DEGREE PROJECT IN ELECTRICAL ENGINEERING, SECOND LEVEL
STOCKHOLM, SWEDEN 2017

Make yourself at home

A comparative study of VoLTE Roaming architectures

IOANNIS KALTSAS

KTH ROYAL INSTITUTE OF TECHNOLOGY
INFORMATION AND COMMUNICATION TECHNOLOGY

Make yourself at home

A comparative study of VoLTE Roaming architectures

Ioannis Kaltsas

2017-01-22

Master's Thesis

Examiner
Gerald Q. Maguire Jr.

Supervisor
Anders Västberg

Industrial adviser
Jori Hämäläinen

KTH Royal Institute of Technology
School of Information and Communication Technology (ICT)
Department of Communication Systems
SE-100 44 Stockholm, Sweden

Abstract

While the data traffic has increased through the years, the average revenue per user (ARPU) remains flat. Thus, mobile network operators need to find a solution for how to support the growing amounts of traffic at fixed revenue per user. Similarly, the number of roaming users has increased, but according to recent European Union (EU) regulations, mobile network operators have to lower their charges for roaming to zero by June 2017. While this decrease in roaming charges will benefit European roaming users, mobile network operators have to cover their expenses for their own roaming subscribers, thus they have to find a way to lower their operational expenses (OPEX). Additionally, it is important for operators to consider how they might actually benefit from the removal of roaming charges.

This project will focus on roaming in 4th generation (4G) mobile networks. A common roaming scenario would include three different networks: the Home Public Land Mobile Network (HPLMN) a transit network, and a Visited Public Land Mobile Network (VPLMN). Normally, both the signaling traffic and the media payloads traverse these networks, thus causing additional latencies and increasing OPEX. However, in recent years, a new mechanism, called local breakout (LBO), was introduced that can lower the costs of roaming and avoid unnecessary traffic while meeting a roaming user's needs.

The goal of LBO is to decrease the operator's OPEX when supporting roaming subscribers. A secondary goal of LBO is to reduce the latencies experienced by roaming subscribers during their sessions. Achieving both of these goals will satisfy both operators and consumers.

This thesis project analyzes Voice over Long Term Evolution roaming with the aim of presenting the various alternative architectures for Voice over LTE roaming, compare them in different scenarios, and evaluating them based on criteria defined during this project. The conclusion is that the best solution that is applicable to all the mobile network operators for all the possible roaming scenarios does not exist yet. The various VoLTE roaming architectures can be chosen by the mobile network providers according to their needs.

Keywords

Local breakout, RAVEL, S8HR, VoLTE roaming

Sammanfattning

Datatrafiken har ökat genom åren men den genomsnittliga intäkten per användare (ARPU) är fortfarande oförändrad. Mobilnätsoperatörer bör hitta en lösning för att kunna stödja den växande mängden trafik på fasta intäkter per användare. Samtidigt har antalet roaming användare ökat. Enligt de senaste reglerna från Europeiska unionen (EU) måste mobiloperatörer sänka sina roamingkostnader till noll senast juni 2017. Denna minskning av roamingavgifterna kommer att gynna europeiska roaminganvändare, men mobilnätoperatörer måste samtidigt täcka sina kostnader för sina egna roamingabonnenter. Detta medför att mobilnätoperatörer måste hitta ett sätt att sänka sina driftskostnader (OPEX). Dessutom är det viktigt för operatörerna att fundera över hur de faktiskt kan utnyttja denna uteslutning av roamingavgifter.

Detta projekt kommer att fokusera på roaming i den 4:e generationens (4G) mobila nät. Ett vanligt roaming scenario skulle omfatta tre olika nätverk: Home Public Land Mobile Network (HPLMN), transitnätverk, och en Visited Public Land Mobile Network (VPLMN). Vanligen är det både signaleringstrafik och media som passerar dessa nätverk. Det leder till ytterligare latenstider och ökande driftskostnader. Under de senaste åren har en ny mekanism som kallas för Local BreakOut (LBO) införts. Detta används för att sänka kostnaderna för roaming och undvika onödig trafik, samtidigt som den bemöter roaming användarens behov.

Målet med LBO är att minska Operatörens OPEX när den stödjer roaming abonnenter. Ett sekundärt mål av LBO är att minska latensen för roamingabonnenter under sina sessioner. Uppnående av båda dessa mål kommer att tillfredsställa både operatörer och konsumenter.

Detta examensarbete analyserar Voice over Long Term Evolution roaming i syfte att presentera de olika alternativa arkitekturer för Voice over LTE roaming, jämföra dem i olika scenarier, och utvärdera dem utifrån kriterier som fastställs under detta projekt.

Nyckelord

Local breakout, RAVEL, S8HR, VoLTE roaming

Acknowledgments

I would like to express my gratitude to Professor Gerald Maguire who introduced me to this interesting topic and provided his prompt assistance in every step of the way. Furthermore, I would like to thank my supervisor Jori Hämäläinen for the insightful comments and remarks through this master thesis. Also, I would like to thank Tero Jalkanen for providing me with very useful information regarding this thesis topic, for informing me about the latest news around the area of this project and his valuable feedback. As well as my opponent who revised my thesis Efthimios Neroutsos. Last but not least, I would like to thank my family, friends and my manager Anna Demchenko who expressed her support in every way possible and helped me to complete this thesis.

Stockholm, 10 -01-2017
Ioannis Kaltsas

Table of contents

List of Figures

List of Tables

List of acronyms and abbreviations

AAA	Authentication, Authorization, and Accounting
ADSL	Asymmetric Digital Subscriber Line
AS	Application Server
A-SBG	Access Session Border Gateway
ARPU	average revenue per user
BGCF	Breakout Gateway Control Function
CAMEL	Customized Applications for Mobile networks using Enhanced Logic
CAP	CAMEL Application Part
CAPEX	capital expenses
CBO	Carrier Breakout
CDR	Call Data Record
CS	Circuit-Switched
CSCF	Call Session Control Function
E-CSCF	Emergency- call Session Control Function
eSRVCC	enhanced SRVCC
EU	European Union
4G	4^{th} generation
FNO	Fixed Network Operator
FTTH	Fiber to the Home
GGSN	Gateway GPRS Support node
GPRS	General Packet Radio Service
GSMA	GSM Association
GTP	GPRS Tunneling Protocol
HPLMN	Home Public Land Mobile Network
HSDPA	High-Speed Downlink Packet Access
HSS	Home Subscriber Server
IBCF	Interconnection Border Control Function
I-CSCF	Interrogating- CSCF
IMS	IP Multimedia Subsystem
IMS-ALG	IMS-Application Level Gateway
IM-SSF	IP Multimedia Service Switching Function
IP-CAN	IP carrier access network
IPX/GRX	IP exchange/ GSMA GPRS Roaming eXchange
ISP	Internet service provider
ISUP	Integrated Services Digital Network (ISDN) User Part
LBO	local breakout
LRF	Location Retrieval Function
LTE	Long Term Evolution
MBO	Mid Breakout
MGCF	Media Gateway Control Function
MGW	media gateway
MME	Mobility Management Entity
MMS	Multimedia Messaging Service
MNO	Mobile Network Operator
MRFC	Multimedia Resource Function Controller
MRFP	Multimedia Resource Function Processor
NAT	Network Address Translation
OMR	Optimal Media Routing
OPEX	operational expenses
OSA	Open Service Architecture
P-CSCF	Proxy-CSCF

PDN	Packet Data Network
PGW	PDN Gateway
PoC	Push-to-talk over Cellular
PoPs	Points of Presence
PS	Packet-switched
PSAP	public-safety answering point
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RAVEL	Roaming Architecture for Voice over IMS with Local Breakout
RADIUS	Remote Authentication Dial-In User Service
RAN	Radio Access Network
RCS	Rich Communication Services
RTP	Real-time Transport protocol
RTT	Round Trip Time
SBC	Session Border Controller
SCS	OSA Service Capability Service
S8HR	(LTE interface) S8 231Home Routing
S-CSCF	Serving CSCF
SG	Session Gateway
SGSN	Serving GPRS Support Node
SGW	Serving Gateway
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SLF	Subscriber Location Function
SRVCC	Single Radio Voice Call Continuity
SS7	Signaling System Number 7
3GPP	Third Generation Partnership Project
TrGW	Transition Gateway
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
ViLTE	Video over LTE
VoLTE	Voice over LTE
VPN	virtual private network
VPLMN	Visited Public Land Mobile Network
WLAN	Wireless Local Area Network

1 Introduction

This chapter defines the problem that addressed in this Master's thesis project along with its corresponding context and goals. Finally, an outline of the structure of this thesis is given.

1.1 Background

The first commercial LTE networks were established in Stockholm, Sweden and Oslo, Norway on December 14, 2009 [1]. Today, Long Term Evolution (LTE) mobile networks have been deployed in most urban areas of the world [2] [3]. Furthermore, users are used to having a good quality of service when they are located in their home network and they expect to receive a similar quality of service when roaming. To make this possible, the roaming environment has to adapt to meet the demand for new services, such as Voice over LTE (VoLTE), thus requiring the operators to deploy new network architectures or modify their existing network architecture to support these services. However, LTE roaming remains a challenge for the mobile network operators who want to expand the area where their network is available.

To improve LTE roaming, network operators are offering international roaming services to attract more subscribers and to increase the satisfaction and loyalty of their existing subscribers. In this way, operators can increase their revenue, while offering their services their subscribers *despite* the fact that these users are utilizing an access network operated by other operators. However, implementing roaming services can be cumbersome due to the adoption of different standards by each of the different networks and different pieces of equipment (both of the users and the network operators). In practice, this means that network deployments and mobile services are not the same in all countries; hence, achieving the appropriate interconnection via various networks is an important roaming issue.

Roaming operates in a large and complex network environment with various network architectures and different implementations. Moreover, different architecture and implementation combination each fulfill a certain purpose – but not all of these purposes are aligned. Furthermore, different protocols and interfaces are used to interconnect the various networks. In a typical roaming scenario, many network entities must participate in order for the roaming subscribers to be able to use their desired services when they have roamed to another network. For example, when the subscribers are located outside of their home network (i.e., their Home Public Land Mobile Network – HPLMN) and they have roamed into another operator's service area, the Visited Public Land Mobile Network (VPLMN) has to communicate with the HPLMN. Usually, this communication occurs via an IP exchange/GSM Association (GSMA) General Packet Radio Service (GPRS) Roaming eXchange (IPX/GRX) network. The communication includes signaling messages using protocols, such as the Session Initiation Protocol (SIP) and Diameter. In addition to the signaling messages, there is also a need to transport the user's media, typically using a protocol such as the Real-time Transport protocol (RTP). With the recent launch of LTE roaming, another framework, called IP Multimedia Subsystem (IMS), is needed. IMS provides the User Equipment (UE) with all the services to which a subscriber has subscribed. Chapter 2 describes each of these networks and protocols.

1.2 Problem definition

On the 8th of July 2015, the council of the European Union (EU) reached a decision that "roaming surcharges in the European Union will be abolished as of 15 June 2017" [4]. While this agreement within the EU will benefit European roaming subscribers, it remains a challenge for the network operators who must find a solution to cover their expenses for this roaming. In order to realize this regulation, mobile network providers have to find a way to lower their operational expenses (OPEX) as their average revenue per user (ARPU) remains flat. Moreover, network operators need to consider how they can benefit from this new regulation.

Local breakouts (LBOs) can potentially offer such a solution. LBO enables the VPLMN to *locally* forward the user's data (typically their media stream), thus avoiding the need to send traffic back and forth between the VPLMN and HPLMN. However, in this approach it is not entirely clear which of the involved roaming entities "roles" (i.e., VPLMN, HPLMN, IPX providers, etc.) benefit from LBO and which ones will not. Certainly, subscribers will benefit from LBO since they will receive better service (as the delay will potentially be lower, especially if the subscriber is calling a party that is local to the VPLMN) and do so at a lower cost. In contrast, an IPX provider might lose revenue due to the introduction of LBO, since the user's data traffic will be forwarded locally by the VPLMN, hence this traffic will not be sent to the HPLMN via an IPX network. It is clear that the VPLMN gains from LBO, as they do not have to forward all of the user's traffic back to the HPLMN. Unfortunately, the HPLMN lacks the ability to monitor the Quality of Service (QoS) provided by the VPLMN to the HPLMN's roaming subscribers. The situation becomes even more complex if the VPLMN charges the HPLMN for locally delivering the HPLMN's subscriber's traffic.

Despite these concerns, a survey presented on 2 May 2016 states that 40% of providers and network operators plan to deploy VoLTE roaming by 2017 [5]. Although, 57% of respondents are not quite sure which roaming model (LBO or S8HR[*]) they will choose.

1.3 Purpose

Although the EU regulation will take effect in 2017 and VoLTE is beginning to take off, network providers and IPX carriers are not yet sure which roaming architecture will prevail and they are not confident in choosing a single roaming solution. The purpose of this project is to present the different VoLTE roaming architectures, compare them, and evaluate them based on criteria that will be defined in this thesis (see Section 4.1 on page 37). This aim is to provide the information necessary to decide in which situations and from which point of view a given VoLTE roaming architecture is the desired solution. Additionally, this thesis suggests several ways that a network operator can benefit from a given VoLTE roaming infrastructure.

1.4 Goals

The goals of this project are to present different approaches for implementing VoLTE roaming, evaluate them, and suggest the most suitable mechanisms that can be applied to the current network architectures of mobile network operators.

1.5 Research Methodology

In this Master's thesis, qualitative research is utilized to achieve the stated goals. A quantitative method was not chosen since it was infeasible to test different VoLTE roaming architectures in a test environment in order to evaluate these roaming models. However, surveys and tests have been performed by GSMA, Third Generation Partnership Project (3GPP), and various network providers and these will be discussed in this thesis and their results are used to compare the VoLTE roaming models. Furthermore, the analytical research method was used to analyze and evaluate the existing literature. More significantly, the literature study identified essential background information about all the relevant network architectures, network elements, and the roles of the different network operators who take part in VoLTE roaming. All of the existing VoLTE roaming architectures were analyzed using secondary research. The different VoLTE roaming architectures are compared in terms of criteria identified in this thesis project. Finally, all of this research leads to a conclusion that may offer further insights into this subject.

[*] Chapter 3 describes these two roaming models.

1.6 Delimitations

As stated in the previous section, quantitative research upon this topic is out of the scope of this Master's thesis. The reason for this is the complex environment of a VoLTE roaming infrastructure. Setting up a testing environment requires collaboration between two roaming partners and an IPX provider. In addition, the configuration process can be quite complex. Finally, the purpose of this Master's thesis project was *not* to choose **the** *best* available solution for VoLTE roaming as each network operator has different needs and capabilities. Moreover, each network operators has different relationships with other network operators. Instead, the aim was to pinpoint the important advantages and disadvantages of *all* the available VoLTE roaming models. This thesis should be of interest to readers who are interested in gaining a clear view of which model a network operator should use in various network scenarios.

1.7 Structure of the thesis

The remainder of this thesis is divided into four chapters. The second chapter provides essential background information about all the relevant network architectures, network elements, and the roles of the different network providers that take part in a VoLTE roaming environment. Chapter two also presents related research work about this topic. The third chapter analyzes and discusses all the available VoLTE roaming architectures, so that the reader will be able to comprehend and become familiar with these models. In the beginning of chapter four, the criteria used for the evaluation of the different models are introduced and categorized into two groups: general criteria and the criteria depending on the call scenario. Next, the VoLTE roaming architectures are compared based on these criteria. Finally, chapter five concludes this thesis and describes some suggested future work.

2 Background

This chapter provides basic background information about roaming. Additionally, this chapter describes all the information required to understand the different roaming architecture models (specifically, Home Routing, LBO, and S8HR) and the protocols and systems that are necessary for all of them. The chapter also describes related work on LBO.

2.1 LTE

The Long Term Evolution (LTE) project was started by 3GPP in 2004 [6]. LTE is commonly referred to as 4G system, although formally only LTE-Advanced meets the ITU's criteria for being an IMT-Advanced (4G) system. LTE is the successor of Universal Mobile Telecommunication System (UMTS), which in turn was derived from the Global System for Mobile Communications (GSM). LTE was initially defined in Release 8 of 3GPP [7] and is the access part of the Evolved Packet System (EPS) [8].

According to [8]–[10] the main goals of LTE is to deliver:

- An all-IP network,
- Low latency (the Round Trip Time RTT is 5 ms under ideal radio conditions),
- High data rates (300Mbps peak downlink and 75Mbps uplink),
- High spectral efficiency,
- Flexible frequency,
- Flexible bandwidth (channel bandwidths for the 1900 MHz frequency band include: 1.4, 3, 5, 10, 15, and 20 MHz),
- Seamless mobility
- Adequate quality of service.

The LTE architecture consists of three main entities (shown in Figure 2-1):

User Equipment (UE)	Subscribers use a UE to access the LTE services to which they have subscribed. A UE can be either a mobile phone that supports LTE connectivity or a device (such as a laptop) that has an LTE interface.
Evolved UMTS Terrestrial Radio Access Network (E-UTRAN)	E-UTRAN is the radio access network and is responsible for the communication between the UEs and the evolved packet core. The E-UTRAN is comprised of one or more evolved base stations (abbreviated as eNodeB or eNB). An eNodeB interconnects UEs located in one of the cells that it realizes to the EPC via the S1 interface, while it uses the Uu interface to communicate with the UE. An eNodeB can be interconnected to other eNBs via the X2 interface (to support handovers).
Evolved Packet Core (EPC)	The EPC is the core network that connects the E-UTRAN to Packet Data Networks (PDNs). PDNs can be private networks, the public Internet, or an IP Multimedia Subsystem.

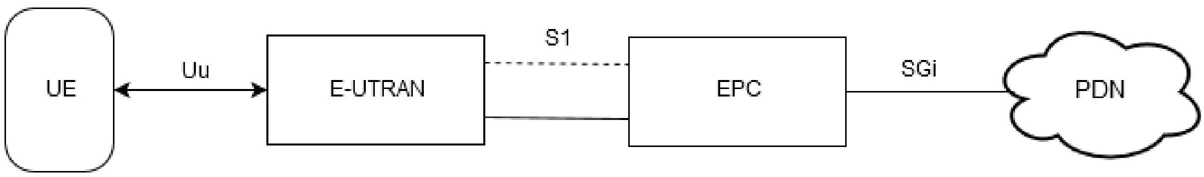

Figure 2-1 High level architecture of an LTE network

A closer look into the EPC reveals that it consists of the following main components (shown in Figure 2-2):

Home Subscriber Server (HSS)	The HSS is a database that contains all the relevant information of the subscribers that belong to a provider's network. It uses the S10 interface to connect with the MME. For further reference, see the Section 2.4.2.1
PDN-Gateway (P-GW)	The P-GW interconnects the EPC with an external IP network (PDN) through the SGi interface. This gateway transports user plane traffic from UEs to/from the PDNs.
Mobility Management Entity (MME)	The MME is the main signaling node within the EPC [11]. Its main functions are paging initiation and the authentication of the UE. Moreover, it is responsible for storing information about the location of the subscriber and contributes to the selection of the appropriate gateway in the process of the initial registration. The MME uses the S1-MME interface to communicate with the eNodeB and the S11 interface to connect to the S-GW. Lastly, the MME contributes to the handover procedure within LTE and
Serving Gateway (S-GW)	Similar to the P-GW, the S-GW is responsible for transporting IP data traffic between the UE and the PDNs. The S-GW acts as a router by forwarding user plane traffic between the eNB and the P-GW using the S5/S8 interface. Additionally, the S-GW uses the S1-U interface to communicate with the eNB.
Policy Control and Charging Rules Function (PCRF)	The PCRF lies within the P-GW and has two main functions: "Flow Based Charging, including charging control and online credit control" [12] and policy control (such as gating control, QoS control, and QoS signaling).

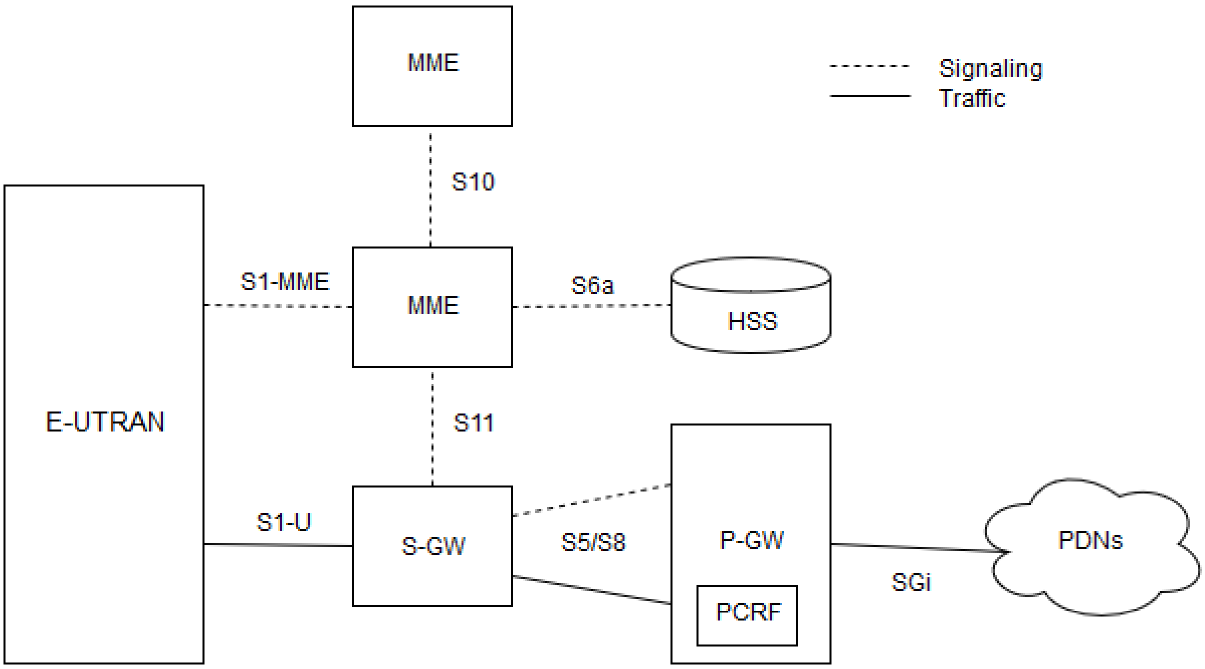

Figure 2-2 The Evolved Packet Core (adapted from figure 3 of [13])

2.2 Roaming Interconnection Types

Network operators can connect with each other directly or connect via a GPRS Roaming Exchange (GRX) or an IP Packet exchange (IPX) network. These interconnections are typically accompanied by roaming agreements that enable the operators to apply policies, control network access for roaming subscribers, and operate their services.

A direct interconnection is a fast and straightforward solution and it can be established in two different ways. One solution is to implement tunnels (using protocols such as Internet Protocol Security (IPsec)) over the public internet. However, this solution is not recommended since the QoS generally does not meet the carrier industry's standards. Another method is to use private links (for example via a leased line or a third party provided virtual private network (VPN)). While this later method can provide better performance in terms of QoS and security, it is not recommended as this approach is not scalable – as each operator has to have a direct connection with every other operator. The number of direct connections grows as N^2 with N operators, which makes this approach difficult to manage and it rapidly increases the operator's OPEX and capital expenses (CAPEX).

Alternatively, GRX/IPX networks operated by third party carriers can be used. A GRX/IPX carrier has connections with multiple network operators, thus enabling each network operator to connect with other operators simply by connecting to a single GRX/IPX network. This type of interconnection is less costly, has better scalability, and when used together with end-to-end encryption is more secure than a simple leased line. These reasons make this option preferable, hence today this is the recommended solution for interconnecting network providers [14].

2.2.1 GRX

The GRX network was initially defined in 2000 in order to support GPRS roaming. However, only Mobile Network Operators (MNOs) were allowed to connect to it. Over the years, additional services have been added, such as Universal Mobile Telecommunications System (UMTS) roaming, Multimedia Messaging Service (MMS) interworking, and Wireless Local Area Network (WLAN) (with authentication) data roaming [15].

In simple terms, GRX can be described as a hub that interconnects mobile networks over GPRS roaming networks owned by GRX service providers. The main advantages of a GRX network are:

- Offers access to multiple network providers with a single connection,
- Shares network resources within the connected networks,
- Provides easier access to the central root DNS (a separate DNS for a GRX),
- Roaming services can be rapidly implemented,
- Less expensive than having dedicated connections to all the desired networks,
- Offers redundancy
- It is highly secure (since it is a private network that is separate from the public internet).

Despite the above advantages, according to GSMA "the GRX offers a transport-only interconnection service between mobile operators on a bilateral basis with no guarantees of QoS end-to-end" [15]. This means that the GRX does *not* provide guaranteed QoS.

Figure 2-3 illustrates the high-level architecture of a GRX network. GRX includes those GRX providers that are interconnected with each other using a peering interface. These peering interfaces either can be direct connections or can utilize a common peering point. GRX peering partners generally sign a Service Level Agreement (SLA) to define the services and QoS that both partners promise to deliver. Inside a GRX network, there is a DNS root database for this GRX. This DNS supports domain name resolution and all of the GRX providers have access to it.

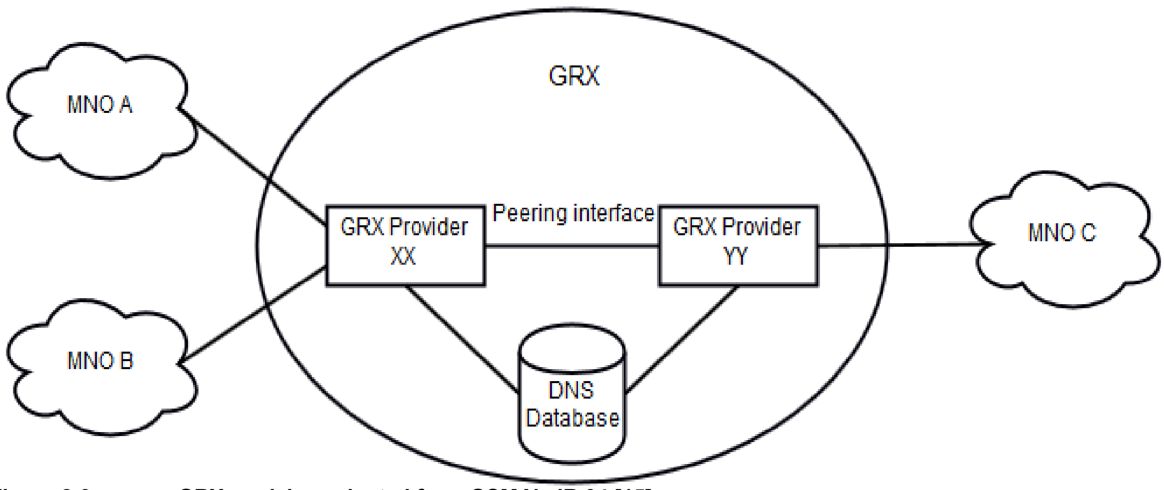

Figure 2-3: GRX model as adapted from GSMA's IR.34 [15]

2.2.2 IPX

IPX was defined in 2006 [17]. The final phases of the trials of voice services over IPX were completed in 2008[18]. IPX is a private IP backbone network and represents the evolution of GRX. Compared to GRX, IPX offers a better environment in terms of flexibility and compatibility. This flexibility is due to the interconnections using a single protocol for bilateral and multilateral connection services. Compatibility comes from the fact that in addition to MNOs, Fixed Network Operators (FNOs) can also connect to an IPX network. In general, any type of service provider (for example, an Internet service provider (ISP)) can connect to an IPX. As with its predecessor, IPX is a private network separate from the public internet. However, in contrast with GRX, IPX is able of delivering end-to-end QoS. In order to deliver this feature, IPX is required to be service aware. Regarding these services, IPX must support standardized services, such as IP Voice Telephony, IP Video Telephony, Push-to-talk over Cellular (PoC), Instant Messaging, Multimedia Messaging Service (MMS), Presence, and Video Sharing. IPX can also provide service for mobile network signaling, mobile data roaming, and Rich Communication

Services (RCS). Furthermore, security is a very important feature of IPX, hence the data of each provider is isolated and the IPX network is invisible to the end users.

Network operators can establish a connection with an IPX carrier by signing an agreement and connecting using a local tail. For redundancy, service providers can connect to more than one IPX provider. It is also possible for the network operator to select one IPX carrier to transfer signal plane traffic and another IPX carrier to transfer their user-plane traffic. As Figure 2-4 shows, an IPX network consists of several competing IPX providers with a common DNS root database (similar to the DNS root database in a GRX architecture). An additional feature is the IPX proxy that can support interworking of specific IP services, thus making it feasible to use a “cascading interconnect billing and a multilateral interconnect model” [16]. In simple terms, IPX’s model is able to handle international interconnections and correlate the Call Data Records (CDRs) of the traffic that traverses a cascade of interconnections.

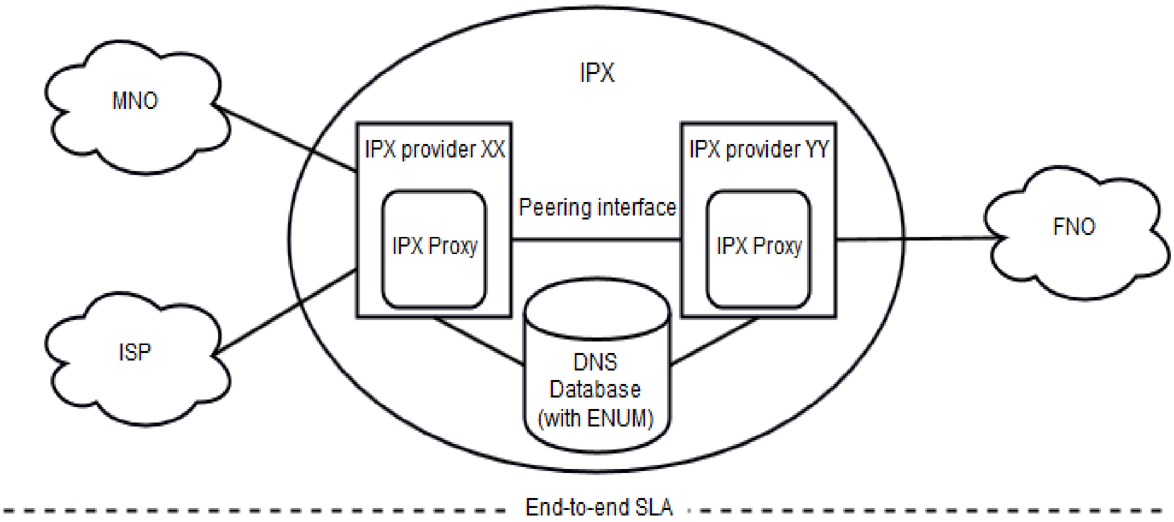

Figure 2-4 **IPX network with defined end-to-end SLAs**

2.3 Signaling in LTE Roaming – Diameter

Usually, in LTE roaming, the HPLMN and the VPLMN interconnect over an IPX network. The Diameter protocol is used for the signaling messages between the MME and the HSS. Additionally, Diameter is widely used by IMS entities. Diameter was defined in 1998 and as its name implies, is the evolution of Remote Authentication Dial-In User Service (RADIUS) [19]. IETF RFC 6377 states that Diameter is designed to “provide an Authentication, Authorization, and Accounting (AAA) framework for applications such as network access or IP mobility in both local and roaming situations” [20]. Diameter exchanges messages between peers; hence, it is a peer-to-peer protocol. Diameter uses TCP or SCTP as its transport layer protocol, rather than UDP as used in RADIUS. As a result, Diameter relies on the underlying transport protocol for reliable delivery of messages.

Depending upon the network deployment, the nodes that implement the Diameter protocol can act as either a server or a client. Therefore, the term Diameter node can refer to a Diameter Server, a Diameter Client, or a Diameter agent. For example, a Diameter node that acts as a Diameter client receives a user’s request for a connection along with her/his credentials. Similarly, a Diameter client can send an access request message to a Diameter Server. The Diameter Server will authenticate the user based upon the information provided. If the authentication is successful, then the subscriber’s access rights are sent in a response message to the Diameter client. Otherwise, the Diameter server sends an access-reject message. Although this example seems similar to a client-server architecture, it is important to clarify that depending upon the situation, a given Diameter node can act as a client and/or a server.

A Diameter agent is a special type of Diameter node. Diameter agents are typically classified into one of three kinds of Diameter agents [21]:

Relay Agent A relay agent forwards a message to an appropriate destination according to the information that the message contains. An advantage is that the relay agent can aggregate requests from various regions and forward these requests to a specific region. This functionality avoids the need to configuration network access for each Diameter server even with changes in network topology.

Proxy Agent A proxy agent also forwards messages (similar to a relay agent), but the proxy agent can alter the contents of the message before forwarding it. This enables the proxy agent to offer value-added services, perform administrative tasks, or enforce rules on messages.

Redirect Agent A redirect agent is useful when there is a need to store all of the Diameter routing information in a centralized location (a routing database for the Diameter nodes). When the redirect agent receives a request from a node, the redirect agent searches the routing table and then responds with a message containing the relevant redirection information. This type of agent is useful when the network architecture is composed of many nodes and it is desired to keep the routing information in a centralized node (which is accessible to all the other nodes), rather than every node needing to store and maintain a local routing table.

2.3.1 Translation agent

In addition to the above three types of agents, there is another type of agent called a translation agent. The function of this agent is to translate messages from one AAA protocol to another. More specifically, it can support backward compatibility for other protocols, such as RADIUS. This type of agent is useful when migrating from RADIUS or other AAA protocols to Diameter.

2.3.2 Diameter Message

A Diameter message has two parts: a Diameter header and Attribute Value Pairs (AVPs) in the payload. Each AVP consists of a header and data. AVPs contain information about authorization, authentication, accounting, routing information, and other types of information. Figure 2-5 shows the structure of a Diameter message.

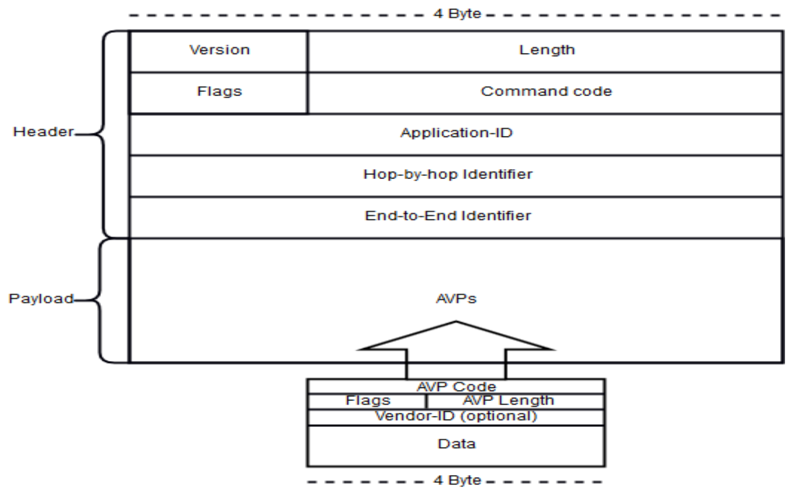

Figure 2-5: Diameter message and a Diameter attribute within such a message

2.4 The IP Multimedia Subsystem

The IP multimedia Subsystem (IMS) is a framework that provides IP-based multimedia services. IMS is comprised of IP-based technologies that enable access to multimedia services from the user's equipment (mobile phone, landline phone, or a personal computer) at any time and everywhere. This means that users are able to use their services (voice, video, messaging data, and web-based technologies) wherever they are roaming in an automated way *without* requiring the users to configure anything in their terminals. From a telecommunication provider's perspective, IMS is a way to offer value-added services to subscribers, while at the same time reducing transaction costs. IMS is an open architecture that makes use of the SIP and Diameter protocols in order for its components to communicate with each other. The first formal document that specified the IMS network was issued by 3GPP in Release 5 [22]. After that, the following releases added more functionalities to IMS. A thorough analytic description of IMS can be found in 3GPP technical specifications (TSs) [23].

2.4.1 The IMS architecture

Figure 2-6 shows the IMS network architecture and its complexity as a system. As can be seen, IMS has many entities and functions that interconnect with each other. While this figure might seem difficult to comprehend, after learning what these entities are – it is easy to understand how they work together.

An IMS network can be separated into the following parts: User Equipment (UE), access network, core network, and application server layer. In the context of this thesis, the user plane (shown in red in the figure) and control planes (shown in blue in the figure) are separate. The IMS architecture will initially be described at a high level and then we will focus on those parts that are related to VoLTE roaming.

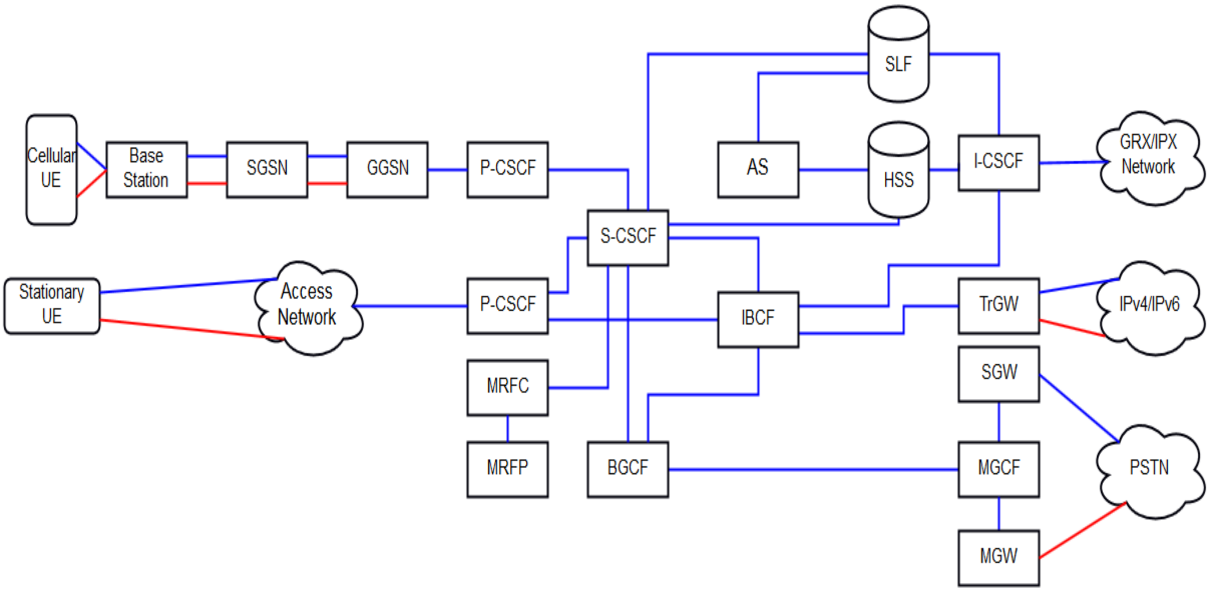

Figure 2-6: **IMS Architecture, the labels in the boxes represent the different types of logical entities that can be present in a IMS.**

2.4.2 IMS Core Network

The IMS core network is comprised of two main types of entities: Call Session Control Function (CSCF) entities and the HSS.

2.4.2.1 Home Subscriber Server (HSS)

The Home Subscriber Server (HSS) is a database within the IMS*. This database contains all the relevant information about the subscribers (i.e., the network operator's own users). This information is needed by the system to be able to handle each subscriber's communication sessions. HSS provides the following functions: "identification handling, access authorization, authentication, mobility management (keeping track of which session control entity is serving the user), session establishment support, service provisioning support, and service authorization support" [24]. The HSS sends the subscriber's profile to a particular CSCF node when the subscriber is registering with the IMS network or when a session needs to be established. When there is more than one HSS, a Subscriber Location Function (SLF) is used to find the HSS that stores this subscriber's data. The SLF and HSS use the Diameter protocol over the Cx and Dx interfaces to communicate with other entities within the IMS.

2.4.2.2 The Call Session Control Function (CSCF)

The different types of CSCFs are the main components of the IMS architecture. Each CSCF is basically a SIP server that processes SIP signaling messages. A CSCF provides session control for applications and terminals that want to make use of the IMS. CSCFs communicate with the HSS to access all the required information. Currently, there are four types of CSCFs: Proxy-CSCF (P-CSCF), Interrogating-CSCF (I-CSCF), Serving CSCF (S-CSCF), and Emergency-call Session Control Function (E-CSCF) [24].

* This HSS can be part of an LTE's HSS or an independent HSS.

The Serving – CSCF (S-CSCF)	As can be seen in Figure 2-6, the S-CSCF is the heart of IMS and it is responsible for provisioning SIP signaling. The S-CSCF forwards the relevant information required during a session to all the involved entities. The S-CSCF is responsible for the maintenance of sessions, routing, translation, interaction with services, and initiating the creation of CDRs [25]. Lastly, the S-CSCF utilizes Diameter over the Cx and Dx interfaces to communicate with the HSS to access the subscriber's service profile when there is a need to authenticate the subscriber during the registration process [24].
Interrogating Call Session Control Function (I-CSCF)	An I-CSCF communicates with peer IMS networks. Its main function is to communicate with the HSS using Diameter over the Cx and Dx interfaces in order to locate the relevant S-CSCF and to find where the user is registered. If a user is not currently registered, then the I-CSCF will choose a new S-CSCF to handle this user.
Proxy Call Session Control Function (P-CSCF)	The P-CSCF is a SIP proxy. It is the first point of contact when a UE wants to access the services to which this subscriber has subscribed. Its main functions are: (1) it forwards the UE's request to the subscriber's home network and (2) it filters SIP messages to ensure the safety of the IMS network. The P-CSCF makes use of IPSec tunnels to provide confidentiality and security to the information exchanged between the UE and the IMS network. In addition, the P-CSCF can check whether the correct QoS policies are applied. Last but not least, the P-CSCF is responsible for locally routing emergency calls. In all cases, the P-CSCF can compress those SIP messages that are being exchanged between itself and the UE (this technique is called "SigComp") [26] [27].
Emergency Call Session Control Function (E-CSCF)	The purpose of the E-CSCF is self-explanatory. It handles some attributes of emergency sessions initiated by the user. When the E-CSCF gets an emergency request from either the P-CSCF or the S-CSCF, then if the UE does not possess any credentials, a non-dialable callback number will be offered to the UE. Additionally, if the UE's location information needs to be clarified, then the E-CSCF will query the Location Retrieval Function (LRF). The LRF can also be queried if there is a need for routing information. Finally, the E-CSCF routes the emergency requests to the correct destination (typically a public safety answering point (PSAP)).

2.4.2.3 Other key IMS network entities

There are a number of additional IMS network entities:

Media gateway (MGW)	The MGW handles the media processing that is required to process calls to and from the Public Switched Telephone Network (PSTN).
Media Gateway Control Function (MGCF)	The MGCF controls the MGW to send or receive calls to and from the Circuit-Switched network (CS). To contact the CSCF and BGCF MGCF uses SIP messages and to communicate with the MGW it uses the gateway control protocol (H.248).
Breakout Gateway Control Function (BGCF)	The BGCF chooses the network over which a connection to the PSTN will be made. The call will be forwarded either by the BGCF using SIP to an alternative BGCF in order to be processed more thoroughly or to an MGCF that controls access to the corresponding PSTN.
Multimedia Resource Function Processor (MRFP)	The MRFP in conjunction with the Application Servers (ASs) is responsible for all the media processing. For example, to provide services such as voice mail, conferencing, recording, voice processing, etc.

Multimedia Resource Function Controller (MRFC)	The MRFC regulates the MFRP in order to provide the media processing required by the ASs [28].
Interconnection Border Control Function (IBCF)	The IBCF is an element that lies on the edge of an IMS network. Its role is to surpass any differences with other IMS networks and IP-CAN concerning their preferred way of communication and connection. Therefore, the IBCF performs actions such as Network Address Translation (NAT), topology hiding, SIP screening and signaling interconnection selection. The IBCF's task can also be performed by a Session Border Controller (SBC)[29].
Transition Gateway (TrGW)	The TrGW is a part of the role of the IBCF. The TrGW acts as a NAT and checks if the addresses of inbound and outbound media streams are correct [30].
Transit Routing Function (TRF)	TRF was firstly defined in TS.23.228 [31] where it was referred to as "transit function". Later in TR 23.850 [32], the acronym TRF was introduced. The basic action of this function is to analyze the destination address and decide where to route a session. Also, TRF anchors the control and the user plane, thus both signaling and the media will be transferred over the same path from the VPMN to the HPMN of the terminating user. This applies when the signaling was first transported to the HPMN of the caller and back to the VPMN through a particular loopback route. Although this is non-optimal routing of traffic, the main reason for this is the business need for maintaining the existing charging rules without any modifications regardless of an upgrade in technology from circuit-switched (CS) networking to IP. There are many possible destinations for the TRF to forward a session. For example, the session may be routed to an MGCF, BGCF, another IMS entity inside the network or to another IMS network, to a CS domain, or PSTN. The address resolution may use DNS and ENUM or private database lookups. Finally, it is worth mentioning that the TRF may be a stand-alone entity or may be located in the following IMS network nodes: MGCF, BGCF, I-CSCF, S-CSCF, or IBCF.

2.4.3 IMS Access Network

This subsection gives a short description of the IMS access network. For a more detailed and comprehensive description see Chapter 8 of the *IMS Application Developer's Handbook* [23] by R. Noldus, U. Olsson, C. Mulligan, I. Fikouras, A. Ryde, and M. Stille.

One of the advantages of the IMS core network is that is independent of the access network; thus, UEs connected via any type of access technologies (fixed, cable, WLAN, or mobile access) can use all the services. Originally, Rel-5 of 3GPP introduced IMS for use only by 3G mobile networks. Later, when High-Speed Downlink Packet Access (HSDPA) and long-term evolution (LTE) became available, they too could use IMS. Unfortunately, UMTS access is unable to offer many functions for mobile voice based on IMS, hence 4G networks are preferred. With the evolution of IMS, other access networks, such as Asymmetric Digital Subscriber Line (ADSL) and Fiber to the Home (FTTH) are also supported.

Access to an IMS system can be established over any IP carrier access network (IP-CAN) which provides the UE with IP connectivity. Thus, the UE's control-plane and media transfer are delivered to IMS over IP-CAN. This is achieved with two separate interfaces for signaling and media: Gm and Mb respectively. The Gm interface is used for communication between the UE and the P-CSCF, while the Mb interface connects the UE with the IMS access gateway.

In order to better understand the architecture of an IMS access network, an example will be given. Although this example does not present all possible implementations, it provides a simple (and typical) example. A UE can connect to a Radio Access Network (RAN), such as LTE. The first IMS entity that the UE encounters when it uses SIP signaling is the P-CSCF. The user plane and the control plane are separate from each other and they are not necessarily routed over the same path*. As mentioned earlier, SIP messages will be sent to the P-CSCF and the user data will be sent to the IMS access gateway. The P-CSCF controls the route the media follows to the IMS access gateway. Access to the IMS network can also traverse an Access Session Border Gateway (A-SBG) which includes a Session Border Controller (SBC) and a Session Gateway (SG). The SBC is responsible for handling the SIP signaling and the SG is responsible for handling the media. The P-CSCF can be included in the A-SBG. In this case, the IMS-Application Level Gateway (IMS-ALG) function is performed by the SBC, rather than the P-CSCF.

2.4.4 IMS Application Server Layer

As stated by H. Khlifi and J. C. Grégoire in their paper "IMS Application Servers: Roles, Requirements, and Implementation Technologies" [33], the role of an AS is to host and run services. The subscriber's services in an IMS network can be messaging, video conferencing, content sharing, online gaming, etc. An AS can support one or more IMS applications and the AS is connected with the control layer either directly or over the OSA Service Capability Service (SCS) with the S-CSCF. Communication between these two entities is done with SIP signaling. SIP messages are sent to the AS, which converts the "end-user service logic" into a sequence of SIP messages and sends these message back to the S-CSCF to forward them to the involved nodes. One of the benefits of an IMS is that a network operator can have a number of different application servers. These ASs can be SIP application servers, Open Service Architecture (OSA) application servers, or a Customized Applications for Mobile networks using Enhanced Logic (CAMEL) service environment. The CAMEL service environment has functions that allow end-users to execute services that are offered by their provider even when the end-users are roaming. The IP Multimedia Service Switching Function (IM-SSF) enables interoperability between the CAMEL application part and SIP by translating CAMEL Application Part (CAP) signaling to SIP signaling and vice versa.

2.5 The SIP protocol

The Session Initiation Protocol (SIP) is a control protocol that operates at the application layer. SIP is designed to be independent of the underlying transport layer. SIP is a text-based protocol and it is used for the creation, modification, and termination of sessions that include one or many participants [34]. These sessions can be voice and video calls or instant messaging over an IP network, i.e. Voice over IP (VoIP). For a session to be successfully established and delivered to the recipients, SIP cooperates with other protocols in the application layer. These protocols include the Session Description Protocol (SDP), for media identification and negotiation, and the Real-time Transport Protocol for the media transmission. If the session has to be secure, Secure RTP (SRTP) can be used and the SIP messages can be encrypted with Transport Layer Security (TLS). The first draft of an early version of SIP was issued in 1996 by IETF [35] and the latest version was specified in 2002 in RFC 3261 [34]. Two years earlier, in 2000, the SIP protocol was recognized as a 3GPP signaling protocol and is a significant element of IMS [36]. In the following subsections, the network elements, SIP requests and responses, and the Registration and Call flows, will be briefly described so that a reader who is not familiar with SIP has sufficient background to understand the rest of this thesis.

* This is one of the key enablers for LBO.

2.5.1 Network Elements

Although it is possible for two SIP user agents to exchange SIP messages *without* the help of any additional elements, this approach is impractical and scales poorly. This is due to the lack of directory services that can retrieve the required information from the available nodes of the network. In contrast, a normal SIP network will make use of more SIP entities beyond the two SIP user agents. The most common SIP components (shown in Figure 2-7) are [37]:

User Agent — A user agent (UA) is located in a SIP end node. A UA is responsible for handling a SIP session. A UA can have two roles: the User Agent Client (UAC) and the User Agent Server (UAS). The UAC sends SIP requests and the UAS receives these requests and sends back a SIP response that can either accept, reject, or redirect the request [38]. A UA can be a hardware device or software, i.e., a softphone.

Proxy server — A proxy server is an intermediate node that receives a request from a UA and forwards it to a terminating UA. In other words, the proxy server acts as a router. UAs can reside bilaterally to a proxy server and the maximum number of proxy servers that can lie between the originating end and the terminating end is 70 [37]. There are two types of proxy servers: a Stateless Proxy Server and a Stateful Proxy Server. A stateless proxy server does not keep any information regarding sessions and simply routes the received SIP messages. In contrast, a stateful proxy server stores all SIP requests and the corresponding responses for potential future use. Additionally, a stateful proxy server can resend a request if there is no response from the end-point. Moreover, proxies have the ability to apply policies. For example, they can check if the end user is allowed to initiate a call. Lastly, when required, a proxy can interpret and rewrite a request message before forwarding it.

Registrar — The Registrar is a server that receives requests from a UA and can authenticate the users. The Registrar updates a record of the Uniform Resource Identifier (URI)* along with the location of the UA in the location server's database to assist other SIP servers that are located within the same domain. Following a successful SIP registration, the user's SIP URI can be used by a caller to initiate calls to this user's registered SIP UAs.

* A URI is a string of characters used to identify a resource in a network. For example, the URL that is used for web addresses is a type of URI. For further reference see the RFC 2396 [39]

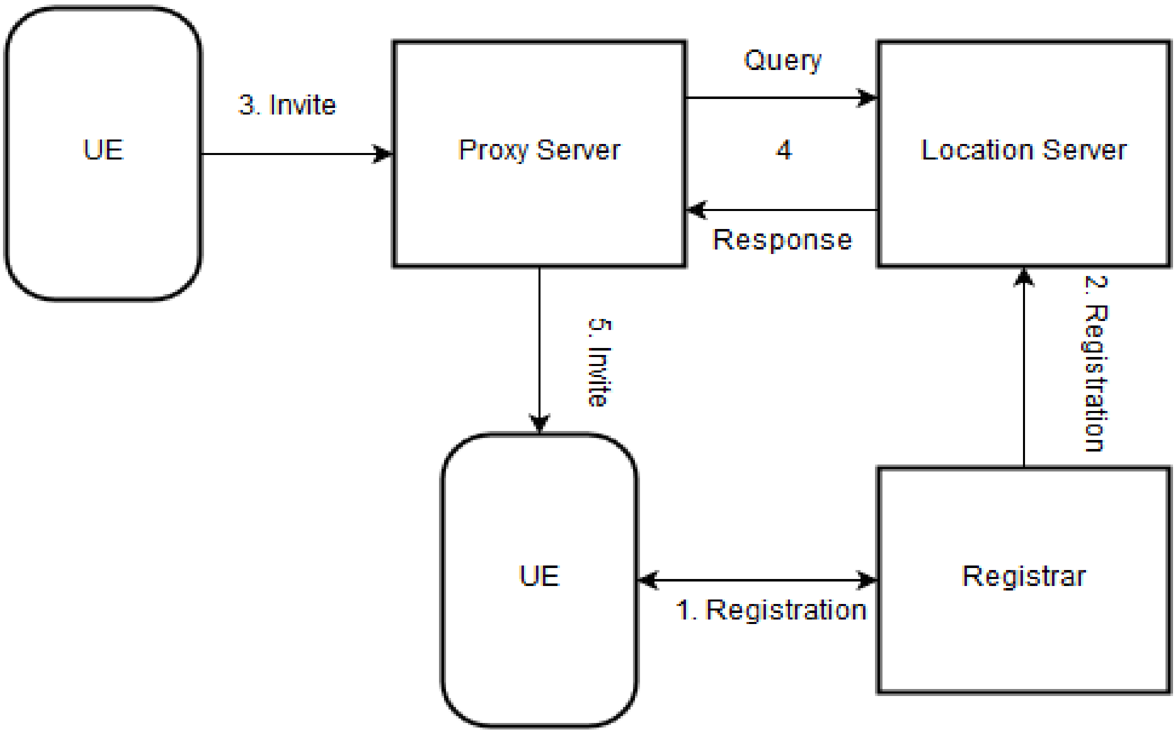

Figure 2-7: Functions of each network element in a SIP Session establishment (adapted from figure of [37])

In addition, two additional SIP servers can be used:

Redirect Server	The redirect server is a UA server that helps with the session initiation process. It receives requests and searches for the corresponding called device(s) in the database that is updated by the registrar. Then, the redirect server sends this information to the calling device and directs the UAC to connect with an alternative URI.
Location Server	The redirect server or the proxy server access the location server to retrieve the location of the callee. For this reason, the location server manages a database that stores IP addresses associated with each SIP URI.

2.5.2 SIP Messages

Two types of SIP messages (requests and responses) are needed to establish a call (i.e., session). Every request is characterized by a method and uses a Request-URI to specify where the request is to be sent. A response message is characterized by a response code. In the following sections, SIP requests and responses are analyzed on a high level.

2.5.2.1 SIP Requests

SIP requests are used to initiate, modify, or terminate a communications session. The SIP responses sent in answer to these requests clarify whether the request was successful or not. The method, the Request-URI, and the SIP protocol version form a Request-Line [34]. There are two types of methods: core methods and the extension methods [40].

There are six core methods:

INVITE	The INVITE method is used to initiate a session between UAs.
BYE	The BYE method is used to end an established session. Only UAs are able to send a BYE message. This means that proxy servers or any other SIP servers cannot send such a message. BYE methods are routed directly from one end-point to the other bypassing any proxy servers.
REGISTER	A REGISTER message is sent from a UA to a registrar server to request registration of a UA or to terminate an earlier registration.
CANCEL	A CANCEL request is used to terminate initiation of a session that has not been established. This request can be sent by UAs or any proxy server to abort a pending session establishment.
ACK	The ACK request is used to acknowledge an INVITE request.
OPTIONS	The OPTIONS method is used to transfer information regarding the capabilities of a UA to other SIP nodes.

Additionally, there are extension methods. These methods include:

SUBSCRIBE	A SUBSCRIBE request is sent by the UAs subscribe for notifications about a particular service.
PUBLISH	The PUBLISH message is sent upon the occurrence of an event that is reported by a service.
NOTIFY	The NOTIFY message is sent to those UAs who subscribed to be notified about a specific event.
REFER	The REFER message requests that the receiver refers to a resource that is provided in the request [41].
INFO	The INFO method is used by a UA to send signaling information to another UA when there is an established media connection.
UPDATE	A UPDATE request modifies the state of a session. For example, during a session the user can change their choice of CODEC.
PRACK	Provisional Acknowledgement (PRACK) [42] is similar to ACK, but its role is to acknowledge *provisional* responses (see Section 2.5.2.1).
MESSAGE	The MESSAGE method is used to send an instant message using SIP.

2.5.2.2 SIP Responses

A SIP response is a reply from the UAS or any SIP server to a SIP request generated by a UAC. These messages may carry some extra header fields that contain additional info that is needed by a UAC. SIP responses are categorized into six classes. The first five, have been taken from HTTP and the last category is unique to SIP [43]. These classes are distinguished by the first digit of a three digit number. All of the classes but the first one, are final responses. The first class 1xx is considered to be a provisional response. Table 2-1 shows a summary of these classes with a short description of their functionality.

Table 2-1: SIP Categories referenced from [34] [43] [44]

Class	Description	Functionality
1xx	Informational	These provisional responses act as indicators of a specific action in progress that is performed by a server (which does not have a final response at this time).
2xx	Success	Success indicates that the request was successfully delivered, understood, and accepted.
3xx	Redirection	Redirection responses return information regarding the UA's new location or additional services that may be able to complete the call.
4xx	Request Error	This class of responses concerns requests that have failed due to an error caused by the client. The client can resend the request if the request is modified according to the response. Note that if the same request was sent to another server it might be successful.
5xx	Server Failure	This response informs the sender that the request failed because of an error in the server. The same request might be sent to an alternative server.
6xx	Global Failure	The request is unable to be satisfied at any server.

2.5.3 SIP Registration Flow

According to RFC 3665 [45] and Andrew Prokop [46], a successful new SIP registration (shown in Figure 2-8) can be described through the following steps:

1. The UA sends a SIP REGISTER message request to the registrar. The headers To and From include the user's identification, or the Address of Record (AOR). In the REGISTER, and more specifically, in the Expires header, there is a value indicating the time during which the registration will be valid.
2. The Registrar responds to the UA with a 401 unauthorized response along with a WWW-Authenticate header. This header is used to send data that will help in the encryption of the user's password and it contains a nonce* and information regarding the encryption algorithm that the UA must use.
3. The UA sends another REGISTER request to the SIP server that has an authorization header. This header contains the user's encrypted password.
4. Finally, assuming that the password is correct, a 200 OK response is sent back to the UA as a confirmation of a successful registration.

* A nonce is usually a long random or pseudo-random number that is used for encryption purposes

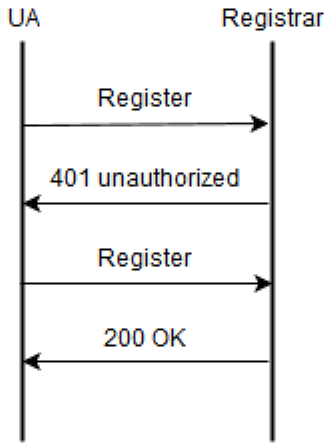

Figure 2-8: A successful new SIP Registration flow

2.5.4 Successful Session Establishment

As discussed previously, users can establish a session without any additional network entities (such as proxies). In this section, the demonstration of a basic call flow (shown in Figure 2-9) will involve a proxy server to present a more realistic scenario. According to RFC 3665 [45] and [47], a successful SIP session establishment involves the following steps:

1. The caller sends an INVITE request to a proxy server to initiate a session.
2. In response, the proxy server sends a 100 trying message back to the caller. This provisional response avoids any unnecessary re-transmissions of an INVITE request to the proxy server.
3. Additionally, the proxy server sends a query to the location server in order to find the address of the callee. Afterward, the INVITE is forwarded to the callee's UA.
4. In response, the callee sends an 180 ringing response back to the caller via the proxy server.
5. As soon as the callee answers the call, a 200 OK response is sent to the caller via the proxy server.
6. Once the caller receives the 200k OK response, it sends an ACK back to the callee.
7. In the meantime, the session is being established and the media is transmitted using the RTP protocol.
8. When the conversation reaches an end, either the caller or the callee sends a BYE request to terminate the session.
9. Note that the callee receives the BYE request directly from the caller without passing through the proxy server.
10. Lastly, the callee sends a 200 OK response to the caller confirming the BYE requests and after that, the session is terminated.

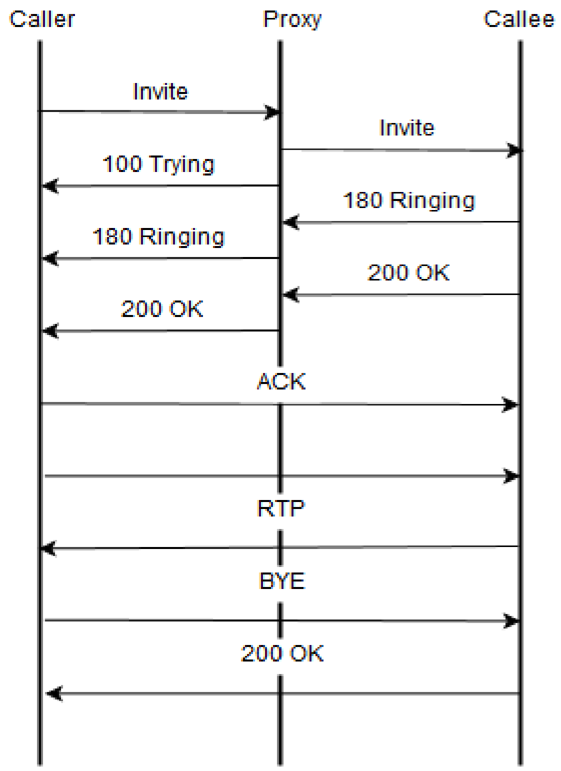

Figure 2-9: SIP Basic Call Flow

2.6 The RTP protocol

The Real-time Transport Protocol (RTP) operates in the application layer along with SIP. RTP is independent of the underlying transport and network layer. The fact that RTP is defined as a transport protocol can be misleading since RTP uses other transport protocols, such as UDP. RTP offers end-to-end transport of real-time data, such as audio, video, or simulation data. These media streams can be multicast or unicast. RTP is used in conjunction with the Real-time Transport Control Protocol (RTCP). RTP handles the multimedia streams, while RTCP is responsible for monitoring transmission statistics and measuring the QoS. However, RTCP does not guarantee QoS and make resource reservations. In 1996, RTP was defined in RFC 1889 [48] and the latest version was defined in 2003 in RFC 3550 [49].

RTP supports four main functions: identification of payload type, source identification, sequence numbering, and timestamping [50]. An RTP session is normally comprised of an RTP stream, commonly using a UDP port with the port number (n), and corresponding RTCP traffic, commonly using UDP port number (n+1). The datagrams are sent from/to a user's IP address. Figure 2-10 illustrates the structure of an RTP packet.

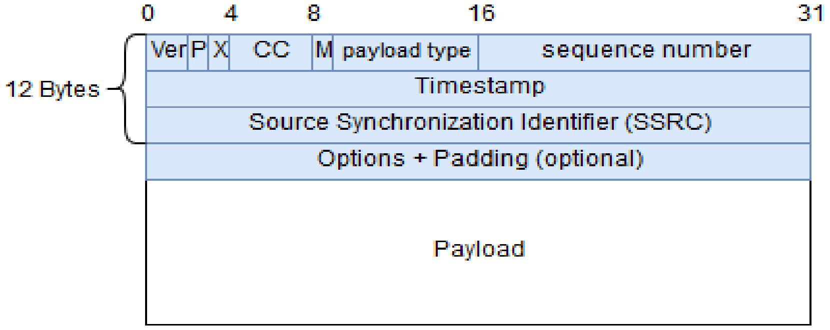

Figure 2-10: RTP packet

2.7 GSM CS Voice roaming architectures

Before analyzing the different roaming models designed for voice service delivery over LTE networks, the roaming architectures for Global System for Mobile communications (GSM) are worth mentioning. Therefore, in this section, three methods for roaming in GSM networks will be described at a high level. Currently, in circuit switched (CS) GSM networks, there are mainly two options for roaming: VPMN routed [51] and CAMEL-enabled HPMN routed [52]. There is also a third option called Optimal Routing (OR) [53], but it does not seem to have been applied very often.

2.7.1 CS calls handled by the VPMN

Usually, in GSM roaming, the VPMN network is responsible for routing voice calls for inbound roamers towards the terminating network. Because of this, the VPMN and the interconnection providers need to be service aware and cannot simply act as a “bit pipe”. Certainly, the visited network is unable to route calls unless it has the necessary information from the home network. The Visited Location Register (VLR) in the visiting Mobile Switching Center (MSC) communicates with the Home Location Register (HLR). The VPMN can receive the authentication and the service profile parameters using the Mobile Application Part (MAP) protocol according to the Location Update (LU) procedure [54]. Also, in most call scenarios, the routing is optimized (to a certain degree) and a shorter path is chosen towards the destination network. However, this is not always the case in VoLTE roaming. Chapter 3 describes the equivalent roaming approaches for VoLTE.

2.7.2 CS calls assisted by CAMEL

CAMEL is a set of standards that offer many services on a GSM/UMTS network. While deep analysis of CAMEL is out of the scope of this Master’s thesis, CAMEL will be mentioned in order to describe another voice roaming architecture which uses the assistance of CAMEL. For further information, the reader should refer to the ETSI TS 123 078 [55]. In most cases, the CAMEL Service Environment (CSE) resides in the home network. Therefore, the VPMN will be aware that the roamer is making use of CAMEL triggers according to the call handling specified in the VLR record. Based upon the call origination trigger for a specific subscriber, service logic will be invoked in the visited network by the visited MSC (V-MSC) with the CSE giving routing information to the V-MSC to route the call. In HPMN routing, the V-MSC will transfer the call to a CSE call processing application that routes the subscriber’s call to the terminating destination.

2.7.3 Optimal Routing

In Optimal Routing (OR), CAMEL origination triggers are used in order for the HPMN to be able to make routing decisions for its subscriber’s calls. Depending on the interrogation of the HLR of the called party, the HPMN of the subscriber who initiated the call determines whether the visited network is a better option than itself to route the call. One example of the use of OR would be a call scenario where there are two subscribers of the same HPMN who are roaming in the same visited network and one subscriber user calls the other. In order to avoid unnecessary charges from a GRX provider for this call, OR can be determined by the HPMNs of both the caller and the callee *provided* that the CAMEL phase 3 is supported. The decision as to whether or not OR is applied to a specific call is made by the HPMN of the calling party. Regardless of the promising routing optimization, this architecture is not implemented in commercial networks due to tariff issues and there are no specific means to charge differently for the calls that OR has been applied to. A study that addresses these issues has been made by GSMA in IR.20 [56].

2.8 VoLTE

Voice over LTE (VoLTE) refers to voice carried over the IMS network and through an LTE access network. VoLTE was firstly defined by GSMA in PRD IR.92 [57] and the first deployment took place in Singapore on 19th of May of 2014 [58]. In more detail, the voice traffic (signaling and media) are delivered in the form of data flows using LTE data bearers. Since voice services are sensitive to latency, these bearers are prioritized above other data to provide a better QoS. If a user wants to make use of VoLTE this user needs to be authenticated using the Diameter protocol when using the SIP protocol for registration, call setup, etc. Some of the main advantages of VoLTE are: (1) it offers "High Definition Voice" and (2) it is possible to achieve a greater capacity for data and voice while using existing spectrum allocations. This is possible because the VoLTE's network capacity is three times greater than 3G UMTS and consequently, approximately six times more than a GSM network [59]. Additional advantages include the possibility for the operators to switch off their legacy CS domain and move to a cheaper "all-IP" domain, from the user's perspective call setup time is noticeably quicker in VoLTE, and finally, the same IMS can be reused to offer other services, such as VoWiFi, Video over LTE (ViLTE), and RCS. Nowadays, one major challenge for telecommunication operators is to provide VoLTE roaming. Chapter 3 describes the different approaches that are capable of delivering such a service.

2.9 Related work

In recent years, 3GPP and GSMA have defined two main roaming architectures: RAVEL and S8HR. These roaming architectures will be described in Chapter 3. According to the author's research, there is very little related work that analyzes these roaming models in order to distinguish whether or not a particular technique is better suited than another in a specific scenario. Also, in quite a few articles, there is no distinction between the variations of an LBO architecture. This can lead to misconceptions of the VoLTE roaming architectures. For example, the majority of reports that cover the topic of this thesis only mention one of the three available LBO options, typically LBO-VR or RAVEL. Despite the fact that there are two additional options: LBO-HR and the LBO-OMR. These are described in Sections 3.2.1 and 3.2.3, respectively. Moreover, there is no specific distinction between the LBO options: LBO-VR and LBO-OMR. Furthermore, the existing literature is not fully updated on later studies of VoLTE roaming architectures. For example, the S8HR model is still presented as inadequate to fulfill certain requirements, despite recent studies by 3GPP and GSMA that have shown progress in resolving these issues. Consequently, there have been no criteria stated that could be used to guide an operator when selecting a roaming architecture. Nevertheless, in the literature, there are many papers and a few books that analyze these roaming models on a high level. Chapters 3 and 4 are based upon these previous studies. For example, the book "The LTE Advanced Deployment Handbook" [60] mentions the 3 different roaming techniques in its Chapter 4. Towards the end of that chapter, there is a summary of the main advantages and drawbacks of Home Routing and Local Breakouts. This project will analyze this topic more thoroughly and will attempt to define criteria that can be used in order to be able to select the most suitable roaming model for each situation.

3 VoLTE roaming architectures

Roaming allows subscribers to use their services when visiting a network other than their home network. One important prerequisite for roaming is that the roaming partners have to use the same network standards and have a roaming agreement with each other. Roaming can be national or international. National roaming occurs when a subscriber can use other network provider‘s network(s) or services even while they are located in their home (network operator's) country. In contrast, international roaming occurs when the mobile subscriber can connect via networks and services when they are abroad (i.e., not in their home (network operator's) country). This thesis will focus on international roaming, rather than national roaming.

Based upon Huawei's LTE International Roaming Whitepaper [14], there are two general techniques for roaming: home-routed roaming and local breakout. Home-routed roaming is commonly used in 2G/3G networks. This method has been successful and all of its issues have been resolved. Home-routed roaming supports LTE roaming and was the preferred means of implementing roaming in early deployments of LTE roaming. Unfortunately, it comes at the cost of sending all signaling and user traffic from the visited network to the home network, even if the ultimate destination of the user's traffic is in the visited network! One of the advantages of LBO compared to home-routed roaming, is that mitigates unnecessary media traffic loops between the visited and the home network, hence it uses fewer transmission resources. Additionally, there is a lower delay, hence with LBO the provider offers a better user experience and can do so at a lower cost. However, there are some difficulties in deploying LBO. For example, functions such as policy control, service control, and charging are complex since the user plane is handled in the visited network, hence the home network operator does not have direct access to this traffic or even information about it.

GSMA and 3GPP currently define two architectures for realizing Voice over LTE (VoLTE): Roaming Architecture for Voice over IMS with Local Breakout (RAVEL) [32] and S8HR [61]. A common misconception is that the LBO and RAVEL can be used interchangeably. LBO has three variants: Home Routing, Visited Routing, and optimal media routing. Using one architecture or the other is controlled by the HPMN and may vary on a call by call basis.

3.1 Home Routing

According to Jyrki T. J. Penttinen's book, *The LTE-Advanced Deployment Handbook The Planning Guidelines for the Fourth Generation Networks* [60], home-routed roaming is based upon an architectural model that transfers both the control plane traffic and the user data from the VPLMN to the HPLMN. This means that the HPLMN provides all the services, while the VPLMN simply acts as a pipe for all traffic. More specifically, the roaming data are routed inside a GPRS Tunneling Protocol (GTP) tunnel through the GRX network, starting from the Serving GPRS Support Node (SGSN) located in VPLMN to the Gateway GPRS Support node (GGSN) located in the HPLMN. The same technique can be applied in LTE roaming with the traffic accessing the Serving Gateway (SGW) and the Mobility Management Entity (MME) in the VPLMN on its way towards the Proxy Gateway in the HPLMN. Figure 3-1 shows the situation when subscriber A calls subscriber B when both are visiting different VPLMNs. This figure shows that there are various optimizations that could be advantageous. For example, when A and B are both visiting the same VPLMN (as shown in Figure 3-2) it is clear that the media traffic between A and B should go directly from one to the other *within* the VPLMN, rather than being tunneled back to two different home networks and then tunneled between these two HPLMNs and then tunneled back to the VPLMN.

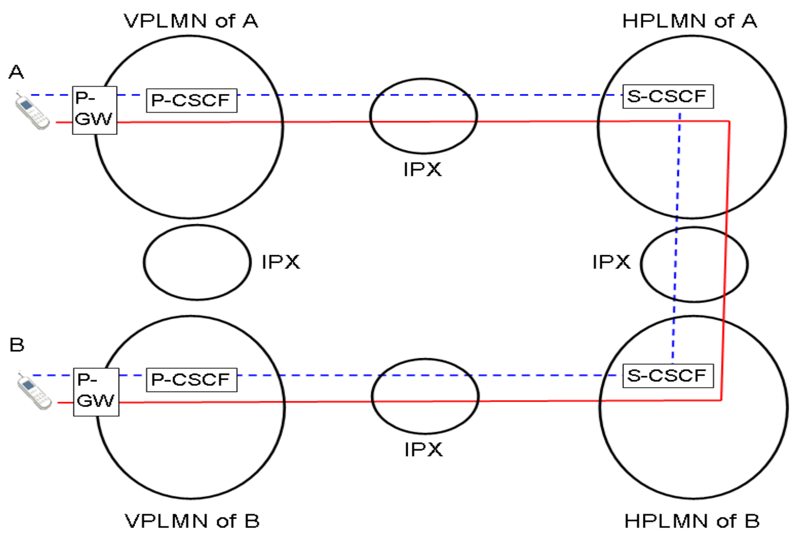

Figure 3-1: **Home-routed roaming when the VPLMN of A and VPLMN of B are different and both subscribers are outside of their HPLMNs**

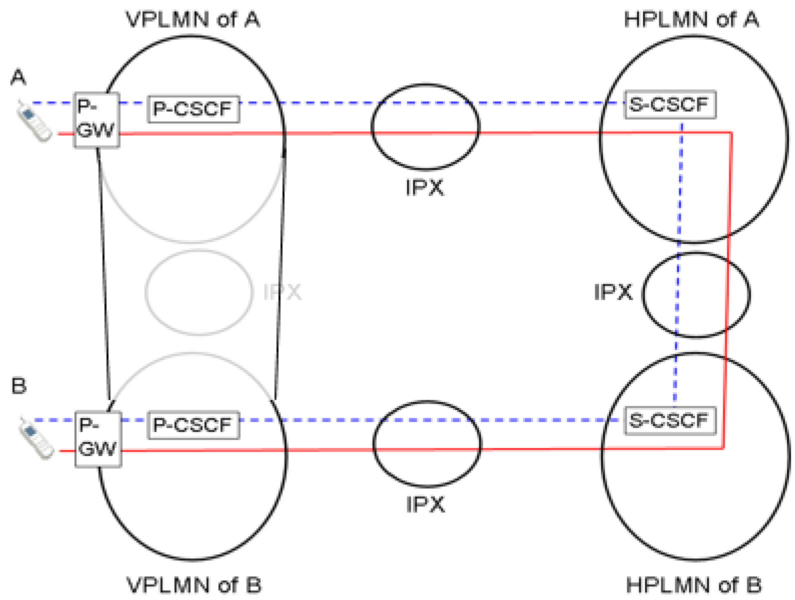

Figure 3-2: **Home-routed roaming when the VPLMN of A and B is the same VPLMN**

3.2 LBO

GSMA in IR.65 [62] identified three basic requirements for IMS voice services that are based on local breakout:

1. Firstly, the media routing regarding Voice and Video over IMS when the caller is in a roaming state should be at least as optimal as that already used in a CS domain. This means that the media should not be routed towards the HPMN of the originating party (unless this has been requested

by the HPMN). This requirement is met with the help of Local Breakout VPMN Routing (LBO-VR).

2. The charging model for roaming that is used in a CS domain should not be altered in voice over IMS. This requirement is fulfilled with the implementation of a P-CSCF in the visited network along with the help of TRF.
3. The HPMN should have the ability to choose according to its "service and commercial considerations & regulatory obligations" to steer the roaming subscriber's traffic to itself. The fulfillment of this requirement is met by enabling the home routing option in an LBO infrastructure (LBO-HR).

The local breakout architecture offers three different options:

- LBO Home Routing (LBO-HR)
- LBO Visited Routing (LBO-VR)
- LBO Optimal Media Routing (LBO-OMR)

In spite of the different implementations, some general principles and technical specifications remain the same for the three options of LBO, specifically:

- The UE connects to the IMS Access Point Network (APN) through the PGW located in the VPMN. Other services regarding data use other APNs and can still use the S8 interface.
- The P-CSCF is located in the visited network.
- The VPMN's IMS User Network Interface (UNI) and IMS roaming Network-to-Network Interface (NNI) between the VPMN and HPMN are used to support IMS voice calls.
- Apart from voice, traffic related to Rich Communication Services (RCS) such as SMS, Video calls, and enhanced address book, that makes use of the IMS APN is broken out locally at the visited network.
- An important attribute of local breakout is that is service aware because it allows traffic distinction and thus more efficiently provides QoS.
- When an emergency call is being transferred over the IMS APN, the P-CSCF located in the VPMN identifies the dialed number and rejects the call with a SIP response code of 380. This will cause the UE to make another call attempt according to the emergency procedures defined in 3GPP TS 23.167 [63].
- The session control of the calls is handled by the S-CSCF in the HPMN of the *calling* party.
- Similarly, the Telephony Application Server (TAS) is also located in the HPMN of the *calling* subscriber.

3.2.1 LBO-HR

LBO Home Routing (LBO-HR) can be used when both the control and the user plane are to be routed to the HPMN. The architecture is similar to that described in Section 2.7.2 regarding the GSM CS Voice roaming architectures. Figure 3-3 illustrates the architecture of an LBO-HR. As can be seen, the media and the signaling follow the same path. This LBO option makes it possible for the home network to control the routing of calls to the destination. Depending on the call scenario, LBO-HR can be either efficient or deficient. For example, when calls have as a destination the HPMN, then LBO-HR is a preferable option since both control and user planes must reach the roamer's home network. However, LBO-HR can have a relatively long media latency when the calls are not terminated in the home network or in an alternative PMN in the home region. Latency is especially high when the calls are between areas that are geographically distant from the home network.

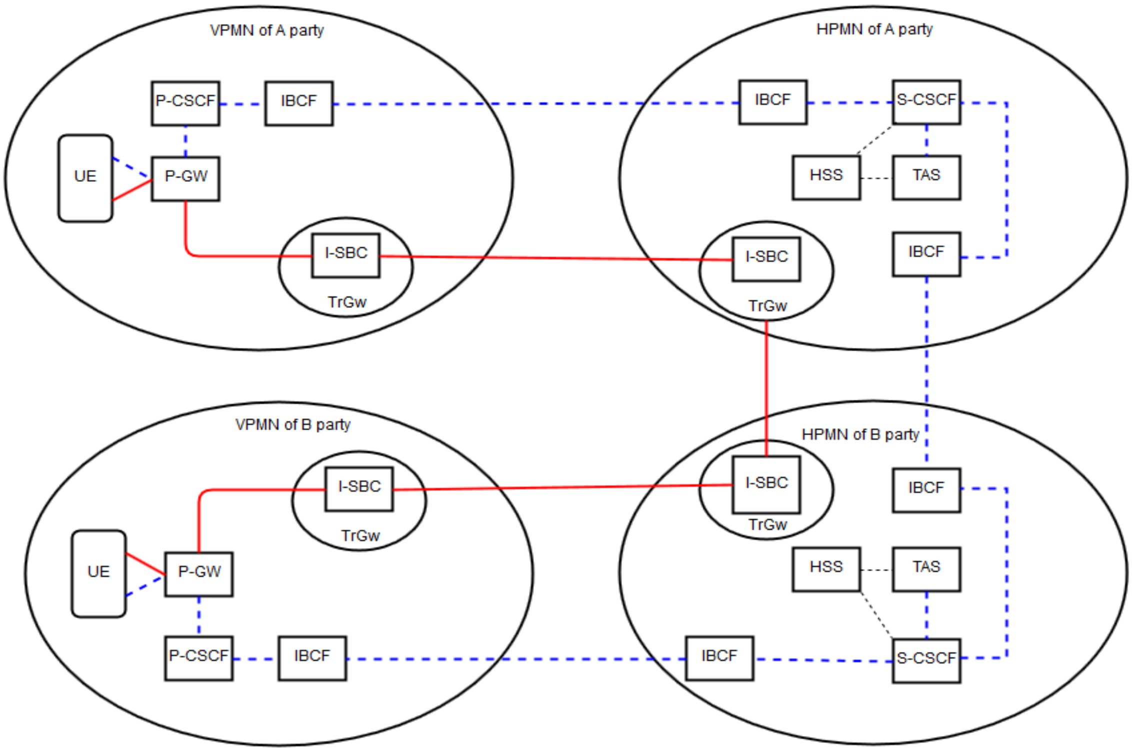

Figure 3-3: LBO-HR adapted from IR.65

To make this option more comprehensible, the equivalent session origination procedure (described earlier) is explained below, as described in the 3GPP TR 23.850 [32]. The invocation of this procedure is controlled by the home network. This procedure specifies the standard IMS routing case where the terminating network is reached over IMS routing from the home network. However, this flow also applies to all other types and different locations of terminating networks, such as CS breakout and when the HPLMN is a destination network. Figure 3-4 shows the flow of the origination procedure:

1. Firstly, the roamer's UE will send an INVITE request to the P-CSCF located in the VPMN.
2. Then, the INVITE is forwarded from the P-CSCF to the IBCF in the visited network.
3. The IBCF will allocate a TrGW for the user plane traffic and forward the INVITE request further.
4. The INVITE request has now left the VPLMN and goes through an intermediate network (such as an IPX network) to the IBCF in the home network.
5. Once the request has reached the IBCF in the HPLMN, the IBCF forwards the INVITE request to the S-CSCF.
6. The S-CSCF invokes the corresponding ASs in order to perform service control and invokes a TRF.
7. The TRF executes the standard IMS routing procedures and forwards the INVITE request towards the destination network.
8. Finally, in the last step, the establishment of an end-to-end user plane traffic has been made from the roamer's UE over the IBCF/TrGWs towards the destination network.

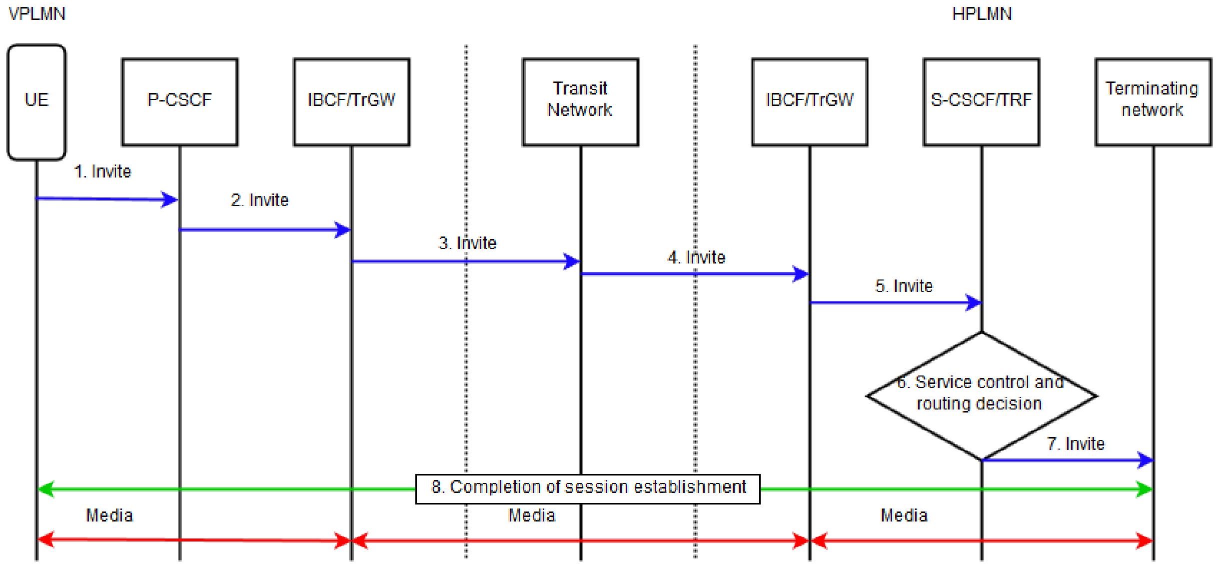

Figure 3-4: **Session origination procedure for home routing**

3.2.2 LBO-VR

LBO-VR (shown in Figure 3-5) can be compared with the architecture described in the Section 2.7.1 where the CS calls are handled by the VPMN. LBO-VR is a roaming model where signaling is handled in the same way as it is in home-routed roaming, i.e., all signaling is sent back to the subscriber's home network. However, the user-data is handled by the visited network. LBO-VR was mentioned by GSMA for GPRS roaming under the name "Visited GGSN roaming" [64]. Despite the advantages of such an implementation, home-routed roaming was preferred by many operators. The main reasons they cited for this decision were (1) the HPLMN's inability to monitor the QoS for the subscriber's traffic, (2) their inability to charge their subscribers for this traffic, and (3) because home-routed roaming was believed to be simpler [60].

As mentioned above, one of the main reasons why an LBO-VR architecture is considered to be better than home-routed roaming is because the user plane traffic is handled more efficiently with LBO-VR. For example, for services where low latency is a desirable QoS factor (such as for VoLTE), it is highly preferable for the user traffic to be routed using the shortest possible path to the destination. If a subscriber is roaming in an area far from the HPLMN and initiates a call via VoLTE to a local subscriber who has the same HPLMN, then as we saw earlier in Figure 3-2 the user's media traffic is routed from the VPLMN to the HPLMN and back. According to IR.34 [16], a typical Round Trip Time (RTT) between southern Europe and East Asia through an IPX/GRX network is 414 ms. However, a delay greater than 500 ms for voice is considered "unacceptable" [65] and delays greater than 177 ms (a one-way delay which corresponds to mouth-to-ear delay) rapidly deteriorate in perceived quality with increasing delay [66]. Therefore, it is understandable that LBO can provide a better QoS. The increasing adoption of VoLTE is the main driving force for the adoption of LBO.

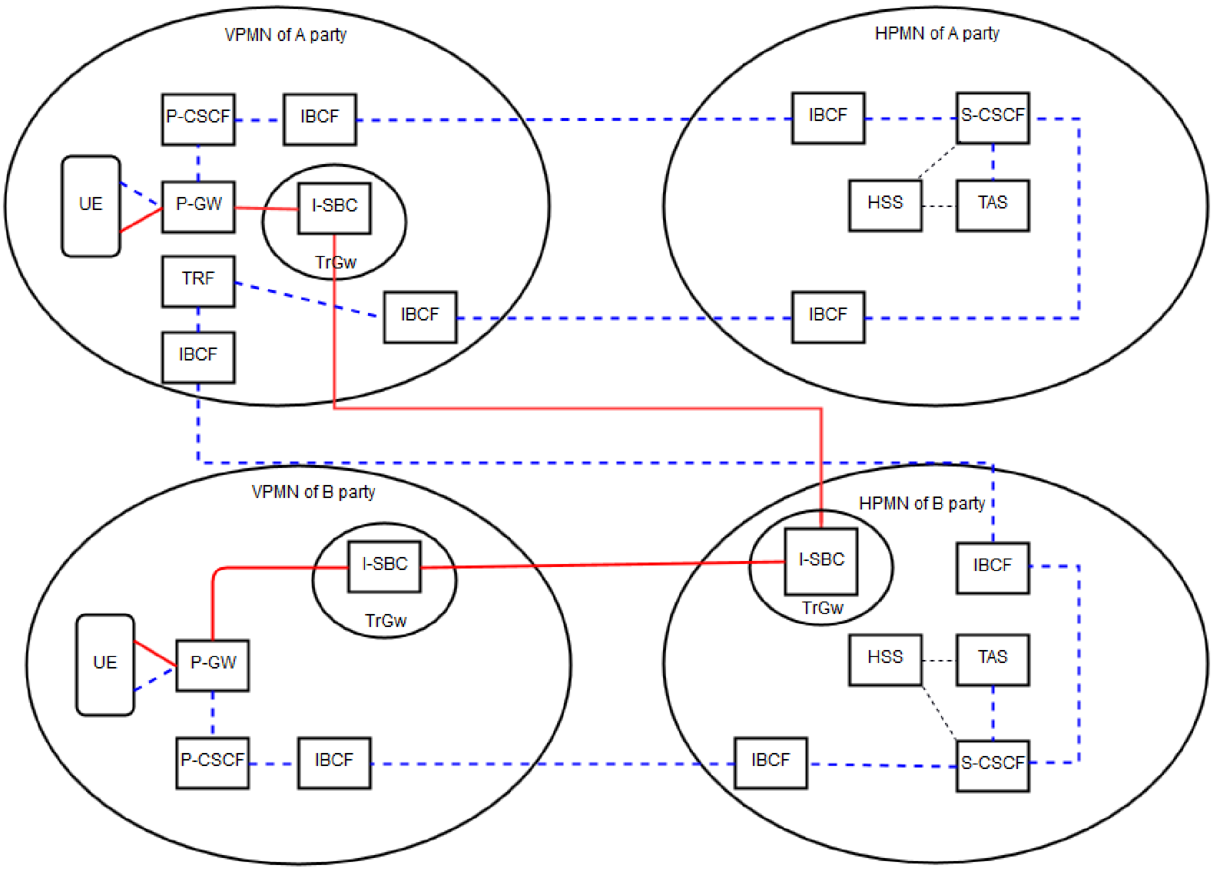

Figure 3-5: LBO-VR adapted from IR.65

The same method as in the Section 3.2.1 is used to describe the session origination procedure for routing in the visited network. Again, the home network decides whether or not to invoke this procedure. The flow of the origination procedure is shown in Figure 3-6. This procedure specifies the standard IMS routing case over which the terminating network is reached over IMS routing from the visited network. However, this also applies to all other types and different locations of terminating networks, such as CS breakout and when the VPLMN is a destination network.

1. First, the roamer's UE sends an INVITE request to the P-CSCF located in the VPMN.
2. Then, the INVITE is forwarded from the P-CSCF to the IBCF in the visited network. This allows TrGW bypass using Optimal Media Routing (OMR). This request can potentially contain location information.
3. The IBCF allocates a TrGW for the user plane traffic and will apply the standard OMR procedures to process and forward the request to enable this TrGW to be bypassed if the INVITE returns to the visited network and no other interconnecting nodes anchor the payload traffic before this request returns.
4. In this step, the INVITE request has left the VPLMN and goes through a transit network (such as an IPX network) to the IBCF in the home network.
5. Once the request has reached the IBCF in the HPLMN, the IBCF forwards the INVITE request to the S-CSCF.
6. Afterward, the S-CSCF invokes the corresponding ASs in order to perform service control and invokes a TRF in order to execute a modified routing procedure. The TRF routes the INVITE request back to the visited network according to its local policy, the identity of the VPMN, the status of the availability of any Multimedia Resource Functions (MRF) in the visited network that might be needed, and whether or not there is a need for MRF resource allocation in the VPMN.

7. The TRF obtains the address of the I-CSCF in the VPMN from the identity of the visited network and sends the request towards the I-CSCF in the VPLMN. A loopback indicator is included in the INVITE request from the TRF that advises the visited network that the current request is being sent back to the VPMN for transit routing. The TRF maintains the UE's location information in the INVITE request, provided that it is available. Afterward, the TRF passes on the request to an IBCF to permit TrGW bypass by using OMR.
8. Another IBCF in the HPLMN will transfer the INVITE request to the transit network.
9. The transit network in turn forwards the INVITE request towards the I-CSCF in the visited network.
10. Once the INVITE request reaches the VPMN, it passes through an IBCF that will notice that the SDP contains a media address *inside* the visited network - thus the reserved TrGWs can be bypassed. Next, the IBCF forwards the INVITE request to the I-CSCF.
11. According to the local policy and the existence of a loopback indicator in the INVITE request, the I-CSCF invokes the TRF to route the request toward the terminating network depending on the Request URI. The request URI contains the information about the terminating UA.
12. Since the routing decision already took place in the previous step, E.164 Number to URI mapping (ENUM) provides the domain of an onward network and the INVITE request is routed toward the destination network through an IBCF in the visited network. Then, the TRF forwards the INVITE to an IBCF that "anchors media". Provided that the location information of the UE is included in the request, the TRF can choose an IBCF near the point of call origination.
13. An IBCF in the visited network assigns a TrGW for the payload, whilst the INVITE request is being forwarded towards the terminating network. Interconnecting network entities towards the terminating network are neither mentioned nor shown in Figure 3-6.
14. Finally, in the last step, the establishment of an end-to-end user plane traffic has been made from the roamer's UE over the IBCF/TrGWs towards the destination network.

Regardless of the benefits of an LBO, one attribute that needs to be considered before deployment, is that the visited network is required to have an IMS or at least have the Packet Data Network (PDN) Gateway (PGW), P-CSCF, and PCRF nodes. In this requirement, an opportunity lies for IPX carriers, as instead of the VPLMN being responsible for breaking out the user plane, an IPX carrier could perform this action. This technique is called Carrier Breakout (CBO) or Mid Breakout (MBO) and it will be analyzed in Section 4.4 starting on page 45.

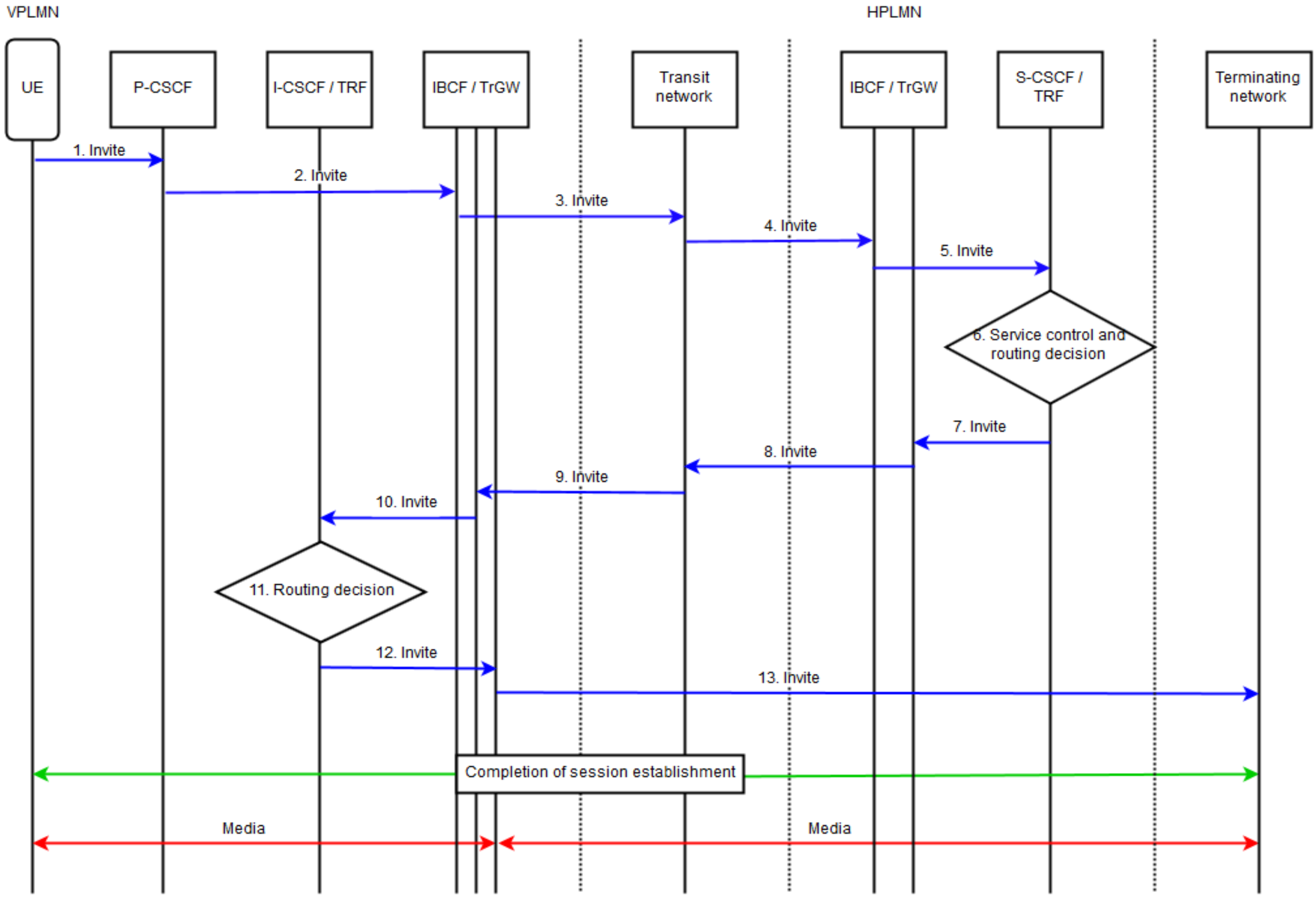

Figure 3-6: **Session origination procedure for VPMN routing**

3.2.3 LBO-OMR

The third option for LBO establishment is a local breakout with Optimal Media Routing (LBO-OMR). This approach can be compared with the CS approach mentioned in the Section 2.7.3. In this implementation, the control and user plane can be routed over different paths as shown in Figure 3-7. Therefore, the signaling is being handled in the same manner as of that of LBO-HR, while the media is routed on the most optimal path between the IMS end users. Another benefit of this model is that whether or not OMR will be applied is decided on a call by call basis, as the HPMN of the caller executes an HSS inquiry as part of the call routing procedure to choose the recipient's roaming destination and then determines whether or not OMR would be advantageous. If OMR is selected by the HPMN for the call routing, then signaling information is sent through all the relevant networks that will take part in this call. In this way, it is possible to enable an independent media connection within the networks that serve the subscribers. It is worth mentioning that the OMR option can not be used if the terminating party is in a CS domain. Also, a concerning fact about the ease of use of OMR depends upon all of the involved networks allowing OMR and agreeing upon choosing OMR. If one or more network does not agree to use OMR, then the route will not be fully optimized. Furthermore, there is no service awareness in the interconnection networks and the usual charging model for calls does not apply in this situation. Therefore, a new model for charging between the networks is needed.

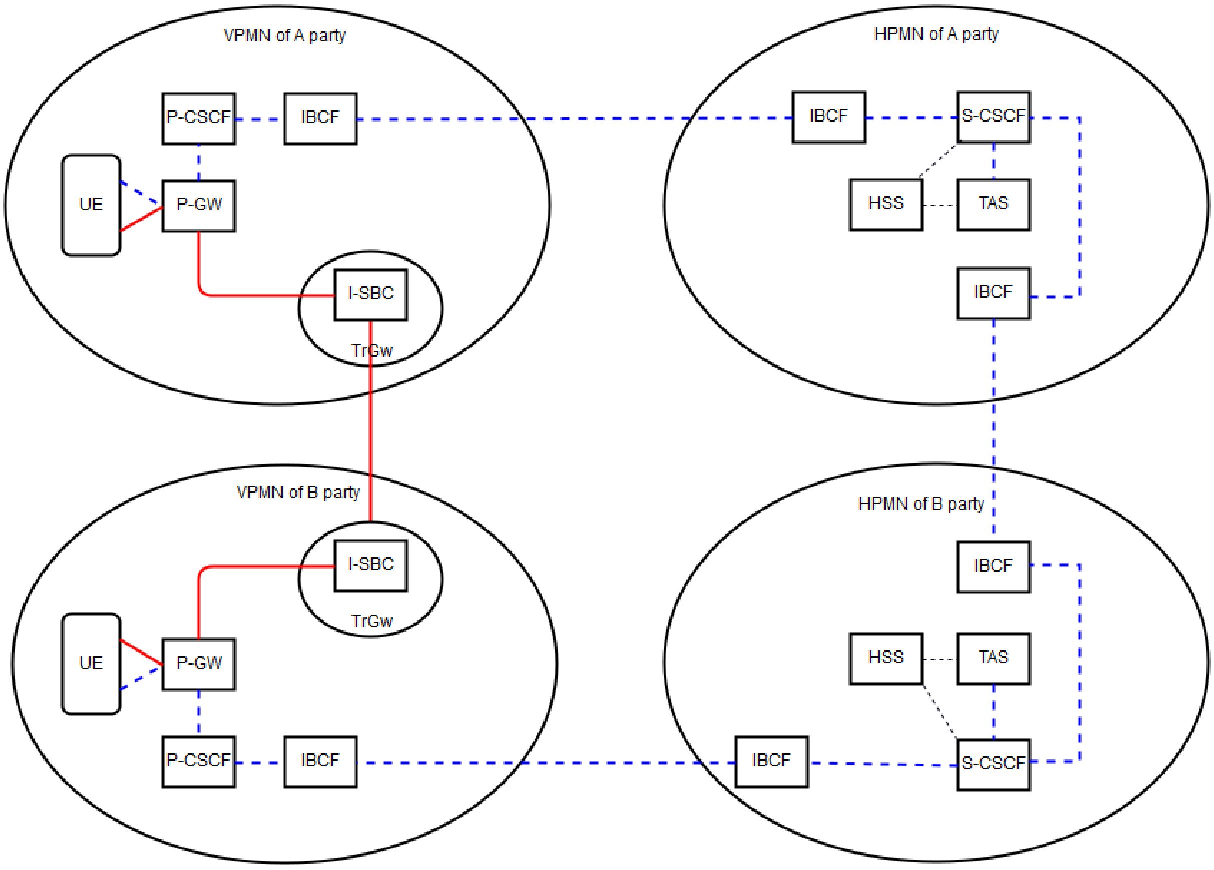

Figure 3-7: LBO - OMR

3.3 S8HR

Despite the efficiency of the RAVEL architecture, this model's requirements have been found to be rather demanding and the time-to-market is long. Additionally, the implementation of an LBO model is considered to be complex. Due to these reasons, GSMA introduced a new candidate for VoLTE roaming: S8HR. The name S8HR is derived from the LTE interface S8, while HR stands for Home Routing. The S8 interface lies between the SGW in the VPLMN and PGW in the HPLMN and uses GTP. The first VoLTE roaming implementation using the S8HR concept was made in 2015 [67]. In this method, the GTP protocol tunnels the signaling and the user-data between the visited network and the home network as shown in Figure 3-8 (where provider B is the HPLMN of the subscriber with UE 1). The roamer's call is handled by the IMS of their home network.

Home Routing is quite similar to mobile data scenarios and one of its advantages is that roamers can use VoLTE *regardless* of whether the VPLMN has an IMS architecture or not (as long as their home network does) [68]. Therefore, the PGW, PCRF, and P-CSCF are located in the home network. This makes S8HR a strong candidate for VoLTE roaming since the time-to-market is shortened because the HPLMN simply deploys S8HR, as there is no dependence upon the visited network.

The S8HR architecture is presented in Figure 3-8 and Figure 3-9 on a high level. In general, this architecture is similar to the Over The Top (OTT[*]) architecture. Since the routing for the traffic of both models is transferred in a quite similar way. The basic architectural characteristics as described in [69] and in TR 23.749 [70] are summarized below and done in a way to justify the argument that the time-to-marker for offering VoLTE roaming services is shorter than the RAVEL architecture:

- All the IMS services are routed towards the home network by using the "well-known"[†] IMS APN over the S8 interface (in the same way as regular data roaming traffic). Both the control and the user plane use the same well-known IMS APN implemented in the HPMN, but with different QoS Class Identifier (QCI) values. Apart from this distinction, there is no other difference between the signaling and the media traffic.
- In this architecture, the HPMN has the full control of the call handling.
- Since it is not a prerequisite for S8HR deployment for the VPMN to have an IMS, therefore network entities such as P-CSCF are not mandatory to implement as the visited network is *not* service aware. Differentiation among different services can be accomplished with the use of APNs or QoS levels. Although the QoS rules are produced in the HPMN, they are applied starting from the visited network according to the roaming agreement between the HPMN and VPMN. Thus, the VPMN could potentially downgrade or even reject the requested QoS [71].

[*] Over The Top (OTT) is the term that describes the delivery of voice, video or other media contents through an ISP that simply routes the related traffic but it does not have the control over the content nor any responsibility for it.

[†] In order to enable roaming for IMS services as well as VoLTE roaming, a well-known IMS APN is used. IR.88 [71] states that the APN name must be "IMS", which is also the APN network identifier that is part of the full APN.

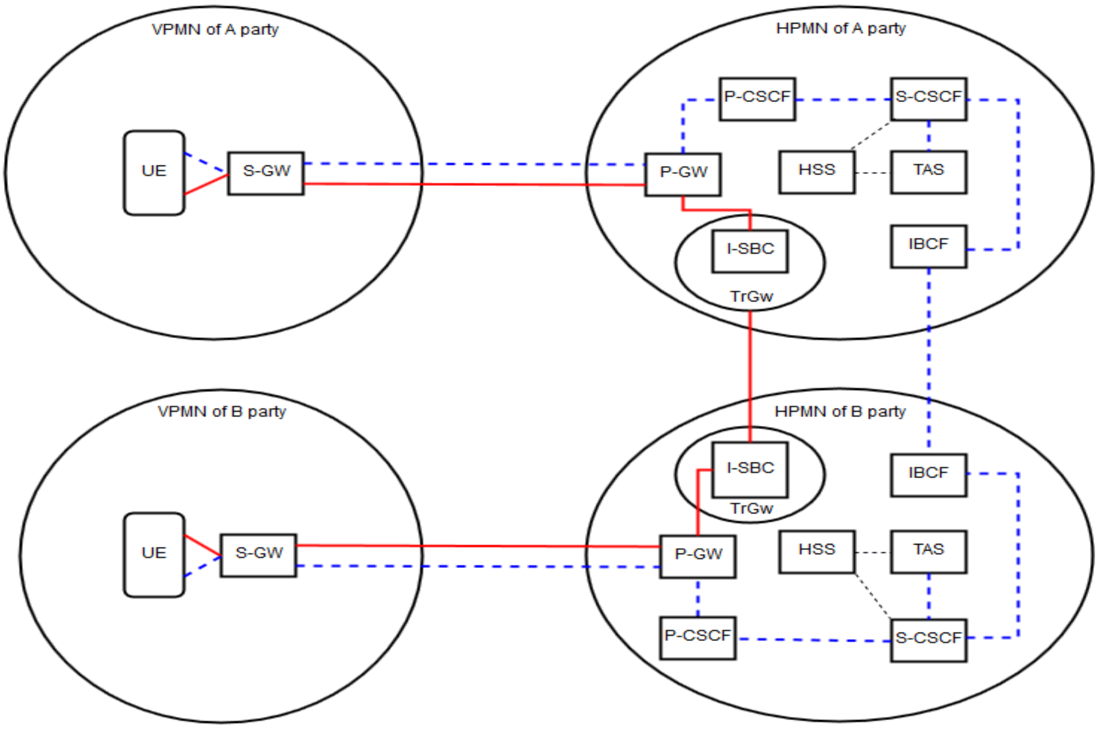

Figure 3-8: S8HR Architecture

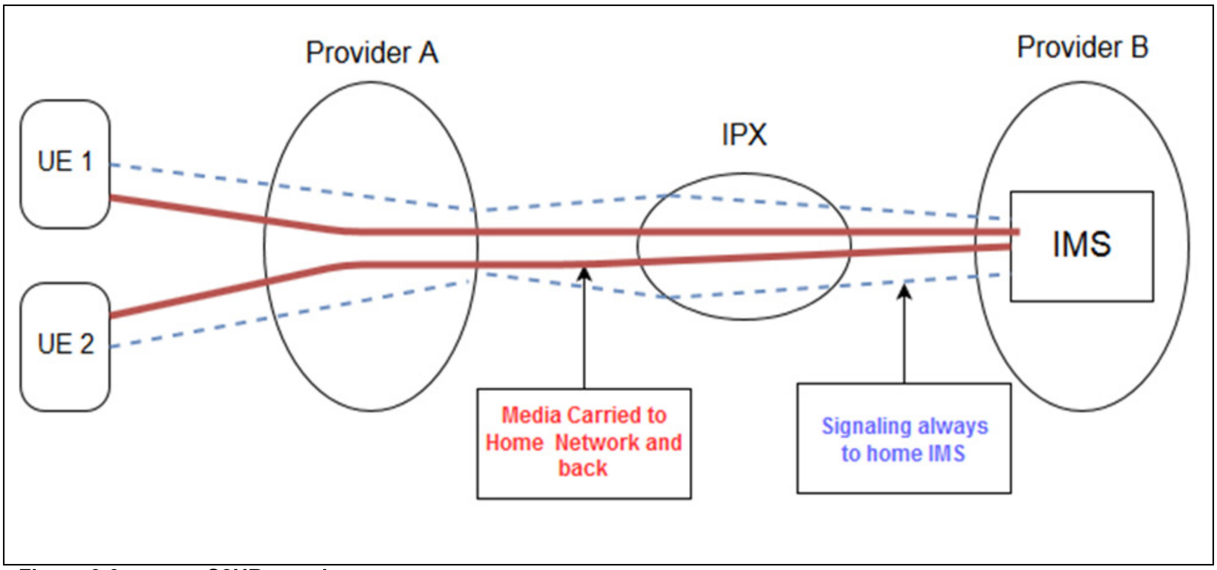

Figure 3-9: S8HR roaming

3.4 S8HR versus Local Breakout

In this section, the main advantages and drawbacks of the S8HR (summarized in Table 3-1) and Local Breakout summarized in Table 3-2) roaming architectures are summarized to give the reader an overview of the strengths and weaknesses of each of these roaming architectures. Chapter 4 compares these architectures more thoroughly according to the criteria defined in Section 4.1.

Table 3-1: Summary of S8HR

Advantages	Drawbacks
No requirement for VPLMN to deploy an IMS	The user plane is not routed in an optimal way to the destination
Reuse of the established packet switched roaming model	Inability to deliver latency-sensitive services, e.g., Voice on top of LTE in a worst case scenario.
No significant changes are needed to the routing arrangements, agreements, charging models, etc.	Increase in OPEX due to extra traffic that has to be carried via IPX carriers
Handover is feasible within LTE and 2G/3G without altering the working procedure of packet switched roaming	Difficulties in initiating an Emergency Call.

Table 3-2: Summary of LBO

Advantages	Drawbacks
User-data are routed in a more efficient manner	Extra CAPEX is required due to many changes that have to be made in the roaming networks
It is possible to deliver services such as VoLTE even in worst case scenarios	VPLMN is required to deploy an IMS architecture in order to serve the roamers
Lower OPEX, since the user plane is not routed back to the HPLMN	HPLMN is unable to monitor if the subscribed services of the roamers are being appropriately served by the VPLMN
Potential new source of revenue for the VPLMN	HPLMN has to fully trust the VPLMN to provide accurate CDRs

4 Comparison of VoLTE roaming architectures

This chapter compares the four different VoLTE roaming models described in the previous chapter: (1) LBO-HR, (2) LBO-VR, (3) LBO-OMR, and (4) S8HR. These models are compared based on criteria that will be defined in Section 4.1.

4.1 Criteria

Based upon literature study of all the major network infrastructures, standards, and protocols that have a role in this thesis; and having described all the solutions proposed for VoLTE roaming – this section defines criteria upon which a comparison between these architectures can be made. The criteria have been select by considering the requirements discussed in Section 3.2 and the service quality delivered to the end user. Additional criteria were prompt after studying the survey [5] mentioned in Section 1.2.

The criteria are divided into two categories: general criteria and criteria depending upon the call scenario. The reason behind this distinction is that some particular call scenarios are more likely to occur than others are. For example, according to statistics and the author's personal experience, the most common roaming pattern is that the terminating destination of a roamer's call is the home network. Thus, it would be beneficial to compare these roaming architectures depending upon the call scenarios as the selection of a roaming option based on the specific call scenario may lead to a more desirable outcome. For example, an HPMN whose subscribers when they are roaming tend to call other subscribers located in their home network, rather than calling subscribers in another HMPN or VPMN, might choose a roaming model that favors this scenario.

4.1.1 General Criteria

This section presents a set of general criteria that are independent of the call scenario.

4.1.1.1 Level of compliance with CS roaming models

The first two requirements for an IMS-based roaming model (Section 3.2) refer to the CS model and state that the architecture should have same performance or better and comply with the current business model so that no adaptations are required in accounting. Hence, the intermediate providers (IPX carriers) have to be service aware. Additionally, the billing process should remain the same as for CS calls. Another thing to take into account with regard to preserving the same billing process as CS is that both the control and the user plane should be transferred along the same path.

4.1.1.2 Seamless Handover

This criterion is based on the support of Single Radio Voice Call Continuity (SRVCC) and the ability of the roaming architecture to be able to handover the call from an LTE network to a 2G/3G network using SRVCC. This procedure is currently important to mobile network operators and preserves the QoS that is provided to their subscribers. Although LTE deployment is on the rise, its coverage and availability do not yet reach the same levels as of that of CS networks. Therefore, it is vital that the continuity of the calls continues even when the UE moves out of LTE coverage.

4.1.1.3 Support for future IMS services

With the rapid technological changes, advancements, and new features, it is reasonable to consider the sustainability and the readiness of an infrastructure to support new IMS services. For this reason, it is important for the VoLTE roaming models to support new IMS services, thus operators can rapidly offer these new services to their subscribers *without* being forced to make any additional investments.

4.1.1.4 Complexity and service equality

One of the main attributes of VoLTE roaming is that the roamers are able to use the same services in the visited network as they could on their home network. Depending upon the roaming architecture, this ability is provided with either more or less complexity and if the roaming option depends upon many other networking systems there might be a limitation of the available services that the subscriber can use when in a visited network. Using this criterion, an effort will be made to distinguish the level of complexity and independence from other systems of a given roaming architecture.

4.1.1.5 Regulatory compliance

As CS roaming calls must comply with various regulations, it is expected that VoLTE roaming will have to follow these same regulations. When introducing VoLTE roaming, two specific challenges arise. Support for emergency calls and support for lawful intercept. For many operators, these regulations are mandated by the local regulators. Since these obligations are considered vital for dealing with emergency situations and security is important, the VoLTE roaming models will also need to comply with these regulations. More significantly, local emergency services should be available to roamers and they should be able to make use of them without any issues arising. Regarding lawful intercept, the involved PLMNs should be able to deliver voice calls to the local authorities when given a court order to do so.

4.1.2 Criteria depending on the call scenario

This subsection introduces a set of criteria that depend upon the call scenario.

4.1.2.1 Deployment

This criterion will be used to compare the ease of deploying the VoLTE roaming architectures in order to actually implement VoLTE roaming.

4.1.2.2 Routing Efficiency

Routing efficiency will address the way the user plane is routed and considers the latency introduced. This criterion will also consider the possibility for route optimization.

4.1.2.3 Call Quality

Call quality is considered to be an important factor for VoLTE and is one of the key arguments from operators to persuade their subscribers to make use of VoLTE [72]. The same argument also applies to VoLTE roaming. In a roaming environment, the quality of the call depends on the scenario and the roaming implementation; hence, this criterion is included in the comparison.

4.2 Comparison of VoLTE roaming models according to defined criteria

In this section, an effort will be made to compare the different roaming implementations with the help of the criteria defined in Section 4.1. This section will follow the same structure as Section 4.1.

4.2.1 General Criteria

This section applies the criteria that are *independent* of the calling scenario when comparing the different VoLTE roaming models.

4.2.1.1 Level of compliance with CS roaming models

S8HR does *not* follow the same charging model as in CS roaming. In CS roaming, charging is proportional to the call's duration. In contrast, in the S8HR solution charging is based on QoS charging and the billing model is similar to charges for data roaming. This difference may create difficulties for accounting since a new billing model would have to be adapted for VoLTE roaming. In addition, due to the fact that the visited network is not service aware, it will be cumbersome to apply different charges for signaling and media. Another difference between the CS roaming model and S8HR is that the charges for call termination will be applied in the HPMN, rather than the VPMN.

Additionally, when the requirements were firstly set for VoLTE roaming, it was stated that the routing should be at a minimum as optimal as that already being used in a CS domain. In S8HR this does not always occur since the media will always be transferred to the HPMN *regardless* of the call scenario.

Both LBO-HR and LBO-VR share the same business models as CS roaming. Although LBO-OMR does not have any specified standard for a business model yet, hence this needs to be addressed in a further study by the relevant organizations. As far as the similarity between CS roaming and LBO options regarding route optimization, if all three options are able to be used on per call basis, then this requirement is met. Otherwise, if for example, only LBO-HR is available, then this model does not fulfill the route optimization requirement.

4.2.1.2 Seamless handover

According to TR 23.749 [70] in an S8HR architecture, the interface within the V-MSC and the IMS in the home network are based on Integrated Services Digital Network (ISDN) User Part (ISUP). The ISUP protocol is part of the Signaling System No.7 (SS7) protocol suite and it is used to determine the protocols and procedures that are needed to setup, manage, and release circuits that transfer voice and data calls in a CS network [73]. The use of ISUP in CS networks plays a comparable role to SIP in Packet-switched (PS) networks. As this interface uses ISUP, the improved version of SRVCC, called enhanced SRVCC (eSRVCC), is not supported in an S8HR implementation. Nevertheless, S8HR supports the releases of SRVCC before the 10th release. Although the deterioration of the end user's experience is non-negligible, the handover delay using SRVCC is greater than the delay in eSRVCC [74]. When using SRVCC, the HPMN takes part in the handover procedure, while in eSRVCC the VPMN performs the handoff. Therefore, eSRVCC offers shorter handovers, leading to a seamless handover that enhances the user's experience. One of the VoLTE quality metrics is that the voice interruption time during a handover should be less than 300 msec [75]. More specifically, this value is also mentioned in TS 22.278 [76] as a *satisfactory* SRVCC handover performance. Unfortunately, this threshold is unlikely to be met when using SRVCC in a VoLTE roaming scenario with S8HR, especially when the VPMN and HPMN are at a significant distance from each other.

Unlike S8HR, the LBO options support eSRVCC [32], therefore they can exploit all the advanced features of this version of SRVCC and provide a better QoS and a better user experience by minimizing the voice interruptions that caused by the handover delay. This is realized by two additional entities introduced in eSRVCC: the Access Transfer Control Function (ATCF) and the Access Transfer Gateway (ATGW). The role of ATCF together with ATGW is to anchor the signaling and the media in the VPMN. Given that LTE coverage is not as extensive as the earlier generation networks, a soft handover from an LTE network to a 2G/3G network is necessary; however, this will gradually become unnecessary with increasing LTE deployment.

4.2.1.3 Support for future IMS services

A VoLTE roaming architecture based on S8HR is *independent* of the visited network since the visited network (i.e., the VPMN) is not required to have an IMS and only the HPMN must have the ability to handle VoLTE roaming. For this reason, the home network can make new upgrades and thus offer new

IMS services *without* the need to wait for the visited network to make any upgrades or adjustments to support those new services that the home network wants to provide to its subscribers. Nevertheless, in the case of the emergence of new services that require different QoS handling, the home network may be dependent upon the visited network.

In contrast, LBO-VR, LBO-HR, and LBO-OMR make use of IMS UNI and NNI interworking as described in Section 3.2, thus making it possible for the visited network to be service aware - therefore the home network will have the ability to steer user plane traffic in a more efficient way for specific new services. However, an LBO architecture is considered to be more attached to the visited network, thus creating greater dependencies between the HPMN and the VPMN. These dependencies might have a large impact on the ability to support new services via the home network.

4.2.1.4 Complexity and service equality

In some ways, both S8HR and the LBO options share the same dependencies. For example, all the VoLTE roaming implementations require QoS to be supported by the home network, the visited network, and the IPX interconnecting networks. In the case of an S8HR implementation, apart from IMS emergency calls and eSRVCC enabled calls, there is no other aspect of the implementation that increases the complexity of this architecture. Thus, the HPMN's subscribers that are roaming in a visited network will be able to make use of the same services as they would in their home network. Additionally, since S8HR only uses the IMS network nodes in the home network, roamers can use their services independently from the type of their access network. In addition, unlike LBO, S8HR does not require any tests for IMS UNI and NNI interoperability in the visited network.

In the case of a roaming VoLTE architecture based on the LBO options, the roaming end users might be unable to receive all the IMS services and features that they would when in their home network. The reason for this is that the P-CSCF located in the visited network along with the NNIs may not support the same functionalities as the home network. Furthermore, the LBO options require additional configurations and tests to be made for a VoLTE roaming environment. Some of these are configurations of the UNI and NNI to ensure interoperability between the involved networks. In addition, in order to be able to provide VoLTE services to the roamers, the LBO-VR option needs to have additional network entities, such as the TRF. Another attribute that might impact the quality of experience that end users experience, whether they are roaming or not, is the fact that the ability of an end-to-end "HD Voice" call over an LBO-VR implementation will depend upon the visited network; therefore, roamers will be limited to the same QoS as the subscribers in the VPMN.

4.2.1.5 Regulatory compliance

In general, there are two significant concerns regarding regulations that need to be considered by the VoLTE roaming architectures: support of emergency calls and support for lawful intercept. Each of these is addressed in the following paragraphs.

4.2.1.5.1 Support for emergency calls

In order to support the regulations for supporting emergency calls, all the VoLTE roaming implementations must use the local emergency services provided by the visited network. The S8HR architecture has been proven to be inadequate for supporting emergency calls. In general, there have been two key issues that 3GPP described in 3GPP TR 23.749 [70]. These two challenges are the execution of a successful IMS emergency session for a detected and non-detected UE. Authenticated emergency calls depend on the existence of an IMS NNI *within* both the visited network and the home network. In the case of an S8HR implementation, an IMS NNI is not provided; hence, it is infeasible to authenticate a subscriber in the IMS domain using the specifications given in TS 23.167 [77, p. 167] and TS 23.228 [31]. Thus, a solution is needed in order for an IMS emergency registration to be successful so that the end user can be authenticated. The latter issue is due to the P-CSCF in the home network being incapable of detecting and managing a non-UE detectable emergency session from an end user located in a visited network.

The solution proposed by 3GPP in TR 23.749 to address the first challenge, is to bypass the authentication that is required on the IMS-level. In this situation, authentication at the EPS level will suffice; more specifically the IMS network will accept a SIP REGISTER message *without* attempting to authenticate the end user and *without* creating a UE registration in the IMS. Even if this can resolve this issue, it needs a security study - since there is a potential security threat from unauthorized or malicious users who could exploit this vulnerability. A solution that addresses the second issue of a non-UE detectable emergency session can be achieved with a help of a database. More specifically, the P-CSCF located in the IMS of the HPMN can send a query to a database that contains the local emergency call numbers in the visited network. This database containing all the emergency numbers of its roaming partners could be provided by the HPMN or it can be located in the VPMN and accessible to its roaming partners. Alternatively, 3GPP stated in TR 23.749 that GSMA could maintain a *global* database for this purpose. With this latter solution, the P-CSCF can compare a request URI with the emergency numbers contained in the database and detect an emergency call. Afterward, the P-CSCF has two options: it can either forward the INVITE to the S-CSCF or reject the request by sending an error message to the subscriber and suggest alternative actions. For example, it could suggest making an emergency call using a CS network (if such a network exists).

On the other hand, a VoLTE architecture that makes use one or more LBO options does not have any problem supporting emergency calls, due to the fact that all LBO options make use of IMS UNI and IMS roaming NNI for IMS emergency calls. At the time this Master's thesis was being written, no incidents have been found that would indicate any difficulties for any of the LBO options with regard to supporting emergency calls. In summary, S8HR is not fully capable of handling emergency calls and there are issues that have not been resolved yet (beyond the suggestions of 3GPP on the matter). In contrast, the LBO options can fully support emergency calls *without* any modifications or improvements required of the existing architecture models.

4.2.1.5.2 Support for Lawful Intercept

Lawful intercept is subject to the local regulation and the Law Enforcement Authority (LEA) and the requirements for lawful intercept vary from country to country. However, in general, there is some information that is needed by the LEA in the case of various emergencies related to national security that more or less remain the same. According to [78], this information can include: (1) identification of the involved users, (2) information about the location of the involved parties, (3) call logs that can be found in the CDRs, and (4) access to the media traffic.

As far as the S8HR architecture is concerned, it does not support lawful interception as has been mentioned in IR. 65 [62]. However, a recent study by 3GPP in TR 33.827 [79] addresses this problem by identifying fifteen key issues related to the lack of support for lawful interception and proposes four solutions that address these issues. Some of these issues include: the operations needed to obtain the UE's location, to find the procedure that will enable the targeting and reporting of the communication of a specific International Mobile Station Equipment Identity (IMEI)*, ways to target and find the GTP tunnels that carry the desired traffic that need to be sent to the LEA, and retention of relevant data in the VPMN.

In an attempt to resolve all these challenges, the first two solutions proposed by 3GPP introduce a procedure that is based on GTP tunnel extraction. The second solution complements the first solution by providing an enhanced architecture; more specifically, this method is a mixture of two mechanisms: a Lawful Intercept "tap" and a UNI based IMS/SIP service state machine. The "tap" is an operation that implemented in interfaces and monitors all traffic that is passing through the specific interface in a passive manner while intercepting the targeted traffic based on selection criteria. Afterward, the tap

* The International Mobile Equipment Identity (IMEI), is a unique number that is used to identify mobile phones.

sends a copy of the intercepted traffic to a LEA for further analysis. This lawful Intercept tap can either work as an individual network entity or be embedded in other network entities, such as the S-GW. The UNI based IMS/SIP service state machine create events that can be reported to the LEA in the form of lawful intercept messages. The third and fourth solutions mainly focus on ways to forward the targeted data from the HPLMN to the VPMN in order to create CDRs that can be retained in the visited network.

To summarize, in order to support lawful interception the S8HR architecture needs to be able to identify the GTP tunnel over which the desired traffic is being transmitted and extract the data and send the call logs and the payload to the LEA. 3GPP concluded that the second solution is the most appropriate one and would comply with the regulations. Although after the study in TR 33.827 S8HR can support lawful interception, it is important to mention that this conclusion is based on the assumption that the home network at the point where the interception is taking place does *not* enable confidentiality protection. This means that will affect the confidentiality of both the signaling and the user plane; therefore, the encryption normally provided by the IMS security association has to be *disabled* and there should be no end-to-end payload security.

As for the LBO options, no issues have yet been identified that would not allow lawful intercept procedures to be applied. Nevertheless, there are two main requirements in order for the LBO options to support lawful interception. The first requirement regards the support for network provided location information about the targeted UE serviced by a specific access network. A thorough study of Network Provided Location Information (NPLI) can be found in the TR 23.842 [80]. The second requirement, that matches the same requirement as in the S8HR architecture, is that payload security should be disabled in order for the visited network to intercept the traffic flows.

4.2.2 Criteria depending on the call scenario

This subsection applies a set of criteria that are dependent upon the call scenario when comparing the different VoLTE roaming models.

4.2.2.1 Deployment

According to the analysis of the available VoLTE roaming architectures in Chapter 3, it can be inferred that in order to be deployed the LBO models require more modifications and upgrades of the established network infrastructures than S8HR. One of the biggest requirements of an LBO deployment is the presence of an IMS network in the VPMN. This prerequisite alone was one the main reasons why the LBO options did not meet the expectations of some providers as they did not want to rely on the capabilities of the VPMN and they preferred a solution with a short time-to-market. Therefore, 3GPP along with GSMA proposed the S8HR architecture. Furthermore, an NNI infrastructure is necessary for an LBO deployment, whereas S8HR does not require one. In addition, to realize any of the LBO options, new network elements, such as the TRF, have to be introduced and implemented in the IMS network; whereas S8HR does not require any additional entities. In spite the fact that S8HR is more feasible to be deployed, there is still has a noticeable impact on the deployment an S8HR architecture due to its inability to fully comply with the regulations concerning emergency calls and lawful interception.

4.2.2.2 Routing Efficiency

This criterion concerns the media traffic efficiency of the different VoLTE roaming architectures and does not compare the signaling traffic efficiency (since the signaling routing is always the same regardless of the architecture). As mentioned earlier, the most common roaming scenario is for roamers calling back to their home network. In this case, both LBO options and S8HR are equally efficient when it comes to media routing since all of them forward the user plane traffic towards the home network. In contrast, when the roamer initiates a call from the visited network towards a destination other than the home network, then LBO-OMR is the most efficient option because the media breakout occurs directly to the terminating destination *without* the need to send the user plane back to the home network of the roamer or to the recipient's home network (when one roamer calls another roamer who is located in a different VPMN). In addition, in a roaming scenario that a roamer initiates a call session towards a home subscriber who is roaming in the same VPMN, then all the VoLTE roaming models (except LBO-OMR) have to send the media traffic back to the HPMN. In contrast, LBO-OMR will transfer the media directly to a recipient in the same VPMN as the roamer. The next best option is LBO-VR since the media traffic is not sent back to the home network unnecessarily. The least efficient architectures in a scenario where the HPMN is not the terminating destination, are LBO-HR and S8HR - due to the fact that these models require the media traffic to always being sent to the home network along with signaling, regardless of the destination of the media traffic.

Although LBO-OMR can be considered the best model in terms of routing efficiency, since it can either be equivalently efficient with other architectures or better, it is still not being commercially used. One of the main reasons for this lack of adoption is that the optimization of the route is highly dependant upon *all* the involved networks that take part in a call session. This means that all these networks must support an OMR solution and include this option in their roaming agreements with each other. In conclusion, an adequate solution that could cover all the possible scenarios in an efficient way would be to use the most appropriate LBO option depending on the call scenario for each call. However, the provider's selection of the correct architecture depends upon to their own needs but should consider their roamers' calling patterns. For example, if the vast majority of the calls have as a destination the roamer's home network, then an S8HR solution might be preferable.

4.2.2.3 Call Quality

All the VoLTE roaming architectures support the same QoS levels. Additionally, these models meet the Key Performance Indicators (KPIs) for voice call setup time as they have been tested in trials made among providers with the help of international carriers who provided the IPX infrastructure [81]. The main KPIs that were tested in these trials were the call success rate, the voice quality measured by the Mean Opinion Score (MOS)* metric, the media delay, and call setup time. All these measurements have proven that the call quality remains the same regardless of the VoLTE roaming architecture with negligible differences. Nevertheless, the end-to-end latency of the calls is associated with the routing efficiency as described in Section 4.2.2.2. Therefore, depending upon the VoLTE roaming model and the geographical distance between the end users, the media delay will vary. For example, if there is a call between subscribers who reside in different continents, then the media delay will be smaller when using LBO-OMR than using S8HR since the LBO-OMR routes the media traffic more efficiently in comparison with S8HR.

4.3 VoLTE roaming architecture suitability depends on the provider's needs

The discussed earlier, VoLTE roaming deployments have different functionalities and they support different regulations and needs. VoLTE roaming with LBO-HR requires an IMS infrastructure in the VPMN and an IMS interconnection with the corresponding HPMN. Some significant attributes of LBO-HR are that it supports voice charging for both originated and terminated calls, IMS emergency calls (which is a major issue for S8HR), SRVCC, and QoS over interconnection networks [62]. SRVCC enables a vertical handover between LTE and 2G/3G connections to support call continuity if LTE coverage is unavailable due to poor reception or any other reason. For further information regarding SRVCC, the reader is referred to TS 123.216 [82]. LBO-HR is a suitable solution in call scenarios where the terminating destination is in the HPMN since all the media will eventually reach the HPMN. In addition, LBO-HR is a convenient option for operators who want to comply with local regulations

* Mean Opinion score (MOS) is a metric over which the quality of the voice is measured as it is perceived by the end user. MOS has five values ranging from one to five. The value one indicates that the voice quality is unacceptable whereas the value five corresponds to a very satisfactory voice quality.

(regarding emergency calls and lawful interception). However, LBO-HR is incapable of supporting geo-local services (i.e., services that are available in the visited IMS network) nor any level of OMR for calls originated in the VPMN.

LBO-VR, on the other hand, complements LBO-HR by supporting geo-local services in the visited network and it offers OMR for originated calls. As mentioned in Section 3.2.3, in order for OMR to be used, all the involved networks that are responsible for establishing calls between end users must support OMR. Therefore, LBO-VR is suitable for network operators that wish to make use of LBO-HR and require the use of both geo-local services and OMR.

One of the main reasons why network operators had second thoughts about the RAVEL solution, was the fact that the LBO options require an IMS to be implemented in the visited network. Currently, not all PLMNs have an IMS, and even when there is an IMS, the HPMN has to trust that the VPMN will handle the calls in accordance with the HPMN's information. In addition, for an LBO approach, the network operators have to increase their CAPEX. Thus, S8HR was proposed because is similar to data roaming but handles voice services. Section 3.3 describes S8HR and stated that there is no need for an IMS in the visited network. However, IMS interconnection with the HPMNs might be needed when both subscribers are roaming and belong to different home networks. Hence, S8HR is better suited for network operators who do not want to rely on the visited network to provide more than access services to its subscribers. Moreover, S8HR is suitable for network operations who want a shorter time-to-market period and want to avoid any additional deployment issues. Nevertheless, those network providers who implement S8HR have to compromise due to the limitations of this option. The most important challenges with this architecture, as mentioned in the IR.65 [62], are the following:

1. The services offered only by the VPMN are impossible to be provided when using S8HR.
2. Since the VPMN is not obliged to use an IMS in an S8HR implementation, this means that the visited network is *not* service aware; hence, the VPMN simply transfers both signaling and media towards the HPLMN acting as a "bit-pipe".
3. Due to the lack of service awareness in the visited network, S8HR is unable to support lawful interception for roamers (although this is a requirement from many local regulators).
4. Another issue caused by service unawareness is that S8HR encounters difficulties in establishing emergency calls.
5. The subscriber's user experience might be degraded since there is no support for SRVCC, thus call continuity is difficult to achieve since handover between PS and CS networks is not fully supported.
6. The architecture of the S8HR clearly states that there is no possibility for OMR since both the control and the user plane are always being transferred back to the subscriber's HPMN regardless of the roaming scenario.
7. Furthermore, S8HR does not support QoS bearer charging and it needs the IPX providers to support it.

Considering all the above about the different roaming architectures, a network provider who wishes to use not just one option (for example, S8HR with LBO-HR and LBO-VR) has to support all of them. In reality, this is unlikely to happen because it requires complex configurations, increases time-to-market, and increases CAPEX. Therefore, each operator has to decide upon the most suitable roaming architecture according to their own needs. Due to the diversity of these needs and different capabilities of operators, there is a possibility that more than one type of VoLTE implementations will be established and must co-exist. In theory, a solution to overcome this diversity is for IPX providers to provide interoperability by taking over the cumbersome task of supporting the available VoLTE roaming options. Moreover, this can potentially give them a new revenue stream. Alternatively, the network operators will have to wait for 3GPP and GSMA to provide a single solution for VoLTE roaming. However, a decision from these organizations does not oblige an operator to use a particular standard. Currently, the default suggestion for a VoLTE roaming architecture is RAVEL, while S8HR is proposed as an alternative. Currently, S8HR seems to be preferred more than RAVEL [83]. In fact, all commercial launches of LTE roaming are using S8HR.

4.4 VoLTE roaming from a carrier's perspective

Currently, neither GSMA nor 3GPP has reached a solid conclusion regarding a default VoLTE roaming architecture. In addition, there is great diversity between the network operators' needs, their network infrastructure capabilities, their willingness to upgrade their network, and their business plans. Unless organizations such as GSMA propose that only one standard can be used, multiple roaming implementations will need to coexist. Moreover, it is possible for a network operator to choose an alternative architecture other than that suggested by these organizations as the default roaming architecture. Consequently, it can be reasonable to speculate that not all the providers will choose the same VoLTE roaming model. The difference in the VoLTE roaming approaches by various providers will have an impact on compatibility among them. This raises a concern about how will the network operators will interconnect with each other when they are using different VoLTE roaming architectures.

For this reason, there is a challenge for the IPX providers to support all these different implementations, while providing their services in a seamless way. Therefore, in order for the IPX providers to remain competitive in the international market, it is important if they can support more than one VoLTE roaming architecture. In this way, the intermediate network providers can assist more network operators, rather than only those who have the same roaming model as the IPX provider. Unfortunately, this requires more expenditures by the IPX operators, but with a market research and some CAPEX/OPEX estimations it might be proven to be a profitable solution.

In addition to the above, a bigger challenge for the carriers is to offer a hybrid solution via which network operators with different VoLTE roaming architectures are able to interconnect with each other. In the following paragraphs, an approach for an IPX carrier will be discussed. The different models that will be preferred by the network operators can lead to a scenario where a visited mobile operator (O-VPMN) chooses the S8HR architecture, while a home operator (O-HPMN) chooses an LBO implementation. In order for the IPX carrier to assist both operators and provide an interconnection between them regardless of the roaming model that they are using, the IPX provider has to implement network entities, such as the P-GW and the P-CSCF/TRF, in their hybrid method. In this way, "Hub BreakOut" (HBO) [84]can take advantage of the P-GW and the P-CSCF by breaking out calls from the O-VPMN without the voice call being transferred to the S8 interface and routed towards the O-HPMN. In a model such as HBO, the TRF acts as an anchor for the control plane and media plane for calls that are routed locally within the IPX's HBO. Hence, calls can be transferred directly from the HBO location with the assistance of the O-HPMN that provides policy control. Alternatively, calls can be sent to the O-HPMN with the help of the IMS NNI for routing.

With this hybrid solution, another service that an IPX carrier could provide can emerge. This service could be offered to VPMNs who have not yet deployed an IMS network but want to support VoLTE roaming for their roaming partners who have already implemented an LBO solution. In this approach, the VPMN is not required to implement an IMS network to provide VoLTE roaming nor does it need to implement any of the LBO options. Instead, it is only necessary for the visited network to acquire only those network entities that take part in VoLTE roaming calls, such as the P-CSCF. Therefore, an HBO solution could assist the visited network by offering all of the necessary network elements so that the VPMN can support VoLTE roaming – but without the VPMN needing to actually add these nodes to their network. In addition, in call scenarios where the roamer calls an end user who is located in the same network, with the help of HBO it is feasible to enhance the VoLTE roaming experience for regional calls. This enhancement could be realized by the HBO since it can route the media back to a local user of the VPMN *without* the need to transfer the user plane to the HPMN. This model is shown in Figure 4-1.

Another term that could be used instead of HBO is Carrier BreakOut (CBO). CBO follows the same principles of HBO. CBO can be established in the Points of Presence (PoPs) of the IPX carrier by upgrading the PoP with the required network elements: P-CSCF, TRF, and IBCF. Then, the CBO can offer LBO services to the VPMNs and can provide a hybrid method that interconnects roaming partners with different VoLTE roaming architectures.

In conclusion, it is worth mentioning that as promising a solution as CBO/MBO may seem, from the GSMA's perspective there is not yet any interest. Unless GSMA proposes a model such as CBO/HBO, it will not be an official model supported by 3GPP or GSMA. However, this does not necessarily mean that an IPX carrier would be unable to offer such a service, especially if it were transparent to the HPLMN and the VPLMN. This concept is proposed on a high level and there are not yet clarifications of how to implement such a model in a way that both LBO and S8HR needs are satisfied. Moreover, it is important that there is no need for a network operator to change anything in their own network. For basic call scenarios, this might be easy enough, but it will probably be increasingly difficult to implement when there is a need to implement functions regarding SRVCC and emergency calls.

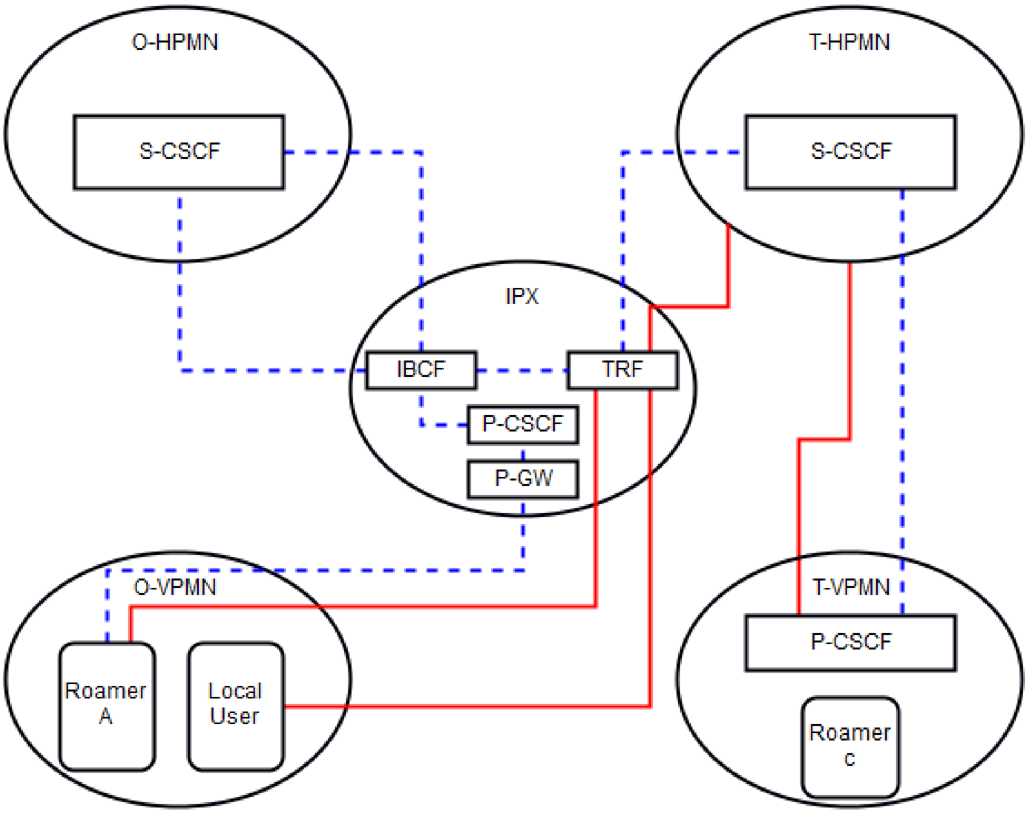

Figure 4-1: High-level Architecture of an HBO model. Adapted from [84]

5 Conclusions and Future work

This chapter contains a final discussion regarding VoLTE roaming and what has been achieved during this thesis. In addition, Section 5.2 suggests some further study of related topics that could provide further insights into the subject of this thesis. Finally, the chapter concludes with some reflections on ethical and sustainability issues relevant to this thesis project.

5.1 Conclusions

The world of telecommunications is moving towards to achieving always being connected anywhere with the best possible service. One aspect of this achievement is VoLTE roaming. End users may find VoLTE roaming more appealing than OTT services that deliver voice at affordable prices. Since roaming surcharges in the European Union will be abolished as of 15 June 2017, roaming traffic may increase if subscribers prefer using an LTE network rather than an OTT service. Certainly, these are speculations about the foreseeable future of VoLTE roaming. Moreover, VoLTE implementations have just begun to take place in home networks and it remains to be seen how subscribers will start using this service, especially when roaming. From a personal experience, as a VoLTE tester, my expectations regarding the advantages of VoLTE (see Section 2.8) have been met so far. Thus, there is still time to further investigate VoLTE roaming and to decide which architectural solution is most appropriate for each carrier (apart from those network operations who want to attract the early adopters in the market). This is why GSMA and 3GPP are still releasing technical reports on this area to provide further information that would help the network operators with their decision.

This thesis presented several different approaches for implementing VoLTE roaming and evaluated them according to the criteria defined in Section 4.1. In addition, in Section 4.3 suggestions were made about the most suitable mechanisms to apply to the current network architectures of network operators (according to their needs).

In summary, this project offered some insights into the architectures, entities, interfaces, and protocols that together form a telecommunications environment. After acquiring sufficient basic knowledge about all these aspects of telecommunication networks, a thorough study was made on VoLTE roaming. This made it possible to compare the different VoLTE roaming architectures and to identify the strengths and weaknesses of each model. Consequently, this resulted in an estimate of what VoLTE roaming might look like in the near future. With the help of surveys made on this topic, it can be inferred that many providers will or already have started using the S8HR architecture to support VoLTE roaming. One of the main driving influences regarding this decision is the fact that the operators are unwilling to spend too much OPEX or CAPEX to establish a VoLTE roaming infrastructure. In addition, it is desirable to avoid relying on the visited networks that offer LTE roaming service to other than their own subscribers. Furthermore, as the most common roaming scenario is that the majority of originated calls from a visited network have the subscriber's home network as a destination or the other way around. Therefore, if a provider has subscribers that follow this call pattern, it will be advantageous for the network operator to choose the S8HR architecture, rather than LBO. Although it may seem more appropriate to use the S8HR model for the near future, it is interesting to see what network operators will do when the majority of local network operators have implemented an IMS network. At such a point in time, LBO will be easier to establish and it will improve the QoS in certain roaming scenarios. Unfortunately, it is not yet clear whether it is desirable to deploy an S8HR network for the time being, while afterward changing the architecture into an LBO model.

This thesis provided the necessary information about VoLTE roaming and the latest updates on this topic from organizations such as GSMA and 3GPP. For readers who are interested in this subject, this project should serve as a good starting point since it explains VoLTE roaming in a simple manner. Nevertheless, if the readers want to dig deeply into this area, they can read all the related material cited as references in this thesis. In addition, because the pace of changes and new information regarding VoLTE roaming is rather rapid, readers should follow the latest updates and technical reports of GSMA and 3GPP, while keeping track of surveys and other related articles about VoLTE roaming.

5.2 Future work

A complementary topic that has not been addressed in this thesis is OTT services as they have a quite substantial amount of market share. Thus, a comparative study of VoLTE roaming versus OTT voice services would be interesting both on a technical and a business oriented level. Currently, many subscribers outside their home network use OTT services because they are more affordable than 3G/4G roaming. However, since roaming will become cheaper within Europe, it is important for the network operators to speculate what fraction of potential end users will switch from using OTT services to VoLTE roaming.

In addition to another comparative study, another interesting research effort would be to establish a testing environment in which the different VoLTE roaming architectures could be tested in real scenarios with live traffic. This would enable researchers to see if the theory applies in a real environment and to test the different models and compare their results to see which ones are more appropriate for certain providers and to understand which specific scenarios will lead to large difference in the impact of different selections by taking into account the results of these measurements.

Finally, thorough research regarding an HBO/CBO solution would be appreciated by IPX carriers since this has the potential to show whether such a hybrid method is feasible, the complexity of this method and whether adoption of this method would require substantial OPEX and CAPEX expenditures. As mentioned earlier, this concept is not standardized by any organization. Therefore, this area remains an unexplored and technical and business research is needed to see if such a solution would be profitable for interconnect carriers and also for the network operators.

5.3 Required reflections

The first driving force that started this master's thesis was the regulation of the European Union which decided that mobile network operators have to lower their charges for roaming to zero by June 2017. Due to this ethical regulation, the thesis presented the available VoLTE roaming architectures that potentially can help the mobile network operators to adapt to it while lowering their opex on certain occasions. Apart from that, with the emergence of VoLTE roaming, the roamers will be able to have a better quality of service on a much lower cost. This can mean that the roamers might start using mobile networks and its services rather than any alternative solutions like OTT services with the use of a WiFi access.

References

[1] “First Commercial LTE Network Goes Live,” *Network Computing*, 14-Dec-2009. [Online]. Available: http://www.networkcomputing.com/networking/first-commercial-lte-network-goes-live/752107374. [Accessed: 03-Jun-2016].
[2] “Mapping LTE Deployments: LteMaps.” [Online]. Available: http://ltemaps.org/home/. [Accessed: 03-Jun-2016].
[3] OpenSignal.com, “The State of LTE September 2015.” [Online]. Available: http://opensignal.com/reports/2015/09/state-of-lte-q3-2015/. [Accessed: 24-Jun-2016].
[4] “End of roaming charges: Council confirms agreement with EP - Consilium.” [Online]. Available: http://www.consilium.europa.eu/en/press/press-releases/2015/07/08-roaming-charges/. [Accessed: 10-Mar-2016].
[5] iBasis, “VoLTE Roaming – Let’s take away the confusion | iBasis.” .
[6] tutorialspoint.com, “LTE Tutorial,” *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/lte/index.htm. [Accessed: 07-Jul-2016].
[7] E. U. T. R. A. Network, “3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access,” 2011.
[8] “LTE.” [Online]. Available: http://www.3gpp.org/technologies/keywords-acronyms/98-lte. [Accessed: 07-Jul-2016].
[9] tutorialspoint.com, “LTE Overview,” *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/lte/lte_overview.htm. [Accessed: 07-Jul-2016].
[10] “3GPP Long Term Evolution (LTE): Overview of LTE 3GPP releases.” .
[11] “What is MME? - FAQs - Ask the Expert! - LTE University.” [Online]. Available: http://lteuniversity.com/ask_the_expert/b/ltefaqs/archive/2008/11/03/what-is-mme.aspx. [Accessed: 08-Jul-2016].
[12] E. U. T. R. A. Network, “S1 Application Protocol (S1AP)(Release 10),” 2011.
[13] tutorialspoint.com, “LTE Network Architecture,” *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/lte/lte_network_architecture.htm. [Accessed: 09-Jul-2016].
[14] “LTE International Roaming Whitepaper.” [Online]. Available: http://carrier.huawei.com/en/technical-topics/core-network/LTE-roaming-whitepaper. [Accessed: 02-Jun-2016].
[15] G. S. M. Association and others, “Inter-Service Provider IP Backbone Guidelines 5.0,” *Official Document IR*, vol. 34, 2011.
[16] G. S. M. Association and others, “Official document IR. 34,“Guidelines for IPX Provider networks (Previously Inter-Service Provider IP Backbone Guidelines), Version 12, January 2016.” .
[17] G. S. M. Association and others, *IPX White Paper*. Version, 2007.
[18] Dialogic, “Interconnecting Networks with Dialogic’s Global Multimedia Exchange Platform,” no. 1.
[19] A. Lior and A. DeKok, “Remote Authentication Dial In User Service (RADIUS) Protocol Extensions.” [Online]. Available: https://tools.ietf.org/html/rfc6929. [Accessed: 17-Jun-2016].
[20] J. Arkko, G. Zorn, V. Fajardo, and J. Loughney, “Diameter Base Protocol.” [Online]. Available: https://tools.ietf.org/html/rfc6733. [Accessed: 17-Jun-2016].
[21] “Introduction to Diameter,” 24-Jan-2006. [Online]. Available: https://www.ibm.com/developerworks/library/wi-diameter/. [Accessed: 17-Jun-2016].
[22] 3GPP, “Release 5.” [Online]. Available: http://www.3gpp.org/specifications/releases/75-release-5. [Accessed: 27-Sep-2016].

[23] R. Noldus, U. Olsson, C. Mulligan, I. Fikouras, A. Ryde, and M. Stille, "Chapter 8 - Introduction to the IMS Network," in *IMS Application Developer's Handbook*, Oxford: Academic Press, 2011, pp. 223–328.
[24] A. B. Ericsson, "Introduction to IMS," *White Paper, maaliskuu*, 2007.
[25] "3GPP specification: 32.240." [Online]. Available: http://www.3gpp.org/DynaReport/32240.htm. [Accessed: 20-May-2016].
[26] C. Bormann, Z. Liu, R. Price, and G. Camarillo, "Applying Signaling Compression (SigComp) to the Session Initiation Protocol (SIP)," 2007.
[27] O. Hersent, "The 3GPP IP Multimedia Subsystem (IMS) Architecture," in *IP Telephony*, John Wiley & Sons, Ltd, 2010, pp. 269–311.
[28] M. A. Qadeer, A. H. Khan, J. A. Ansari, and S. Waheed, "IMS Network Architecture," 2009, pp. 329–333.
[29] "IBCF - Interconnection Border Control Function -," 19-Jan-2016. .
[30] "TrGW - Transition Gateway -," 19-Jan-2016. .
[31] 3GPP, "Stage 2 (Release 9), 3GPP TS 23.228, v9.3.0."
[32] "3GPP TR 23.850 - Study on roaming architecture for voice over IMS with local breakout." [Online]. Available: http://www.tech-invite.com/3m23/tinv-3gpp-23-850.html. [Accessed: 07-Apr-2016].
[33] H. Khlifi and J. C. Grégoire, "IMS Application Servers: Roles, Requirements, and Implementation Technologies," *IEEE Internet Computing*, vol. 12, no. 3, pp. 40–51, May 2008.
[34] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson,R. Sparks, M. Handley,E. Schooler, "RFC-3261." [Online]. Available: https://www.ietf.org/rfc/rfc3261.txt. [Accessed: 09-Jul-2016].
[35] D. 26 and 2002 Kieran McCorry | Windows IT Pro, "A Short History of SIP." [Online]. Available: http://windowsitpro.com/systems-management/short-history-sip. [Accessed: 09-Jul-2016].
[36] H. Chaouchi and M. Laurent-Maknavicius, *Wireless and Mobile Networks Security*. John Wiley & Sons, 2013.
[37] tutorialspoint.com, "SIP Network Elements," *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/session_initiation_protocol/session_initiation_protocol_network_elements.htm. [Accessed: 16-Jul-2016].
[38] "The Session Initiation Protocol - The Internet Protocol Journal - Volume 6, Number 1," *Cisco*. [Online]. Available: http://www.cisco.com/c/en/us/about/press/internet-protocol-journal/back-issues/table-contents-23/sip.html. [Accessed: 16-Jul-2016].
[39] T. Berners-Lee, R. Fielding, L. Masinter, "Uniform Resource Identifiers (URI): Generic Syntax," *rfc2396*, Aug-1998. [Online]. Available: https://www.ietf.org/rfc/rfc2396.txt. [Accessed: 17-Jul-2016].
[40] tutorialspoint.com, "SIP Messaging," *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/session_initiation_protocol/session_initiation_protocol_messaging.htm. [Accessed: 21-Aug-2016].
[41] R. Sparks, "The Session Initiation Protocol (SIP) Refer Method," Apr-2003. [Online]. Available: https://www.ietf.org/rfc/rfc3515.txt. [Accessed: 21-Aug-2016].
[42] J. R. H. Schulzrinne, "Reliability of Provisional Responses in the Session Initiation Protocol (SIP)," Jun-2002. [Online]. Available: https://www.ietf.org/rfc/rfc3262.txt. [Accessed: 21-Aug-2016].
[43] tutorialspoint.com, "SIP Response Codes," *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/session_initiation_protocol/session_initiation_protocol_response_codes.htm. [Accessed: 25-Aug-2016].
[44] "SIP response codes - voip-info.org." [Online]. Available: http://www.voip-info.org/wiki/view/SIP+response+codes. [Accessed: 25-Aug-2016].
[45] S. Donovan, A. Johnston, K. Summers, C. Cunningham, and R. Sparks, "Session Initiation Protocol (SIP) Basic Call Flow Examples." [Online]. Available: https://tools.ietf.org/html/rfc3665. [Accessed: 18-Aug-2016].

[46] “Understanding SIP Registration,” *SIP Adventures*, 29-May-2014. .
[47] tutorialspoint.com, “SIP Basic Call Flow,” *www.tutorialspoint.com*. [Online]. Available: http://www.tutorialspoint.com/session_initiation_protocol/session_initiation_protocol_basic_call_flow.htm. [Accessed: 19-Aug-2016].
[48] H. S. <hgs@research.att.com>, “RTP: A Transport Protocol for Real-Time Applications,” Jan-1996. [Online]. Available: https://tools.ietf.org/html/rfc1889. [Accessed: 26-Aug-2016].
[49] V. Jacobson, R. Frederick, S. Casner, and H. Schulzrinne, “RTP: A Transport Protocol for Real-Time Applications.” [Online]. Available: https://tools.ietf.org/html/rfc3550. [Accessed: 26-Aug-2016].
[50] T. Koistinen, “Protocol overview: RTP and RTCP,” *Nokia Telecommunication*, 2000.
[51] V. Bajaj, “Roaming between GSM/UMTS and IMS networks,” in *3rd International Conference on Communication Systems Software and Middleware and Workshops, 2008. COMSWARE 2008*, 2008, pp. 488–494.
[52] R. Noldus, *CAMEL: Intelligent Networks for the GSM, GPRS and UMTS network*. John Wiley & Sons, 2006.
[53] “IR.37 Guidelines and Testing for Optimal Routing - Service definition (Stage 1 Testing) v3.0.0,” *Newsroom*, 07-Apr-2005. .
[54] T. ETSI, “Mobile Application Part (MAP) specification 100 974 V7. 6.0 (Sep. 2000),” *Part*, vol. 1, pp. 1–285.
[55] C. B. Barring, “Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Anonymous Communication Rejection (ACR) and Communication,” 2009.
[56] “IR.20 Technical Implications of the Use of Optimal Routing v3.4.0,” *Newsroom*, 27-Sep-2012. .
[57] G. S. M. Association and others, “IMS Profile for Voice and SMS,” *Official Document IR*, vol. 92, 2013.
[58] “Singtel, Samsung and Ericsson unveil world’s first full-featured Voice over LTE service.” [Online]. Available: http://www1.singtel.com/about-us/news-releases/singtel-samsung-and-ericsson-unveil-worlds-first-full-featured-voice-over-lte.html. [Accessed: 08-Sep-2016].
[59] “The Secret Value of VoLTE,” *Next Generation Communications*. [Online]. Available: http://blog.tmcnet.com/next-generation-communications/2014/04/the-secret-value-of-volte.html. [Accessed: 08-Sep-2016].
[60] Jyrki T. J. Penttinen, *The LTE-Advanced Deployment Handbook The Planning Guidelines for the Fourth Generation Networks*. : Wiley, 2015.
[61] “3GPP specification: 23.749.” [Online]. Available: http://www.3gpp.org/DynaReport/23749.htm. [Accessed: 24-Jun-2016].
[62] G. S. M. Association and others, “IMS Roaming & Interworking Guidelines Version 20.0,” *Doc. IR*, vol. 65.
[63] I. D. Specifications, “Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS) emergency sessions (3GPP TS 23.167 version 9.6.0 Release 9),” 2006.
[64] G. O. Association and others, “GPRS Roaming Guidelines,” *PRD IR*, vol. 33, Jun. 2014.
[65] I. T. S. Sector, “ITU-T Recommendation Z. 120,” *Message Sequence Charts (MSC96)*, 1996.
[66] R. G. Cole and J. H. Rosenbluth, “Voice over IP performance monitoring,” *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 2, pp. 9–24, 2001.
[67] “Press Releases : DOCOMO to Offer VoLTE International Roaming in South Korea-Japan’s first mobile operator to provide VoLTE international roaming service-,” *NTT DOCOMO Home*. [Online]. Available: https://www.nttdocomo.co.jp/english/info/media_center/pr/2015/0930_00.html. [Accessed: 24-Jun-2016].

[68] "VoLTE Roaming – Are We There Yet? - Corporate Blog - Dialogic - Dialogic Exchange Network." [Online]. Available: http://www.dialogic.com/den/d/b/corporate/archive/2015/09/22/volte-roaming-are-we-there-yet.aspx. [Accessed: 24-Jun-2016].
[69] M. A. Abe Itsuma Tanaka, Shin ichi Isobe, "Standardization of New VoLTE Roaming Architecture," vol. 17, 2015.
[70] 3GPP, "TR 23.749: Study on S8 Home Routing Architecture for VoLTE (Release 14), Version 2.0.0." [Online]. Available: https://portal.3gpp.org/ngppapp/CreateTdoc.aspx?mode=view&contributionUid=SP-160325. [Accessed: 22-Sep-2016].
[71] "IR.88 LTE and EPC Roaming Guidelines v14.0," *Newsroom*, 16-Feb-2016. .
[72] "FUTURE OF VOICE SPECIAL REPORT: LTE roaming: The rise of a roaming empire." [Online]. Available: http://www.capacitymedia.com/Article/3456817/FUTURE-OF-VOICE-SPECIAL-REPORT-LTE-roaming-The-rise-of-a-roaming-empire.html. [Accessed: 05-Oct-2016].
[73] "SS7 Protocol Suite," *Protocols*. .
[74] T. Koshimizu, I. Tanaka, and K. Nishida, "Improvement on the VoLTE (Voice over LTE) Domain Handover with Operator's Vision," in *World Telecommunications Congress (WTC), 2012*, 2012, pp. 1–5.
[75] S. Gautam and D. P. Sharma, "Solution to Reduce Voice Interruption Time during Handover of VoLTE Call in Enhanced Single Radio Voice Call Continuity," in *2015 Fifth International Conference on Advanced Computing Communication Technologies*, 2015, pp. 627–633.
[76] 3GPP, "Service requirements for the Evolved Packet System (EPS) TS 22.278." .
[77] 3GPP, "IP Multimedia Subsystem (IMS) emergency sessions TS 23.167." [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=799. [Accessed: 20-Oct-2016].
[78] R. Evripidis, "Lawful Interception and Countermeasures: In the era of Internet Telephony," 2008.
[79] 3GPP, "Study on providing for Lawful Interception (LI) in the S8 Home Routeing (S8HR) architecture for Voice over LTE (VoLTE) TR 33.827." [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2986. [Accessed: 20-Oct-2016].
[80] 3GPP, "Study on network provided location information to the IP Multimedia Subsystem (IMS) TR 23.842." [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=922. [Accessed: 21-Oct-2016].
[81] A. Aittokallio, "China Mobile, KPN, iBasis claim first VoLTE roaming implementation," *Telecoms.com*, 14-Oct-2014. [Online]. Available: http://telecoms.com/292992/china-mobile-kpn-ibasis-claim-first-volte-roaming-implementation/. [Accessed: 03-Nov-2016].
[82] E. U. T. R. A. Network, "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommu nications System (UMTS); LTE; Single Radio Voice Call Continuity (SRVCC); Stage 2 (3GPP TS 23.216 version 10.1.0 Release 10)," 2011.
[83] Itsuma Tanaka and Ian Crawford, "Interconnected-Mobile-All-IP-Communications-and-the-benefits-for-Operators."
[84] R. Xu, "Bridging the gap between VoLTE Roaming Models | iBasis," [Online]. Available: http://ibasis.com/blog/volte-roaming-lets-take-away-the-confusion/ .

TRITA-ICT-EX-2017:1