliance, accessed Aug 21, 2013.
[98] ZigBee Resource Guide, accessed Aug 21, 2013.
[99] About Z-Wave, Z-Wave Alliance, accessed Aug 21, 2013.
[100] M. Stefanov and T. ErmesjØ, Application Note AN121, Texas Instruments, 2013.
[101] 169 MHz-Wireless M-Bus, Radiocrafts,
[102] SmartLEWIS TRX, Infineon Technologies, p. 5, Apr 6, 2012.
[103] “Wireless M-Bus Documentation: Reference Manual”, Steinbeis Transfer Center, pp. 6-8, Jan 17, 2011.
[104] “Telit Wireless M-Bus 2010 Part 4 User Guide”, Rev. 7, Telit, p. 9, May 5, 2013.
[105] 6lowpan – Charter for Working Group, IETF, accessed Aug 22, 2013.
[106] G. Mulligan, “The 6LoWPAN architecture”, in *Proc. 4th workshop on Embedded networked sensors,* 2007.
[107] Fast Facts, Bluetooth, accessed Aug 22, 2013.
[108] A Look at the Basics of Bluetooth Wireless Technology, Bluetooth, accessed Aug 22, 2013.
[109] Specification Adopted Documents, Bluetooth SIG, accessed Aug 22, 2013.
[110] “Bluetooth Low Energy”, Litepoint, pp. 2-3, Jun, 2012.
[111] Connect the smart way to the broadest range of Bluetooth devices, Bluetooth, accessed Aug 22, 2013.
[112] Near Field Communication, Android Developers API Guides, accessed Aug 22, 2013.
[113] About NFC, NFC Forum, accessed Aug 22, 2013.
[114] The NFC Forum Certification Program, NFC Forum, accessed Aug 22, 2013.
[115] This is ANT, accessed Aug 22, 2013.
[116] Dash7 Alliance, accessed Aug 22, 2013.
[117] ISA 100 Wireless, accessed Aug 22, 2013.
[118] MiWi Protocol, Microchip, accessed Aug 22, 2013.
[119] One-Net, accessed Aug 22, 2013.
[120] HART Communication Foundation, accessed Aug 22, 2013.
[121] “CableLabs Issues DOCSIS 3.0 Specifications Enabling 160 Mbps”, Cable Labs, 2006.
[122] The M-Bus: Documentation - 4 Physical Layer, M-Bus, accessed Aug 22, 2013.
[123] Technical Questions, Z-Wave, accessed Aug 22, 2013.
[124] “Wireless M-Bus Security”, Compass Security, Jun 30, 2013.
[125] “Telephone line”, Wikipedia, accessed Aug 22, 2013.
[126] “Defining 4G: Understanding the ITU Process for the Next Generation of Wireless Technology”, 3G Americas, p. 5, Aug, 2008.
[127] WiMAX Forum, accessed Aug 22, 2013.
[128] P. Rysavy, “Transition to 4G”, Rysavy Research, pp. 6-8, 2010.

[129] W. Stallings, *Data and Computer Communications*, 8th ed. New Jersey, NY: Pearson Prentice Hall, 2007, p. 267.
[130] "DSL White Paper", Allied Telesis, p. 9, 2007.
[131] "iCON-32333 Series", Axiomtek Co, 2010.
[132] VDSL2 Vectoring, Alcatel-Lucent, accessed Aug 23, 2013.
[133] "Docsis 3.1 Targets 10-Gig Downstream", Light Reading, Oct 18, 2012.
[134] A. M. de Oliveira Duarte, D. C. Carrilho, "Redes Híbridas de Acesso Fixo sem Fios: Análise Tecno-Económica", Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal, pp. 11-20, 2010.
[135] "FTTx Solution White Paper", ZTE.
[136] A. M. de Oliveira Duarte, "Access Network: Tecnologias e Meios de Transmissão", class notes for 41498 – Redes de Acesso, Department of Electronics, Telecommunications and Informatics, University of Aveiro, Aveiro, Portugal, p. 22 , 2012.
[137] Mobile technologies GSM, ETSI, accessed Aug 23, 2013.
[138] M. C. Chuah and Q. Zhang, Design and Performance of 3G Wireless Networks and Wireless LANs, New York, NY: Spinger, 2006, pp. 3-223.
[139] GPRS, GSMA, accessed Aug 23, 2013.
[140] UMTS, 3GPP, accessed Sep 12, 2013.
[141] "HSPA+ Advanced", Qualcomm, p. 2, Feb 14, 2013.
[142] "LTE - Qualcomm Leading the Global Success", Qualcomm, p. 2, Aug, 2013.
[143] "Guidance for use of the LTE logo", 3GPP.
[144] "WiMAX and the IEEE 802.16m Air Interface Standard", WiMAX Forum, Apr, 2010.
[145] WiMAX Open Retail Initiative, WiMAX Forum, accessed Aug 23, 2013.
[146] "DVB-RCS2 Published", DVB, Mar 14, 2011.
[147] "Introduction to Free Space Optical Wireless Networking or FSO", fSONA Systems, July 30, 2010.
[148] D. A. Rockwell and G. S. Mecherle, "Optical Wireless: Low-Cost, Broadband, Optical Access", fSONA Systems.
[149] SONAbeam® Security, fSONA Systems, accessed Jan 5, 2013.
[150] "WHAT ARE OPTICAL LINKS?", fSONA Systems, accessed Aug 24, 2013.
[151] "Heterogeneous Network (Hetnet)", Ericsson.
[152] "Heterogeneous Networks", Ericsson, Feb, 2012.
[153] "The 1000x Mobile Data Challenge", Qualcomm, Jun, 2013.
[154] "AT&T: Wi-Fi will be in all of our small cell deployments", Fierce Broadband Wireless, Jan 9, 2013.
[155] "Wi-Fi in Heterogeneous Networks", Ericsson, Jun, 2013.
[156] "Ericsson Mobility Report", Ericsson, June, 2013.
[157] "Wi-Fi & Packet Core Integration", Ericsson/LG, Oct 19, 2012.
[158] Innovative Wi-Fi Offload Features, Aptilo, accessed Oct 2, 2013.
[159] "Digital Subscriber Line Access Multiplexer (DSLAM)", International Engineering Consortium, 2007.
[160] Cable Modem to Customer Premise Equipment Interface Specification, Cable Television Laboratories, Vers. 3.0, p. 8, 2008.
[161] ERM1000 Edge Resource Manager, ARRIS, accessed Aug 29, 2013.
[162] "The Definition of OSS and BSS", OSS line, Dec 5, 2010.
[163] J. U. Sverzut, *Redes GSM, GPRS, EDGE e UMTS*, 1st ed. São Paulo, SP: Editora Érica, 2005, pp. 70-394
[164] G. Heine, *GSM Networks: Protocols, Terminology and Implementation*, Norwood, MA: Artech House, 1998, pp. 5-125.
[165] O. Waldhorst and J. Mückenheim, "UMTS System Architecture and Protocol Architecture", Oct, 2011.
[166] GSM 03.40, GSM/ETSI, Vers. 5.3.0, pp. 11-12, Jul, 1996.
[167] ETSI TS 132 296, 3GPP/ETSI, Vers. 11.5.0, pp. 11-14, Apr 18, 2013.
[168] Y. Ouyang and M. H. Fallah, "An Analysis of Traffic and Throughput for UMTS Packet Core Networks", Howe School of Technology Management, Stevens Institute of Technology, Hoboken, NJ, USA.

[169] Y. Ouyang and M. H. Fallah, "A Performance Analysis for UMTS Packet Switched Network based on Multivariate KPIs", Howe School of Technology Management, Stevens Institute of Technology, Hoboken, NJ, USA.

[170] "Lawful Interception for 3GPP: Cisco Service Independent Intercept in the GGSN", Cisco Systems, accessed Sep 18, 2013.

[171] G. Atkinson, S. Strickland and M. Chuah, "UMTS Packet Service Infrastructure: Distributed is Cheaper", Networks 2002, Jun, 2002.

[172] 3GPP TS 23.002, 3GPP, Vers. 10.2.0, Mar, 2011.

[173] D. Flore, "LTE RAN architecture aspects", IMT-Advanced Evaluation, Beijing, P. R. China, Dec 17-18, 2009.

[174] "The LTE Network Architecture", Alcatel- Lucent, 2009.

[175] S. Gupta, "Home Gateway", Wipro Technologies.

[176] Conference "Portugal Telecom Technology & Innovation", Lisbon, Oct 30, 2012.

[177] "AT&T will shutter its 2G network by 2017", Fierce Wireless, Aug 03, 2012.

[178] S. Z. Hosain, "2G GSM & GPRS — The Shutdown Begins", Aeris, May 29, 2012.

[179] "LTE Offers M2M Applications Longevity, Heavy Reading Says", M2M World News, Mar 18, 2013.

[180] "M2M executives speak: the shift from 2G to 3G, 4G in M2M", Telecom Engine, Oct 13, 2011.

[181] SIGFOX, accessed Nov 13, 2013.

[182] NEUL, accessed Nov 13, 2013.

[183] WEIGHTLESS, accessed Nov 13, 2013.

[184] "IPv4 address exhaustion", Wikipedia, accessed Nov 13, 2013.

[185] R. D. Woolley, "Web Performance Measurement & Capacity Planning: Briefing Paper", Salt Lake City, Utah, USA: Chief Information Officer's Section - Office of the Governor, Jan, 2000.

[186] Personal Communication with Ricardo J. Ferreira, M.Sc. Eng., PT-Inovação, Jun, 2013.

[187] ETSI TS 123 060, 3GPP/ETSI/GSM, Vers. 4.11.0, pp. 27-150, Dec, 2006.

[188] ETSI TS 129 060, 3GPP/ETSI/GSM, Vers. 11.6.0, pp. 13-14, Apr, 2013.

[189] Personal Communication between A. M. de Oliveira Duarte, University of Aveiro, and B.R Jarmo, NOKIA, May 22, 2001.

[190] N. K. Shankaranarayanan, Z. Jiang, and P. Mishra, "User-perceived performance of Web-browsing and interactive data in HFC cable access networks", in *IEEE International Conference on Communications*, Vol. 4, 2001.

[191] J. T. Chapman, "Multimedia Traffic Engineering - The Bursty Data Model", in *SCTE Emerging Technologies*, 2002.

[192] Voice Over IP - Per Call Bandwidth Consumption, Cisco, accessed Nov 23, 2013.

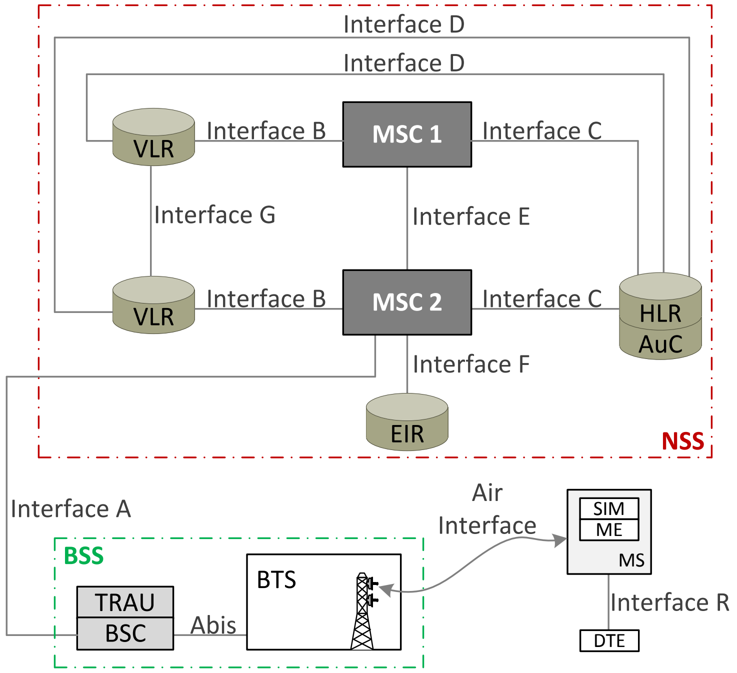
[193] Personal Communication with João G. R. de Figueiredo, M.Sc. Eng., Department of Electronics, Telecommunications and Informatics, University of Aveiro, 2013.

[194] ETSI TS 23.236, 3GPP/ETSI/GSM, Vers. 6.1.0, p. 11, Sep, 2005.

[195] M. Corici, J. Fiedler, T. Magedanz, and D. Vingarzan, "Evolution of the Resource Reservation Mechanisms for Machine Type Communication Over Mobile Broadband Evolved Packet Core Architecture", Fraunhofer FOKUS Institute, Berlin, Germany, 2011.

Appendices

Appendix A – GSM Interfaces



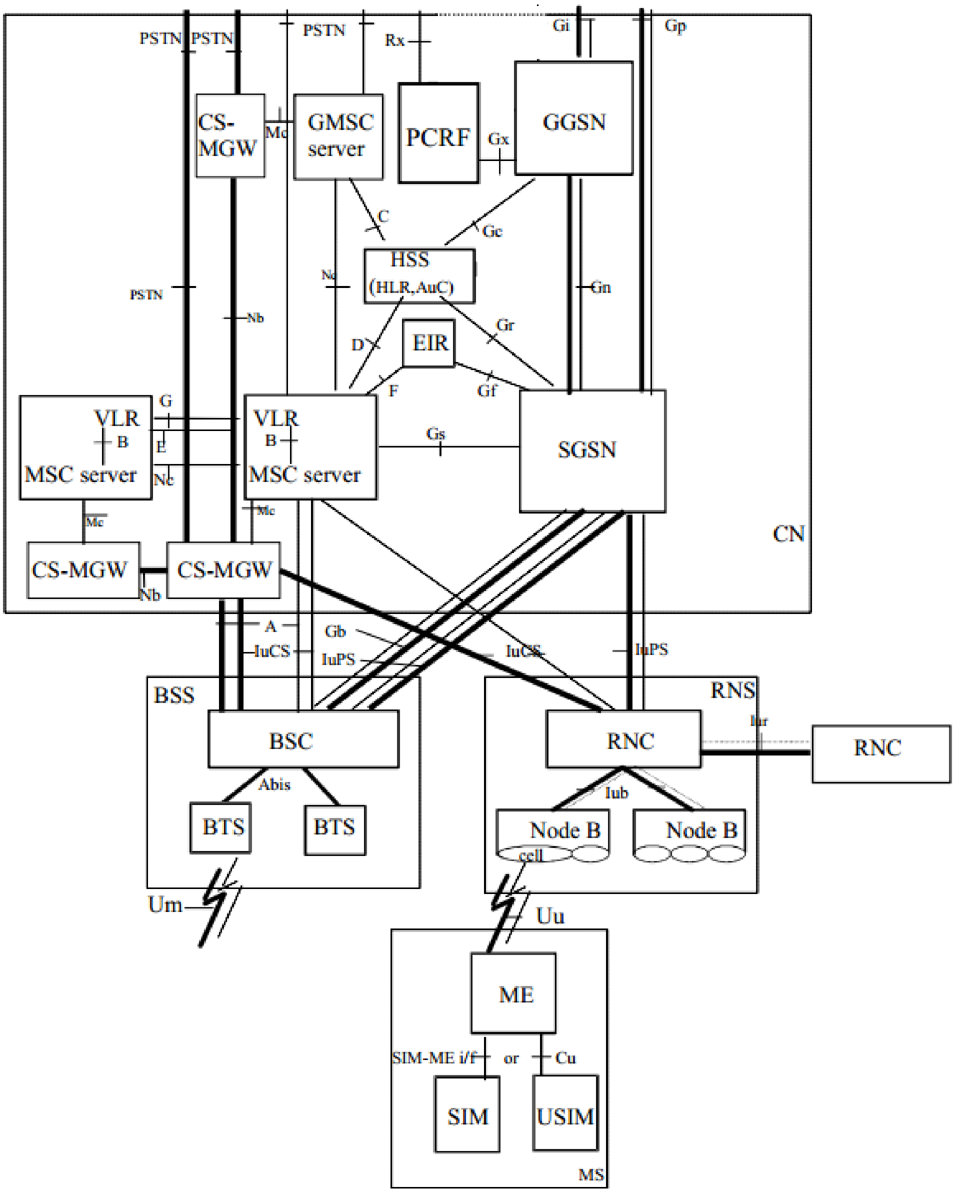
Interface	Connection
Air interface	MS - BTS
Abis	BTS - BSC
Interface A	BSC - MSC
Interface B	MSC - VLR
Interface C	MSC - HLR
Interface D	HLR - VLR
Interface E	MSC - MSC
Interface F	MSC - EIR
Interface G	VLR - VLR
Interface R	MS - DTE

AuC – Authentication Centre
BSC – Base Station Controller
BSS – Base Station System
BTS – Base Transceiver Station
DTE – Data Terminal Equipment
EIR – Equipment Identity Register
HLR – Home Location Register
ME – Mobile Equipment
MS – Mobile Station
MSC – Mobile services Switching Center
NSS – Network Switching System
SIM – Subscriber Identity Module
TRAU – Transcoding Rate and Adaptation Unit
VLR – Visitor Location Register

Figure 92 – GSM interfaces (based on [163])

Appendix B – Basic Configuration of a PLMN supporting CS and PS services

3GPP TS 23.002 [172] should be consulted for more detailed information.



Legend:
Bold lines: interfaces supporting user traffic;
Dashed lines: interfaces supporting signalling.
NOTE 1: The figure shows direct interconnections between the entities. The actual links may be provided by an underlying network (e.g. SS7 or IP): this needs further studies.
NOTE 2: When the MSC and the SGSN are integrated in a single physical entity, this entity is called UMTS MSC (UMSC).
NOTE 3: A (G)MSC server and associated CS-MGW can be implemented as a single node: the (G)MSC.
NOTE 4: The Gn interface (between two SGSNs) is also part of the reference architecture, but is not shown for layout purposes only.

Figure 93 – Basic Configuration of a PLMN supporting CS and PS services (using GPRS), and interfaces [172]

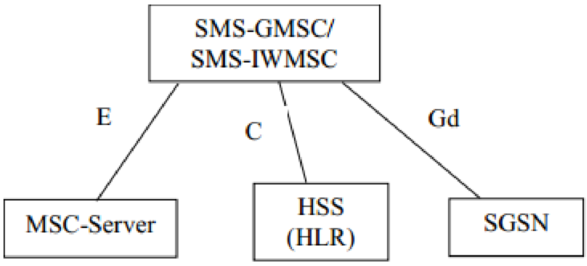


Figure 94 – Configuration for Short Message Service [172]

Appendix C – UMTS Protocol Architecture

Figure 70 depicts the UMTS protocol architecture for Data Plane, while Figure 71 depicts the UMTS protocol architecture for Control Plane. For detailed information, please refer to ETSI TS 123 060 [187].

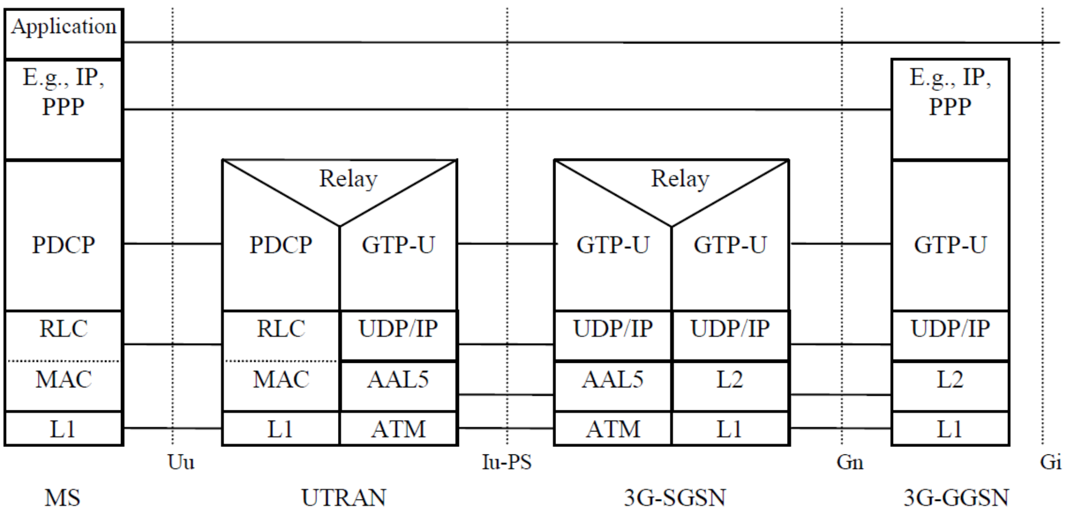


Figure 95 – UMTS Data Plane protocol architecture [187]

- **GTP-U (GPRS Tunneling Protocol for User plane):** Tunneling of user data between UTRAN and the 3G-SGSN; Tunneling between the GSNs in the backbone network; Encapsulation of all PDP PDUs [165].

- **PDCP (Packet Data Convergence Protocol):** Provides protocol transparency for higher-layer protocols; Support for e.g., IPv4, PPP and IPv6; Compression of control information; No user data compression in Iu mode [165].

- **RLC (Radio Link Control):** RLC protocol provides logical link control over the radio interface; There may be several simultaneous RLC links per MS; each link is identified by a Bearer Id [165].

- **MAC (Medium Access Control):** MAC protocol controls the access; Signaling (request and grant) procedures for the radio channel [165].

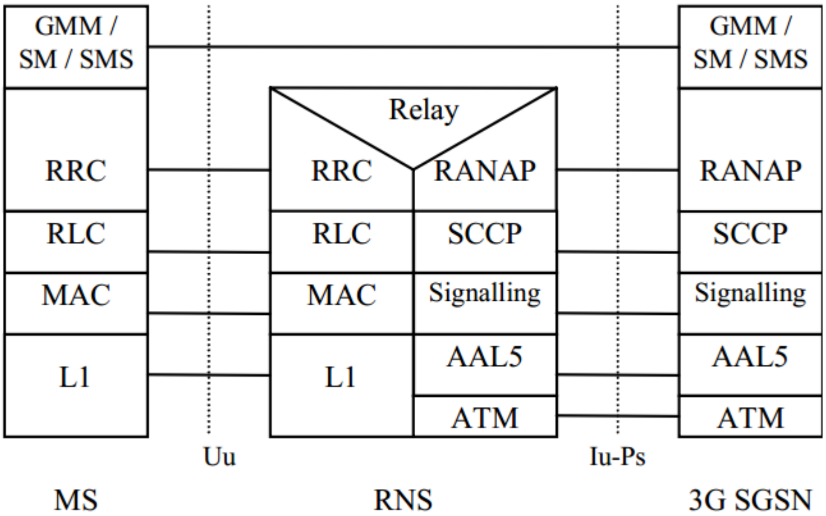


Figure 96 – UMTS Control Plane protocol architecture [187]

- **GMM (GPRS Mobility Management):** GMM supports mobility management functionality such as attach, detach, security, and routing area update [165].

- **SM (Session Management):** SM supports PDP context activation and deactivation [165].

- **SMS:** Supports short message service [165].

- **GTP-C (GPRS Tunneling Protocol for Control plane):** Establish, manage and release GTP tunnels [165].

- **RANAP (Radio Access Network Application Protocol):** Transport of higher-layer signaling; Handling of signaling between the 3G-SGSN and UTRAN; Management of the GTP connections on the Iu interface [165].

- **RRC (Radio Resource Control):** Information Broadcast (AS and NAS); RRC connection management (setup, release, reconfiguration); Radio Bearers management (setup, release, reconfiguration); Management of radio resources for the RRC connection; RRC connection mobility functions; Paging/notification [165].