UNIVERSITÉ DU QUÉBEC EN ABITIBI-TÉMISCAMINGUE

ANALYSIS AND DESIGN OF A NEW INTEGRATED MOBILE SIP PROXY TO ENHANCE THE SCALABILITY IN MOBILE NETWORK OPERATORS

THÉSE

PRÉSENTÉ

COMME EXIGENCE PARTIELLE

DE LA MAÎTRISE EN INGÉNIERIE

Par

LINA ABOU HAIBEH

Jury d'évaluation

Examinateur Externe: Prof. Gilles Y. Delisle, Université Laval

Examinateur Interne: Prof. Nahi Kandil, Université du Québec en Abitibi-Témiscamingue

Directeur de Recherche: Prof. Nadir Hakem, Université du Québec en Abitibi-Témiscamingue

UNIVERSITÉ DU QUÉBEC EN ABITIBI-TÉMISCAMINGUE

ANALYSIS AND DESIGN OF A NEW INTEGRATED MOBILE SIP PROXY TO ENHANCE THE SCALABILITY IN MOBILE NETWORK OPERATORS

A THESIS PRESENTED TO THE UNIVERSITE DU QUEBEC EN ABITIBI-TEMISCAMINGUE IN FULFILLMENT OF THE THESIS REQUIREMENT FOR THE DEGREE OF MASTER IN SCIENCE AND INFORMATION TECHNOLOGY

BY

LINA ABOU HAIBEH

COMMITTE MEMBERS

External Examiner: Prof. Gilles Y. Delisle, Université Laval

Internal Examiner: Prof. Nahi Kandil, Université du Québec en Abitibi-Témiscamingue

Research Supervisor: Prof. Nadir Hakem, Université du Québec en Abitibi-Témiscamingue

ABSTRACT

The emergence of the two new technologies, namely Software Defined Network (SDN) and Network Function Virtualization (NFV) have radically changed the development of computer network functions and the evolution of mobile network operators (MNOs) infrastructures. These two technologies bring to MNOs the promises of reducing costs, enhancing network flexibility and scalability to handle the growth in the number of mobile users and the need to extend its coverage to rural areas.

The aim of this thesis is to exploit the advantages of the NFV concept to support the implementation of fully integrated solution with an external Session Initial Protocol (SIP) proxy application to enhance the scalability in MNOs. The proposed solution offers a hosted SIP proxy application installed on a virtual machine (VM) environment. The SIP proxy provides full Private Branch Exchange (PBX) and Switch (SW) functionality with Interactive Voice Response (IVR) capabilities. It maximizes the capacity in the existing servers and value-added services (VAS) data centers within the MNOs. The proposed solution enhances the usage of the existing bandwidth by using the unlicensed radio frequency (RF) spectrum bandwidth instead of the licensed RF spectrum to support a larger number of smartphones and data plans.

In the initial experimental testbed, TeleFinity IP PBX, which is an external SIP proxy, is deployed on a virtual platform and integrated with the mobile network. The integration is realized by establishing a point to point protocol (PPP) SIP trunk connection between TeleFinity IP PBX and the Gateway Mobile Switch Center (GMSC). Several testing scenarios were carried out over a local area network (LAN) and a wide area network (WAN) using different voice codecs: G.711 u-law, G.723, and G.729 to validate the voice call quality offered by the proposed solution. The Network analyzer software solutions: 1) Startrinity SIP tester, 2) Commview and 3) Resource Monitor are used to measure several Quality of Service (QoS) metrics. These include voice jitter, delay, packet loss, and MOS. This procedure ensures that the proposed solution can handle voice communications with acceptable quality compared to LTE standards.

Keywords: Network Function Virtualization (NFV); SIP Proxy; Quality of Service (QoS); Scalability; Mobile Network Operators; Virtual machine

Dedicated to my little son, Ryan, who adds more colors and joy to my life

ACKNOWLEDGEMENTS

Working towards my master's degree has been a deeply enriching experience; at times it's been elating, but it has always been very rewarding. Looking back, many people have helped me to shape this journey. I would like to extend them my thanks.

I would like to express our deepest gratitude to my advisor, Prof. Nadir Hakem, for his invaluable guidance, encouragement, faith and advice during this research work. I would also like to thank my husband, Ousama Abu Safia, for his continuous support and kindness help. I have greatly benefited from observing how he dissects research problems.

I would also like to greatly appreciate and thank my family members and friends for their supporting and encouragement throughout my master's study.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

LIST OF FIGURES ix

LIST OF TABLES xii

1 CHAPTER I INTRODUCTION 1

1.1 Background 1

1.2 Literature Review 4

1.3 Research Objectives 9

1.4 Research Methodology 10

1.5 Thesis Structure 11

2 CHAPTER 2 NETWORK FUNCTION VIRTUALIZATION 13

2.1 NFV Overview 13

2.2 NFV Key Benefits for MNOs 14

2.3 NFV and Software Defined Networking (SDN) relation 15

2.4 NFV architectures and key modules 18

2.5 NFV Design Considerations 22

2.6 Vendor Specific NFV Implementations 23

3 CHAPTER 3 VIRTUAL MOBILE SIP PROXY INTEGRATION 25

3.1 Virtual Network Functions Concept in MNO 25

3.2 What is SIP Proxy 26

3.3 TeleFinity IP PBX Overview 26

3.3.1 TeleFinity IP PBX Benefits 28

3.4 TeleFinity IP PBX Modules 28

3.4.1 Call Center Module 29

3.4.2 Video Conference Module 30

3.4.3 Interactive Voice Response (IVR) Module 31
3.4.4 Call Recording Module..32
3.4.5 Call Accounting Module 32
3.4.6 Fax Module 33
3.5 Mobile SIP Proxy Integration Approach...34
3.6 Proposed Architecture 37
3.6.1 Solution Design 37
3.6.2 VM Specifications 39
3.7 Signaling Stream Scenarios...40
3.7.1 User Registration Scenario 41
3.7.2 Mobile Call Originating Scenario 42
3.7.3 Mobile Call Terminating Scenario 45
4 CHAPTER 4 EXPERIMENTAL TESTBED IMPLEMENTATION 48
4.1 Introduction 48
4.2 Experimental Testbed Setup..48
4.2.1 Hardware Parts 49
4.2.2 Software Solutions 49
4.3 Experimental Testbed Configuration..51
4.3.1 TeleFinity IP PBX Configuration...51
4.3.2 AcFinity IVR Configuration..52
4.3.3 Startrinity SIP Tester Configuration 54
4.4 Network Topology...54
5 CHAPTER 5 RESULTS AND DISCUSSION 57
5.1 Voice Codecs Properties..57
5.1.1 G.711 u-law 57
5.1.2 G.729 58

5.1.3 G.723.1 58

5.2 Measured Parameters 59

5.2.1 Voice Quality of Service (QoS) Parameters 60

5.2.2 Signaling Performance Parameters 62

5.2.3 Network Utilization Parameters 63

5.2.4 Hardware Utilization Parameters 64

5.3 Testing Scenarios and Results Analysis 65

5.3.1 Scenario 1: WAN Topology- RTP Passing Through SIP Proxy 65

5.3.2 Scenario 2: WAN Topology- RTP Not Passing Through SIP Proxy 73

5.3.3 Scenario 3: LAN Topology- RTP Passing Through SIP Proxy 80

5.3.4 Scenario 4: LAN Topology- RTP Not Passing Through SIP Proxy 86

6 CHAPTER 6 CONCLUSION AND FUTURE WORK 92

6.1 Conclusion 92

6.2 Future Work 94

7 REFERENCES 95

8 APPENDIX A 101

LIST OF FIGURES

Figure 1.1 LTE Network Entities [2] 2

Figure 2.1 Traditional network layers versus logical layers in a SDN Network [30] 16

Figure 2.2 Indicative SDN architecture and mapping to NFV [30] 17

Figure 2.3 High Level NFV Domains [32] 18

Figure 2.4 End to end service Graph representation [32] 19

Figure 2.5 Architectural description of NFV [32] 20

Figure 3.1 TeleFinity IP PBX Architecture © 27

Figure 3.2 Call Center Module Architecture © 29

Figure 3.3 Supported Multi Media Services © 30

Figure 3.4 IVR Module Call Flow Designer [36] 31

Figure 3.5 Call Recording Module [34] 32

Figure 3.6 Call Accounting Module [34] 33

Figure 3.7 Fax Module © 34

Figure 3.8 Proposed Solution Diagram © 36

Figure 3.9 Proposed Solution Architectural Design © 37

Figure 3.10 User Registration Scenario © 42

Figure 3.11 Mobile Call Originating Scenario © 44

Figure 3.12 Mobile Call Terminating Scenario © 46

Figure 4.1 Windows Resource Monitor © 51

Figure 4.2 IVR System Settings © 53

Figure 4.3 Startrinity Configuration © 54

Figure 4.4 LAN Topology © 55

Figure 4.5 WAN Topology © 56

Figure 5.1 SRD Time [36] 62

Figure 5.2 SIP Proxy Services CPU Load © 64

Figure 5.3 SIP Proxy Services Memory Utilization © 65

Figure 5.4 WAN Topology- RTP Passing through SIP Proxy Scenario © 66

Figure 5.5 Scenario 1- Voice QoS Results 69

Figure 5.6 Scenario1. SIP Performance Results 70

Figure 5.7 Scenario 1- Network Utilization Results 71

Figure 5.8 Scenario1- Hardware Utilization Results 72

Figure 5.9 WAN Topology- RTP without Passing through SIP Proxy Scenario © 73

Figure 5.10. Scenario 2- Voice QoS Results 76

Figure 5.11. Scenario 2- Signaling Performance Results 77

Figure 5.12. Scenario 2- Network Utilization Results 78

Figure 5.13. Scenario 2- Hardware Utilization Results 79

Figure 5.14. LAN Topology- RTP Passing through SIP Proxy Scenario © 80

Figure 5.15. Scenario3- Voice QoS Results 82

Figure 5.16. Scenario3- Signaling Performance Results 84

Figure 5.17. Scenario 3- Network Utilization Results 85

Figure 5.18. Scenario3- Hardware Utilization Results 86

Figure 5.19. LAN Topology- RTP Not Passing through SIP Proxy Scenario © 87

Figure 5.20. Scenario4- Voice QoS Results 88

Figure 5.21. Scenario 4- Signaling Performance Results 89

Figure 5.22. Scenario 4- Network Utilization Results 90

Figure 5.23. Scenario4- Hardware Utilization Results 91

LIST OF TABLES

Table 2.1 Key differences between SDN and NFV 17

Table 2.2 Industry NFV projects 24

Table 3.1 Minimum VMs Specifications 39

Table 4.1 VMs Specifications used in the test bed 49

Table 4.2 SIP Extension Configuration Parameters 51

Table 4.3 SIP Trunk Configuration Parameters 52

Table 4.4 IVR Configuration Parameters 53

Table 4.5 Comparison between LAN and WAN 55

Table 5.1 VoIP Codecs specification 59

Table 5.2 ITU-T Recommendation G.114 for Delay Specification [44] 60

Table 5.3 MOS Rating [46] 61

CHAPTER I

INTRODUCTION

1.1 Background

Recently, we have witnessed an explosion in the number of mobile phone users which accompanies to the appearance and emergence of new types of applications and services. As a result, an exponential increase in the data traffic has been addressed. This enormous data traffic leads mobile network operators (MNOs) to upgrade their systems, invest in their existing infrastructure, or look for new technologies and solutions that help to satisfy their customers' demands. It is expected that these demands will be fulfilled in the next generation of mobile networks; the so-called fifth generation (5G) network, which will achieve an extremely high data rate, ultra-low latency, high user mobility and ultra-reliable communication [1]. Moreover, it is expected that MNOs will be able to manage easily the huge number of mobile terminals and extend their coverage to rural areas without any additional cost.

MNOs have been evolved through four generations. Starting from being a circuit-based analog telephony system in the first generation (1G). Then, it became a partially packet-based system in 2G and 3G, and finally, it became all-IP packet-based system in 4G/LTE. The 3GPP (3rd Generation Partnership Project) introduced a new mobile core network architecture called Evolved Packet Core (EPC) in LTE. This architecture is able to interoperate with the legacy of 2G and 3G systems [2].The introduction of EPC (Evolved Packet Core) allows mobile users to access multimedia resources in external packet data networks such as the Internet. On the other hand, the current LTE mobile network architecture has several limitations that impose challenges to MNOs toward upgrading their infrastructures. These limitations are explained in details in the following points:

1) The LTE network entities, shown in Figure 1.1, include: 1) Mobility Management Entity (MME), 2) Serving Gateway (SGW), 3) Packet Data Network Gateway (PGW), 4) Home Subscriber Server (HSS) and 5) Policy control and Charging Rule Functions (PCRF). These components are typically based on a customized standard hardware which is usually configured, deployed and provisioned in a static and cost-ineffective manner. This leads to excessive capital expenditure (CAPEX) and operational expenditure (OPEX) losses, besides an increased complexity in the management of the hardware platforms.

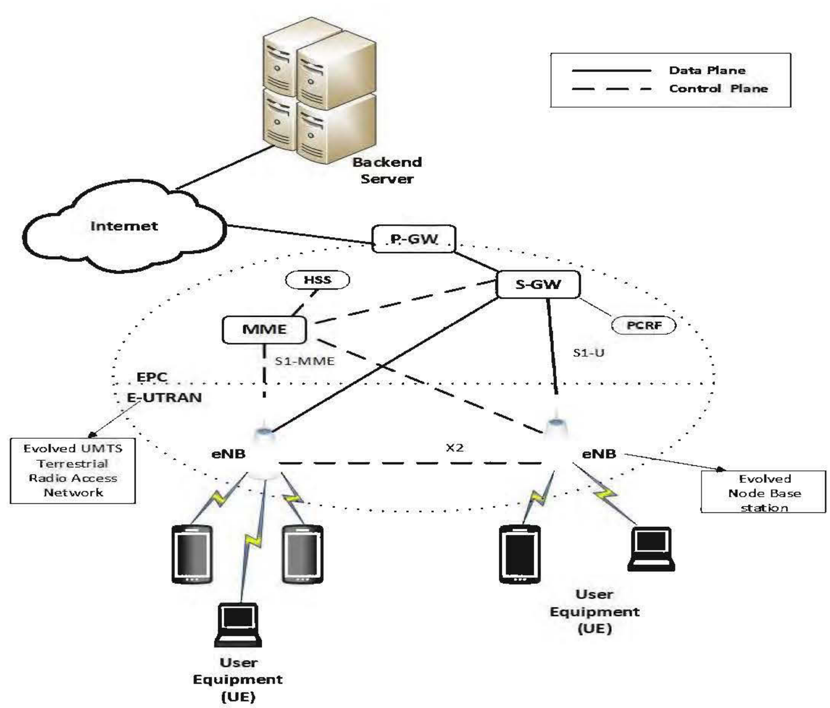


Figure 1.1 LTE Network Entities [2]

2) The control plane, which is responsible for the signaling between the EPC and the user for the users in the mobile network equipment (UE), and the data plane, which is assigned for the actual voice and data communication architecture, are tightly coupled at SGW and PGW. This coupled design results the inflexibility of network management, and limits the scalability of the network. Since the control plane and data plane have different performance requirements, the control plane requires low latency to process signaling messages, whereas the data plane requires high throughput to process the users' data traffic, it is necessary to decouple these planes to get them scaled independently and efficiently during the provisioning process.
3) The data plane of the current mobile architecture is centralized. Indeed, all uplink traffic from user equipment's (UEs) has always to traverse along a path that begins from the radio access, passing through the mobile backhaul, and then enters the IP networks via a small number of centralized PGWs. Even if some UEs are just communicating with local Application Services. Such a hierarchical architecture results an inefficiency in the data packet forwarding and the mobility management, in addition to a high latency. Thus, it is not suitable to fit the aforementioned 5G requirements.

Recently, the advent of some cutting-edge technologies such as cloud computing, mobile edge computing, network virtualization, Software Defined Networking (SDN) [3], and Network Function Virtualization (NFV) [4] have changed the way network functions and devices are implemented. Also, they changed the construction of network architectures. More specifically, the network's equipment or device is now changing from closed, vendor specific to open and generic with SDN and NFV technologies. This enables the separation of control and data planes, and allows networks to be programmed by using open interfaces.

With NFV, network functions which previously realized in costly hardware platforms are now implemented as software appliances placed on low-cost commodity hardware or running in the cloud computing environment. NFV technology together with cloud computing and network virtualization, bring to MNOs the promises of reducing the Capital Expense (CAPEX) [5] and the Operation Expense (OPEX) [6].It is also shortening the end-to-end network latency, enhancing network flexibility and scalability.

By taking the advantage of NFV implementation in MNOs, this thesis proposes an original software-based solution that allows having a scalable mobile network operator. The new architecture includes the integration with an external SIP proxy, and it is deployed on a virtual machine (VM). The VM can be defined as an emulation for computer system and provides the punctuality of physical computer. Moreover, the SIP proxy controls the call management function in MNOs for newly added mobile users. More details will be discussed later in the thesis.

1.2 Literature Review

The challenges and limitations of mobile network architecture require that many operators are being compelled to resort to iterative schemes of network planning. These include adding new infrastructure and backhaul to accommodate the growing demand for capacity and coverage. However, limited spectrum availability and the physical characteristics of radio waves do not allow a complete overhaul of the system beyond their capacity limits. MNOs, which are already CAPEX and OPEX constrained, are seeking for affordable, easily deployable, scalable and agile alternatives for network expansion. Many researches have been conducted to provide smart mobile network expansion approaches that can resolve the aforementioned MNOs' limitations and challenges. Some of these approaches are discussed below:

Basta *et al.* in [7] proposed an architecture of virtualized mobile packet core (MPC) gateways and NFV-based transport network elements. The control plane was not described in that work. The data plane entities are virtualized running on a data center platform. They are managed by a data center orchestrator. The NFV-based transport network is used to interconnect these virtualized gateways to the radio access and external IP networks. As a main contribution, the authors proposed several solutions to find the optimal data center location to host these virtual gateways so that the network load is minimized under a time-varying traffic pattern and a given data plane delay budget.

In [8], the same authors improved their work in [7] by addressing the function placement problem. They grouped the four deployment models in [7] into two main categories: a virtualized gateway (NFV) and a decomposed gateway (SDN). The first category refers to fully virtualizing SGW and PGW into a data center, and an off-the-shelf network element (NE) is used to direct the data traffic from the transport network to the data center. The second category refers to decomposing gateway functions, meaning that only the control plane function is shifted to the data center and integrated with SDN controllers while the data plane is processed by enhanced SDN network elements. To find the most optimal deployment solution, the authors in [8] formed a model by taking the control-plane load and data-plane latency into account, and then, tried to minimize these parameters. By doing so, the operators will have a tool to make their own deployment decision: virtualizing all gateways or decomposing all gateways or a combination of the two.

The authors in [10][11] similarly adopted SDN and NFV into EPC S/PGWs. In these articles, the control function of an S/P integrated gateway (S/PGW) is decoupled from the user plane. While virtualized S/PGW control is realized as VMs in a cloud computing system. The S/PGW user plane can be realized either by VM or dedicated hardware. The dedicated hardware is usually located close to the access network and it is responsible for fast path processing.

SoftEPC in [12][13] presented a virtual network of EPC functions over a physical transport network topology. SoftEPC followed the concept of NFV by decoupling the network services and functions from the special purpose hardware. SoftEPC is composed of a collection of General Purpose Nodes (GPN) that are core-class commodity servers running hypervisor [12][13]. Hypervisor is a virtualization software that creates multiple virtual environments based on the physical hardware. The SoftEPC method can lead towards achieving significant saving in network bandwidth and processing load, especially in the core, by load-aware dynamic instantiation of P/S-GW functions on the GPN nodes. The SoftEPC shows the flexibility and elasticity compared with the conventional EPC. However, the authors did not discuss in detail how the GPNs are managed. Taleb *et al.* in [13][13] envisioned an end-to-end carrier cloud architecture . All the related EPC entities are virtualized as VMs running in a distributed manner at different data center (DC) locations. The VMs and their locations are launched on a carrier cloud service platform based on requirements of the number of the served subscribers at each location [13][13]. In order to achieve an optimal end-to-end connectivity for UEs, the Follow-Me-Cloud (FMC) concept is introduced. The main idea of FMC is to allow contents and services to follow the user during his/her movement. Thus, enabling the service continuity and reducing the end-to-end network latency. However, the shortcoming of this work is the lack of how it works in reality. In addition, the detailed design of each functional unit and the interfaces used to communicate between them are not provided.

Similarly, the concept of having distributed data centers (DCs) to accommodate EPC functions is also introduced in Klein [13][15]. Klein is disruptive design for an elastic cellular core by combining network functions virtualization with smart resource management. . Compared to [13][13] ,Klein also enables the placement of the data plane entities in a distributed manner. In addition, an orchestrator is introduced to allocate the network resources and to assign UE’s data and network traffic to correct locations. By using a data driven analysis, the authors proved that Klein can almost optimally achieve the benefits of “clean-slate” approaches such as SoftCell [13][16] and SoftMoW [13][17] while working within the operational constraints of existing 3GPP standards. As an attempt to cope with the increase of Machine-to- Machine (M2M) or Machine Type Communications (MTC), the authors in [13][18] proposed a multi-vEPC architecture, which is able to provide optimized mobile communication service according to various requirements of M2M services. M2M services enable networked devices to exchange information and perform actions without the manual assistance of humans. The M2M services are classified based on their requirements such as policy-based service, mobility required service or IP reachability required service. However, the authors did not describe in details the design of the EPC selector, how it works and how they classify the M2M services into different groups and in a static or a dynamic manner.

Another NFV-based EPC is presented in [13][19]. In this paper, the authors introduced the concept of EPC as a service (EPCaaS) where each EPC entity is virtualized as an individual VM, communicating to each other using 3GPP standard interfaces. As a practical realization of EPCaaS, Jain *et al.* [13][20] developed an open source software, which implements most of the conventional EPC functions and run them as VMs in a cloud system. Although these are the simplest ways to virtualiz EPC, Hawilo *et al.* [13][21] argued that such design can significantly impact the performance, for example, result in a longer communication delay between EPC VNFs. In order to solve that problem, the authors in [13][21] have grouped several VNFs together on the basis of their interaction and workload and internalize communication between these VNFs, thus reducing the network latency.

While all presented works assume the use of VMs to implement EPC VNFs without considering the performance aspect(S), Kiess *et al.* [13][22]provided a comparison of different implementation models of PGW (also applicable to other VNFs) such as device model, cloud-aware model, and software-as-a-service model. Through a cost-based evaluation, they find that the two last models have cost advantages in terms of OPEX saving.

Another approaches adopted the NFV concept in external networks such as IP Multimedia Subsystem (IMS) and internet to provide improve the services provided for mobile users. In [23],applying NFV concept for IMS system was introduced. The network functions of the IMS system was deployed in VMs, which have scalable hardware resources and can be dynamically relocated in cases of a VM's overload or failure [24].Thus, the operator desired service continuity and service availability can be obtained. Similarly, Lu *et al.* [25] proposed a cloud-based IMS architecture. The architecture contains a load balancing algorithm and a mechanism for dynamic resource allocation [24].Ito *et al.* [24][26] proposed a new EPC/IMS system based NFV concept where each service can be processed by a particular virtual network to reduce the signaling load.

The aforementioned papers presented some of up-to-date solutions that adopted NFV into the core network in LTE and external networks such as IMS. However, there are still a number of issues need to be addressed in order to make these solutions feasible solutions and can be deployed in MNOs. Some open challenges raised in this domain such as the backward compatibility issue. The goal of the aforementioned issue is to find a flexible, effective and compatible design strategy with the existing mobile networks. Another challenge is related to the deployment model issue, which is concerns about finding a trade-off among various deployment models based on careful evaluations.

In this thesis, the advantage of the NFV concept is utilized to develop a novel solution having a simpler integration architecture, lower cost and better customizability features compared to standard solutions. The proposed solution relays on the integration between external SIP proxy and the Gateway Mobile Switch Center (GMSC) without changing the current mobile network architecture. The new solution can be integrated with existing mobile network systems (2G, 3G, 4G and 5G) systems. The SIP proxy is virtualized on VM and is responsible for the call management functionality of the new added mobile users. Moreover, the performance and quality of SIP calls carried over a proposed solution are analyzed in this thesis. The testing campaigns are carried out in a real environment for different voice codecs, G.711, G.727 and G.723.1.

1.3 Research Objectives

The primary research objective of this thesis aims to integrate the Session Initiation Protocol (SIP) proxy with the mobile operator network to resolve the aforementioned scalability related limitation.

The proposed solution allows achieving the following objectives:

1- Maximizing the use of existing bandwidth in order to to support more smartphones and data plans. Each virtualized SIP proxy is able to handle 2,000 concurrent calls and 12,000 users depending on virtual machine specifications;

2- Eliminate the need to install new proprietary hardware appliances. This would reserve more space and power resources in MNOs when there is a demand to increase the bandwidth;

3- Extend the coverage of MNOs in rural areas with minimum operation cost using the unlicensed RF spectrum;

4- Ensuring high levels of quality of service for all subscribers wherever they are located;

5- Managing large traffic volumes securely across multiple data centers;

6- Utilizing the Network Function Virtualization (NFV) concept to enhance the functionality of the proposed solution;

7- Proposing an original solution that can integrate with any telephony system generations (2G, 3G, 4G and 5G).

1.4 Research Methodology

The methodology steps related to this research project are introduced in the following points.

1- Implementing the proposed solution experimental testbed. It is divided into the following two parts:

a) Hardware Parts which include powerful servers to install virtual machines needed to implement a SIP proxy called TeleFinity IP PBX, IVR solution named ActFinity IVR and Voicemail and SIP tester to generate SIP calls.

b) Software Solutions which include: 1) a TeleFinity IP PBX that serves as external SIP proxy to handle clients SIP calls, 2) Startrinity SIP call generator, ActFinity IVR to receive the generated calls from SIP tester and play music file,

3) Network analyzer software solutions such as real-time reports in Startrinity SIP tester, 4) Commview and Resource Monitor which is used to measure several Quality of Service (QoS) metrics such as voice jitter, end-to-end delay, packet loss, and MOS.

2- Integrating the mobile network and the TeleFinity IP PBX by configuring the SIP trunk between them. Next, installing the SIP client applications on the mobile devices. Then, creating SIP extensions on the SIP Proxy application. Next, selecting the type of voice codec from the main system settings of TeleFinity IP PBX. The Voice Codecs, G.711, G.729 and G.723.1, will be selected for the tests.

3- Testing the calls' quality provided by the proposed solution. The calls' quality will be evaluated using four QoS metrics: mean opinion score (MOS), jitter, delay, and packet loss to ensure that the proposed solution can handle voice communications with comparable quality to LTE standards.

1.5 Thesis Structure

This thesis is composed of six chapters including the first introduction chapter. Chapter II presents a general overview on NFV and its infrastructure. As well, the roots of NFV, key benefits and relation with Software Defined Networks (SDN) is the focus area in Chapter 2. Also, key benefits to telecommunication service operators through NFV are described briefly .Chapter 3 provides a detailed description for the proposed solution architecture, TeleFinity IP PBX modules and its main features. Signaling messages that are exchanged between MNO network components to handle any update in the registered subscribers information and status are discussed in Chapter 3. Chapter 4 describes the hardware specifications and software solution used to implement the experimental testbed and evaluate the performance SIP proxy under virtualization environment for different voice codecs, G.711, G.727 and G.723.1. Moreover, the SIP proxy configuration and the evaluation metrics are introduced. The solution network topology implemented over these two networks is described. Testing scenarios that carried on LAN and WAN network topologies are described in chapter 5. Parameters description which are measured during testing are also introduced. The testing results for the three voice codecs G.711, G729 and G.723.1 are presented followed by analyzing the performance of the SIP proxy under a virtualization environment. Finally, the thesis's conclusion in addition to the future work are presented in Chapter 6.

CHAPTER 2

NETWORK FUNCTION VIRTUALIZATION

2.1 NFV Overview

NFV concept was introduced in 2012, when a number of leading Telecommunication Service Providers around the world introduced a white paper [27] calling for industrial and research activity. In November 2012, seven of these operators (AT&T, BT, Deutsche Telekom, Orange, Telecom Italia, Telefonica and Verizon) selected the European Telecommunications Standards Institute (ETSI) [28] to be the home of the Industry Specification Group for NFV (ETSI ISG NFV). Now, a vast community of experts are working deeply to develop the required standards for NFV as well as sharing their experiences of its development and early implementation.

The ETSI has proposed a number of use cases for NFV [29] such as Customer Premises Equipment (CPE), Virtual Evolved Packet Core (V-EPC), Open-BTS and Open- MSC.

As the NFV concept becomes mature, it is important to note that it is not only sufficient to deploy specific network functions over virtualized infrastructures. Network users are generally not concerned with the complexity of the underlying network. All users require seamless network service to allow them access to the multi-media applications they need, anytime and anywhere. Therefore, NFV will only be an acceptable solution for telecommunication providers if it meets the following key considerations: 1) Acceptable Network Architecture and Performance, 2) Security and Resilience, 3) Support for Heterogeneity, 4) Reliability and Availability, 5) Legacy Support, and 6) Network Scalability and Automation.

2.2 NFV Key Benefits for MNOs

NFV solutions can bring to telecommunication service operators many benefits. One of the most important benefits is the flexibility to evolve and deploy new network services within MNO platform in a more cost-effective manner. The aspects of the flexibility feature are presented in the following key points [30]:

- Hardware and software decoupling: The NFV eliminates the need for the combination of the hardware and software network modules to develop specific network service. Therefore, a separate planning and maintenance activities needed to be done.
- Flexibility in the network functions operations and design: The software and hardware decoupling leads to the redesign of hardware resources and elements and their usage for multiple concurrent network operations. Due to this, operators can deploy network services faster for their clients running on the same hardware entities.
- Effective network scaling: The dynamic scaling capabilities that are provided with NFV enable the usage of NFV instances with different granularity. For instance, NFV solves the heavy traffic issue that is faced from the operators by deploying specific NFV instance based on the traffic application scenario used.

It is worth remarking that the decoupling between software and hardware does not mean that the resource virtualization of all network elements is obligatory. Operators can still develop software and run it on the existing physical servers. The difference is that when running the software on virtual machines, this will result better performance and CAPEX/OPEX profits. Finally, hybrid scenarios where functions running on the virtualized resources can mutually operate and coexist with functions running on standard physical resources are suggested, till a full transition to virtualization takes place.

2.3 NFV and Software Defined Networking (SDN) relation

NFV is tightly coupled with SDN (Software Defined Networks). SDN is a networking approach which leads to the independence between controlling network functions and data forwarding functionality. For example, right now a packet that arrives in a standard network element such as a router or switch must follow a set of forwarding or routing rules related to error correction, NAT (network address translation), QoS (Quality of Service) as well as standard routing protocols rules [30]. SDN and NFV have started being developed independently, however it can be said that the former acts as complementary to the latter.

With the use of SDN, the above-mentioned operations of data forwarding, routing and network control functions are decoupled. To put it in a simpler way, data plane and control plane are decoupled, leading to centralized management of control plane and de-centralized data plane management [31].

SDN focuses on Layer 2 and 3 network elements and operations. SDN requires re-policing or re-configurations in the network infrastructure. As shown in figure 2.1, SDN decouples the network control and forwarding functions. This allows network control to become directly programmable via an open interface such as OpenFlow. As well, the underlying infrastructure becomes simple packet forwarding devices (the data plane) that can be programmed **Error! Reference source not found.**[30].

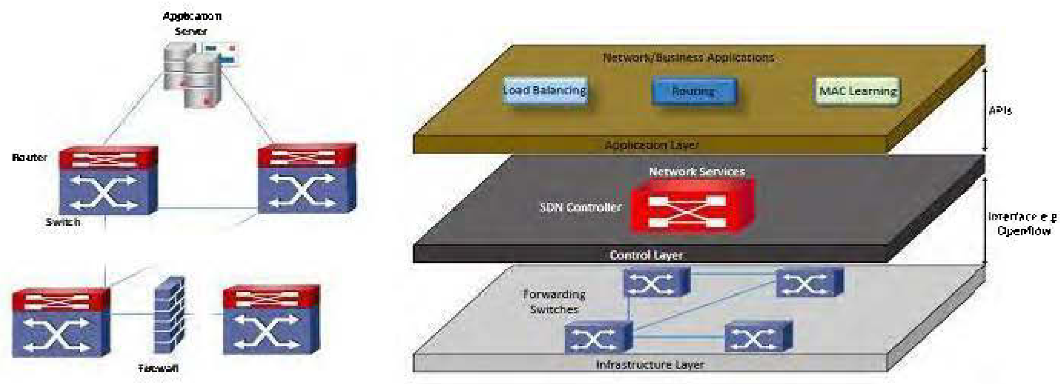


Figure 2.1 Traditional network layers versus logical layers in a SDN Network [30]

The key point in the relation of SDN and NFV is to understand that many IT services already run on cloud services. In the NFV case, the telecommunication providers focus on the real-time performance requirements which are the more stringent **Error! Reference source not found.**[30]. Figure 2.2 maps the cloud service models to the NFV architecture and the layered resources associated with each model. Cloud service models include software as a service (SaaS) model, platform as a service (PaaS) model, and infrastructure as a service (IaaS) model .It can be observed that IaaS corresponds to both the physical and virtual resources in NFV. While, the services and virtual network functions (VNFs) in NFV are similar to the SaaS service model in cloud computing [30].

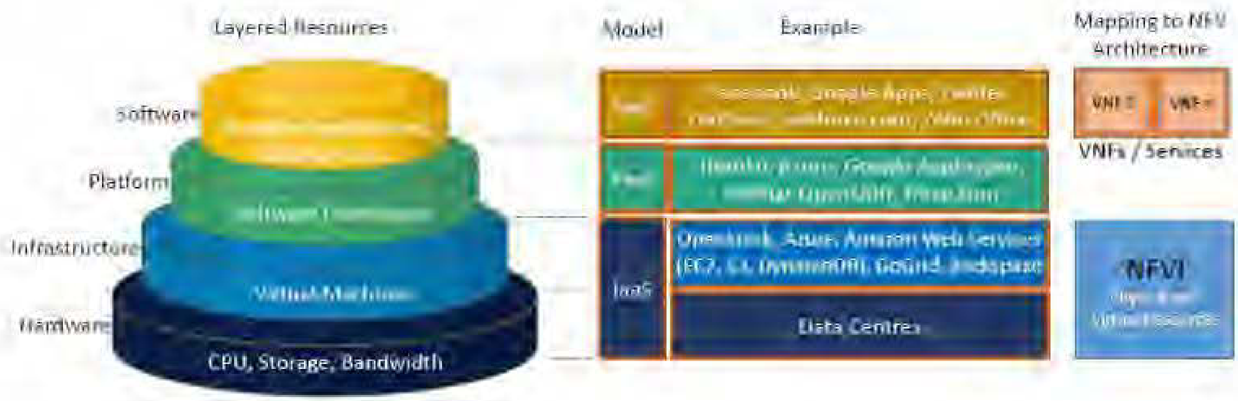


Figure 2.2 Indicative SDN architecture and mapping to NFV [30]

It can be said that there are a lot of similarities between SDN and NFV are represented by using similar open source software and standard network infrastructure with non-proprietary protocols. Their combination can lead to even better results, since SDN can chain several network functions in a NFV deployment and provide further automation of functions **Error! Reference source not found.**[31].

The main difference between the two is that NFV aims at the decoupling of software with hardware while SDN aims at the separation of packets and interfaces from the network control plane [30]. Some additional key differences between SDN and NFV are described by the following table as well.

Table 2.1 Key differences between SDN and NFV

Issue	NFV(Telecom Networks)	SDN
Approach	Service/Function/Abstraction	Networking Abstraction
Formalization	ETSI	ONF
Advantage	Flexibility and cost reduction	Unified programmable control
Protocol	Multiple control protocols	Openflow
Application runs	Commodity servers and switches	Commodity servers control plane

Leaders	Telecom service providers	Networking software and hardware

2.4 NFV architectures and key modules

Figure 2.3 presents a high-level design approach where NFV is based on the following domains [24][28]:

1. A software implementation of a network function (Virtualized Network Function) that is able to run over a NFV infrastructure.
2. The NFV Infrastructure (NFVI), consisting of the hardware resources where virtual networks can reside on and run.
3. A NFV Management and Orchestration module (MANO) which is assigned to handle and orchestrate the lifecycle of virtual network functions.

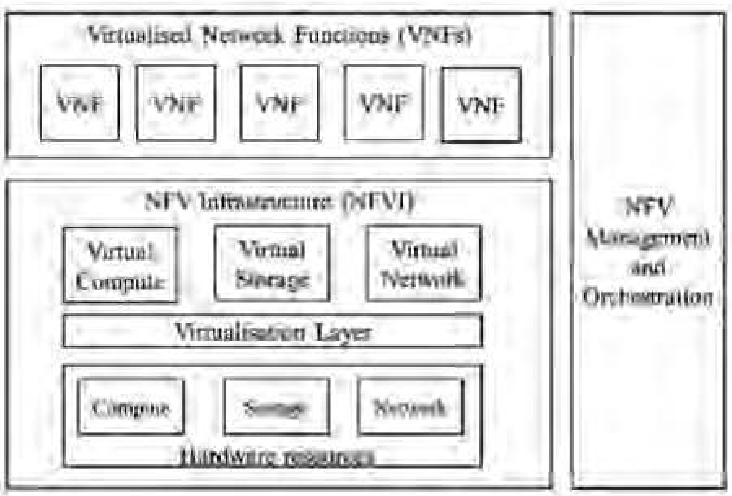


Figure 2.3 High Level NFV Domains [24][32]

Figure 2.4 shows a description of an end-to-end network service (e.g. mobile voice / data, Internet access) which consists of VNFs and end points. The dotted lines represents the logical interfaces that can be created between the NFVI and the NFVs. It can be either wired or wireless connection [32].

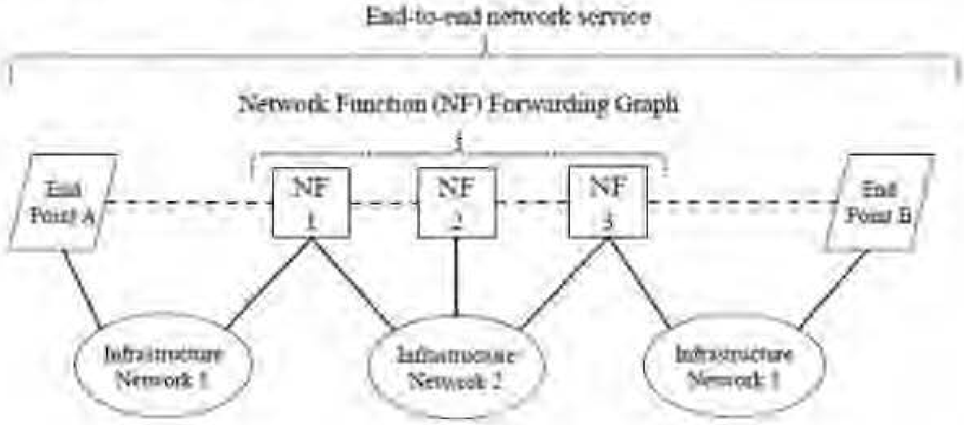


Figure 2.4 End to end service Graph representation [32]

NFV is based on the idea that the physical infrastructure deployment of a VNF is not visible from the end to end service perspective while redundant infrastructures can reside in different locations. Based on this a VNF can run on totally different geographically dispersed physical appliances [31].

Referring to Figure 2.5, the basic architectural functional blocks of NFV are the following [31]:

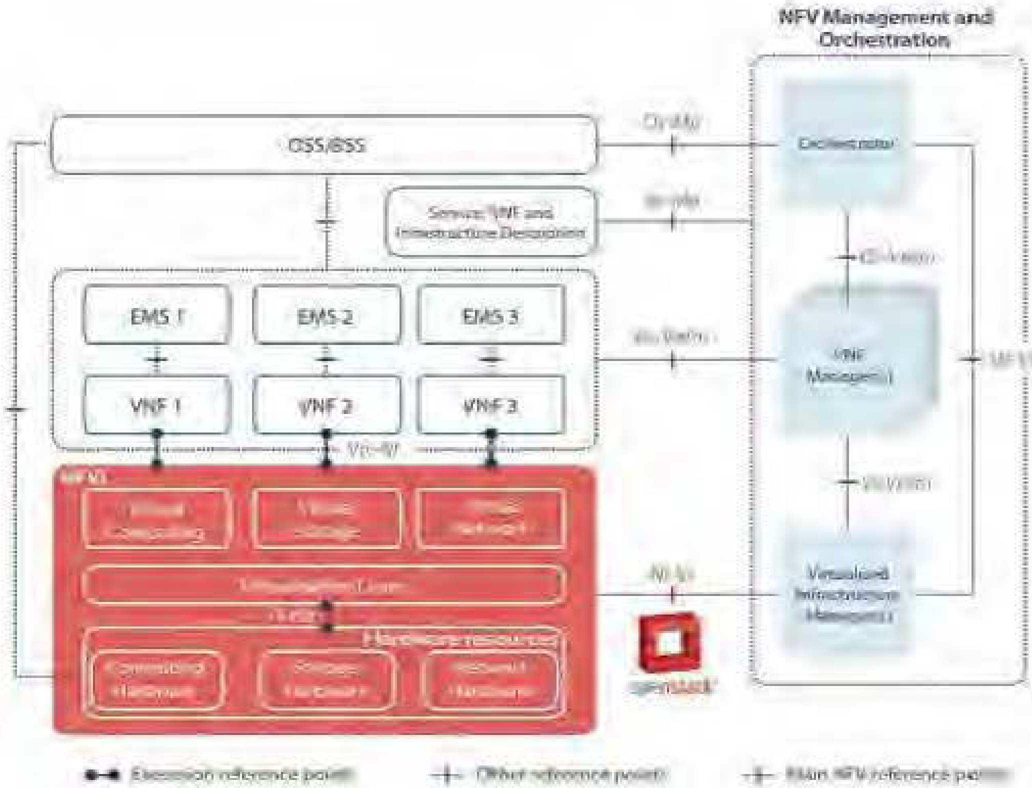


Figure 2.5 Architectural description of NFV [32]

- Virtualized Network Function (VNF): The Virtualized Network Function refers to a standard network function which can be fully or partially virtualized. Typical examples might be the core network elements of LTE - EPC (Evolved Packet Core) such as the MME (Mobility Management Entity), the SGW (Serving Gateway), or even the eNodeB through the C-RAN virtualization.

- Element Management (EM): Refers to the management operations strategy of the VNFs.

- NFV Infrastructure (NFVI): VNFI refers to the common hardware resources (servers, storage) that host the NVFs and can be physically located in various data centers across a city or a country supporting a pool of resources for the NFVs. Routers, switches and wireless links interconnecting the main servers can be regarded as part of the NFV Infrastructure as well.

- Virtualization layer: It has the role of abstracting and decoupling the virtualized network functions from the hardware resources. Whereas, each VNF shall use dedicated hardware resources to ensure the highest performance characteristics, while the software running the NFV is running on multiple hardware resources simultaneously [24][32]. Through the latter virtual access to the underlying compute resources is ensured while standard actions like starting, stopping or migrating VMs can be deployed.

- Hypervisor software: It is able to manage several guests' operating systems. As well, it enables consolidation of physical servers onto a virtual stack on a single server. CPU, RAM, and storage are flexibly allocated to each VM via software deployments [13].

- Virtualized infrastructure managers: VNFI managers provide resource management for the hypervisors, allocate resources to NVFs and provide fault management capability for the NFVI **Error! Reference source not found.**[31].

- NFV Manager: Provides the NFV lifecycle management and either multiple VNF managers operating per VNF, or one handling all VNFs can be deployed.

- NFV Orchestrator: Provides the orchestration and management of NFVI and combined with the VNF manager and VNF Infrastructure manager. This layer provides the connectivity and interaction with the other virtualized networks or standard network infrastructures. Open stack protocols and Software Defined Networking (SDN) functions can be deployed on the orchestrator layer. Also, it can considered a connection point to the Operations Support System and Business Support System (OSS / BSS) of a service provider or operator [31].

2.5 NFV Design Considerations

Some consideration should be taken into account while designing the NFV solutions. The primary goal of these consideration is to keep the same quality constraints for the end network users with the existing architectures when using NFV. Some of these design consideration are explained in the following points [13]:

1. Performance: The performance of network services must be the same as with existing network services running on dedicated hardware. This shouldn't lead to bottlenecks at all layers and should keep low latency characteristics. As an example, consider a scenario where NFVs providing the same service reside in different VMs, then the interconnection of the latter must ensure high bandwidth and low latency.
2. Security: Current security policies applied to the network services must be able to operate at the same way in NFVs. The most important point as far as security is concerned is that the NFVI must be protected from the services delivered to the end users through firewalls included in the NFV solution architecture.
3. Availability and reliability, Disaster recovery compliance: It must be ensured that outages are within the same timeframes described by today's

SLAs (Service Level Agreements), while in case of a failover there must be a redundancy solution.

4. Heterogeneous Support: Currently operators have the option of sharing network elements and selecting among different vendors since all platforms can communicate through standardized interfaces. The same rationale must be able to be deployed with NFV through keeping open interfaces and ensuring interoperability of multiple vendors.
5. Legacy systems support: The transition to NFV is not yet mature, since right now operators are still evaluating the solutions and very few commercial deployments exist. Due to this, it is expected that during the period hybrid NFV solutions being able to support current network architectures with legacy hardware and software systems shall prevail till full virtualization becomes a reality.

2.6 Vendor Specific NFV Implementations

A variety of NFV implementations have already been deployed by multiple vendors mainly in the telecommunications field. Some examples are listed below [32]:

- CISCO Open network strategy: This platform provides a solution to MANO NFV layer while the NFV manager deployed by CISCO is able to support third party NFVs as well. Additionally; CISCO developed a virtual EPC (LTE core network) solution which was recently deployed in NTT DOCOMO, largest mobile network provider in Japan and announced on 11th of March 2016 [31].
- HUAWEI NFV Open Lab: HUAWEI launched an open NFV lab aiming to test various deployment scenarios like a virtual EPC (LTE core network) PoC (Proof of Concept) trial in NTT DOCOMO in 2014.

- NEC: NEC has already deployed a vEPC solution.
- HP Open NFV: HP developed Open NFV solution based on the ETSI standards.
- ClearWater: ClearWater in cooperation with Metaswitch developed an open source implementation of IMS (IP Multimedia System) network.
- Alcatel Lucent (ALU): ALU in cooperation with RedHat developed a solution called Cloud Band.

The following table summarizes some of the industry projects on NFV up to this point:

Table 2.2 Industry NFV projects

	Functionality	Platform	Driving Standards
HP OpenNFV	Open standards-based NFV reference architecture in which carriers vendors can test vEPC	OpenStack	ETSI
NFV Open Lab	Support the development of NFV infrastructure, platforms and services	OpenStack, OpenDaylight	ETSI
CloudNFV	Provides a platform for virtual network service deployment and	OpenStack	ETSI and TMF
Alcatel CloudBand	Can be used for standards IT needs as well as for service providers who moving mobile networks into the	Red Hat Linux, Openstack Platform	ETSI
Cisco ONS	Automated services delivery, improved network and data center use, and fast deployments	OpenStack, OpenDaylight	ETSI
BroadBand NFV	Migrate virtual functions between platforms based on various vendor solutions		ETSI

CHAPTER 3

VIRTUAL MOBILE SIP PROXY INTEGRATION

3.1 Virtual Network Functions Concept in MNO

A network function (NF) is a functional block within a network infrastructure which has well-defined external interfaces and well-defined functional behavior**Error! Reference source not found.**[33] . A virtual NF is the implementation of the NF on the virtual resources such as a virtual machine (VM). A single virtual NF may be composed of multiple internal components, and hence it could be deployed over multiple VMs. Whereas, each VM hosts a single component of the virtual NF [32]. In the case of NFV, the NFs that make up the service are virtualized and deployed on virtual resources such as a VM. However, in the perspective of the users, the services whether based on functions running on a dedicated equipment or on VMs should have the same performance. The number, type and ordering of virtual NFs for each service are determined by the service's functional and behavioral specification. Therefore, the behavior of the service is dependent on the constituent virtual NFs **Error! Reference source not found.**[32].

In this chapter, a solution to resolve the scalability issue in MNOs is proposed. The solution uses the NFV concept to make up the virtualized call management NF in MNOs. TeleFinity IP PBX, which serves as SIP proxy, is the virtualized software solution that is responsible for controlling the call management function in MNO and manage users SIP calls' quality. The main benefits to virtualize the SIP proxy are the programmability, minimal operation cost , unified management, routing, efficient load balancing, and optimized management of server resources (including CPU and memory).

Next section presents an overview about the SIP proxy software, its benefits and modules which are used as a part in the proposed solution in the thesis.

3.2 What is SIP Proxy

SIP proxy, in telecommunication, is one of the main components of an IP PBX that is used by SIP to perform many of the call set-up functions. As described in RFC3621 standard ,SIP makes use of elements, called proxy servers (SIP proxy) to route requests to the user's current location, authenticate and authorize users for the services, implement provider call-routing policies, and provide features to users.

Within the SIP networks, the SIP proxy actually manages the setup of calls between SIP devices including the controlling of call routing and it also performs necessary functions such as registration, authorization, network access control and in some cases it also handles network security.

TeleFinity IP PBX is used as SIP proxy server within the solution proposed in this thesis because of its key benefits described in the subsection 3.3.1.

3.3 TeleFinity IP PBX Overview

TeleFinity IP PBX is a complete telephony software solution that provides telephone calls over IP data networks. It boasts robust PBX features, high reliability and scalability, and multi-protocol support. It is an intelligent IP Call Manager solution that provides contextual, continuous, and high capability journeys for end users [34][34]. It can be also connected to traditional PSTN lines, E1 lines, Q-SIG lines, and GSM lines via an optional gateway.

TeleFinity IP PBX has open platform architecture which is designed to handle a very high load of calls with maximum performance using minimum hardware requirements. The TeleFinity IP PBX platform is built on standard protocol components and operating systems providing MNO with flexible easy-to-use interface. Its architecture is easy to customize and integrate with other systems. TeleFinity IP PBX Development Kit (SDK) provides developers and system integrators with an option to integrate with any mobile telecommunication system. It enables MNO to route and monitor calls, make robust audio calls upon pre-defined flexible call flow designer and generate advanced statistical reports. Below figure shows typical architecture for TeleFinity IP PBX.

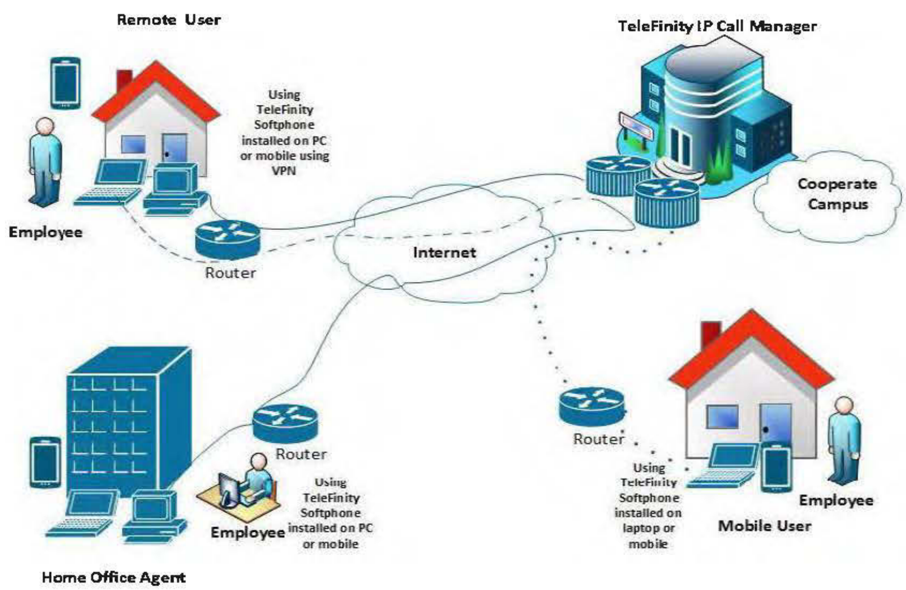


Figure 3.1 TeleFinity IP PBX Architecture ©

3.3.1 TeleFinity IP PBX Benefits

The following points present the key benefits of TeleFinity IP PBX solution that make it the best choice to serve the SIP proxy role in the proposed solution:

- It is IP Call Manage Solution that operates under SIP standards.
- Integrates with any telephony environment (2G, 3G,4G)
- Can handle up to 2,000 concurrent calls on single platform
- It can be supported and activated on any mobile device.
- No need for SIM swaps or on-device applications.
- High quality services
- Much easier to install & configure than a proprietary phone system, and easier to manage because of web/GUI based configuration interface.
- Significant cost savings using VoIP providers.
- Eliminate vendor lock in.
- Scalable Solution.
- Works efficiently in VM environment.
- Allow hot-desking & roaming (Mobile Extensions).
- Better phone usability: SIP phones are easier to use.
- Use existing network infrastructure without additional hardware.

3.4 TeleFinity IP PBX Modules

This section lists the modules supported by TeleFinity IP PBX solution. These modules present several NFs in MNO such as call accounting function, call monitor function, IVR function, sending fax function, etc. Moreover, these modules can be installed on VMs and used by the operators to offer new web services to their clients.

These web services can be managed by the end user himself using specific portal application installed on their mobile devices.

3.4.1 Call Center Module

Contact Center is SIP based solution that routes calls to queues, analyze and generate advanced statistical reports that helps in optimizing the contact center performance.

Contact Center is a cutting-edge PC based solution with open platform architecture. It is robust, reliable, scalable, and affordable contact center solution appropriate for small-scale to large operators.

Contact Center module runs up to 12,000 seats with unlimited number of agents. It can be expanded up to unlimited number of seats, making Contact Center extremely scalable, meeting MNO needs.

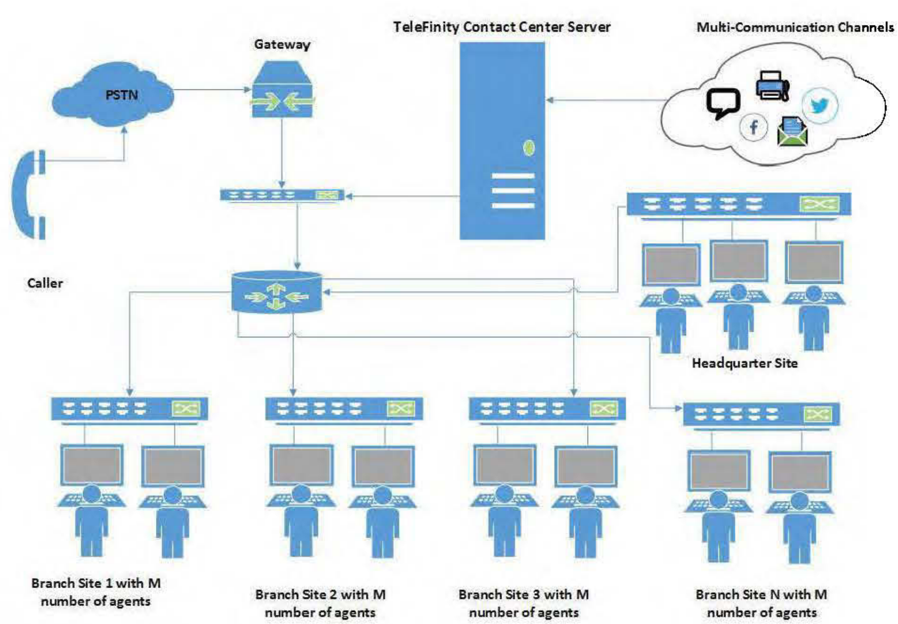


Figure 3.2 Call Center Module Architecture ©

Moreover, Contact Center can unify all customer interactions on single platform, regardless the media used – phone, email, fax, instant messaging (chat), email and Voice. This enables a unified routing, management, monitoring and reporting.

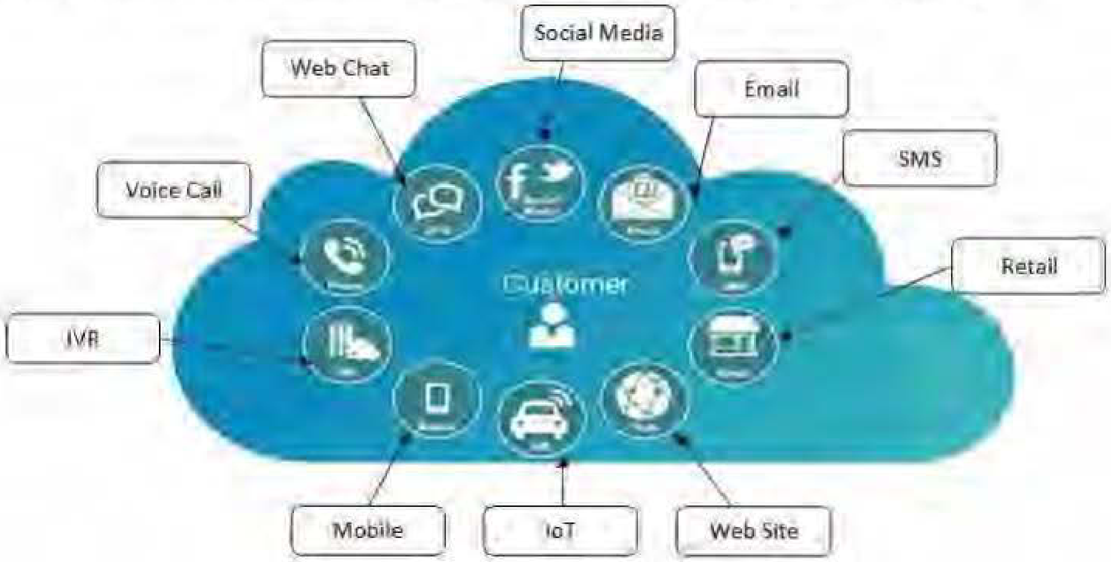


Figure 3.3 Supported Multi Media Services ©

3.4.2 Video Conference Module

TeleFinity IP PBX includes sophisticated video module that offers video calls function for MNO with below features:

- Breakthrough ease of use makes video conferencing as easy as a phone call.
- HD audio and video help participants hear every word and see every expression.
- Intelligent design that simplifies sharing HD content among all participants.
- Supports up to 16 video calls simultaneously.

3.4.3 Interactive Voice Response (IVR) Module

IVR module is a powerful Interactive Voice Response (IVR) Engine and Management Module. It is designed to offer very sophisticated features with a user configurable interface that uniquely provides IVR designer tool in a flow chart design. The IVR module can replace digit-intensive dual tone multi-frequency (DTMF) "Touch-tone" IVR interfaces with speech-enabled service.

Users can develop their own IVR applications using powerful and familiar scripting tools such as visual basic (VB) Script. The user can take full control of the call flow using the embedded object, which provides all the telephony functionality necessary to develop IVR applications. The user can use the existing capabilities of VB Script to connect to other system's database and hosts. Scripting Language supported by IVR module is VXML with .wav supported audio type.

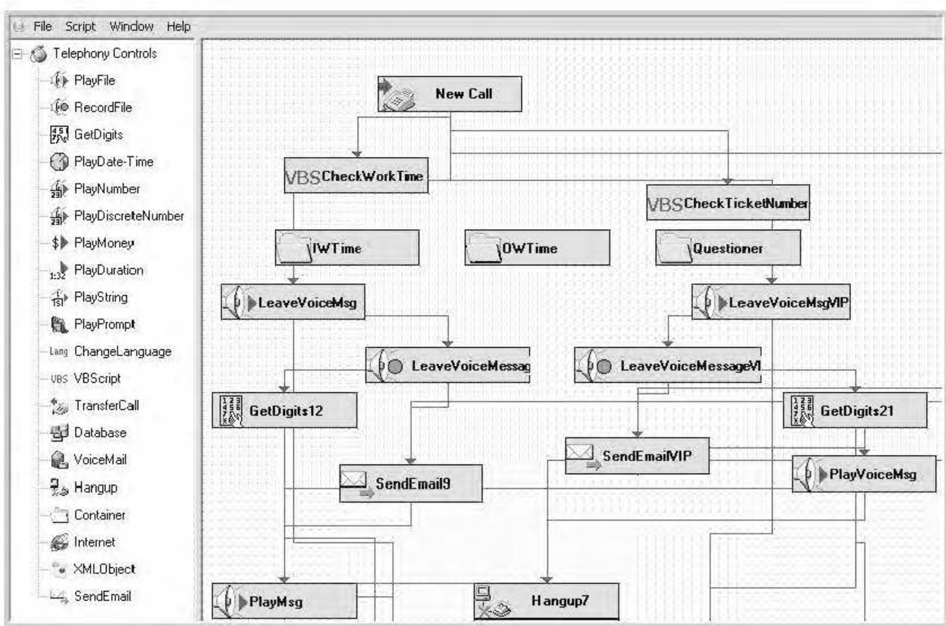


Figure 3.4 IVR Module Call Flow Designer [36]

3.4.4 Call Recording Module

Call recording module provides comprehensive multichannel recording capabilities that help MNO optimizing the customer care, agent performance, and regulatory compliance.

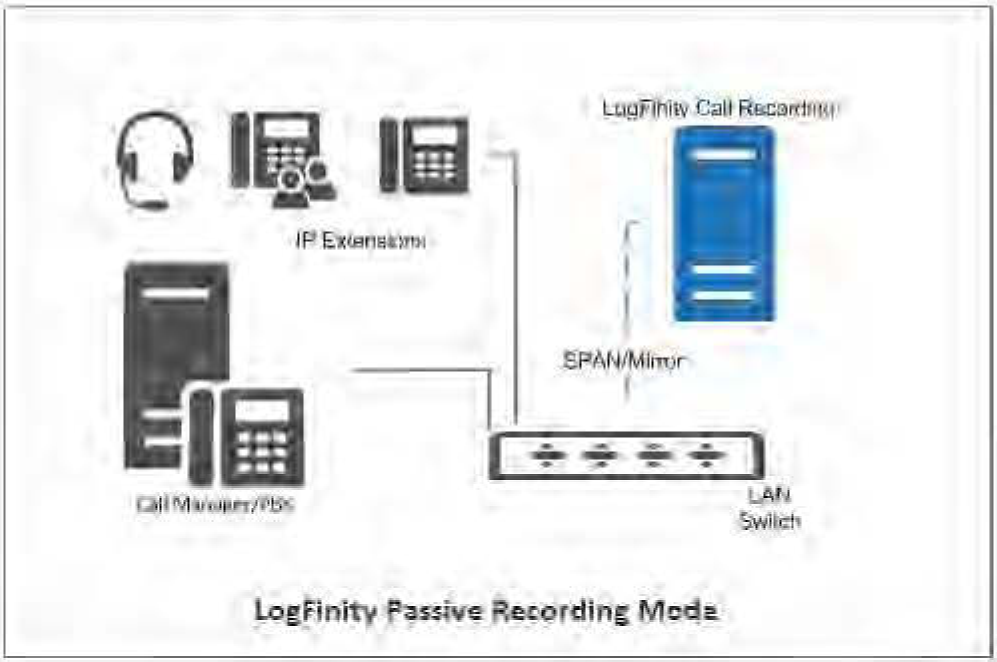


Figure 3.5 Call Recording Module [34][34]

3.4.5 Call Accounting Module

Call accounting module tracks incoming and outgoing phone calls in real time. Additionally, it identifies telecom expenses, fraud, emergency calls, and alarm conditions. It also provides easy to read call volume and trunk usage graphs that help MNO budget on track.

Figure 3.6 Call Accounting Module [34]

3.4.6 Fax Module

The Fax Server Application solution is implemented using the fax over IP (FOIP) technology integrated with IP Call Manager module.

Fax over IP functionality enables sending faxes over the fax line and the Internet at the same time. It can be integrated with the existing IP infrastructure, such as IP-enabled PBXs.

Using the SMTP protocol, the users will be able to send and receive faxes and text messages directly from their email account. To achieve this, the fax module must be integrated with the e-mail server. When the fax is received, it will be sent to the agent email account with fax message attachments (TIFF, RTF, etc…),.As well, the message can be saved easily and the agent has the ability to send, reply and receive faxes easily. The Fax message is stored in the mailboxes of each user

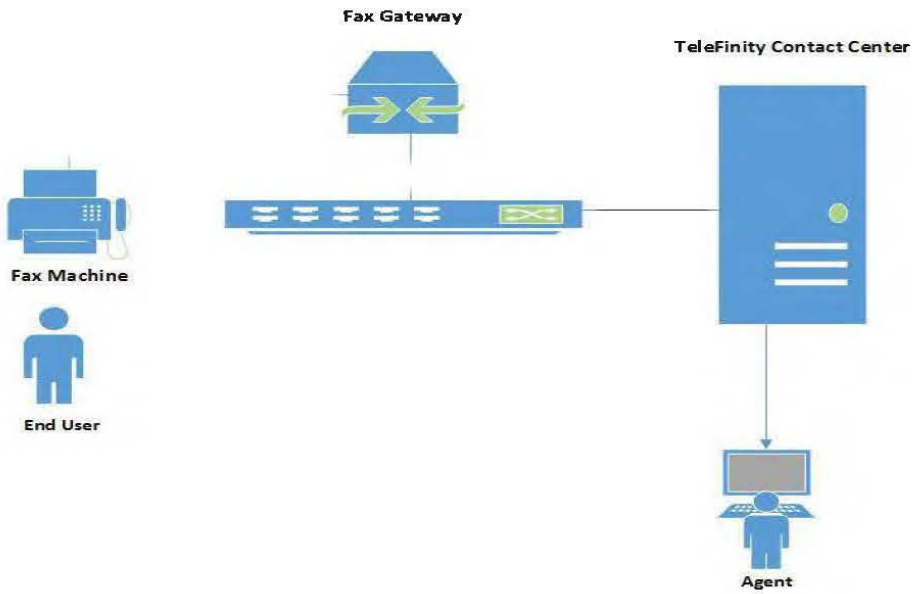


Figure 3.7 Fax Module ©

3.5 Mobile SIP Proxy Integration Approach

The proposed solution offers a hosted SIP Proxy application installed on VM within the MNO platform and it is integrated with the existing telecommunication entities using a SIP trunk. Multiple SIP proxy solutions can be deployed on multiple VMs to allow handling a greater number of subscribers. Each SIP proxy is installed on VM with high resources specifications i.e. CPU is Intel Xeon or better and memory size is 64GB or higher, and it can handle up to 2,000 concurrent calls on a single platform.

In this approach, the mobile terminal is registered as VoIP extension to the SIP proxy via mobile software application installed on the user device. Each mobile terminal can experience the same feature set as any mobile phone device connected to a mobile operator. The mobile terminal routes the calls via a packet-switched PS data connection of mobile operator as VoIP calls using IP connections. The mobile software application updates a presence status of the mobile phone to the SIP proxy. As well, it records the dialing user calls to a predefined number which is routed to the SIP proxy's dial-in service, and sends the user dialing to the proxy using DTMF (Dial Tone Multi-Frequency Signaling) signals.

As shown in figure 3.8, the proposed solution ,which is bordered with dotted blue line) is applicable to any user terminal e.g. a mobile or a wireless user terminal, such as a mobile phone , a user communication terminal with wire, a personal digital assistant (PDA), a game console, an e-reading device, a tablet, a smart phone, a personal computer (PC), a laptop, or a desktop computer .As well, this proposed solution can be integrated with any communications system or any combinations of different communications systems that support 2G, 3G, 4G and/or 5G systems using the SIP trunk without the interfering of their existing infrastructure. The communication systems may be a fixed communications system (e.g. landline networks) or a wireless communications system (e.g. mobile phone networks) or a communications system utilizing both fixed networks and wireless networks.

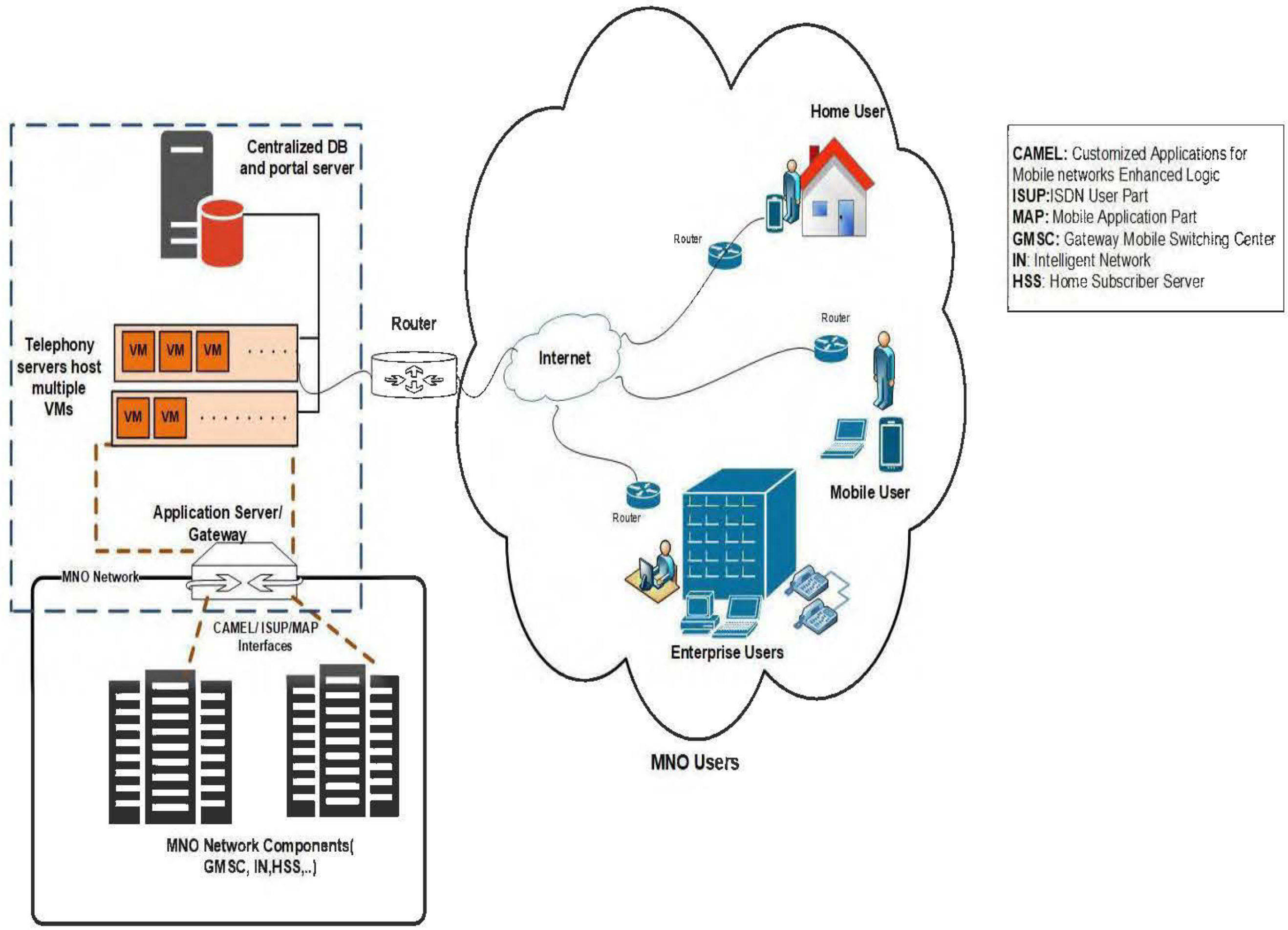


Figure 3.8 Proposed Solution Diagram ©

3.6 Proposed Architecture

3.6.1 Solution Design

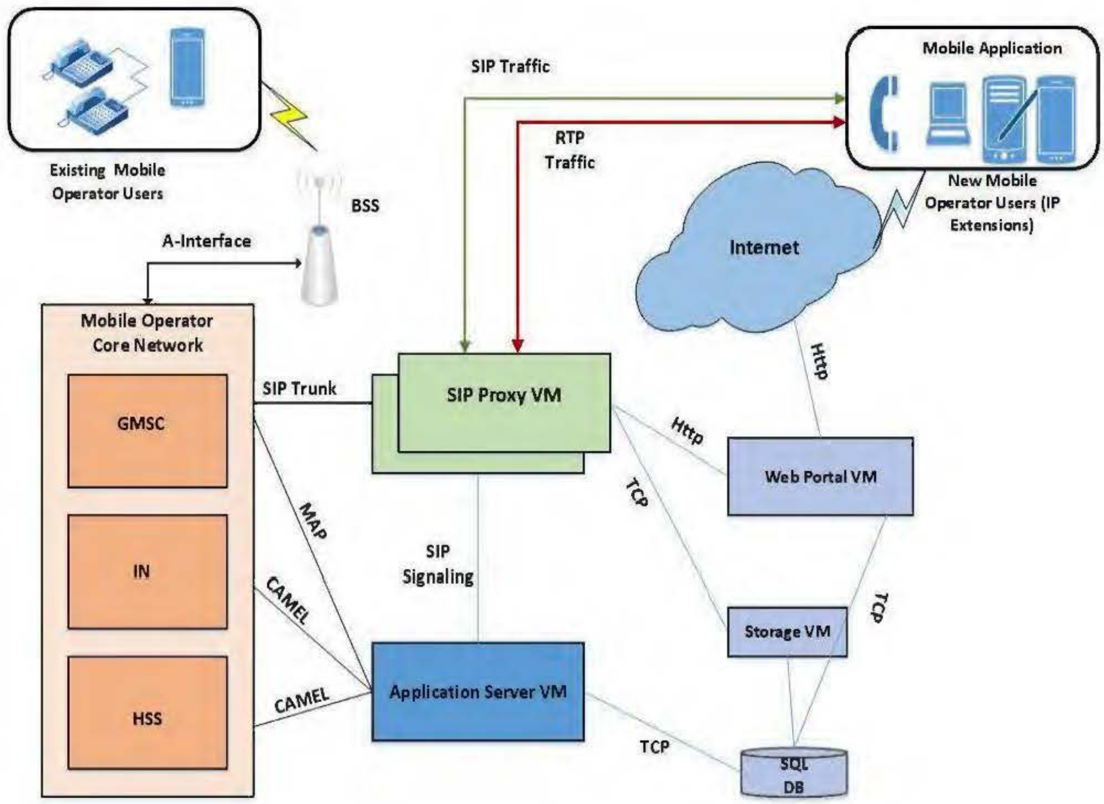


Figure 3.9 Proposed Solution Architectural Design ©

In figure 3.9, high-level solution architectural design is presented. As shown, the existing mobile users can access the communications network via an access network, e.g. via a base station subsystem (BSS) or a radio network subsystem. The core network element in MNO comprises a GMSC element (Gateway Mobile Switching Center element). It is an exchange element controlling the area, in which the mobile subscriber is presently located. As well, it can interrogate the routing of mobile terminating calls. GMSC can send notifications of registration and de-registration events of a mobile subscriber to an SCP (Service Control Point) which is a node that directs services in the intelligent network (IN). GMSC can request routing instructions from SCP on how to route an originating call or a terminating call of a user. SCP can then return a pre-defined prefix which is added before the destination number with which the call can be routed. GMSC can then route or direct the call to a SIP proxy VM. The purpose of using SCP in MNOs is to receive notifications of attach and detach procedures of the subscriber's IMSI (International Mobile Subscriber Identity Code) number and of location update events from GMSC. At least partly based on this information, SCP can determine whether the subscriber is currently registered to the network or not. SCP further delivers this information to the Application Service VM.

The Application Service (AS) VM in the above system architecture is responsible for controlling and managing the service that receives the registration and/or de-registration notifications from the SCP. When AS VM receives a new SIP session from the GMSC, its service recognizes the subscriber from the SIP signaling message, then it begins a new SIP session to the SIP proxy as an SIP extension of the subscriber. The same thing is done for the traffic in the opposite side, when the service on AS VM receives a SIP session from the SIP proxy, it recognizes the recipient from the SIP signaling, then it begins a new session with GMSC to the mobile device, combining the media of the incoming and outgoing sessions.

SIP proxy VM in the above system architecture is responsible for controlling SIP extensions registration and managing the switching of their calls. SIP proxy can receive the registration, and authenticate the user and deliver the traffic according to its configuration. The SIP proxy VM is interconnected with storage VM via TCP connection which is responsible for data processing for client calls history, calls records and clients status and save it in a predefined database schema on SQL Database server. As well, the SIP proxy VM is interconnected with web portal VM via HTTP connection which identifies the incoming requests from web clients. The main components contained in a web portal VM are the virtual directories which receive the inbound HTTP requests and process them.

3.6.2 VM Specifications

This section presents the VM specifications needed to deploy the proposed solution within the MNO platform. Three VMs are required with the below minimum specifications:

Table 3.1 Minimum VMs Specifications

Minimum VMs Specifications

	Telephony VM	DB & Portal VM	Application Server VM
CPU	Intel Xeon E5-2699 v4	Intel Xeon E5-2699 v4	Qaud Core CPU
RAM	64 GB	64 GB	8 GB
HDD Storage	1 TB	1 TB	500 GB
Database	-	SQL Server R2 2008 Standard Edition 64-bits	
OS	Windows Server R2 2008 64-bits SP1	Windows Server R2 2008 64-bits SP1	Ubuntu 14.04.3 LTS server

A- Telephony VM

The SIP proxy solution is implemented on the Telephony VM. The minimum technical specifications of the Telephony VM required to serve up to 12,000 subscribers and 2,000 concurrent calls/VM server are given in table 3.1.

B- Centralized DB & Portal VMs

The minimum requirements for each storage and web portal VMs to handle the aforementioned number of concurrent calls and users are given in table 3.1.

C- Application Service VM

The Application server is responsible for receiving the “Customized Applications for Mobile Networks Enhanced Logic” (CAMEL) triggers for the Mobile Originating Call (MOC) and Mobile Terminating Call (MTC) for/from subscribers who are registered to the SIP proxy. Moreover, it relays the “Originating - CAMEL Subscription Information” (OCSI) and “Terminating - CAMEL Subscription Information” (TCSI) in case of roaming toward the operators’ Intelligent Network (IN) entity. It doesn't require any specific interfaces, it integrates only with the operator using SS7 interfaces and the SIP protocol. Accordingly, the major advantage of using the Application Service over the IMS is that it has more straight forward integration with less operation cost..

The minimum requirements for Application Service VM are given in table 3.1.

3.7 Signaling Stream Scenarios

This section presents different scenarios for the signaling messages which are exchanged between the different units of the proposed solution as follows:

3.7.1 User Registration Scenario

Figure 3.10 is a signaling diagram for a registration procedure in the proposed solution. Starting with (MAP Note Mobile Management (MM) Event message) which means that the Visitor MSC (VMSC) informs the SCP that a mobility management event is occurred. In other words, the event that the subscriber has turned his phone on or off has occurred. SCP acknowledges (MAP Note MM Event result message) this information. If SCP has determined (Presence update message) means that the event is a network registration event, then the SCP will inform the Application Service of this. Next, the Application Service registers (SIP REGISTER message) the terminal apparatus or the mobile device as a SIP endpoint of the user. The Application Service authenticates (message Authentication) the user to the SIP proxy. Finally, the SIP proxy acknowledges (SIP 200 OK message) the registration procedure.

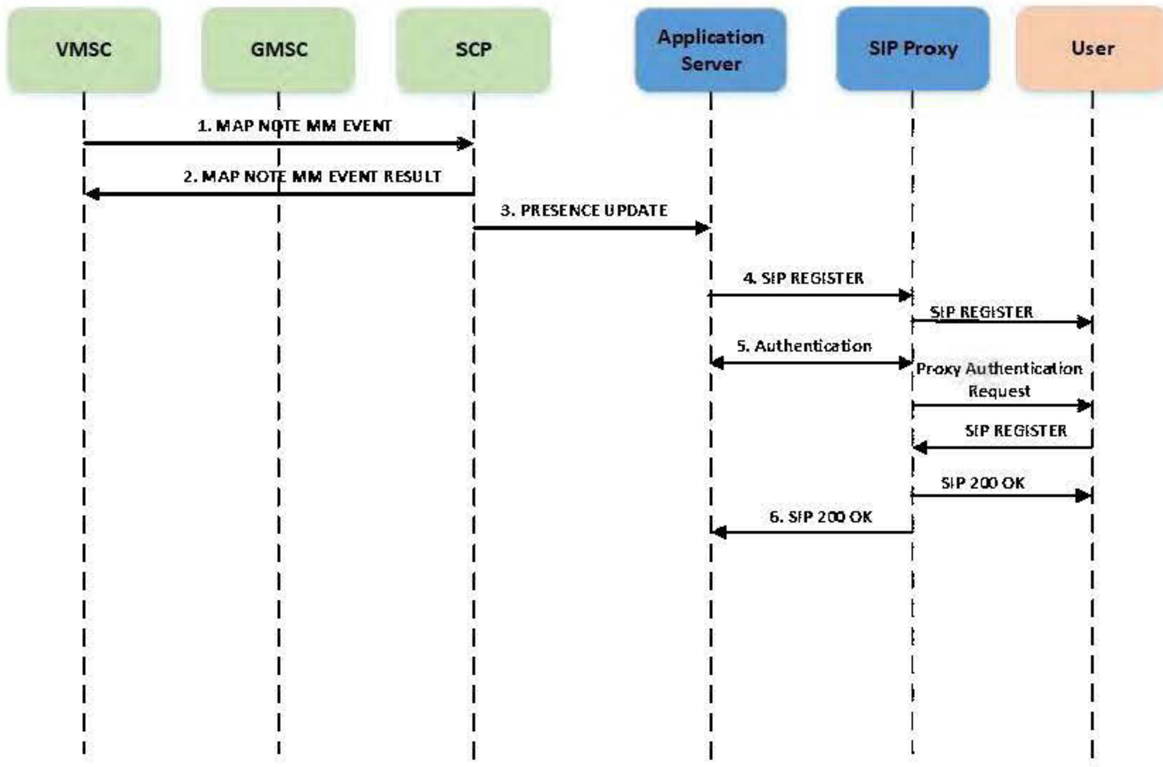


Figure 3.10 User Registration Scenario ©

3.7.2 Mobile Call Originating Scenario

Figure 3.11 shows a signaling diagram for the mobile call (PSTN-SIP) originating signaling flow in the proposed solution. The service user originates (ISUP Initial Address Message "IAM" message) a call using the mobile device. The call is routed to the operator's GMSC. GMSC requests (CAMEL Application Part "CAP" Initial DP message) routing instructions from SCP. Then, SCP returns (signaling CAP Connect) with the destination number and routing prefix. GMSC sends SIP Invite message to the Application Service VM based on the routing prefix returned from SCP .Then, the Application Service recognizes the SIP Invite message received and forms a SIP Invite message to SIP proxy.

The SIP proxy routes the call to GMSC using a SIP trunk interface. GMSC requests (CAP Initial DP message) instructions from SCP. SCP recognizes (CAP ContinueWithArgument message) the user. If the user has defined a caller identification "CLI" number to be shown, SCP sends this number in the generic number parameter of the CAP message. The generic number is then packed into the ISUP IAM Additional Calling Party Number parameter, which has precedence over the Calling Party Number, when showing CLI to the called party.

The call is routed (message ISUP IAM) to its destination. After that the destination node informs (message ISUP Address Complete ACM) GMSC that it has received the message. GMSC can then relay (message SIP 183 Session Progress) this information to the SIP proxy.

Then, SIP proxy relays (message SIP 183 Session Progress) this information to the Application Service, the Application Service receives and transmits or relays (message SIP 183 Session Progress) this information to GMSC, GMSC relays (message ISUP ACM) this information to the originating visitor MSC "VMSC" with ISUP and the destination node informs (message ISUP Call Progress (CPG)) GMSC that the called party is being paged, respectively.

GMSC relays SIP 180 Ringing information or message to SIP proxy, SIP proxy relays (message SIP 180 Ringing) this information to the Application Service which relays (message SIP 180 Ringing) this information to the GMSC and GMSC relays (message ISUP CPG) this information to the originating VMSC with ISUP. The destination node informs (message ISUP Answer ANM) GMSC that the called party has answered the call, and GMSC relays (message SIP 200 OK) this information to SIP proxy. After that SIP proxy relays SIP 200 OK information or message to Application Service. Application Service relays (message SIP 200 OK) this information to GMSC, and GMSC relays (message ISUP ANM) this information to the originating VMSC.

Next, GMSC can trigger the call to CAP (CAMEL Application Part) based forced routing service for originating call (preferably DP2). Prefix returned from the forced routing service routes the call to the Application Service SIP trunk.

The Application Service recognizes calling subscriber from SIP invite and initiates a call to SIP proxy as a user SIP client. SIP proxy routes the call to PSTN via the SIP trunk to GMSC. Incoming call is triggered to optional CAP based CLI change service, which may change additional calling number.

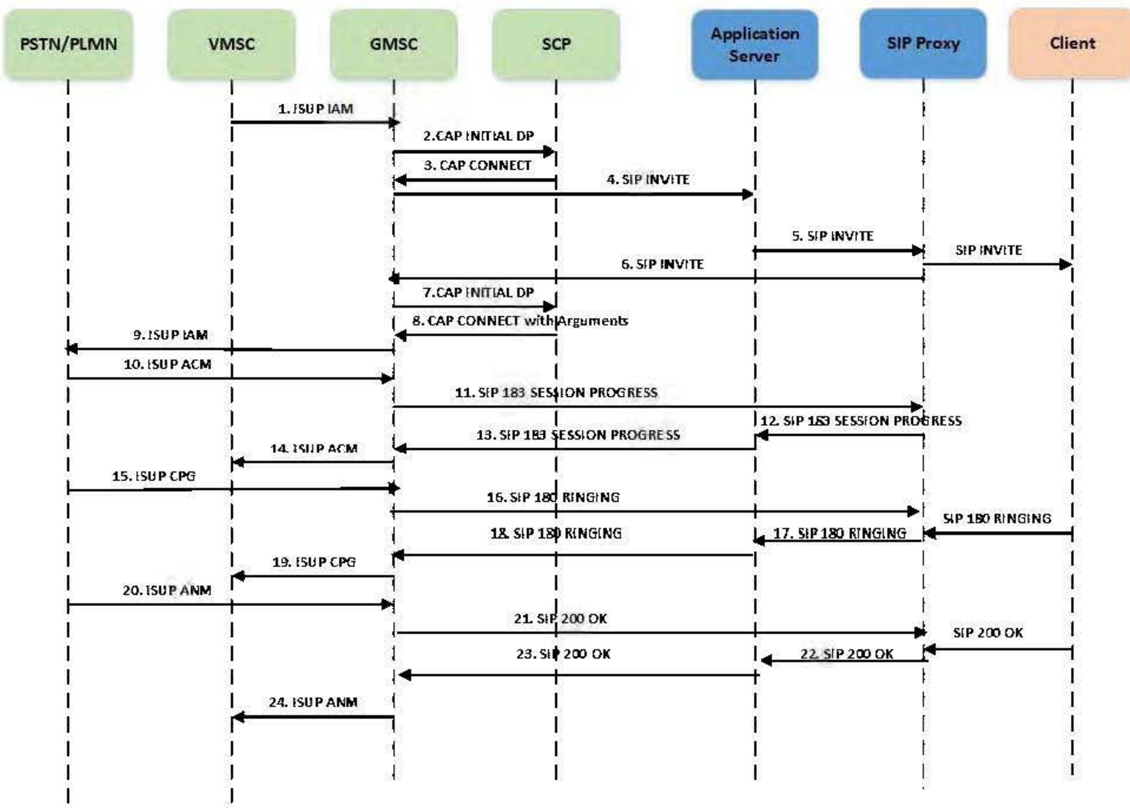


Figure 3.11 Mobile Call Originating Scenario ©

3.7.3 Mobile Call Terminating Scenario

Figure 3.12 shows a signaling diagram for the mobile call (PSTN-SIP) terminating signaling flow according to the proposed solution. The signaling diagram begins with the terminating call arriving (message ISUP Initial Address Message (IAM)) at GMSC. In response to that, GMSC requests (message CAP Initial DP) routing instructions from SCP, and SCP returns (message CAP Connect) a routing prefix before the destination number. Based on the routing prefix, GMSC forms (message SIP Invite) a SIP Invite message, and transmits it to the SIP proxy.
The SIP proxy determines (message SIP Invite) that the call should be connected to the mobile device of service user, and it sends a SIP Invite message to the Application Service. The Application Service initiates (message SIP Invite) a SIP session with GMSC using the SIP trunk interface.

A prefix can be added to the destination number in this part so that GMSC knows to bypass the forced routing service. GMSC routes (message ISUP IAM) in the call to VMSC, where the subscriber is currently registered, and VMSC replies (message ISUP Address Complete (ACM)) that it has received the message. Now, GMSC relays (message SIP 183 Session Progress) this information to the Application Service and it receives and transmits or relays (message SIP 183 Session Progress) this information to the SIP proxy.

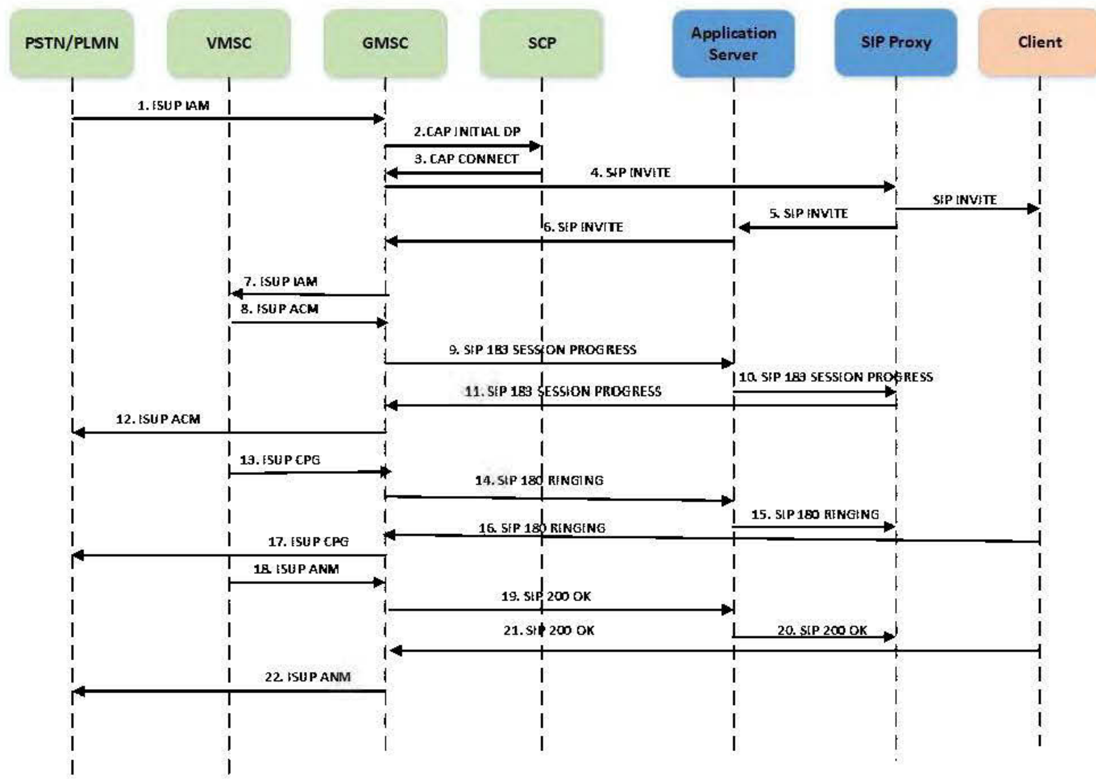


Figure 3.12 Mobile Call Terminating Scenario ©

The SIP proxy can relay (message SIP 183 Session Progress) this information to GMSC, and GMSC relays (message ISUP ACM) this information to PSTN in ISUP. VMSC informs (message ISUP Call Progress (CPG)) GMSC that the subscriber is being paged. Then, GMSC relays SIP 180 Ringing information or message to the Application Service. The Application Service relays (message SIP 180 Ringing) this information to the SIP proxy, the SIP proxy relays (message SIP 180 Ringing) this information to GMSC, and GMSC relays (message ISUP CPG) this information to PSTN in ISUP.

The VMSC informs (message ISUP Answer (ANM)) GMSC that the subscriber has answered the call, GMSC relays (message SIP 200 OK) this information to the Application Service, the Application Service relays (message SIP 200 OK) this information to the SIP proxy, and the SIP proxy relays (message SIP 200 OK) this information to GMSC. Finally, GMSC relays (message ISUP ANM) this information to PSTN in ISUP.

Typically, each SIP endpoint has its own IP address. Because in this case, the number of SIP endpoints on the same server may be in the thousands, this solution does not scale. The SIP client application then appears to the SIP proxy as it was a SIP endpoint behind a SIP supporting NAT (Network Address Translator) firewall. In practice, the SIP REGISTER message's contact header SIP URI (Session Initiation Protocol Uniform Resource Identifier) must include the TCP/UDP (Transmission Control Protocol/User Datagram Protocol) port allocated to the user.

CHAPTER 4

EXPERIMENTAL TESTBED IMPLEMENTATION

4.1 Introduction

This chapter describes the hardware specifications and the software solution used to implement the experimental testbed of the proposed solution and evaluate the performance SIP proxy under virtualization environment for different voice codecs, G.711, G.727 and G.723.1. The SIP proxy configuration and the evaluation metrics are introduced as well. In this thesis, the testing campaigns were carried out over two types of networks: Local Area Network (LAN) and Wide Area Network (WAN). And finally, the solution network topology implemented over these two networks is described.

4.2 Experimental Testbed Setup

In order to implement a real experimental testbed and evaluate the system performance under virtualization, several hardware parts and software solutions are used as below:

4.2.1 Hardware Parts

Three virtual machines on three different physical servers are used in the experimental testbed implementation. The SIP proxy (TeleFinity IP PBX) solution is installed on the first VM, the StarTrinity SIP calls simulator [35][35] is installed on the second VM. And the IVR [36][36] solution is installed on the third VM. In our solution, the IVR system receives the SIP calls which are generated from the StarTrinity SIP simulator and plays a music file (.wav) with one minute duration.

The specifications of VMs used in the testing are shown in table 4.1:

Table 4.1 VMs Specifications used in the test bed

VMs Specifications			
	SIP Proxy VM	**IVR VM**	**SIP Tester VM**
CPU	Intel Xeon E3-1230 v2	Intel Xeon E5-2699 v4	Intel Xeon E5-2699 v4
RAM	8 GB	4 GB	4 GB
HDD Storage	1 TB	1 TB	500 GB
OS	Windows Server R2 2012 64-bits	Windows Server R2 2008 64-bits SP1	Windows Server R2 2008 64-bits SP1

4.2.2 Software Solutions

The software solutions that are used to implement, analyze and evaluate the proposed solution performance under virtualization environment are listed in the below subsections.

A- TeleFinity IP PBX Solution

TeleFinity IP PBX solution is used as a SIP proxy in the proposed method of this thesis. It registers the SIP client applications and setups SIP calls and tears them down. TeleFinity IP PBX features and modules are discussed in sections 3.3 and 3.4.

B- ActFinity IVR and Voicemail Solution

ActFinity IVR and Voicemail solution is used to receive the calls generated by the SIP simulator and play a music file lasts for one minute. Appendix A includes the call script (.vbs) used in the testing.

C- Startrinity SIP Tester

The main role of Startrinity SIP tester is to generate concurrent SIP calls and send them to the IVR system. Both the Startinity and the IVR systems are registered as SIP extensions on the SIP proxy.

Startrinity SIP tester provides the measurement of the key performance parameters for the SIP calls under a high load environment. These key parameters are such as the interval between transmission and reception of some key SIP messages (e.g., Invite, 100 Trying, 180 Ringing). The RTP QoS metrics like packet loss, RTP delay, and MOS are provided as well. These metrics are described in details in chapter 5.

D- Commview Analyzer

Commview Analyzer is used to measure the traffic payload of the proposed solution, in addition to the VoIP channels throughput.

E- Resource Monitor

Resource Monitor is a useful tool in Microsoft Windows that helps in finding the CPU and memory resources used by SIP proxy services to handle the SIP calls.

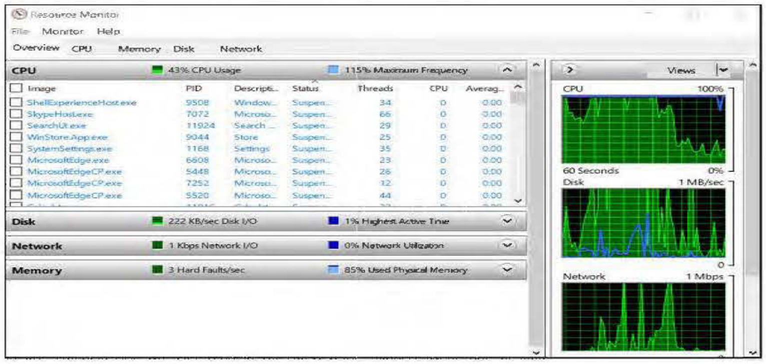


Figure 4.1 Windows Resource Monitor ©

4.3 Experimental Testbed Configuration

4.3.1 TeleFinity IP PBX Configuration

- <u>Step1: Configure the SIP extensions on the SIP proxy</u>: Two extensions are created on TeleFinity IP PBX. The extensions DNs are used to register ActFinity IVR system and Startrinity SIP tester on the SIP proxy. Each extension has 200 maximum concurrent calls allowed. Table 4.2 shows the configuration parameters of each SIP extension. The SIP extensions DNs used in this experimental testbed are extension 20000 and 20001. Extension DN “20000” is used to register the ActFinity system on TeleFinity IP PBX. While, extension DN “20001” is used to register the SIP tester on the proxy.

Table 4.2 SIP Extension Configuration Parameters

Parameter	Description
Terminal ID	Terminal Number
DN	Directory Number of the SIP extension
Codec	Voice codec used to handle extensions’ calls
Password	SIP extension password
Maximum Concurrent Calls	Maximum calls can be handled at the same time
IP Address	IP Address f device that is registered with the SIP extension

- Step 2: Select the voice codec: The voice codec can be selected from the main system settings menu of the TeleFinity IP PBX solution. Voice Codecs G.711, G.729 and G.723.1 are selected respectively during our test.
- Step 3: Configure the SIP trunk on the SIP proxy: Where Table 4.3 shows the SIP trunk settings that are needed to integrate TeleFinity IP PBX solution with the GMSC in MNO.

Table 4.3 SIP Trunk Configuration Parameters

Parameter	Description
Trunk DN and User ID	Trunk number that is used as prefix for the called numbers related to the calls carried on the SIP trunk
IP Address	IP address of GMSC
Codec	Voice codec used for RTP traffic
Port	Signaling port used for SIP traffic
Transport Protocol	TCP or UDP
Password	SIP Trunk password used for authentication purposes

4.3.2 AcFinity IVR Configuration

ActFinity IVR solution is registered as extension on the SIP proxy (TeleFinity IP PBX). Figure 4.2 shows system setting configuration of the IVR system used during the testing.

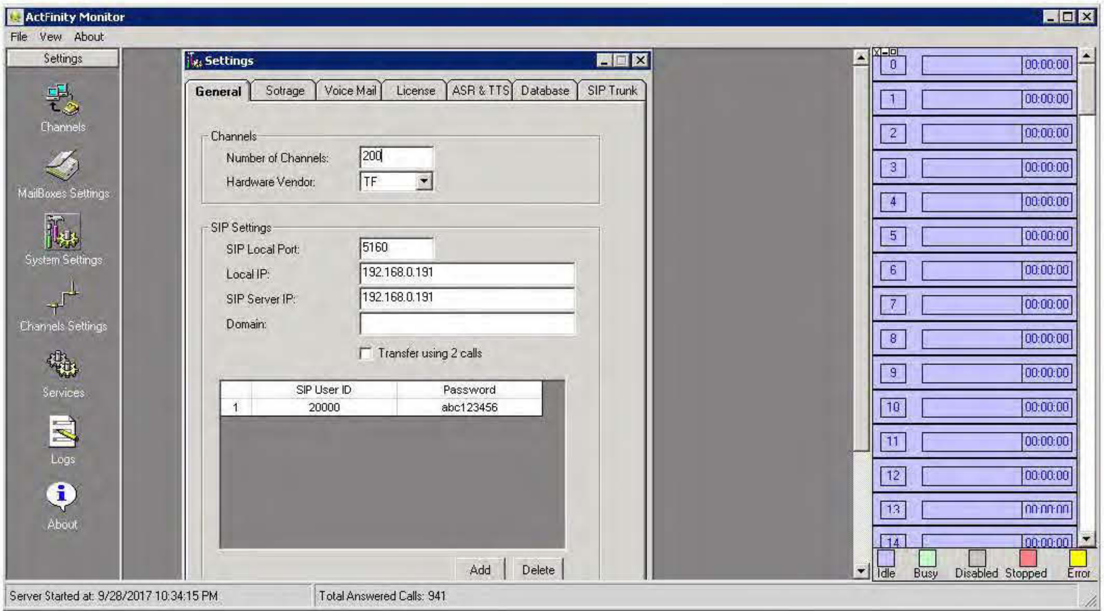


Figure 4.2 IVR System Settings ©

Table 4.4 gives a description for the system setting fields in ActFinity system:

Table 4.4 IVR Configuration Parameters

Parameter	Description
Number of the Channels	Are the number of maximum calls handled by the IVR channels at the same time. In the test, this value is set to 200
Hardware Vendor	Value is set to TF, means the channels work on the SIP standard
SIP Local Port	is the TCP ort used for SIP signaling between TeleFinity IP PBX and IVR system. Port 5160 is used
SIP Proxy IP	is the IP address of the SIP proxy to which IVR is registered
SIP User ID and SIP Password	are the SIP extension DN and password of IVR extension as set in TeleFinity IP PBX

4.3.3 Startrinity SIP Tester Configuration

During the test, the "Outgoing Calls Simulation" section is selected from the SIP tester to simulate the SIP calls. Figure 4.3 shows the system parameters used:

- ✓ Voice Codec;
- ✓ Destination Number which is the SIP extension DN where the generated SIP calls will be sent. In the test, this value is set to the IVR DN;
- ✓ Host IP address where the destination SIP extension is installed;
- ✓ Number of concurrent calls needed to be generated;
- ✓ SIP user ID and password : are used to register Startrinity to TeleFinity IP PBX;
- ✓ SIP proxy IP address to which Startrinity is registered.

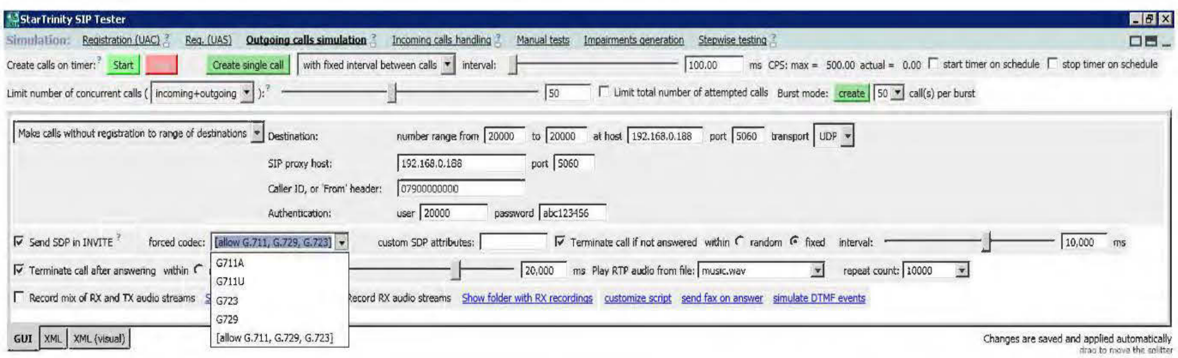


Figure 4.3 Startrinity Configuration ©

4.4 Network Topology

The testing campaigns were carried out over two types of networks: Local Area Network (LAN) and Wide Area Network (WAN). Figure 4.4 shows the LAN topology where all VMs are connected to the same LAN. While figure 4.5 shows WAN topology where the VMs are interconnected over WAN and the RTP traffic is carried over the internet.

Table 4.5 includes a comparison between LAN and WAN features and characteristics to justify the performance results of proposed solution within the two topologies.

Table 4.5 Comparison between LAN and WAN

Points of Comparison	LAN	WAN
Definition	A group of connected devices in a small geographical area	It covers a large area like a country or continent
Congestion	Less	More
Propagation Delay	Short	Long
Bandwidth	High	Low
Data Transfer Rate	High	Low
Transmission Errors	Low	High

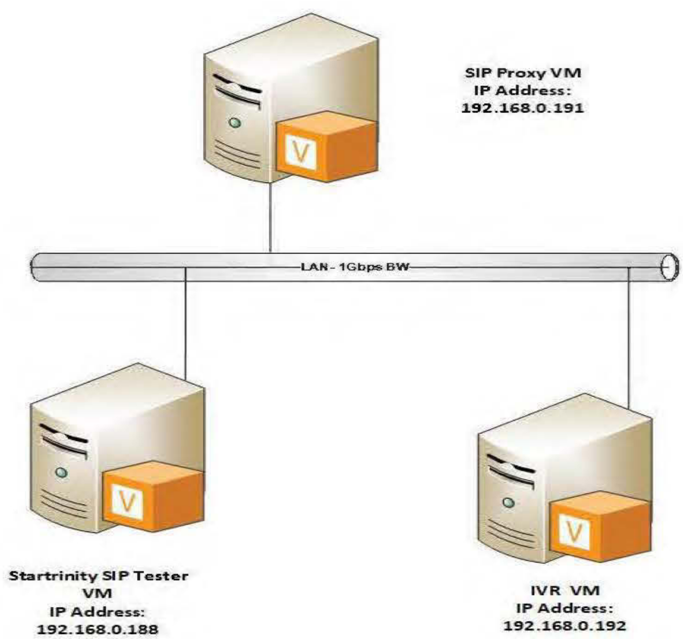


Figure 4.4 LAN Topology ©

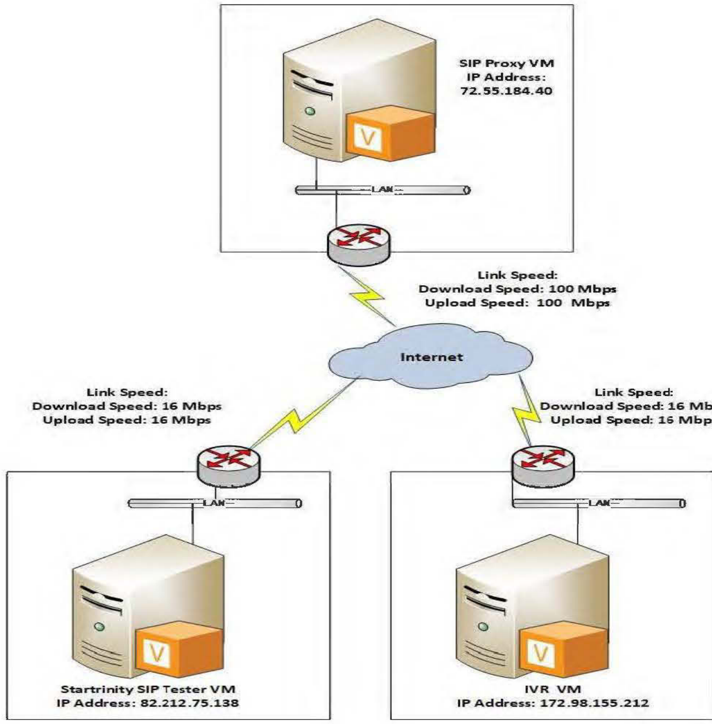


Figure 4.5 WAN Topology ©

CHAPTER 5

RESULTS AND DISCUSSION

This chapter presents the testing scenarios that are carried on the network topologies described in section 4.4. Measured parameters description which are used during testing are also introduced. Testing results for the three voice codecs G.711, G729 and G.723.1 are presented followed by the analyzing for the performance of the SIP proxy under the virtualization environment. The results are compared to LTE QoS standards to ensure the reliability of the proposed solution.

5.1 Voice Codecs Properties

The voice signal must be encoded and compressed in order to be sent over the packet network. For this purpose, voice codecs are used. Each codec has different characteristics concerning the data rate used, the compression level and the associated user-perceived quality. The voice codec characteristics has a direct impact on the overall performance of the VoIP systems.

Below subsections describe the properties of the voice codecs used in the testing.

5.1.1 G.711 u-law

G.711 codec is a suite of the ITU (International Telecommunications Union) standards and it is used for voice communication and telephony networks. Theoretically, G.711 codec provides good quality of voice and requires higher processing, as it has higher bit rate as compared to other ITU CODECS [38]. It employs a logarithmic compression that compresses each 16-bit sample to 8-bits. As a result, it digitizes voice into 64 kbps, which consider the highest bit-rate among the codecs.

5.1.2 G.729

G.729 codec is another ITU-T standard and it has the ability to compress the payload for low bit rate by using an algorithm known as CS-ACELP (Conjugate-structure algebraic-code-excited linear-prediction) with a bit rate 8kbps. Theoretically G.729 codec provides reasonable delay and high speech quality **Error! Reference source not found.**[39].

5.1.3 G.723.1

G.723 codec is another ITU-T standard that was designed for voice & multimedia communication over standard phone system. It is an extension of G.721 codec, which was modified to provide real-time coding and suitable voice quality. Theoretically G.723 codec is not suitable for music and provides lower quality output than other CODECS **Error! Reference source not found.**[40]. It is designed for calls over modem links with data rates of 28.8 and 33 kbps. It operates at 6.3 and 5.3 Kbps. Although this standard decreases bandwidth exhaustion, the voice is much poorer than with G.729 and is not very common for VoIP.

Table below describes the features of three selected ITU-T (International Telecommunications Union) standard voice CODECS such as G.711, G.723 and G.729. The comparison of these CODECS illustrates that each codec has different frame size, payload, speed and processing time, which affect the quality of voice **Error! Reference source not found.**[41].

Table 5.1 VoIP Codecs specification

Codec	G.711	G.723.1	G.729
Coding Speech (kbps)	64	6.3	8
Frame size (ms)	20	30	10
Processing Delay (ms)	20	30	10
Look ahead Delay(ms)	0	7.5	5
License Required	No	Yes	Yes
Payload (bytes)	160	24	20
Algorithm	PCM	MPC-MLQ	CS-CELP

5.2 Measured Parameters

This section includes the parameter which are measured to evaluate and analyze the performance of virtual SIP proxy. These parameters are grouped as below:

1. The first group was measured at SIP proxy side and it includes hardware utilization parameters. At this point, CPU, and the memory utilization were measured.
2. The second group was measured at IVR system side and it includes RTP QoS metrics which are RTP jitter, RTP delay, MOS, and RTP packet loss. As well, signaling performance metrics are measured also which are Session Request Delay (SDR), 100 response time, and Call answer delay.
3. The third group was measured at IVR and Startrinity SIP tester sides and it includes network utilization parameters. At this point, generated RTP traffic and VoIP channels throughput were measured.

5.2.1 Voice Quality of Service (QoS) Parameters

Four metrics are measured during the test to evaluate the voice quality for three codecs (G.711, G.723.1, and G.729). These metrics are measured using "Startrinity SIP tester" software.

1. Packet loss is defined as the percentage of the RTP packets lost while the voice is transmitted from the source to the destination or it the percentage of the voice being discarded in case of the delay. It should be no packet lost in the ideal cases but the packet loss percentage up to 1.5% is still acceptable in the voice communication [37][42]. VoIP is assumed as unreliable real-time communication that uses UDP. The recovery of lost packet is possible by selecting proper voice codec.
2. Delay is the time needed by RTP packets to get from the source to the destination. Voice Delay is the most important factor in determining the voice quality for VoIP. This delay can be caused due to the encoding and decoding delay, the network delay, packets processing delay at source and destination, and propagation delay **Error! Reference source not found.**[43]. One of the main problems caused by the delay is the echo. Below table specifies delay guidelines of ITU-T Recommendation G.114 for an adequately controlled echo **Error! Reference source not found.**[44].

Table 5.2 ITU-T Recommendation G.114 for Delay Specification **Error! Reference source not found.**[44]

Delay	Comments
0-150 ms	Acceptable for most applications
150-400 ms	Acceptable, provided that the impact of delay is known.
Above 400 ms	Unacceptable

3. Jitter is defined as the end to end time variation between RTP packets sent and received between two nodes over the mobile network [45]. Jitter is caused due to many reasons including the network congestion, route changes between the sender and the receiver, and overload in the jitter buffer. Ideally, voice communication systems show zero jitter for the transmitted RTP packets. However, in the real voice communication systems, the acceptable range for the jitter is less than 20ms [43].

   Jitter is measured by using the equation below [45]:

$$\text{Jitter} = \frac{\text{current packet received time} - \text{last packet received time}}{\text{differential of sequence number between two packet}} \tag{1}$$

4. MOS is a subjective measure that shows the overall voice quality as perceived by the end user. Its values range between 1.0 (poor) to 5.0 (best). As defined in the ITU-T Recommendation [46] MOS defines perceived quality in a [0,5] scale as shown in the following table:

   Table 5.3 MOS Rating [46]

Rating	Speech	Description
1	Excellent	Imperceptible errors
2	Good	Perceptible but annoying
3	Fair	Slightly annoying
4	Poor	Annoying
5	Bad	Very Annoying

5.2.2 Signaling Performance Parameters

Three parameters are measured during the test to evaluate the signaling performance for SIP proxy at the three voice codecs. These parameters are measured using "Startrinity SIP Tester" software. All three parameters are affected highly by the lossy channels and the voice jitter.

1. Call Answer Delay: is the full time it takes after party A sends the INVITE message for it to receive the 200 OK response from party B. Call answer delay includes post-dial delay which is the time it takes after party A sends the INVITE message for the phone at party B to ring.
2. 100 Response Delay is the time between sending INVITE request and receiving 100 TRYING message.
3. Session Request Delay (SRD): is the time between first Invite message and its related 180 Ringing messages. Figure 5.1 shows SRD time within SIP signaling flow.

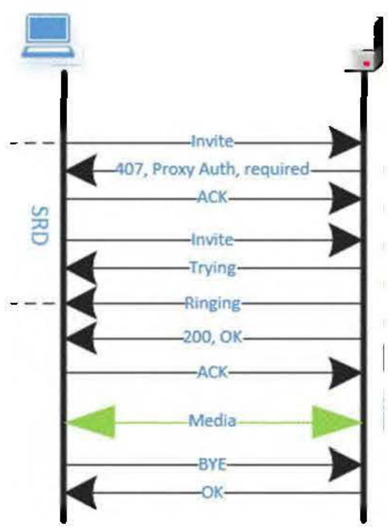


Figure 5.1 SRD Time [35]

5.2.3 Network Utilization Parameters

Two parameters are measured during the test to evaluate the network utilization performance for SIP proxy at the three voice codecs. These parameters are measured using "Commview" software.

1. Throughput: corresponds to the amount of data in bits that is transmitted over the channel per unit time. The throughput for different codec systems is shown in the below table.

   Throughput can be calculated using below equation **Error! Reference source not found.**:

   Bandwidth = total packet size * PPS (2)

   Where;

   PPS (packet per second) = (codec bit rate) / (voice payload size)

   Total packet size = (L2 header: MP or FRF.12 or Ethernet) + (IP/UDP/RTP header) + (voice payload size)

For example, to calculate the required BW by G.729 channel using the above equations with default voice payload equals 20 bytes:

- Total packet size (bytes) = (MP header of 6 bytes) + (compressed IP/UDP/RTP header of 40 bytes) + (voice payload of 20 bytes) = 66 bytes
- Total packet size (bits) = (66 bytes) * 8 bits per byte = 528 bits
- PPS = (8 Kbps codec bit rate) / (160 bits) = 50 pps
- Then, Bandwidth per call = voice packet size (528 bits) * 50 pps = 26.4 Kbps

2. Traffic Payload: It is the IP payload plus the voice payload size that represents the number of bytes (or bits) that are filled into a packet. The voice payload size must be a multiple of the codec sample size. For example, G.729 packets can use 10, 20, 30, 40, 50, or 60 bytes of voice payload size **Error! Reference source not found.**.

5.2.4 Hardware Utilization Parameters

Two parameters are measured during the test to evaluate the hardware utilization performance for SIP proxy at the three voice codecs. These parameters are measured using "Resource Monitor Application" on the Windows server.

1. Average CPU load: It is the Central Processing Unit power allocated for particular application. It is normally calculated as a percentage of the CPU used amount when an application is executed. There are two possibilities; firstly, if large quantity of packets is transmitted over the network, more CPU resources will be used. Secondly, if less quantity of packets is transmitted over the network but more CPU power is used for a specific transition mechanism.

   Figure 5.2 shows SIP proxy services whose CPU load is calculated during the test.

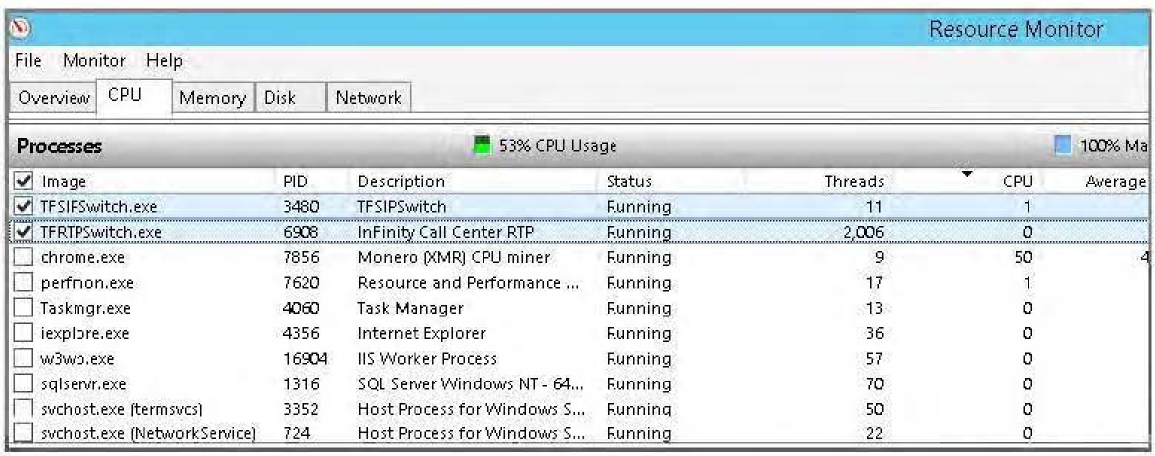


Figure 5.2 SIP Proxy Services CPU Load ©

2. Memory Utilization is the working set of memory that utilized by services of specific application. It is the amount of physical memory currently in use by process of the application services and it is measured in (KB).

Below figure shows SIP proxy services whose memory utilization is calculated during the test: “TFRTPSwitch.exe” and “TFSIPSwitch.exe” services.

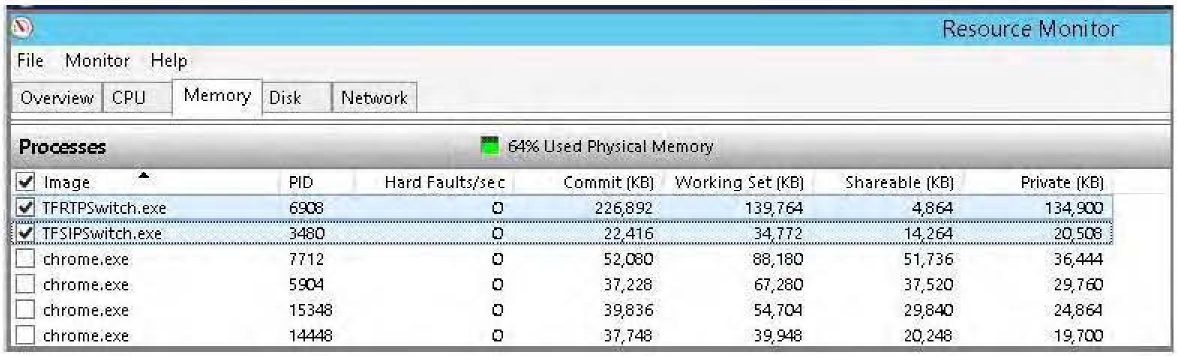


Image	PID	Hard Faults/sec	Commit (KB)	Working Set (KB)	Shareable (KB)	Private (KB)
TFRTPSwitch.exe	6908	0	226,892	139,764	4,864	134,900
TFSIPSwitch.exe	3480	0	22,416	34,772	14,264	20,508
chrome.exe	7712	0	52,080	88,180	51,736	36,444
chrome.exe	5904	0	37,228	67,280	37,520	29,760
chrome.exe	15348	0	39,836	54,704	29,840	24,864
chrome.exe	14448	0	37,748	39,948	20,248	19,700

Figure 5.3 SIP Proxy Services Memory Utilization ©

5.3 Testing Scenarios and Results Analysis

Testing campaigns in this thesis are carried out on the two network topologies: the LAN and WAN networks. For each network topology, two scenarios were performed to evaluate the performance of the proposed solution over different scenarios. In the first scenario, the RTP packets are passing through the SIP proxy for both the LAN and WAN topologies. While in the second case, the RTP packets are not passing through the SIP proxy for both the LAN and WAN topologies. Below subsections include illustration for the testing topologies for each scenario and present the results.

.

5.3.1 Scenario 1: WAN Topology- RTP Passing Through SIP Proxy

Figure 5.4 shows the network topology used in Scenario1.

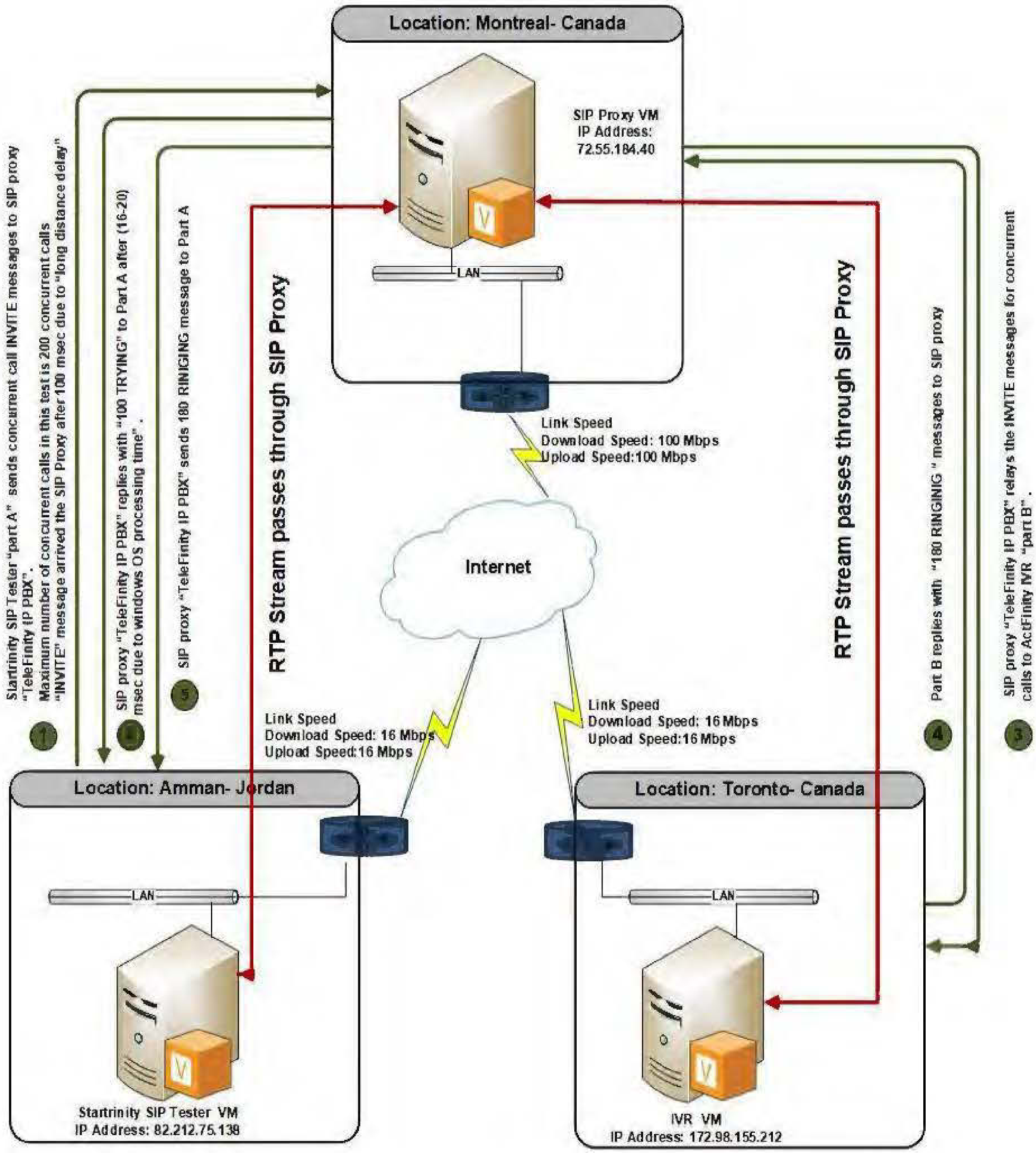


Figure 5.4 WAN Topology- RTP Passing through SIP Proxy Scenario ©

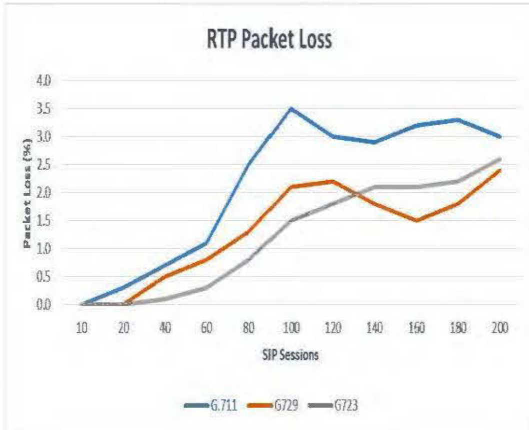
In this scenario, SIP proxy VM, IVR VM, and the SIP tester/ simulator VM are hosted on three different physical servers located at three different geographical locations. The three servers are interconnected over WAN using different internet links speeds as shown in Figure 5.4. The SIP tester generates SIP concurrent calls starting from 10 calls and it is aggregated by 10 till it reaches 200 concurrent calls. Then, the IVR system receives the generated SIP calls from SIP tester side and plays audio wave message with duration equals to one minute. The RTP packets are passing through the SIP proxy VM.

A- Voice QoS Results Discussion

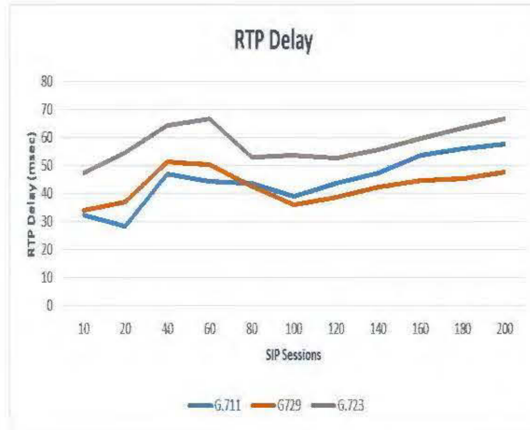
Below figures show the voice QoS results for the voice traffic transmitted over WAN network and passing through SIP proxy "TeleFinity IP PBX". Figure 5.5 (a), shows the RTP packet loss percentages versus the number of concurrent SIP calls transmitted over WAN network for different voice codecs (G.711, G.729, G.723.1). The results present that G.711 codec scheme has the highest voice packet loss. While the other two codec schemes i.e., G.729 and G.723.1 have lower voice packet loss. The packet loss becomes higher as the number of the SIP sessions increases. The packet loss occurs because of the VoIP codec characteristics. The good performance of G.729 and G.723.1 in this test is due to the high packet compression ratio, and the low transmission bit rate. In contrast, the low performance of G.711 codec is due to the built-in framework design which utilizes the A-law/μ-law algorithms in order to deliver precise speech transmission. The codec also produces higher bit rate compare to the others which is about 64 kbit/s. However, the codec is more sensitive towards packet losses due to poor packet loss interpolation. Moreover, all voice codecs exceed the acceptable range for the voice communication compared to LTE standard (<1%) [48]. This issue can be resolved using some techniques described in section 6.1.

Figure 5.5 (b), shows the one- way RTP delay versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 codec scheme has the highest delay. One way delay is happened while transmitting voice packet over WAN due to the complex algorithms that is used in G.723.1 codec. Generally, compressing voice signals reduces the bandwidth requirements at the expense of higher computational time which means more delay and thus degraded voice quality. Another reason for the delay is the time it takes to place bits on the transmitted link, where the higher the link speed, the less time it takes to place the bit. This called the serialization delay. Moreover, all voice codecs show acceptable RTP packet delay values (<100 ms) compared to LTE standard [48]. Figure 5.5 (c), shows RTP jitter versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 codec scheme experienced the highest as many small size packets are generated with variant inter-arrival time and hence the jitter between packets is significant. As the maximum packet size is increased to 120 bytes for G.711, the jitter is less significant as a smaller number of packets with less delay variations are generated. G.711 and G.729 are less than 10 ms and they are within the acceptable range of jitter values according to LTE standards [48]. While G.723.1 exceeds the limits. Figure 5.5 (d), shows MOS versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.711 has the best MOS value with nearly fixed value with the distance averaging around 4.1. The other two codec schemes i.e., G.729 and G.723.1 have lower MOS values averaging around 3.7 and 3.5 respectively. The reason is related to the factors that affect MOS value, which are voice codec, consecutive packet loss, and the range between the minimal and maximal detected delay.

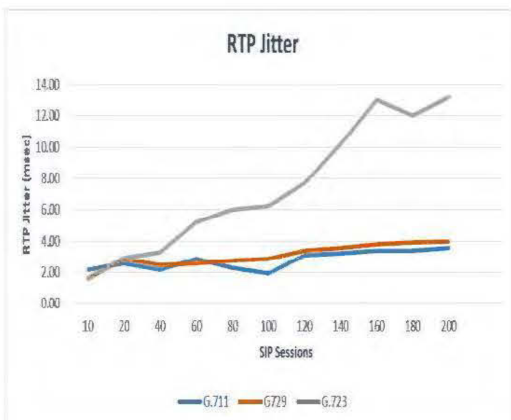
The fluctuations in the results obtained in figure 5.5 is related to the dynamicity of the network.



(a)



(b)



(c)

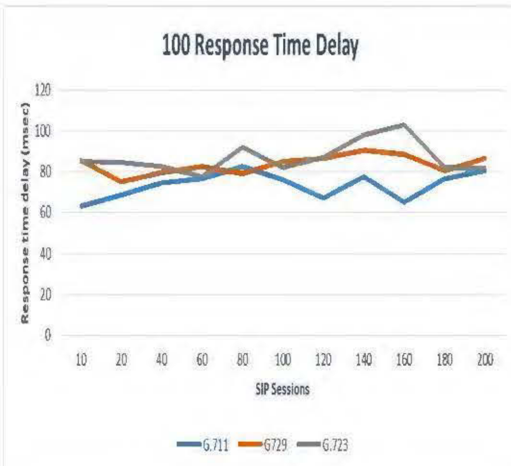
(d)

Figure 5.5 Scenario 1- Voice QoS Results

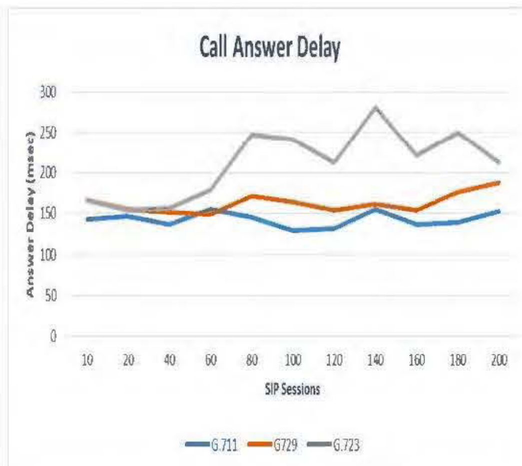
B- Signaling Performance Results Discussion

Below figures show the signaling performance results for the voice traffic transmitted over WAN network and passing through SIP proxy "TeleFinity IP PBX".

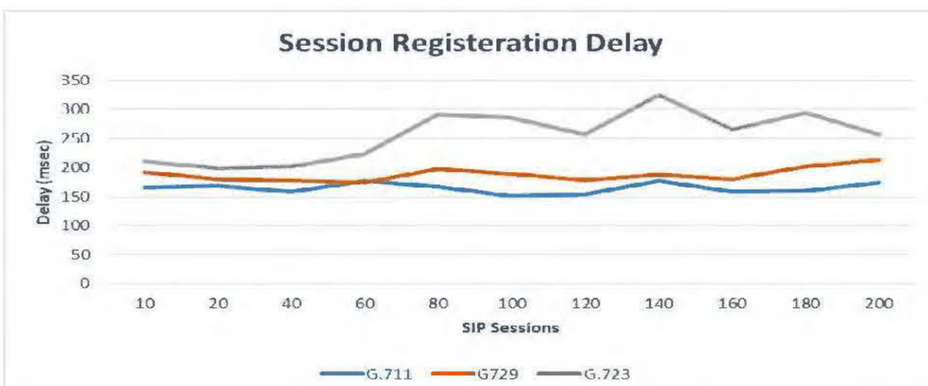
Figures 5.6 (a), (b) and (c) show 100 Response Time Delay, Call Answer Delay, and Session Registration Delay respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 has the worst performance compared to G.711 and G.729. This is because G.723.1 has the highest jitter and delay and this will highly affect these measured signaling timings.



(a)



(b)



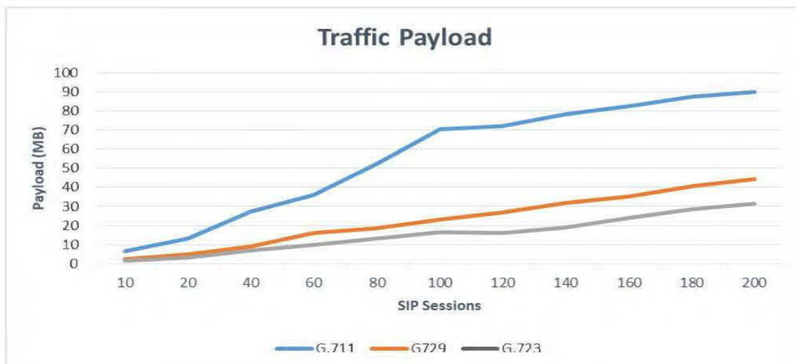
(c)

Figure 5.6 Scenario1. SIP Performance Results

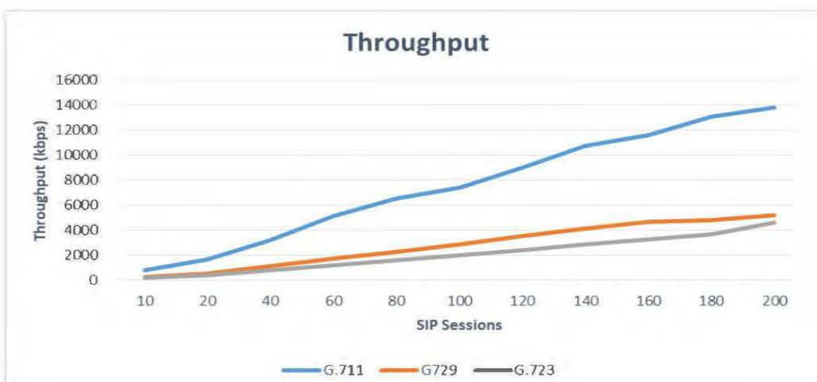
C- Network Utilization Results Discussion

Below figures show network utilization results for the voice traffic transmitted over WAN network and passing through SIP proxy "TeleFinity IP PBX".

Figures 5.7 (a) and (b) show the Traffic Payload and Throughput respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, and G.723.1) and transmitted over WAN network. It is shown that G.711 has the highest throughput and payload as it has the highest codec bit rate (64 Kbps) compared with G.729 (8Kbps) and G.723.1 (6.3 Kbps).



(a)



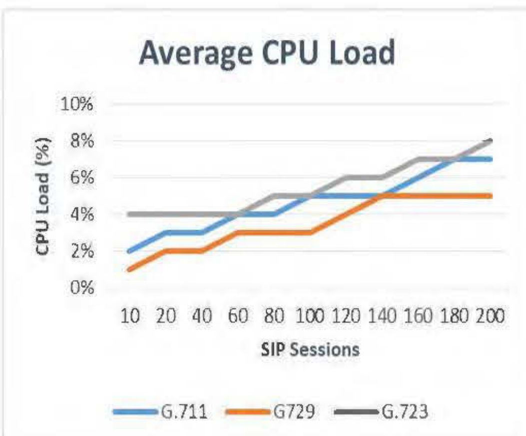
(b)

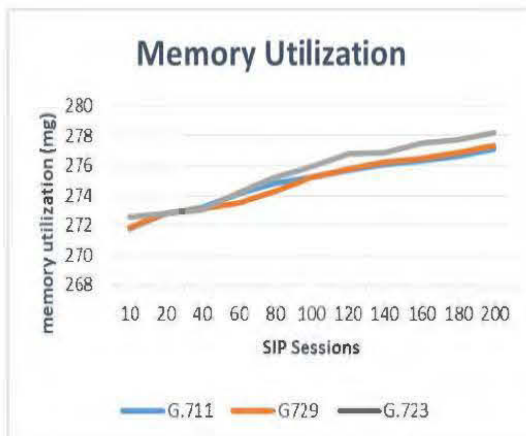
Figure 5.7 Scenario 1- Network Utilization Results

D- Hardware Utilization Results Discussion

Below figures show the hardware utilization results for the voice traffic transmitted over WAN network and passing through SIP proxy "TeleFinity IP PBX".

Figures 5.8 (a), (b) show Average CPU Load and Memory Utilization respectively for the SIP proxy server versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 consumes the highest processing power due to the complex algorithm used, where higher processing power needed for more complex algorithms. CPU and memory utilization are calculated for SIP proxy services that are used for controlling the SIP signaling and switching the RTP packets.





(a) (b)

Figure 5.8 Scenario1- Hardware Utilization Results

5.3.2 Scenario 2: WAN Topology- RTP Not Passing Through SIP Proxy

Figure 5.9 shows the testing topology used for Scenario2 within WAN network and the servers' locations. Signaling messages and RTP stream path also are shown with description in the figure. It is obvious that RTP stream passes between Part A and Part B directly.

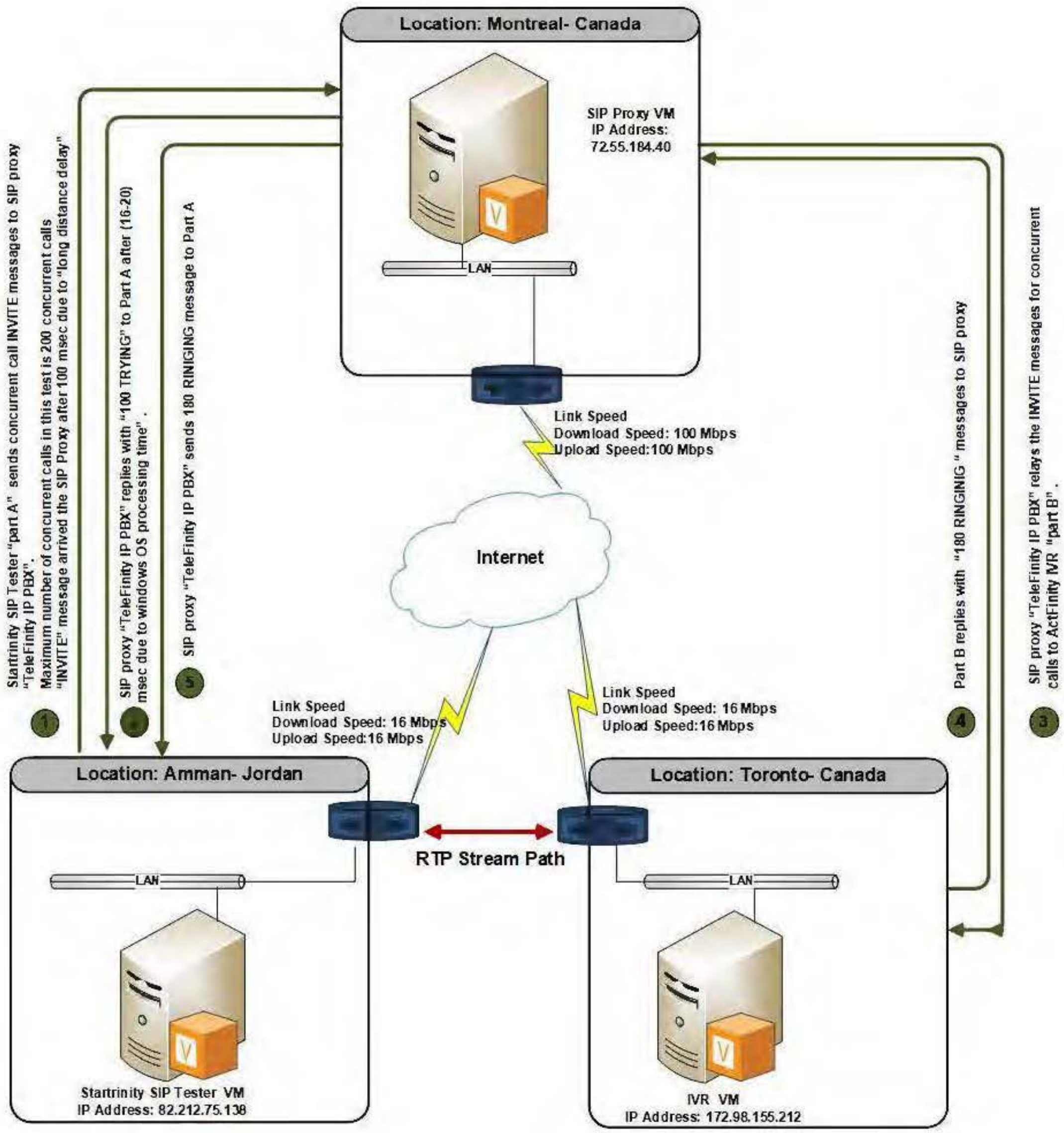


Figure 5.9 WAN Topology- RTP without Passing through SIP Proxy Scenario ©

Like scenario 1, SIP proxy VM, IVR VM, and SIP tester VM are hosted on three different physical servers at three different locations. The three servers are interconnected over WAN with different internet links speeds as shown in the figure.

SIP tester generates SIP concurrent calls starting from 10 concurrent calls and aggregated till 200 concurrent calls. In each test, the IVR system receives the generated SIP calls from SIP tester side and plays audio wave message with duration equals to one minute. The path of RTP or voice stream is not passing through the SIP proxy VM.

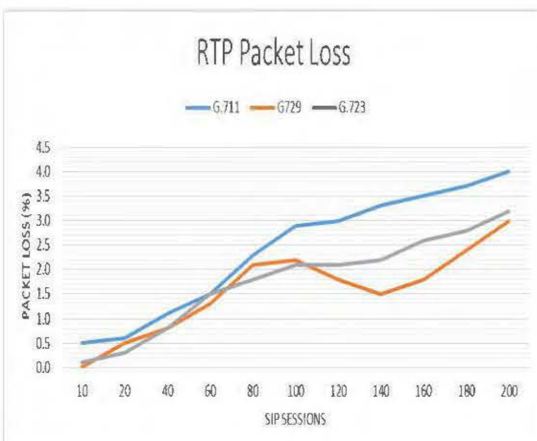
A- Voice QoS Results Discussion

Below figures show the voice QoS results for the voice traffic transmitted over WAN network and are not passing through SIP proxy "TeleFinity IP PBX".

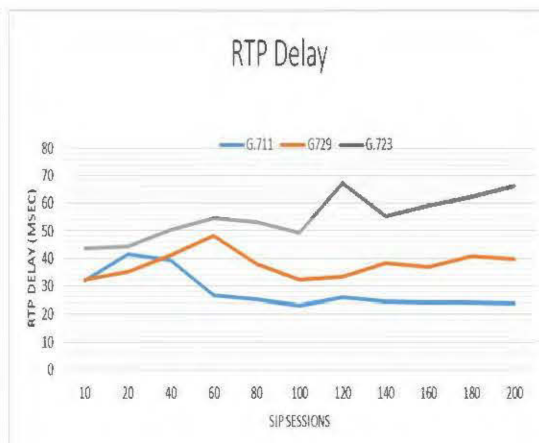
Figure 5.10 (a), shows the RTP packet loss percentages versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. Same as scenario1, the results show that G.711 codec scheme has the highest voice packet loss. In this scenario, the voice packets that are not passing the SIP proxy suffers from higher packet loss compared to scenario one. This is because that the packet loss over WAN depends mainly on the network architecture i.e. internet link BW and link delay, the routes of the transmitted voices packets, and the number of hops it passes to reach the destination. This is called the routing delay. In scenario 1, the routing delay is less than compared to scenario 2 as the packets pass through less number of hops to reach the destination. Moreover, all voice codecs exceed the acceptable range for voice communication compared to LTE standard (<1%) [42]. This issue can be resolved using some techniques described in section 6.1.

Figure 5.10 (b) and (c), shows the one- way RTP delay and RTP jitter versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) respectively. Like scenario 1, the results show that G.723.1 codec scheme experienced the highest delay and jitter. In this scenario, the voice packets that are not passing the SIP proxy suffers from higher delay and jitter compared to scenario one. This is because also the higher routing delay that affects the QoS metrics. All voice codecs show acceptable RTP packet delay values (<100 ms) compared to LTE standard [48]. As well, G.711 and G.729 are less than 10 ms and they are within the acceptable range of jitter values according to LTE standards [48]. While G.723.1 exceeds the limits.

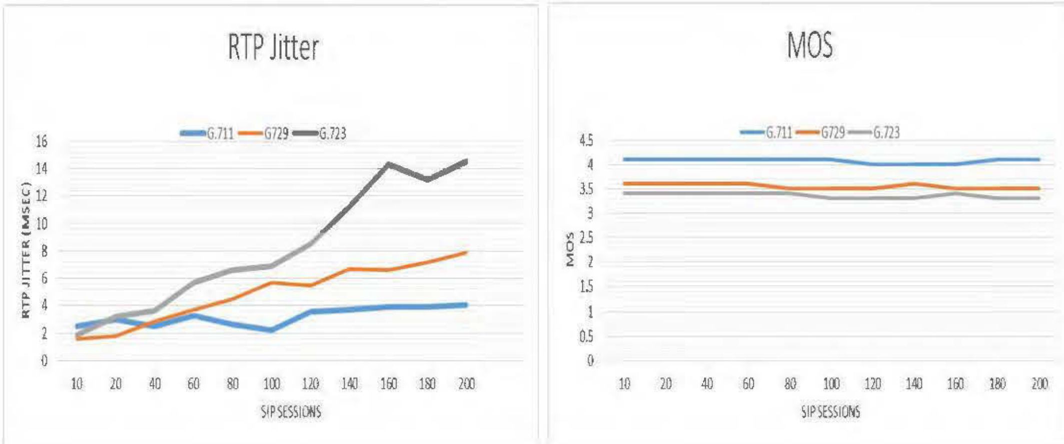
Figure 5.10 (d), shows MOS versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.711 has the best MOS with averaging value around 4.0. The other two codec schemes i.e., G.729 and G.723.1 have lower MOS values averaging around 3.5 and 3.3 respectively.



(a)



(b)



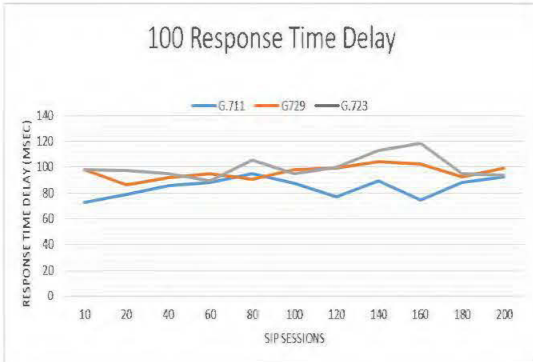
(c) (d)

Figure 5.10. Scenario 2- Voice QoS Results

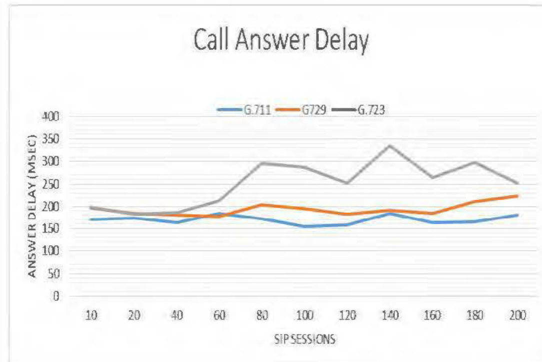
B- Signaling Performance Results Discussion

Below figures show the signaling performance results for the voice traffic transmitted over WAN network and are not passing through SIP proxy "TeleFinity IP PBX".

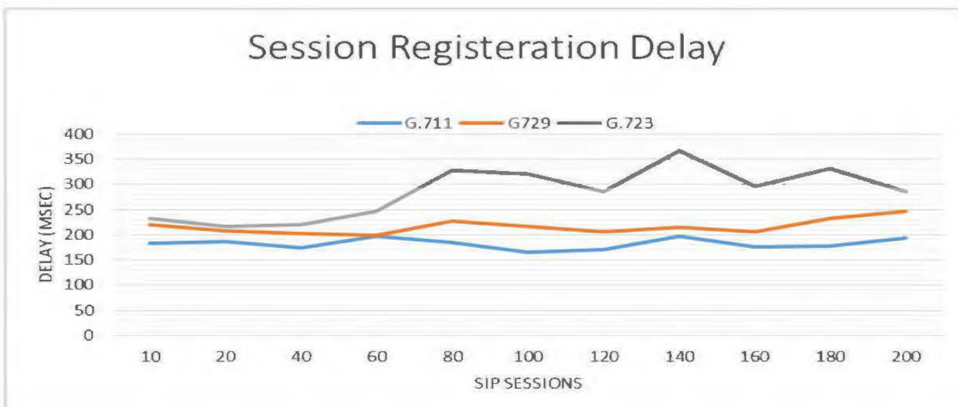
Figures 5.11 (a), (b) and (c) show 100 Response Time Delay, Call Answer Delay, and Session Registration Delay respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 has the worst performance compared to G.711 and G.729. The results obtained in this scenario have slight difference compared with scenario 1 that's because the SIP messages always pass through SIP proxy either if the RTP are passing or not.



(a)



(b)



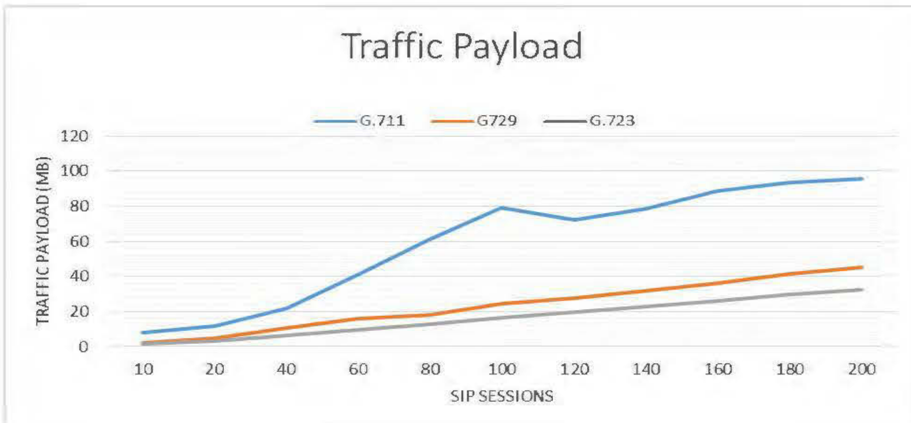
(c)

Figure 5.11. Scenario 2- Signaling Performance Results

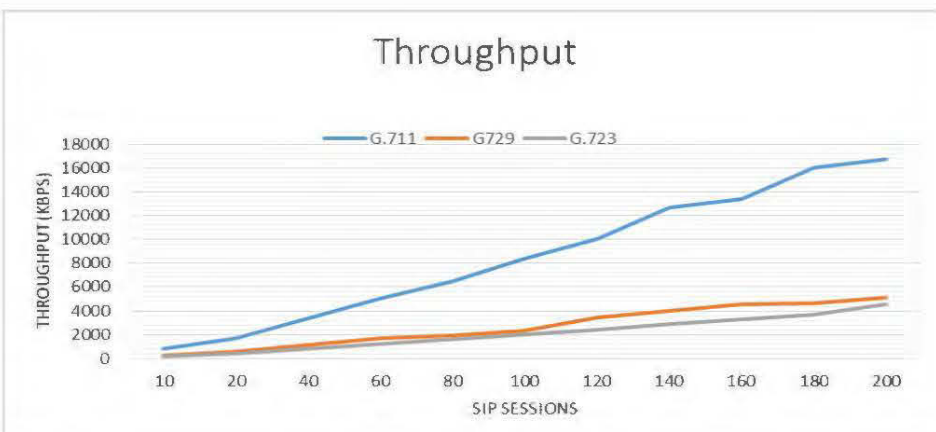
C- Network Utilization Results Discussion

Below figures show network utilization results for the voice traffic transmitted over WAN network and are not passing through SIP proxy "TeleFinity IP PBX".

Figures 5.12 (a) and (b) show the Traffic Payload and Throughput respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, and G.723.1) and transmitted over WAN network. It is shown that G.711 has the highest throughput and payload values that are more than the values measured in scenario 1. The reason is related to the high load experience by the SIP server in scenario 1, which requires higher CPU processing and thus it lowered the payload and throughput.



(a)



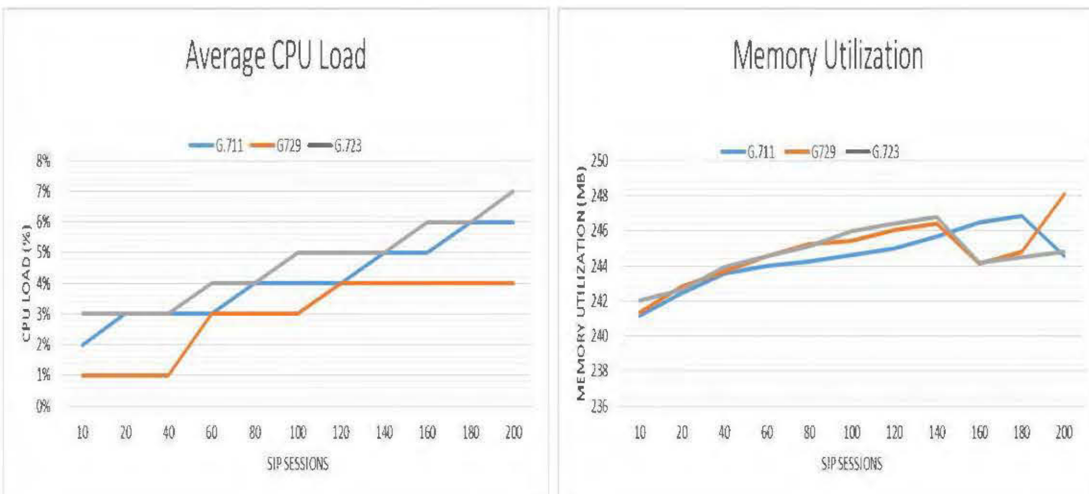
(b)

Figure 5.12. Scenario 2- Network Utilization Results

D- Hardware Utilization Results Discussion

Below figures show the hardware utilization results for the voice traffic transmitted over WAN network and are not passing through SIP proxy "TeleFinity IP PBX".

Figures 5.13 (a), (b) show Average CPU Load and Memory Utilization respectively for the SIP proxy server versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 consumes the highest processing power but it is less than the processing power used in scenario 1. This is because there is no need to transcode and process the RTP packets that are passed through the SIP proxy server.



(a) (b)

Figure 5.13. Scenario 2- Hardware Utilization Results

5.3.3 Scenario 3: LAN Topology- RTP Passing Through SIP Proxy

Figure 5.14 shows the testing topology used for Scenario3 within LAN network and the servers' locations. Signaling messages and RTP stream path also are shown with description in the figure.

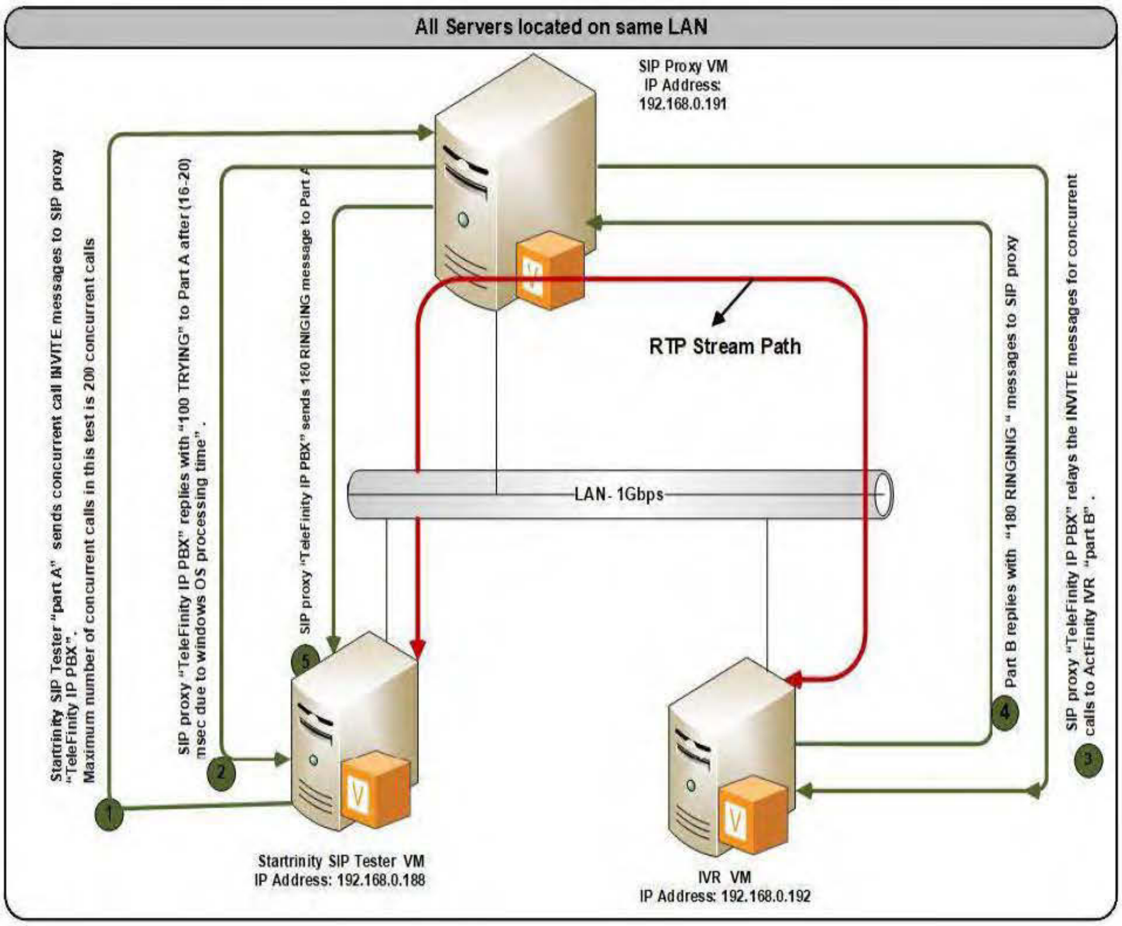


Figure 5.14. LAN Topology- RTP Passing through SIP Proxy Scenario ©

In this scenario, the SIP proxy VM, IVR VM, and SIP tester VM are hosted on three different physical servers on the same network with link speed 1 Gbps.

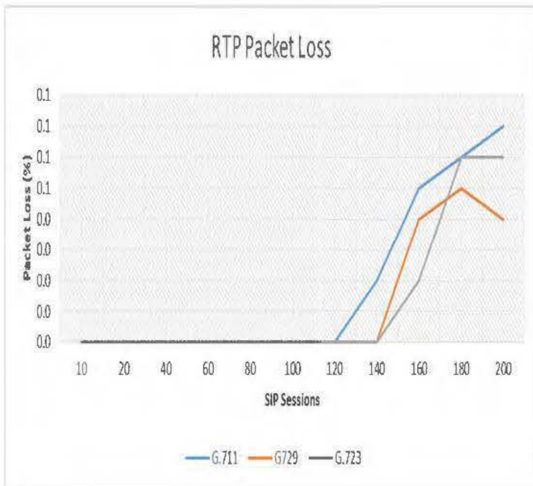
SIP tester generates SIP concurrent calls starting from 10 concurrent calls and aggregated till 200 concurrent calls. In each test, the IVR system receives the generated SIP calls from SIP tester side and plays audio wave message with duration equals to one minute. The path of RTP or voice stream is passing through the SIP proxy VM.

A- Voice QoS Results Discussion

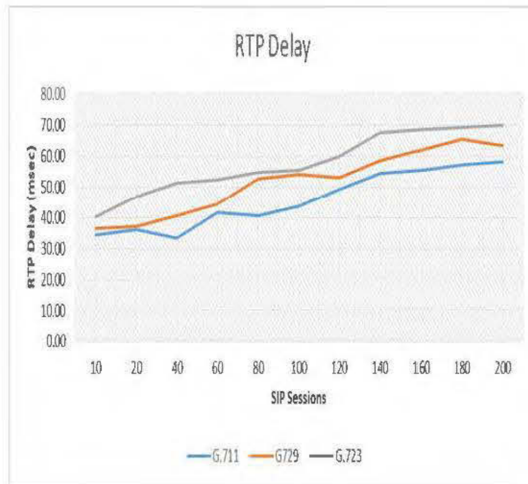
Below figures show the voice QoS results for the voice traffic transmitted over LAN network and passing through SIP proxy "TeleFinity IP PBX".

Figure 5.15 (a), shows the RTP packet loss percentages versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over LAN network. The results show that G.711 codec scheme has the highest voice packet loss. The packet loss percentages for all voice codecs are negligible (less than 0.1%) compared to the WAN scenario as the routing delay and serialization delay are negligible for the packet transmitted over LAN. Similar case for the results shown in figures 5.15 (b) and (c) where the voice packets have less delay and jitter compared to WAN topology and still G.723.1 experienced the highest delay and jitter for the same reasons mentioned in scenario 1. In this scenario, the voice packets that are not passing the SIP proxy suffers from higher packet loss compared to scenario one. Figure 5.15 (d), shows MOS versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over LAN network. The results show better readings compared to WAN scenario where G.711 has the best MOS with averaging value around 4.0. The other two codec schemes i.e., G.729 and G.723.1 have lower MOS values averaging around 4.1 and 3.9 respectively.

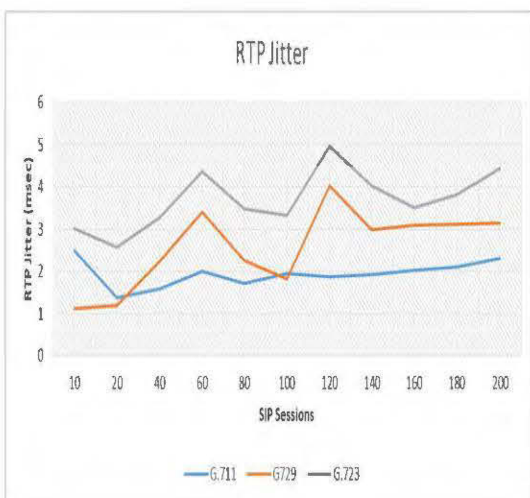
It is obvious that all codecs metrics values are within the acceptable range for voice communication according to LTE standards [48].



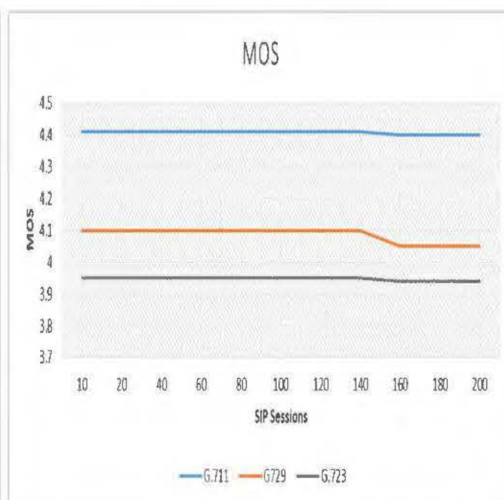
(a)



(b)



(c)



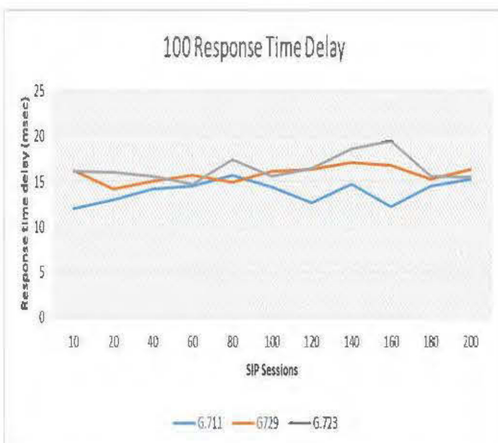
(d)

Figure 5.15. Scenario3- Voice QoS Results

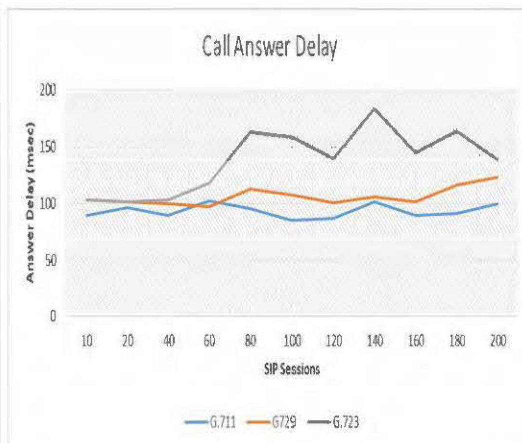
B- Signaling Performance Results Discussion

Below figures show the signaling performance results for the voice traffic transmitted over LAN network and passing through SIP proxy "TeleFinity IP PBX".

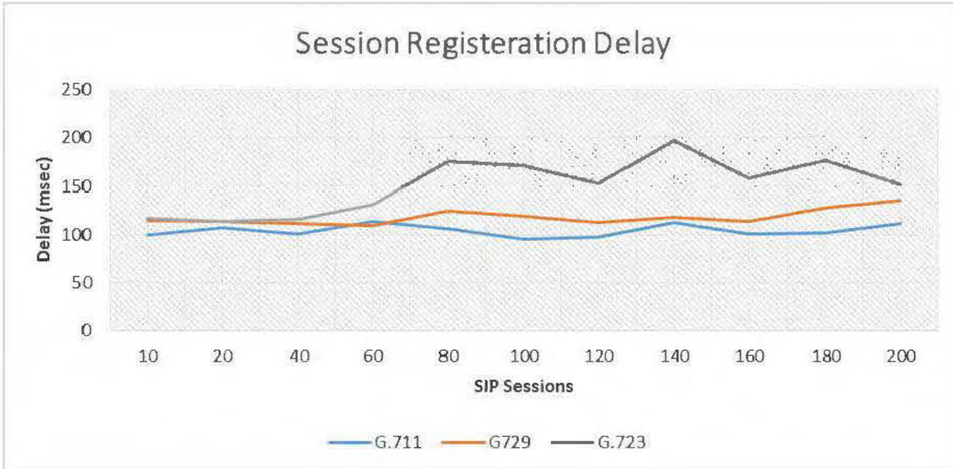
Figures 5.16 (a), (b) and (c) show 100 Response Time Delay, Call Answer Delay, and Session Registration Delay respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over WAN network. The results show that G.723.1 has the worst performance compared to G.711 and G.729. The results obtained in this scenario have less delay compared to the WAN scenario. That's because the delay experienced by the packets in LAN is much less than the WAN where the sterilization and routing delays are eliminated.



(a)



(b)

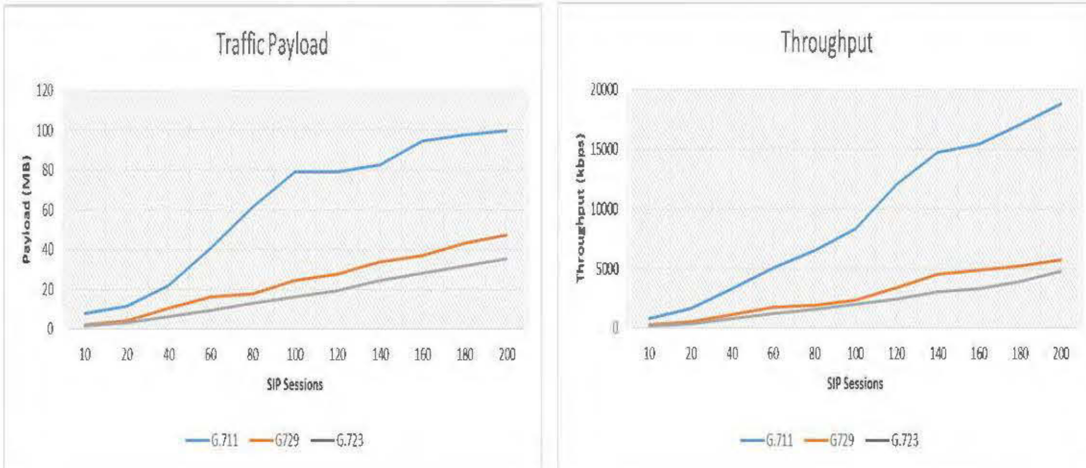


(c)

Figure 5.16. Scenario3- Signaling Performance Results

C- Network Utilization Results Discussion

Figures 5.17 (a) and (b) show the Traffic Payload and Throughput respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, and G.723.1) and transmitted over WAN network. It is shown that G.711 has the highest throughput and payload values that are more than the values measured in WAN scenario. The reason is related to the high load experience by the SIP server in WAN to select optimal routes for the packets and this requires higher CPU processing and thus it lowered the payload and throughput.



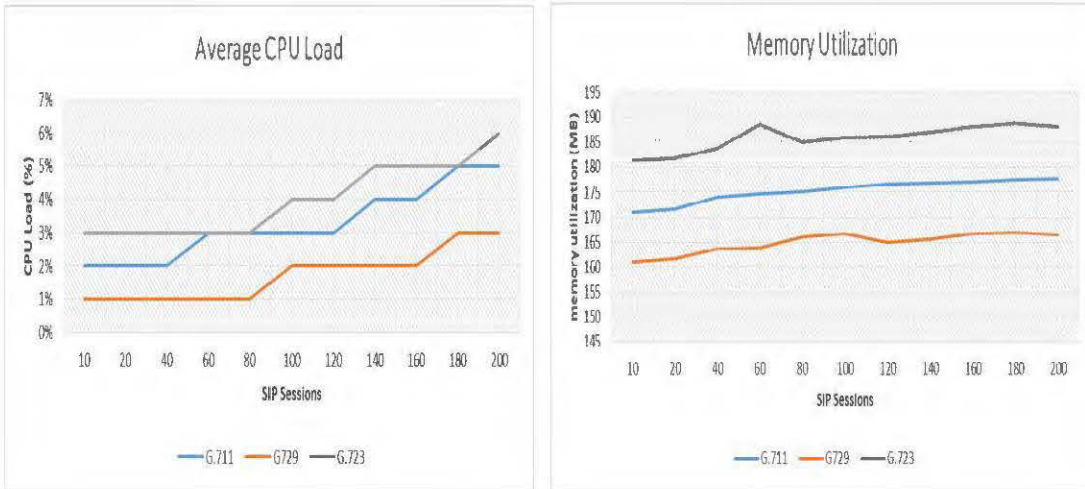
(a) (b)

Figure 5.17. Scenario 3- Network Utilization Results

D- Hardware Utilization Results Discussion

Below figures show the hardware utilization results for the voice traffic transmitted over LAN network and passing through SIP proxy "TeleFinity IP PBX".

Figures 5.18 (a), (b) show the average CPU load and the memory utilization respectively for the SIP proxy server versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over LAN network. The results show that G.723.1 consumes the highest processing power but it is less than the processing power used in scenario 1 and 2 (WAN scenario). This is because of the different network architecture of LAN and WAN. In WAN scenario, the SIP proxy server needs additional processing power to select the optimal route for the RTP packets and this case is not found in the LAN scenario.



(a) (b)

Figure 5.18. Scenario3- Hardware Utilization Results

5.3.4 Scenario 4: LAN Topology- RTP Not Passing Through SIP Proxy

Below figure shows the testing topology used for Scenario4 within LAN network and the servers' locations. Signaling messages and RTP stream path also are shown with description in the figure. It is obvious that RTP stream passes between Part A and Part B directly.

In this scenario, the SIP proxy VM, IVR VM, and SIP tester VM are hosted on three different physical servers on the same network with link speed 1 Gbps.

SIP tester generates SIP concurrent calls starting from 10 concurrent calls and aggregated till 200 concurrent calls. In each test, the IVR system receives the generated SIP calls from SIP tester side and plays audio wave message with duration equals to one minute. The path of RTP or voice stream is not passing through the SIP proxy VM.

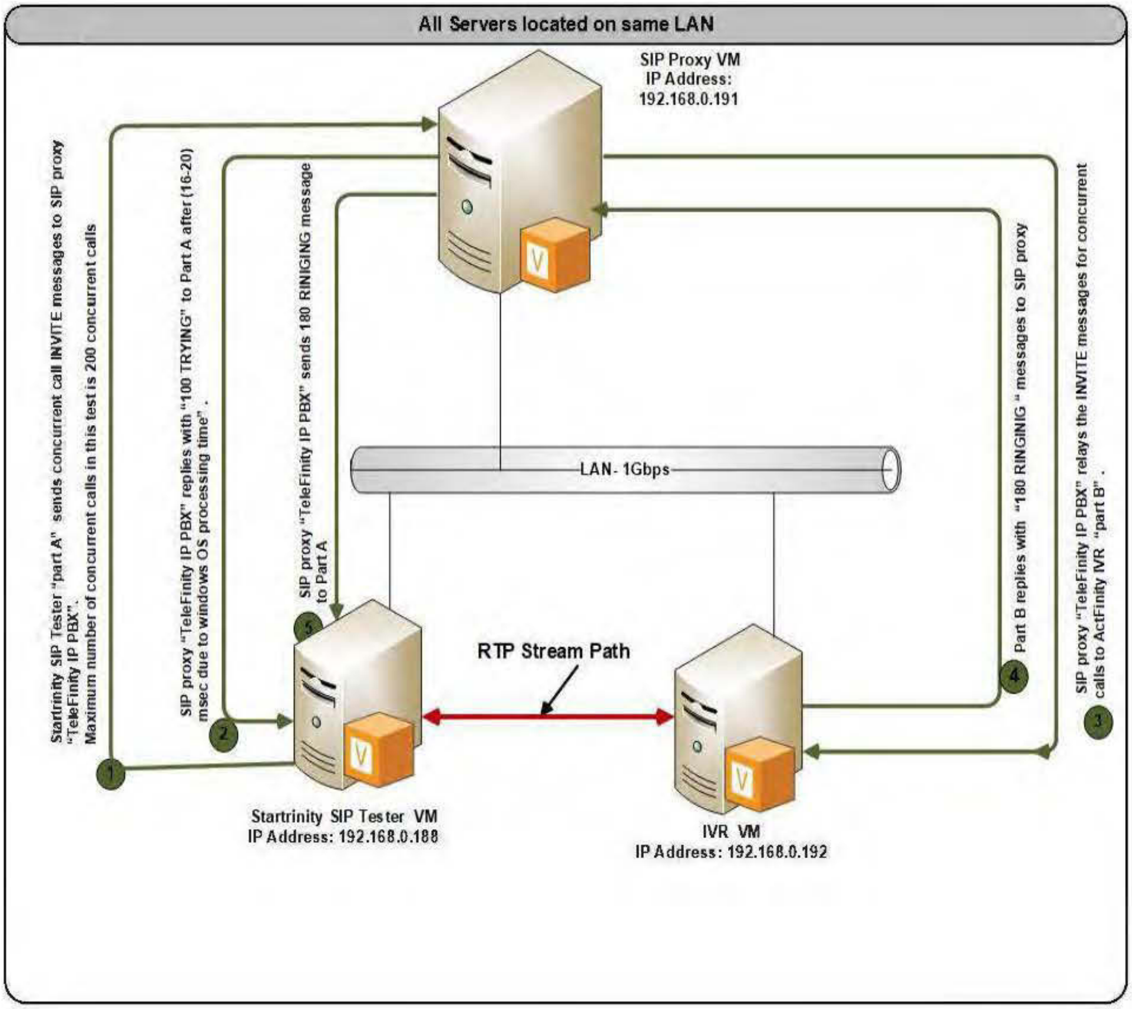
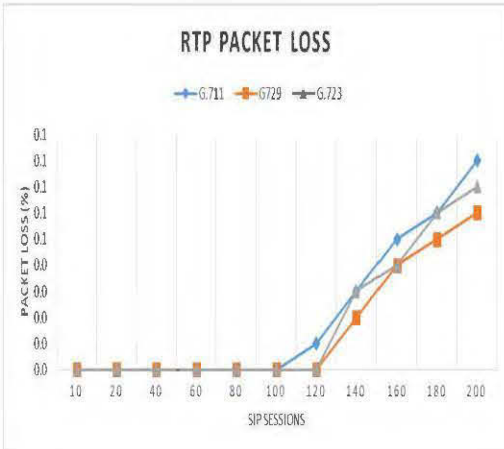


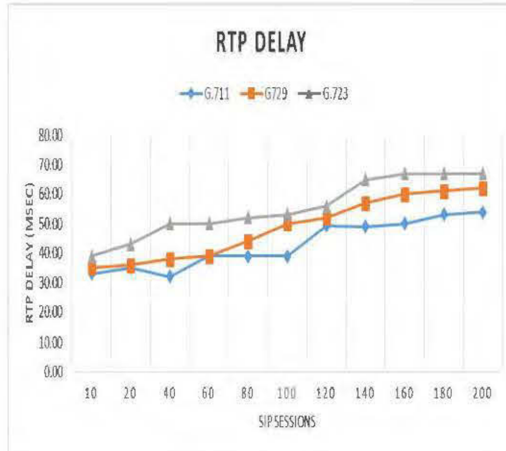
Figure 5.19. LAN Topology- RTP Not Passing through SIP Proxy Scenario ©

A- Voice QoS Results Discussion

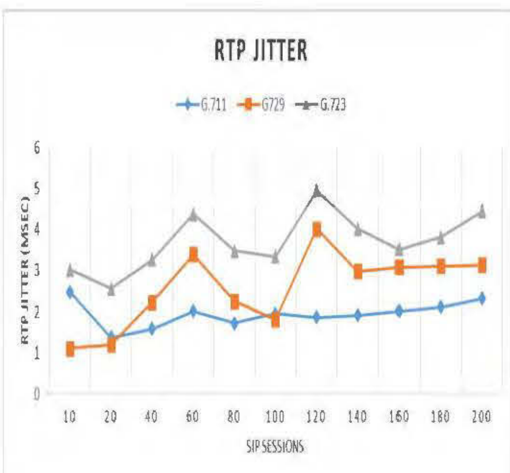
Below figures show the voice QoS results for the voice traffic transmitted over LAN network and are not passing through SIP proxy "TeleFinity IP PBX". QoS metrics shows similar results to the ones get in scenario 3 but with slight less RTP delay as the transcoding delay is eliminated in scenario4. As a result, there is no significant impact for passing the RTP packets through the SIP proxy on the QoS metrics.

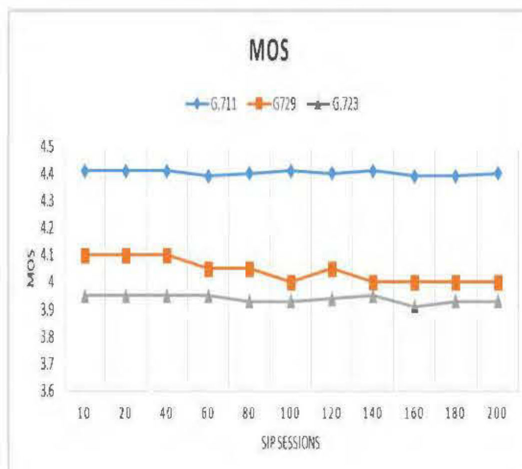
Moreover, all codecs metrics values are within the acceptable range for voice communication according to LTE standards [48].





(a) (b)





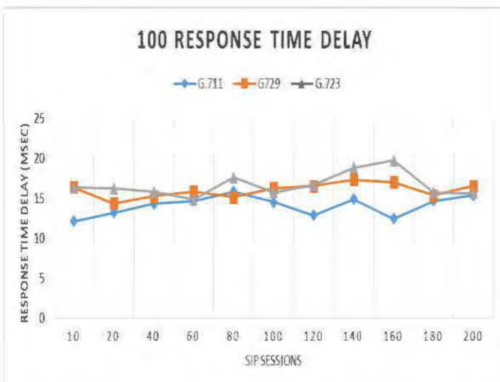
(c) (d)

Figure 5.20. Scenario4- Voice QoS Results

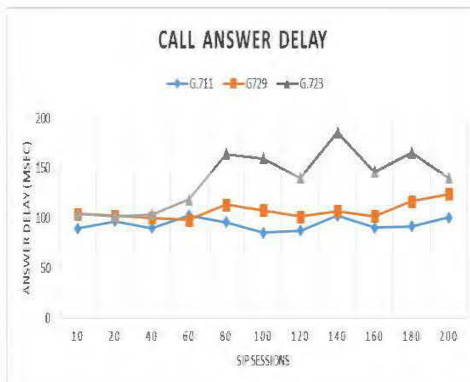
B- Signaling Performance Results Discussion

Below figures show the signaling performance results for the voice traffic transmitted over LAN network and are not passing through SIP proxy "TeleFinity IP PBX".

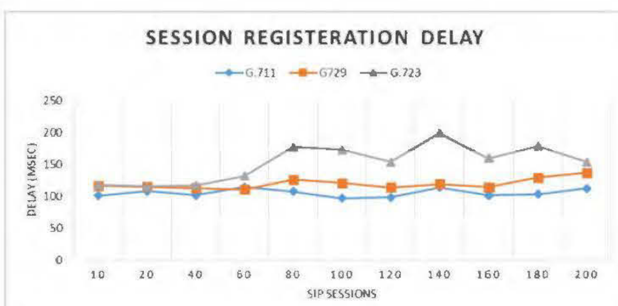
Figures 5.21 (a), (b) and (c) show 100 Response Time Delay, Call Answer Delay, and Session Registration Delay respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over LAN network. The results show that G.723.1 has the worst performance compared to G.711 and G.729. The results obtained in this scenario have slight difference compared with scenario 3 that's because the SIP messages always pass through SIP proxy either if the RTP are passing or not.



(a)



(b)



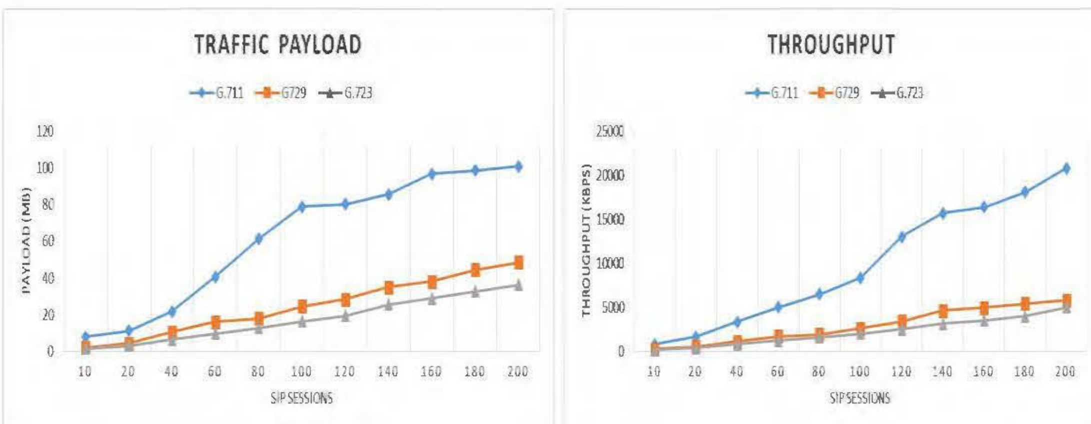
(c)

Figure 5.21. Scenario 4- Signaling Performance Results

C- Network Utilization Results Discussion

Below figures show network utilization results for the voice traffic transmitted over LAN network and are not passing through SIP proxy "TeleFinity IP PBX".

Figures 5.22 (a) and (b) show the Traffic Payload and Throughput respectively versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, and G.723.1) and transmitted over LAN network. It is shown that G.711 has the highest throughput and payload values that are more than the values measured in scenario 3. The reason is related to the high load experience by the SIP server in scenario 3, which requires higher CPU processing and thus it lowered the payload and throughput.



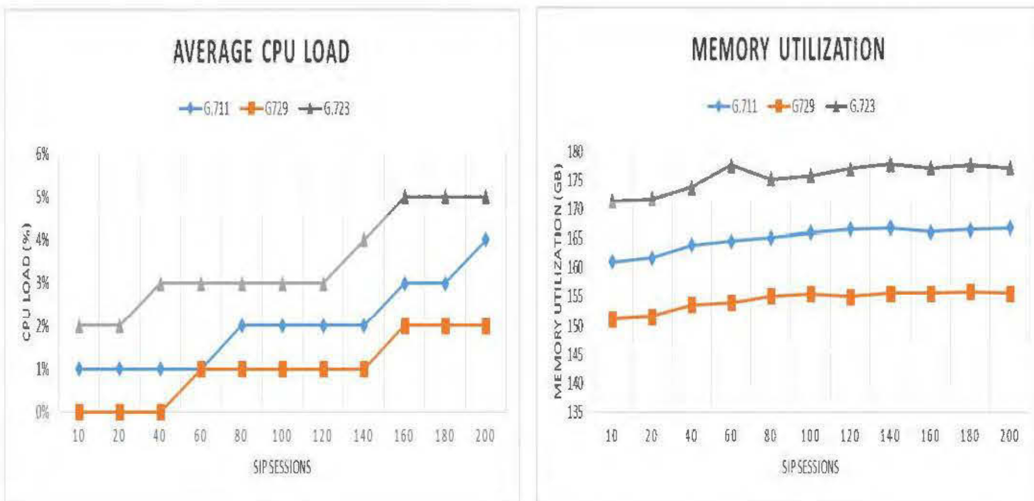
(a) (b)

Figure 5.22. Scenario 4- Network Utilization Results

D- Hardware Utilization Results Discussion

Below figures show the hardware utilization results for the voice traffic transmitted over LAN network and are not passing through SIP proxy "TeleFinity IP PBX".

Figures 5.23 (a), (b) show Average CPU Load and Memory Utilization respectively for the SIP proxy server versus the number of concurrent SIP calls with different voice codecs (G.711, G.729, G.723.1) and transmitted over LAN network. The results show that G.723.1 consumes the highest processing power but it is less than the processing power used in scenario 3. This is because there is no need to transcode and process the RTP packets that are passed through the SIP proxy server.



(a) (b)

Figure 5.23. Scenario4- Hardware Utilization Results

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

The objective of this thesis is to propose a novel integrated SIP Proxy with Mobile Network Operator to overcome the scalability and the coverage expansion challenges. The GMSC entity in the MNO platform can integrate with SIP proxy application (TeleFinity IP PBX) using a SIP trunk method. The new mobile subscribers are registered to the SIP proxy application that manages all users' requests and information by exchanging the signaling messages with different MNO entities like the IN unit, and HSS unit. Several metrics related to the voice QoS, SIP signaling performance, hardware and the network utilization of the virtualized SIP proxy have been evaluated and analyzed on different voice codecs (G.711, G.729, and G.723.1). It has been noticed that the G.711 voice codec demonstrates the best performance among the other two voice codecs. This is related to the G.711 properties that are presented with the lowest one-way delay and jitter values when compared with G.729 and G.723.1. It has the best MOS value among G.729, and G.723.1codecs for the LAN and WAN topology as well. G.711 requires less computational power (utilized CPU and Memory) of the SIP server compared with the other two codecs as it uses simple codec algorithm. Moreover, the SIP proxy that uses G.711 codec scheme has the highest calls capacity compared with the other two codecs. And there is no need for the transcoding process when using G.711 and this reduces the processing delay on the SIP server consequently.

Also, G.711 can handle the fax transactions using the pass- through protocol only. It can be handled by any mobile device phone and there is no need for transcoding. This will eliminate the need to pay additional license fees when using G.711 codec. While G.729 and G.723.1 codecs needed additional licenses fees to implement the codec on the SIP server and on GMSC. And finally, G.711 has the higher BW efficiency due to the low overhead used during the transmission.

Below are some suggestions to enhance the performance of SIP proxy that uses G.711 codec over WAN topology in order to meet LTE QoS standard:

1- Decrease the packet loss percentage by lowering the bandwidth of G.711 calls using Compressed Real-Time Transport Protocol (CRTP) method.

2- Decrease the packets latency using multiple SIP proxy servers' architecture distributed over different location and interconnected by very high-speed links. Each SIP server relays the RTP packets near to its location. This method will minimize the "long distance delay".

3- Minimize the computational power needed by the SIP proxy by configuring the RTP to be exchanged directly between SIP endpoint and no need to pass through the proxy.

4- Enable voice activity detection (VAD).

5- Use echo canceller.

6- Use adaptive jitter buffer.

6.2 Future Work

This thesis focused primarily on testing the performance of the SIP proxy software proposed in the solution to enhance the scalability of MNOs. The performance evaluation was made in four main areas: voice QoS, signaling performance, network and hardware utilization of SIP proxy under virtualization environment. However, there are a number of related topics that need investigation; one interesting topic may be security issues of VoIP calls handled by the proposed solution. Since VoIP uses packet switching, the same kind of vulnerabilities faced by packet networks are valid for VoIP such as Denial of Service (DoS) Attacks. Another topic is the interoperability which is a major problem. VoIP applications which implement different protocols cannot talk to each other and it needs transcoding system. The impact of such systems on the proposed solution performance should be studied. Moreover, even VoIP applications implementing the same protocol such as H.323 may not be able to interoperate because of proprietary differences. Finally, the integration between the Wi-Fi-calling feature and the proposed solution proposed is an interesting topic for further study. As it eliminates the need to install mobile application on end user phone where the user can be registered on SIP proxy directly.

REFERENCES

[1] M. Agiwal, A. Roy and N. Saxena, "Next Generation 5G Wireless Networks: A Comprehensive Survey," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617-1655, thirdquarter 2016.

[2] 3GPP, "3gpp ts 23.401 technical specification group services and system aspects; general packet radio service (gprs) enhancements for evolved universal terrestrial radio access network (e-utran) access, (release 10)." http://www.3gpp.org/dynareport/23401.htm, 2011.

[3] T. Dargahi, A. Caponi, M. Ambrosin, G. Bianchi and M. Conti, "A Survey on the Security of Stateful SDN Data Planes," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1701-1725, thirdquarter 2017.

[4] M. Chiosi *et al.*, ''Network functions virtualisation—Introductory white paper,'' presented at the SDN OpenFlow World Congr., Oct. 2012.

[5] B. Naudts, M. Kind, S. Verbrugge, D. Colle, and M. Pickavet, "How can a mobile service provider reduce costs with software-defined networking?," *International Journal of Network Management*, vol. 26, no. 1, pp. 56–72, 2016.

[6] E. Hernandez-Valencia, S. Izzo, and B. Polonsky, "How will NFV/SDN transform service provider opex?," *IEEE Network*, vol. 29, no. 3, pp. 60–67, 2015.

[7] Basta, A., Kellerer, W., Hoffmann, M., Hoffmann, K., & Schmidt, E. D. (2013). A virtual sdn-enabled lte epc architecture: A case study for s-/p-gateways functions. In *IEEE workshop on software defined network for future networks and services (SDN4FNS)* (pp. 1–7).

[8] Basta, A., Kellerer, W., Hoffmann, M., Morper, H. J., & Hoffmann, K. (2014). Applying nfv and sdn to lte mobile core gateways; the functions placement

problem. In *ACM workshop on all things cellular: Operators, applications, and challenges (AllThingsCellular)* (pp. 33–38).

[9] A. Basta, A. Blenk, M. Hoffmann, H. J. Morper, K. Hoffmann, and W. Kellerer, "SDN and NFV dynamic operation of lte epc gateways for time-varying traffic patterns," in *6th International Conference on Mobile Networks and Management MONAMI*, pp. 63–76, Springer, 2015.

[10] MEVICO project. (2012). D2.2: Architectural epc extensions for supporting heterogenerous mobility schemes. Technical report. http://www.mevico.org/D22.pdf.

[11] 80. Heinonen, J., Partti, T., Kallio, M., Lappanlainen, K., Flinck, H., & Hillo, J. (2014). Dynamic tunnel switching for sdn-based cellular core networks. In *ACM workshop on all things cellular: Operators, applications, and challenges (AllThingsCellular)* (pp. 27–32).

[12] F. Z. Yousaf, J. Lessmann, P. Loureiro and S. Schmid, "SoftEPC — Dynamic instantiation of mobile core network entities for efficient resource utilization," *2013 IEEE International Conference on Communications (ICC)*, Budapest, 2013, pp. 3602-3606.

[13] V. G. Nguyen, A. Brunstrom, K. J. Grinnemo and J. Taheri, "SDN/NFV-Based Mobile Packet Core Network Architectures: A Survey," in *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1567-1602, thirdquarter 2017.

[14] T. Taleb, "Toward carrier cloud: Potential, challenges, and solutions," in *IEEE Wireless Communications*, vol. 21, no. 3, pp. 80-91, June 2014.

[15] Z. A. Qazi, P. K. Penumarthi, V. Sekar, V. Gopalakrishnan, K. Joshi, and S. R. Das, "KLEIN: a minimally disruptive design for an elastic cellular core," in *2016 Symposium on SDN Research (SOSR)*, pp. 1–12, ACM, 2016.

[16] X. Jin, L. E. Li, L. Vanbever, and J. Rexford, "SoftCell: Scalable and flexible cellular core network architecture," in *Proceedings of the ninth ACM*

conference on Emerging networking experiments and technologies (CoNEXT), pp. 163–174, ACM, 2013.

[17] M. Moradi, W. Wu, L. E. Li, and Z. M. Mao, "SoftMoW: recursive and reconfigurable cellular WAN architecture," in *Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies (CoNEXT)*, pp. 377–390, ACM, 2014.

[18] H. Baba, M. Matsumoto and K. Noritake, "Lightweight virtualized evolved packet core architecture for future mobile communication," *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA, 2015, pp. 1811-1816.

[19] T. Taleb *et al.*, "EASE: EPC as a service to ease mobile core network deployment over cloud," in *IEEE Network*, vol. 29, no. 2, pp. 78-88, March-April 2015.

[20] A. Jain, Sadagopan N S, S. K. Lohani and M. Vutukuru, "A comparison of SDN and NFV for re-designing the LTE Packet Core," *2016 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*, Palo Alto, CA, 2016, pp. 74-80.

[21] H. Hawilo, A. Shami, M. Mirahmadi and R. Asal, "NFV: state of the art, challenges, and implementation in next generation mobile networks (vEPC)," in *IEEE Network*, vol. 28, no. 6, pp. 18-26, Nov.-Dec. 2014.

[22] W. Kiess, X. An and S. Beker, "Software-as-a-Service for the Virtualization of Mobile Network Gateways," *2015 IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, 2015, pp. 1-6.

[23] ETSI. (2013). Network function virtualization (NFV): Use cases. ETSI group specification. http://www.etsi.org/deliver/etsi_gs/NFV/001_099/001/01.01.01_60/gs_NFV001v010101p.pdf. Accessed August 2017ETSI, "European Telecommunications Standards Institute, Industry Specification Groups (ISG) - NFV,"

http://www.etsi. org/technologies-clusters/technologies/nfv, 2015, Accessed: August 28, 2017.

[24] Wireless Pers Commun. (2016). SDN and Virtualization-Based LTE Mobile Network Architectures: A Comprehensive Survey.https://link.springer.com/content/pdf/10.1007/s11277-015-2997-7.pdf. Accessed September 2017.

[25] F. Lu, H. Pan, X. Lei, X. Liao and H. Jin, "A Virtualization-Based Cloud Infrastructure for IMS Core Network," *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, Bristol, 2013, pp. 25-32.

[26] M. Ito, K. Nakauchi, Y. Shoji and Y. Kitatsuji, "Mechanism of Reducing Signaling Processing Load in EPC/IMS Using Service-Specific Network Virtualization," *2014 6th International Conference on New Technologies, Mobility and Security (NTMS)*, Dubai, 2014, pp. 1-5.

[27] R. Guerzoni, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges and Call for Action. Introductory whitepaper," *in SDN and OpenFlow World Congress*, June 2012.

[28] ETSI, "European Telecommunications Standards Institute, Industry Specification Groups (ISG) - NFV," http://www.etsi. org/technologies-clusters/technologies/nfv, 2015, Accessed: August 28, 2017.

[29] ETSI Industry Specification Group (ISG) NFV, "ETSI Group Specifications on Network Function Virtualization. 1st Phase Documents,"http://docbox.etsi.org/ISG/NFV/Open/Published/, September 2017.

[30] R. Mijumbi, J. Serrat, J. L. Gorricho, N. Bouten, F. De Turck and R. Boutaba, "Network Function Virtualization: State-of-the-Art and Research Challenges," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 236-262, Firstquarter 2016.

[31] Papidas ,A. G. (2017). *Network Functions Virtualization (NFV) for Mobile Networks* (master's thesis). Athens University of , Athina, Greece.

[32] ETSI GS NFV 002 V1.2.1 (2014-12), "Network Functions Virtualization (NFV), Architectural Framework".

[33] ETSI Industry Specification Group (ISG) NFV. "ETSI GS NFV 002 V1.2.1: Network Functions Virtualisation (NFV); Architectural Framework," http://www.etsi.org/deliver/etsi_gs/NFV/001_099/002/01.02.01_60/gs_NFV002v010201p.pdf. Accesed in Spetember 2017.

[34] http://tele-finity.com/

[35] http://startrinity.com/VoIP/SipTester/SipTester.aspx#features

[36] http://tele-finity.com/Products/ActFinity-IVR-and-Voicemail

[37] Ravi Shankar Ramakrishnan1 and P. Vinod kumar, "Performance Analysis of Different Codecs in VoIP Using SIP", *Mobile and Pervasive Computing* (CoMPC–2008).

[38] L. Roychoudhuri, E. Al-Shaer, H. Hamed and G.B. Brewster. "Brewster audio transmission over the internet: experiments and observations" *IEEE, International Conference on Communications*, June 2003, vol.1, pp. 552–556.

[39] W. Lee, K. Patel & M. Pedram. "Dynamic thermal management for MPEG-2 decoding" *ACM, Proceedings of International Symposium on Low Power Electronics and Design*, October 2006, pp. 316-321.

[40] H. Cui, K. Tang and T. Cheng. "Audio as a support to low Bit-rate multimedia communication" *IEEE, International Conference on Communication Technology*, August 2002, vol. 1, pp. 544-547.

[41] J.K. Muppala, T. Banchherdvanich & A. Tyagi. "VoIP performance on differentiated services enabled network" *IEEE, International Conference on Networks*, 2000, pp. 419-423.

[42] Ravi Shankar Ramakrishnan1 and P. Vinod kumar, "Performance Analysis of Different Codecs in VoIP Using SIP", *Mobile and Pervasive Computing* (CoMPC–2008).

[43] Keagy, Scott, " *Integrating Voice and Data Networks*", pp.208, Cisco Press, 2000

[44] Janssen, Jan and others,"*Delay Bounds for Voice Over IP Calls Transported over Satellite Access Networks*", Mobile Networks and Applications 2002, p. 80.

[45] http://www.cisco.com/c/en/us/support/docs/voice/voice-quality/18902-jitter-packet-voice.html. Accesed in September 2017.

[46] International Telecommunications Union, *Subjective Performance Assessment of Telephone-band and Wideband Digital Codecs ITU-T Recommendation* P.830, 1996

[47] Cisco." *Voice Over IP - Per Call Bandwidth Consumption* ".https://www.cisco.com/c/en/us/support/docs/voice/voice-quality/7934-bwidth-consume.html. Accessed in August,2017.

[48] 3GPP TS 23.203. "*Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 10)*." Accessed in August,2017.

APPENDIX A

IVR .VBS SCRIPT

```
Dim m_ChNum

Dim m_CallerID

Dim m_CallerName

Dim m_CalledID

Dim m_CalledName

Dim m_ContactID

Dim m_HandleNewCall

Dim m_PlayFile2

Dim m_Hangup3

Function HandleNewCall (ChNum, CallerID, CallerName, CalledID, CalledName,
ContactID, Parm1, Parm2, Parm3)

        m_ChNum = ChNum

        m_CallerID = CallerID

        m_CallerName = CallerName

        m_CalledID = CalledID

        m_CalledName = CalledName

        m_ContactID = ContactID
```

```
        Call pIVR.InsertDBCallDetails
(0,"HandleNewCall","","NewCall",m_FileVersion)

        pIVR.Answer()

        If (1<>1) Then

                If err.number = 0 Then

                Else

                End If

        Else

                PlayFile2()

        End If

End Function

Function PlayFile2()

        Call pIVR.InsertDBCallDetails
(1,"PlayFile2","PlayFile2","PlayFile",m_FileVersion)

        Dim m_Result

        m_Result = pIVR.PlayFile("C:\Program Files
(x86)\ActFinity\Prompts\SysGreeting.wav" ,"*",1, 0)

        PlayFile2 = m_Result

        m_PlayFile2 = m_Result

        Select Case m_Result

        Case 0 ' Ok

                Hangup3()

        Case 5 ' Term Digits
```

```
            Hangup3()
        Case 4 ' Max Digits
            Hangup3()
        Case 3 ' Timeout
            Hangup3()
        Case -1 ' Error
            Hangup3()
        End Select
End Function
Function Hangup3()
        Call pIVR.InsertDBCallDetails
(2,"Hangup3","Hangup3","Hangup",m_FileVersion)
        Dim m_Result
        m_Result = pIVR.PlayFile("C:\Program Files
(x86)\ActFinity\Prompts\GoodBye.wav")
        m_Result = pIVR.Disconnect()
        Hangup3 = m_Result
        m_Hangup3 = m_Result
End Function
Function CheckFile(FilePath)
        Dim fso, msg
        fso = CreateObject("Scripting.FileSystemObject")
        If (fso.FileExists(FilePath)) Then
            CheckFile = True
```

```
        Else
                CheckFile = False
        End If
End Function
Function IIf(Expression, TruePart, FalsePart)
On Error Resume Next
        If Expression Then
                IIf = TruePart
        Else
                IIf = FalsePart
        End If
End Function
```