ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

SCHOOL OF INFORMATION SCIENCE

BANDWIDTH OPTIMIZATION FRAMEWORK FOR WIDE AREA NETWORK: THE CASE OF ADDIS ABABA UNIVERSITY

MINILIK ASSEFA

OCTOBER, 2016

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

SCHOOL OF INFORMATION SCIENCE

BANDWIDTH OPTIMIZATION FRAMEWORK FOR WIDE AREA NETWORK: THE CASE OF ADDIS ABABA UNIVERSITY

A THESIS SUBMITTED TO THE COLLEGE OF NATURAL SCIENCES OF ADDIS ABABA UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN INFORMATION SCIENCE

BY

MINILIK ASSEFA

OCTOBER, 2016

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

SCHOOL OF INFORMATION SCIENCE

BANDWIDTH OPTIMIZATION FRAMEWORK FOR WIDE AREA NETWORK: THE CASE OF ADDIS ABABA UNIVERSITY

BY

MINILIK ASSEFA

NAME AND SIGNATURE OF MEMBERS OF THE EXAMINING BOARD

NAME		SIGNATURE	DATE
----------	CHAIRMAN	----------	----------
----------	ADVISOR	----------	----------
----------	EXAMINER	----------	----------
----------	EXAMINER	----------	----------

DECLARATION

I, THE UNDERSIGNED, DECLARE THAT THIS THESIS IS MY ORIGINAL WORK AND HAS NOT BEEN PRESENTED AS A PARTIAL DEGREE REQUIREMENT FOR A DEGREE IN ANY OTHER UNIVERSITY AND ALL SOURCES OF MATERIALS USED FOR THE THESIS HAVE BEEN DULY ACKNOWLEDGED.

MINILIK ASSEFA

OCTOBER, 2016

THE THESIS HAS BEEN SUBMITTED FOR EXAMINATION WITH MY APPROVAL AS UNIVERSITY ADVISOR.

WERKISHET LAMNEW

OCTOBER, 2016

ACKNOWLEGNEMNT

First I would like to thank God for giving me the motivation to start and the strength to complete this work. Next to that, I would like to thank my advisor Workshet Lamenew for his useful comments, guidance and commitment through the learning process of this master thesis. Many individuals have contributed support, idea expertise, finance, insight and time to make this study possible. Special thanks to my family, my former colleagues in AAU ICT office and friends for their support. Thank you all for your support.

Contents

List of Figures

List of Tables

LIST OF ABBREVIATIONS and ACRONYMS

ATM	Asynchronous Transfer Mode
CoS	Class of Service
DifServ	Differentiated Services
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
FTP	File Transfer Protocol
HTTP	Hypertext Transfer Protocol
IntServ	Integrated Services
IP	Internet Protocol
ISO	International Standard Organization
LAN	Local Area Network
MPLS	Multi-Protocol Label Switching
MRTG	Multi Router Traffic Grapher
PRTG	Packet Switched Traffic Grapher
QoS	Quality of Service
RSVP	Resource Reservation Protocol
RTT	Round Trip Time
SARG	Squid Access Report Generator
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SSL	Secure Sockets Layer
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
URL	Uniform Resource Locator
VoIP	Voice over IP
VPN	Virtual Private Network
WAN	Wide Area Network
WCCP	Web Cache Control Protocol
WLAN	Wireless Local Area Network

Abstract

Organizations Wide Area Networks are mostly designed to entertain the needs of several types of users. In the case of academic institution networks, the wide area network is mainly required serve the institutions users in order that it supports the main objective of the institution which is, supporting the teaching and learning process, research process and giving community service purposes. But because of the type of users that the wide area network of academic institutions encompasses, the use of the network resources by the users is highly different from other types of organizations. Academic institutions have a very special type of users, which are students that utilize the network resource in a different way from other organizations. Students mostly abuse network resources for different reasons like entertainment, experiment, to show their creativity by causing cyber-attack on some services in the wide area network or outside. Such type of users always tries to find a way of filling any available bandwidth and always use bandwidth. The act of doing such activities causes unbalanced use of network resources within the network users. This research aims to identify the cause of high bandwidth consumption by using log analysis tools with the view of developing bandwidth optimization framework that will enable the network to provide sustainable quality of service.

In order to achieve the objective of this study the researcher have used proxy log analysis by using SAWMILL to identify the internet usage trend of the users. In addition to that measurement of resource utilization at the core layers of different campuses of the institution has been done in order to find a bottleneck factor at the core layer devices.

The result of the study showed as there is a great resource abuse by some group of users in the network and as there is an unbalanced resource consumption among users of the network. The research result has shown as more than half of the internet bandwidth is getting consumed by less than 10 percent of the network users. To solve the issues mentioned WAN optimization framework was developed by using quality of service building blocks.

In this research we did not incorporate network performance issues at the end devices.

Keyword: Quality of Service, Performance Metrics, Network Traffic data, AAU WAN.

Chapter one

1 Introduction

1.1 Background

Now day, as the communication need to perform day to day activity increases the we are observing as there is parallel growth of network technology. This growth of network technology and communication means has brought several issues that are causing negative impact to the channel that the technology itself uses. Among those problems virus, intrusion, denial of service attack on network services are few of them. On the other hand, the malicious software's and irrelevant data traffics in the network consume majority of the bandwidth. Today's computer networks have gone from typically being a small local area network, to wide area networks, where end devices and servers are interconnected with each other from all over the world.[1] The interconnection technology has its own price for use. This medium of communication technology requires expensive medium to interconnect networks that are located in a very big distance from each other. For a good communication there has to be a good communication medium which can pass enough data traffic through it but, this medium or channel is mostly limited and expensive. This limited and expensive resource is called bandwidth. This development has gradually expanded and the interconnectivity bandwidth has become higher and cheaper. When we deal with less quality of service high congestion in the network, bandwidth is one of the important properties. Delay, jitter and reliability are also important properties for the quality of network connection. This is because different applications have different needs, and therefore require different properties from the network. According to [2] Local Area Network with good bandwidth, with minimal latency and also, if the packets drop rate was minimized, then the throughput will be relatively high. This shows as having a high throughput rate cannot be achieved by only having overprovision of bandwidth, but with its two factors that are latency and drop rate it can be a good network performance indicator.

As described in [2] it is important to pinpoint whether utilization factors, collision rate or bandwidth congestion are responsible for the university wide area network problems. In this study we have started with finding the source of the network performance issue by studying the level of

quality of service metrics (jitter, latency and loss), bandwidth, and analysing the user bandwidth utilization trends. We have also done traffic analysis in order to find out resource utilization fairness, bandwidth usage proportion and identify bandwidth greedy applications in the WAN. In this study we have used techniques for measuring, analysing, presenting and interpreting the different properties for the administration of WAN network. This research aims to investigate the cause of bandwidth congestion in a wide area network and proposes a solution tool for improvement of network QoS and bandwidth utilization.

According to explanation of [3] routing protocols also have impact on network performance and selecting the best routing protocol will have a critical role for network performance of wide area network. In addition to these factors the network devices capacity can also be a problem to the network performance[4]. Devices CPU and memory have to be taken in to consideration while assessing issues of a network. It is obvious that performance of a network degrades through time with depreciation of devices, addition of new routing parameter, addition of access lists and getting change of configuration on the network devices and with increase in number of network hosts. Another reason that may cause problem are a set of rules applied on some devices that do not interoperate with other network routing rules. An ever increasing request by the hosts to the internet and a massive broadcast storm that continuously grow in the network by the hosts are also reasons for congestion.

Most universities in Ethiopia use similar network infrastructure to provide the necessary IT services for the teaching learning, research and administrative processes. In most universities of Ethiopia, the services that the universities provide are mostly similar. The country has only one telecom service provider which is responsible to provide any kind of wide area connectivity service for any service seeking body[5]. This state of affairs has made all the types of organizations and institutions to have only one option. Which is, to register an MPLS-VPN service to interconnect their branch offices and campuses. Considering this situation, we can see that all wide area network interconnectivity services in the country will have similar network quality and features.

For this research we used Addis Ababa university as our target institution, which is one of the biggest university in the country which runs teaching and learning, research and community service processes by using the universities ICT network infrastructure and services. we used Addis Ababa university as a sample for its wide range of services running in the network, highest number

of students compared to other universities, having the highest number of remote campuses that are connected with the countries ISP VPN service and for their highest distance between the campuses physical location. In addition to those criteria's the university holds range of field of study in all level of study that range from undergraduate to PHD level and holds many researches than any university in the country[6].

Similar to other universities Addis Ababa University uses proxy servers for web browsing. This intern enables the network administrators to have a clear understanding of the data traffic through analysis of the log files and implementation of the network traffic management. Network traffic management is important to learn the utilization trend and capacity analysis as well as for troubleshooting[7].

From our observation as a simple user browsing in the university network we can see that there are connection attempts by worms which nag the user by popup to URLs on the Internet that were programmed into the worm's code. This shows how the bandwidth usage of an academic institution can be monopolized by malicious software's and unnecessary files. The academic purpose of the network is not fully reflected in the bandwidth usage, and it is clear that this network is out of control.

1.1.1 Motivation for the Research

As per the explanation of the research conducted by [8] the key function of an ideal network management system is to optimize the operational capability of a network. It is also reported by [9] that the performance problem of applications would be even worse, when the enterprise WANs are not well equipped with the required network resources to run the applications. Other researches' conducted on wide area network performance also indicate as limited network resources like bandwidth are a cause of network performance degradation by causing congestion on devices. As per the study conducted by [10] the researcher has described as there has to be a proper measurement and management of network resources; otherwise the impact of WAN network factors including, latency, throughput, congestion and packet loss will be more significant.

Hence, the previous studies conducted on this area describe as proper network resources utilization is the most critical point for network performance and it has to be properly managed. Since poor network performance highly impacts the performance of any network application running in a wide area network. Research conducted by [9] recommends that solving application performance issues should

be given much attention particularly for WAN environments, which lacks the required network resources. It also describes as many organizations in developing countries like Ethiopia do not have the required financial power to increase their WAN environment with resources like bandwidth and powerful WAN devices.

Based on the research result of [11] and their report about the application of quality of service in WAN environment, it is logical to assume that network performance problem can be optimized by application of similar techniques for wide area networks of institutions in developing countries like Ethiopia. This research is conducted in one of the universities in Venezuela. With this research the researchers were able to develop a framework that can solve the problem of the university wide area network. But when we consider application of the same framework for other institutions, it cannot bring the same result on cases like of universities like Addis Ababa University. Although the target organization has a very big similarity with AAU, the network providers and the scenario of the universities policy and application the ICT policy have a very big difference. Their framework was applicable for cases that users of the network have a good understanding of network usage and proper ICT policy implementation by the university.

As described on the book published by one on the biggest network device manufacturer company [12], implementation of quality of service on wide area network will enable the network administrators to allocate resources based on their need. In addition to that the book also describes as this recent technology enables application of resource allocation by using different criteria's like application type, user type, source address, destination address etc.

Thus, considering the limitation of considering quality of service as a remedy by previous works done and the attractive feature that wide area networks can achieve by applying quality of service features on wide area networks like Addis Ababa University, this research mainly focuses on developing framework that will enable optimized network performance for wide area network.

1.1.2 Common issues in universities network

Problems particularly associated with university campuses include behavior of students which typically have more time, are less supervised, and are under less pressure from work targets than other users of the network like the employees use the network service for entertainment and to download multimedia files which are not related to educational purposes[13]. Such issues make the university network to be one of the most challenging environments to manage[14].

The university community use the internet for so many different purposes, some of which are inappropriate or do not make the best use of the available bandwidth[13]. For example, while it may not generally be a problem if a student downloads a music file, plays on-line games but it becomes a problem when the bandwidth consumed by this activity impacts a researcher from downloading or viewing a scientific article[14]. Such act of using the resources is mostly prohibited by policy of university internet usage in some universities [15, 16].

Even where high amounts of bandwidth are available, monitoring and optimization are the basic operations that the network admins must do in the universities network. Since users, most of the time the students, will always find a way to fill the available amount of bandwidth.

Attacking network services and trying to hack systems by using several tool is mostly common by students' experiment with their computing knowledge, and connect to exposed computer systems both on the universities network or and elsewhere in the world.

In our country the single ISP (Ethio-telecom) gives a shared infrastructure to provide internet service for its customers. So most of the time universities network does not get the amount of bandwidth that the university has subscribed for both the VPN and the internet traffic. In such cases universities may need to police the amount of bandwidth they are getting from the ISP shared system because they might be competing for bandwidth with other customers. For example, if a university gets its bandwidth from an ISP, it is likely that the ISP also sells bandwidth to other users, such as local companies. Organizations should have a clear understanding of the nature of the shared system and how much minimum bandwidth they are paying for.

In other cases, if there is no bandwidth allocation and proper bandwidth management, ICT staff may themselves become part of the problem. In most universities network the ICT VLN gets direct access to the internet without passing through proxy servers. In addition to that ICT staffs have more privilege to the network resources for so many operational reasons. So sometimes some ICT staffs abuse that access by using the bandwidth for their personal interest. by deploying services that benefit only few group of users, like creating ftp server to upload multimedia files or installing proxy servers to provide better privilege for their friends in the university.

So in order to utilize the network bandwidth properly and in order to make informed decisions about how much bandwidth is needed or gets consumed by the users, control and

monitoring is necessary. From our monitoring process if we observe as the bandwidth is used mainly for recreational activities, or is consumed by virus activity and Windows updates, then control is more urgently required than additional bandwidth.

In view of the above and other associated problems, the university bandwidth should be monitored, managed and optimized so as to make university aware of its bandwidth utilization manner and its network capacity. In order to provide fair internet service without facing bandwidth starvation in the campus network and also to avoid unwanted traffic in the network we need to have some way of allocating network resources. For those reasons implementation of quality of service is required. QoS tools are not only useful in protecting desirable traffic, but also in providing deferential services to undesirable traffic such as abuse of the network resource by video streaming and gaming applications. QoS can be used to monitor flows and provide reactions to abnormal flows[7].

1.1.2.1 Infrastructure of Universities

Most universities operate in similar infrastructure design to run services that are most important for their teaching learning, research and community service missions, and their network design mostly are the same.[14] Universities have one main campus and several remote campuses and they deploy all their services in their main datacentre which most of the time will be in the main campus. The main datacentre will accommodate all the network resources and centralized systems to serve the remote campuses. Campuses or the branches are connected with the service provides VPN that the universities will be connected with each other. The network resource requests including internet traffic requests from other campuses will be forwarded to the main datacentre through their VPN.

Networks are designed to fulfil the goal of ICT Network Services, that is to provide a network infrastructure that is highly available, reliable, robust, secure, and easily accessed from anywhere. The ICT backbone data network, managed by the network infrastructure unit, provides the core data network connectivity for all campuses which are located geographically far from the main campus. The campuses though have all network infrastructure built and ready to get every service when connected. The university's network infrastructure is composed of a minimum of 1Gbps Core-to-Distribution throughput and minimum of 100Mbps access level connectivity that

enables them to run multiple multimedia applications smoothly within their campus including VoIP Services, Digital Signage, Video Conferencing, Smart Classroom, and IP TV services.

Network connectivity of university campuses mostly use fibre optics media to connect campuses with a close range with the main campus and other neighbouring campuses. If the distance between any two campuses is more than a limited range, then the universities will subscribe VPN service from the ISP. Using fibre optics media helps the universities to have a better bandwidth.

1.1.3 Addis Ababa University network infrastructure

Since its establishment, Addis Ababa university has been expanding its teaching, research and community services programs. Furthermore, the University has been undertaking various reform schemes in order to cope with and respond to the fast changing national and international educational landscape. At present (following its recent restructuring of institutional setup and governance system), the University has 10 colleges, 4 institutes that run both teaching and research, and 8 research institutes in 15 campuses[17].

Addis Ababa University has a Local Area Network that uses 1Gps speed fibber optics media to connect 8 of its campus. The LAN connects the main datacentre located in the main campus with Social Science Campus (4 Kilo), Technology Faculty Campus (5 Kilo), School of pharmacy, CBE Campus, Yared school, School of fine arts, Dental school, AAIT and each of the departments in these four campuses. The installed ICT Office network architecture is discussed here giving due attention on ideas behind the selection of Layered Network Architecture and the uses mix of the Huawei/Cisco Switches, Routers and Firewall.

The university ICT network uses layered architecture with a Unified Threat Management device (UTM) security device, router for path selection and gateway connectivity for the two links (internet and wereda-net) and core switch for its inter VLAN routing at the core layer of the network. When we look at its distribution layer, the university uses a layer 3 distribution switch for the inter VLAN routing and distribution service for the buildings in the compound with a 1Gbps connectivity. The distribution layer device provides connectivity to campus VPN.

1.1.4 Services in universities

Almost all universities run services like proxy, DNS, DHCP, FTP, E-learning, Student information management, Library, Antivirus, Web, E-mail and network monitoring services. When we look at it from the infrastructure side all universities have deployed wireless network in their compound and all wireless devices can join the universities network without authentication.

Several researches have been undertaken by many scholars on the area of network optimization and bandwidth utilization. So far several frameworks and algorithms have been developed to effectively utilize bandwidth usage on a wide area network. Bandwidth optimization is one of the good solutions for effective quality of service provision in a wide area network[18]. But, one of the biggest problem lies on sustainability of optimum bandwidth utilization as the network grows and in oldest network environments even with an enough bandwidth it is common to experience packet loss, latency and very poor response time. In most cases when measuring the network performance by some tools most figures show a good performance figure but the user experiences a delayed response for his/her request.

In big networks like Addis Ababa university which uses mix of several types of network devices manufactured by several venders it is common to experience very high drop during packet transmit and denial of service at some point in the network path which in turn brings a very high packet loss with a big latency rate at every network node. When we come to performance measurement, the figure shows as the device is operating at a good state and the percentage of utilization the interface capacity shows below its threshold value. But we mostly see the subscribed bandwidth utilization level getting to the top limit. All those problems indicate as either the network performance measurement tools are not reliable or the issue of network is mostly from the local area network than the getaway bandwidth issue.

1.2 Problem Statement

Most organizations and institutions including Addis Ababa university take bandwidth upgrade as a remedy for poor network performance. As a result of that they will be exposed to an expense that is not giving solution for their problem. For example, Addis Ababa university has 600Mbps bandwidth on the main datacentre located at Sidist kilo campus and several other internet gateways on most of its remote campuses which sum up to more than 700Mbps and which in turn

makes the university to use a total of more than 1300Mbps bandwidth but still user satisfaction level of the internet service is very poor.

Previously several researches have worked on developing a model and framework on bandwidth utilization for universities and higher institutions. For the case of African universities, Hailay Weldegebriel, [14] had developed a bandwidth allocation prototype model. In his research he has taken Mekele university as his case study to develop the model. The scope of his study only focused on weblog analysis. But the problem of bandwidth is not only from the utilization perspective but also from lack of consideration from the quality of service end [19]. Another research was conducted by Tsegaye Berhanu [10] to analyze performance of the AAU wide area network. In his research he has taken network performance metrics to measure the performance of the university network on the gateway router only. In his recommendation he has mentioned as a further research is required to be conducted on the metrics like bandwidth utilization, quality of service, error rate, jitter, retransmission rate and reply rate.

Accordingly, this research is done to investigate cause of bandwidth shortage and by studying several metrics of network performance (with quality of service parameters) and analyzing bandwidth utilization data to develop a suitable framework to improve bandwidth usage for AAU network.

The main questions that this research aims to answer are:

- What are the main issues that cause very high response time (response delay) when accessing resources from the internet?
- What are the areas that are causing bottleneck to the network that could make other devices to be idle and where does the loss of packet happen?
- What are the most frequent data packets that are traveling in the network and are not relevant to the majority of the users and not related to the organization or institutions mission (applications eating the bandwidth) and what remedy solution or procedure should be applied?
- What kind of framework should be developed to solve these issues?

In addition to those points this research aims in identification of flow of rules and procedures as well as compatibility and interoperability of network devices, routing and switching

protocols implemented on the network devices as well as network services and quality of configuration.

1.3 Objective

1.3.1 General Objective

The general objective of this work is to develop a framework that can enable higher education networks to achieve quality of service over their WAN by assessing the resource utilization trend.

1.3.2 Specific Objectives

In order to attain the general objective stated above, the researcher has formulated the specific objectives as follows.

Literature review: To review related literatures.

Data collection: To collect the necessary data.

Measure the network performance: To measure the network devices performance.

Analysis of data: To conducting data analysis.

Result Presentation: To present the results.

Proposal: To propose a suitable framework.

Evaluation: To evaluate the performance of the developed framework.

1.4 Significance of the study

This study will help academic institutions to achieve sustainable quality of service from their network and effectively utilize their bandwidth for the intended purpose and makes them to reduce cost of bandwidth. On the other hand, the study will prevail problems associated with shortage of network service quality even with abundant amount of bandwidth. In addition to that the research will help other researchers to study the issues of network performance with the other approaches that are not considered in this study like including impact of end devices performance problems on network performance and considering improvement of network performance by adding performance improvement mechanisms on end devices.

The study supports network administrators for providing a good service to the wide area network users and also indicates major issues that degrade performance of a wide area network. It also addresses approaches to problems that have not been seen in previous studies but heavily improve network performance. Since the study focuses on using quality of service building blocks for wide area network optimization, the study will bring a good motivation for other researchers, ICT firms and network administrators to study on techniques of providing good service through properly allocating resources to proper applications using quality of service framework.

1.5 Methodology

To accomplish the research objectives, the researcher followed statistical data analysis with experimental research methodology and here follow the procedures carried out.

1.5.1 Literature Review

The researcher reviewed several literatures in order to assess the main problems and concepts in the area of network traffic analysis, network protocol distribution, quality of service provision and bandwidth management. A number of books, journals, articles, and research papers from the Internet as well as hardcopies were consulted to assess the proper application of network protocols, proper implementation of network services, bandwidth requirement analysis, bandwidth measurement and main causes of poor quality of service.

Moreover, the researcher reviewed several articles regarding various network traffic analysis tools which are considered to be important to accomplish the research successfully such as SAWMILL, Solar wind, PRTG, SARG and other tools to assess the network performance and identify issues and proposes solution as well as experiment the proposed remedy.

1.5.2 Data Collection Method

In this paper we analysed the online and offline network traffic of AAU wide area network so as to identify the main issues mentioned in the problem statement.

For online traffic analysis the researcher used the device manufactures monitoring system which is Huawei device monitoring tool which clearly shows the interface status and its utilization and PRTG network monitoring tool which are used for measuring bandwidth, delay loss, jitter and availability between two network nodes.

Web browsing access through a proxy server at an academic institution can be analysed using software's like SARG (Squid Access Report Generator) and SAWMILL [6]. Since the web proxy server log keeps a record of every web site visited, as well as the amount of bandwidth consumed, every sites access and bandwidth consumed can be identified.

The offline network traffic analysis is based on the university proxy server logs and the unified threat management device logs. In order to analyse the log files of the proxy servers the researcher used SAWMILL log analyser tool. Proxy log files are records of web server activities and analysing it will provide detail about every request that has passed through the proxy server. They provide details about file requests to the proxy server and the servers response to those requests. The proxy server log files are stored as access log and cache log files. The access log file gives a brief description about the source of a request, the file requested, the date and time of the request, the content type and length of the transferred file, and other data such as errors and the identity of referring pages.

For the purpose of this research the researcher has used proxy server log files of eight consecutive days (April 13,2016 – April 21, 2016) from three proxy servers and seven consecutive days' log of the UTM device. The purpose of using the proxy log and the UTM log is as an attempt to extract useful information in terms of bandwidth usage, browsers usage, website errors, and mainly the proportion of data traffic source and destination and to assess information about web surfing behaviour of the network users.

1.5.3 Data Analysis Method

Since we categorized network traffic analysis in online traffic analysis and offline traffic analysis. The online traffic analysis is implemented using PRTG and Huawei device monitoring tool report to measure the overall resource utilization, bandwidth size and to visualize the network traffic usage at different links of the university intercampus network. The offline network traffic analysis is implemented using SAWMILL so as to analyse the internet traffic and bandwidth usage and to understand the user behaviours in the intercampus network.

In this stage, the raw data are processed in different ways to gather useful information about the measurements. Interesting data can be: minimum value, maximum value, mean value, median value, etc.

1.5.4 Data Presentation Method

After analysing the data, the results will have to be presented in a way that are understandable to the reader. In order to do that we have chosen statistical presentation tools that will show a clear picture of the data analysis result.

1.5.5 Proposing WAN optimization framework

The study mainly focuses on developing a WAN optimization framework for organizational networks with a mix of user types. Developing the framework is guided by the result of analysis of network traffic history and the resource utilization trend of the institution under the case study. For the purpose of attaining best output from this study user involvement on the process of framework development was done.

1.6 Scope and Limitation of the study

Scope of this research aims in proposing solution for Ethiopian higher education institutions network quality of service improvement under the assumption of the institutions network infrastructures and services similarity only. This research mainly studies the network performance of AAU and all the data used for this research is collected only from AAU network. The research data collection has mainly focused on the network traffic, network bandwidth and the network devices performance. Due to the shortage of time and finance this research could not include the end users local machine performance and could not include issues that may cause end users machine local network service problem The study used performance metrics such as response time, packet loss, throughput, bandwidth utilization, traffic received, traffic transferred and network availability.

1.7 Thesis Outline

This thesis paper has five chapters. The first chapter describes general points about existing situation in a wide area network of many wide area networks and their performance status as well as the problem that their network has. This chapter also describes about the problem statement, the points that motivated the researcher to conduct this study and researches that have been conducted on similar issues both local and global studies.

Chapter two describes about concepts on networking and the techniques that network has to be designed and configured. The chapter mainly discuses on the concepts that is several researchers' publications, books and works of scholars reviewed from the WAN optimization and QoS perspective.

Chapter three discusses description of the institutions network under the case study. The chapter describes the existing network performance, wide area network organization and setup of the infrastructure.

Chapter four briefly explains the analysis result of the data analysis that collected from the historical data of the network operational process. The chapter presents the analysis result in graphical and descriptive manner. This chapter contains the proposed optimization framework which is the result of this study.

Chapter five contains the conclusion and recommendation part. In this chapter the researcher has tried to express the overall thesis organization and his recommendation for further work and further study.

Chapter two

2 Literature Review

In this section of the paper we will try to establish a theoretical framework for the research main subject area. We have tried to define the key terms, definitions and terminology by basing on previously conducted research dissertations, books, journal articles and any reliable source of knowledge. In general, we will discuss about the points that are taken in to consideration for the purpose of this research and explain about the results associating them with the reviewed literatures.

2.1 Networking

Computer networks are classified by the range of the network connectivity and coverage of area that the network consist of. If the network ranges a few meters then the network can be classified as personal area networks and networks that range a few hundred meters are classified as local area networks (LAN)[4]. Whereas network that groups several local area networks within a range of some kilometers are classified as a metropolitan area network. And any networks ranging more than some kilometers are classified as Wide Area Networks (WAN).

2.1.1 Local Area Network (LAN)

Devices interconnected within an area of 1000m2, which is the generally accepted maximum size for a LAN, are considered to be in a local area network (LAN). Due to their controlled size and connection media that are used in a local area network low latency and high bandwidth are typically properties which describe them [4]. Now a day's local area network technologies are mostly with 100Mbps and Gigabit Ethernet. These technologies are typically designed to run on either twisted-pair cables or optical fiber cables. Since the bandwidth of a Local Area Network is in gigabit/sec or 100Mbps it will have a minimal latency and also they are mostly with less packet drop rate so their throughput will be relatively high[2]. These new technologies will provide a wider range, higher bandwidth, increased security and quality of service.[20]

2.1.2 Wide Area Network (WAN)

As the number of small LAN networks grow there will always be a need to interconnect local area networks with each other and for that the network technology has brought WAN. A wide area network or WAN is a type of computer network covering a wide geographical area, and are used to connect local area networks together[4]. WAN's can be built to connect several private local area networks in an organization, or built by Internet service providers (ISPs) to provide an organization access to the Internet. Wide area networks are typically built of leased lines, where a router connects the local area network to the private wide area network. An alternative is to use the Internet, which provides a shared infrastructure and a high speed wide area network. Virtual private networks (VPNs) can use encryption and other techniques to make the connection secure and private[4, 21, 22].

2.1.3 VPN

A Virtual Private Network (VPN) is a way of creating a secure connection between two separate networks or user and a network resource or even network and network. VPN uses strong security encryption and restricted, private data access which helps it keep the data secure from the other users of the network that share the media of connectivity which could often be a public network like the internet[4]. A VPN technology provides a secure and encrypted connection over a less secure network. Since VPN uses a shared network private data has to be segmented from other traffic so that only the intended recipient will have access.

VPN is the most effective and versatile form of secure communication across long distances. More bandwidth is required to handle the additional network load. To achieve the security needed for the data VPN uses encapsulated internet packets to transport data in this dynamically created tunnel. Encapsulation means that the VPN application wraps the packet with a header that includes the routing information[23].

The advantage of using a virtual private network is the level of security that it guarantees to the systems connected through it when the underlying network infrastructure alone cannot provide it [21]. The main reason for using a VPN instead of a private network is usually because it reduces down cost and helps to attain feasibility: which means it is either not feasible to install a private

network or it is too costly to do so and in some cases there are regulations that are set by the regulatory body which does not permit to organizations to install networks beyond certain range of distance. In those cases, the countries ISP infrastructure will be used to for VPN connectivity or in other words organizations must subscribe VPN service from the ISP.

2.2 Network Protocols

In a typical network, there are several protocols in use. As the size of the network grows it is obvious that the number of services associated with it will also grow which implies that the number of protocols used in the network will differ with the amount of service and complexity of the network. The most known protocol is the Internet protocol suite, which are the building blocks of the Internet. But there are several important protocols both over and under these protocols. Examples are HTTP, FTP, HTTP, ICMP etc.

2.2.1 Internet protocol suite

The Internet protocol suite in most networking books is explained as it is the set of communications protocols that implement the protocol stack on which the Internet runs. The internet protocol suite has two protocol suits in it. These are the Transmission Control Protocol (TCP) and the Internet Protocol (IP). The internet protocol suite is sometimes called the TCP/IP protocol suite, after the two most important protocols in it, which were also the first two defined[24, 25].

2.2.2 Internet Protocol (IP)

The Internet Protocol (IP) is a connection-oriented protocol used by source and destination hosts for communicating data across a packet-switched internetwork. Data in an IP internetwork are sent in blocks referred to as packets or datagram's which the two terms (packets and datagrams) are basically synonymous in IP. In particular, in IP no setup is needed before a host tries to send packets to a host it has previously not communicated with[24]. The Internet Protocol provides an unreliable datagram service (also called best effort); i.e. it makes almost no guarantees about the packet. Packets in IP may arrive damaged even they may be delivered out of order (compared to other packets sent between the same hosts), it may be duplicated, or it may be dropped entirely[25, 26].

2.2.3 Transmission Control Protocol (TCP)

Transmission Control Protocol (TCP) is a connection-oriented protocol and it provides a reliable delivery of byte streams at the transport layer communication protocol. It does the task of the transport layer in the simplified OSI model of computer networks[24]. In the Internet protocol suite, TCP is the intermediate layer between the Internet Protocol below it, and an application above it. Applications most often need reliable pipe-like connections to each other, whereas the Internet Protocol does not provide such streams, but rather only unreliable packets[24]. TCP connections contain three phases: connection establishment, data transfer and connection termination. A 3-way handshake is used to establish a connection. A four-way handshake is used to tear-down a connection. During connection establishment, parameters such as sequence numbers are initialized to help ensure ordered delivery
and robustness[26, 27].

2.2.4 User Datagram Protocol (UDP)

The User Datagram Protocol (UDP) is a minimal message-oriented transport layer protocol. In the TCP/IP model, UDP provides a very simple interface between a network layer below and an application layer above. UDP provides no guarantees for message delivery and a UDP sender retains no state on UDP messages once sent onto the network. UDP adds only application multiplexing and data check summing on top of an IP datagram[4, 26].

2.2.5 Internet Control Message Protocol (ICMP)

The Internet Control Message Protocol (ICMP) is a set of protocols used by networked nodes to send control data to the network. ICMP differs in purpose from TCP and UDP in that it is usually not used directly by user network applications. One exception is the ping tool, which sends ICMP Echo Request messages (and receives Echo Response messages) to determine whether a host is reachable and how long packets take to get to and from that host [24].

2.3 Network Performance Measurement

Networks must provide predictable, measurable, and sometimes guaranteed services by managing bandwidth, delay, jitter and loss parameters on a network[7]. In order to know the level of performance of any system there is always a practice of monitoring it. In the case of networking network monitoring is an activity of measuring the level of network performance metrics by using

network performance measurement tools[8]. Network performance metrics are measured by various network monitoring technologies. The well-known types of network monitoring methods are three, which are active monitoring, passive monitoring and the method of using SNMP agents[28]. The active monitoring method obtains the current status of the network by setting up the test machine from one side of the network at the point where one wishes to measure, and then sending extra traffic from one machine to another during a specific time. Various network performance measures can be measured by simple and easy tools, such as ping and traceroute. Using such simple tools eases the system load because the size of generated and analyzed traffic is small compared to passive monitoring method. The third method which is using simple network management protocol (SNMP) requires the devices to have SNMP features. This technique helps the network devices to send their SNMP traffic which shows their status to the monitoring system with a scheduled interval[29].

2.4 Quality of Service

Quality of Service (QoS), is a term we use to refer to network resource control mechanisms. Quality of Service is the measure of our ability to provide different priority to different applications, users, or data flows, or to ensure a certain level of performance to a data flow[30]. In other words, it can be explained as it is a term used to describe the capability of the network in handling a traffic in a way that it meets the service needs of applications. The provision of quality of service requires fundamental traffic handling mechanisms in the network, which has the capability identify traffic that is entitled to these mechanisms and the ability to control these mechanisms[7].

QoS functionality tries to satisfy the needs of resource consumer which is network applications and resource allocator which is the network administrators. Most of the time the association of it appears odds, since it is the network administrator duty to limits the resources used by a particular application while the application attempts to seize resources from the network. These apparently conflicting goals can be reconciled by realizing that the network administrator is chartered with maximizing the utility of the network *across the full range* of applications and users.

Dynamic resource control capabilities have become increasingly important for academic networks that must support big scientific research projects at the same time as less data intensive research and educational activities[31]

2.5 QoS Parameters

Different applications have different requirements or parameters regarding the handling of their traffic in the network[7]. Applications generate traffic at varying rates and generally require that the network be able to carry traffic at the rate at which they generate it. In addition, applications are more or less tolerant of traffic delays in the network and of variation in traffic delay. Certain applications can tolerate some degree of traffic loss while others cannot[1]. These requirements are expressed using the QoS-related parameters. The four main properties for a network connections or QoS are:

1. Bandwidth
2. Delay
3. Jitter
4. Throughput.

These four properties define the quality of service (QoS), that the network requires[32]. For some application the QoS may not matter but, but it may be crucial for some other applications.

2.5.1 Bandwidth

Bandwidth describes the maximum data transfer rate of a network or Internet connection. It measures how much data can be sent over a specific connection in a given amount of time[7]. For example, a gigabit Ethernet connection has a bandwidth of 1,000 Mbps.[4] An Internet connection via cable modem may run up to 25 Mbps of bandwidth.

While bandwidth is used to describe network speeds, it does not measure how fast bits of data move from one location to another.[31] Since data packets travel over electronic or fiber optic cables, the speed of each bit transferred is negligible. Instead, bandwidth measures how much data can flow through a specific connection at one time.

Data often flows over multiple network connections, which means the connection with the smallest bandwidth acts as a bottleneck. In most cases, the Internet backbone and connections between servers have the most bandwidth, so it is very rare that they cause bottlenecks. Instead, the most common Internet bottleneck is organizations connection to their ISP. WAN bandwidth remains a constrained resource that is economically infeasible to substantially overprovision.[33] So, it is important to allocate capacity according to service priority and based on the incremental value of additional allocation.

2.5.2 Latency and delay

Latency is a term used for describing the time that data travels in the network or the time it takes from source to the destination. Latency can be expressed as one-way latency or as roundtrip latency. One-way latency, also known as delay, simply means the time it takes for data to travel from the transmitting node to the receiving node[34]. Delay is the time that it takes for a packet or a frame data to travel from its source node or sending device to a destination node or receiving end. Roundtrip latency, also known as Round Trip Time (RTT), measures the time data travels from transmitting node to the receiver plus the time that it takes for the transmitting node to get a response or an acknowledgement from the receiving node. When studying application performance, the most commonly used form of latency is the round trip time (RTT). Latency can be divided into smaller delay components that together generate the overall network latency. These components are propagation delay, serialization delay, processing delay and forwarding delay[9]. The delay is the product of three factors or other delays, which are known as transmission delay, propagation delay and queuing delay.

2.5.2.1 Transmission delay or Serialization Delay

Serialization delay or transmission delay is the amount of time it takes to put the signal onto the cable and it is measured as the time it takes to move bits of a packet into the line[32]. The delay is a result of the three factors which are the size of the packet, network medium and speed of the interface. Usually serialization delay is more significant in lower-speed networks[29].

2.5.2.2 Propagation Delay

Propagation delay is a form of delay that is caused by the distance between nodes or is the amount of time it takes for the signal to travel across the cable[32]. This depends on the type of media used and the distance involved. It is expressed in terms of how fast data can be transferred in the network. Propagation delay is one of the fixed factors affecting the overall network latency. Propagation delay is measured as the time the data packet spends to go through the network. The speed packets can be transferred is called propagation velocity and it is normally around 2 / 3 of the speed or light. Propagation delay becomes significant in long distances, which is usually the case in WANs[9, 34].

2.5.2.3 Processing Delay or Queuing delay

Processing delay is the time it takes to process the packet or frame at intermediate devices such as routers and switches. In other words, it is time that it requires for a network device to process the frame. In a router processing delay is the time that the router takes for comparing a piece of data to its access list in order to forward it to the correct path[9]. The forwarding architecture has to also be counted in processing delay: the node can either wait until the entire packet is received before it makes any decisions what to do with it (store and forward) or the forwarding of the packet can start as soon as the header is received. Depending on the load of traffic in a router the processing delay can vary between less than 1ms to even 10ms when the router is congested.
It is called the queuing delay because most of time is spent in queues within the device[9].
The transmission delay and the propagation delay are quite predictable and stable delays, but queuing delays can introduce considerable variability.

2.5.3 Jitter

Jitter can be explained as the variation in latency or it is the variation in arrival times of successive packet from a source to a destination. And is determined by the difference experienced by subsequent packets, RTTi and RTT i+1. The jitter can be measured by monitoring the round trip time for packets between two nodes[1].

2.5.4 Packet loss

In data transmission systems and data networks things may not always perfect. So, we can do not expect that all packet transmitted is received by the corresponding layer and this is because of there

is an expected packet loss[34]. A relative measure of the number of packets that were not received compared to the total number of packets transmitted. Loss is normally a function of availability. If the network is very much available, then loss during periods of non-congestion would be essentially zero. During periods of congestion, however, QoS mechanisms can determine which packets are more suitable to be selectively dropped to ease the congestion[7]. In the event of a packet loss, the transmitter will not receive an acknowledgment for the lost packet, and it will have to retransmit the packet. Packet loss is not a scenario that can be proactively reported to a transmitter: that is, a router that drops a particular packet generally does not notify a transmitting node that a specific packet has been dropped due a congested queue. Packet loss is generally handled reactively by a transmitting node based on the acknowledgments that are received from the recipient or the lack of recipient of said acknowledgements[10, 28].

2.5.4.1 Transmission Errors

When a transmitter sends a bit of information down a transmission path, it is not always interpreted correctly by the receiving entity on the transmission path. For this reason, protocol suites include error detection or error correction capabilities. Predominantly, error detection methods are employed in most modern WAN networking protocols, due to the relatively low bit error probabilities on transmission systems. These protocols, when detecting an error in a packet, discard the entire data packet. This appears to the higher layer protocols as a packet loss[35].

2.5.4.2 Buffer overflows

Data communications equipment that employs statistical packet multiplexing maintains finite data buffers. Due to the statistical nature of the packet loads, there are occasions when the amount of data to be buffered exceeds the size of the data buffers and packet losses occur. To detect packet losses, protocol suites employ packet sequencing by numbering the packets in the order in which they are sent. A lost packet will result in the receiver seeing a packet number or several numbers skipped over[35].

2.5.4.3 Transmitter time-outs

When a transmitting host sends a packet, it sets a timer. When the transmitter receives the acknowledgment for the packet, it terminates the timer. If the value of the timer exceeds a threshold, that is, the transmitter time-out value, the transmitter assumes that the packet was lost

in transmission, and the transmitter enters into an error recovery state. Time-outs may occur due to packet losses or due to excessively long queuing delays in transit[35].

2.5.4.4 Out-of-sequence receptions

Some network technologies ensure that the order of the packets sent to the network is the same as the order of the packets delivered by the network. Some networks do not maintain packet sequencing, most notably IP router networks. Some transport protocol implementations may discard out-of-sequence packets instead of storing them and reordering the incoming packet stream. In this case, out-of-sequence packets will affect the performance of throughput systems in a fashion similar to other packet loss mechanisms[35].

2.5.5 Throughput

Throughput is the rate of successful data transfer in the network. It can be defined as it is a sum of the three building blocks of quality of service or network performance parameters: network capacity, latency and packet loss[32]. Capacity means the maximum amount of information that can be transferred between two network nodes. The throughput of a network is never more than the capacity of the slowest hop within that network. With the perspective of a system throughput can be described as it is a measure of how many units of information a system can process in a given amount of time[11].

2.6 QoS Models for Congestion control

The network traffic on the internet grows as a consequence of growth of the network traffic and growing number of users and applications. Thus, some points of the network can become congested during the office hours. Network congestion occurs when the network resources like bandwidth, CPU, memory, buffers etc. are not sufficient to support the accumulated demand of users[8]. The consequences of network congestion are low throughput, high delay in delivering data packets, wasting of network resources because of dropped packets and possible network collapse, in which all communications in the entire network stop. The *Internet Engineering Task Force* (IETF), a volunteer organization that sets the standards for the Internet, and other Internet researchers have worked on improving extending the TCP/IP model to Internet congestion control and Quality of Service (QoS) support.

2.6.1 Service Models

A service model includes a set of mechanisms and protocols for managing network resources in order to avoid network congestion conditions which can degrade the agreed service performance level of applications. The IETF group has proposed two Internet Service models. The service models proposed include Integrated Service model (IntServ model) and Differentiated Service models (DiffServ model) along with Resource Reservation Protocol (RSVP).

2.6.1.1 Internet Integrated Services Model

Internet Integrated Services Model (IntServ model) is suggested to support real-time and no real time Internet services. Users are able to explicitly claims some quantitative QoS guarantees, so their applications can run in a suitable way over a certain period of time. Internet Integrated Services Model provides both a mechanism which delivers users' QoS requirements and one which decides if the network can meet those necessities. Traffic control functions are performed by the admission control, packet scheduler, and classifier[11]. The components of the IntServ model interact in order to control the traffic in the network and reserve and negotiate different service classes along the communication path. It comprises a host communicating with an Internet router. The host and router systems are the same except that the application block in the host is replaced by a routing block in the router. Each of those blocks is described below:

a) **Applications**: request specific QoS from the network.

b) **Reservation process**: a set of procedures to reserve resources (eg bandwidth and buffer space) along the path of the data flows.

c) **Classifier**: classifies IP packets according to a set of service classes and assigns them to different queues.

d) **Packet scheduler**: decides which of the set of IP packets will be served next.

e) **Admission Control**: decides whether there are sufficient resources available to grant the requested QoS for a data flow. A data flow is a distinguishable packet stream which results from a single user/application activity and requires the same QoS.

f) **Policy Control**: decides if the user requesting a reservation is permitted to do so. Policy control mechanisms may involve, for example, the identity of the user and application, traffic and data rate requirements, and security considerations.

g) **Routing process**: determines the route along which the packets will be forwarded.

2.6.1.1.1 Classes of Service

The Integrated Services Work Group has defined several classes of service, which are described as follows:

a) **Guaranteed service**: this service is for guaranteed delay-bound real time applications. It provides guaranteed data rate and delay. Also, data packets conforming to their traffic specifications will not be discarded because of queue overflow. The guaranteed service only controls the maximum queuing delay. Other delays which are fixed delays such as transmission delay and propagation delays may be determined by the setup mechanisms. This service is intended for applications which have firm time constraints, such as telephony and medical images[11].

b) **Controlled load service**: corresponds to the predictive real-time service. Nodes (eg routers) which have committed to providing a controlled-load service should offer a service which approximates that provided by a best-effort service under lightly loaded conditions. A high percentage of delivered packets should not exceed a minimum transit delay and should arrive at their destination successfully. Controlled-load service
may be used for applications such as video conferencing and Internet real-audio.

c) **Best effort service**: corresponds to “elastic” applications and is the current service provided by the Internet. [7]

2.6.2 Resource Reservation Protocol

RSVP is designed to be run on network routers and in end hosts to support a QoS application. It reserves resources for a data flow from the sender to one or more destinations (i.e. multicast destination)[36]. Unlike other signaling protocols, RSVP destinations (receivers) request resource reservations. Those requests travel on the reverse path of the data flow by following the pre-established route setup by RSVP. RSVP is also responsible for maintaining reservations on each node associated with the data flow. RSVP uses a soft-state approach where the reservation states must be refreshed periodically; otherwise they are automatically removed. The approach accommodates dynamic route changes, dynamic multicast group membership and dynamic QoS changes. RSVP reserves resources for a session[11].

A session embraces all data flows from one or more senders to the same unicast or multicast destination. RSVP reservation requests are defined in terms of a filter specification (filter spec) and a flow specification (flow spec). A filter spec is used to identify the data flow that is to receive the QoS specified in a flow specification. A flow spec defines the desired QoS in terms of a service class, which comprises a Reservation Specification (RSpec), and a Traffic Specification (TSpec). A RSpec defines the reservation (i.e. desired QoS) characteristics of the flow, for example, the service rate (i.e. the data rate that a data flow can use). A TSpec defines the traffic characteristics of the flow, for example, the peak data rate (i.e. the maximum rate at which the sender is intended to send packets). RSVP uses several messages in order to create, maintain, and release state information for a session between one or more senders and one or more receivers[11].

2.6.2.1 Differentiated Services Model

The main problem of the IntServ proposal is that it is not scalable across large networks. Thus, another working group developed a service model called Differentiated Services (DiffServ). The DiffServ model is intended to solve the scalability problem by aggregating traffic. Large flows with similar service requirements are aggregated. Traffic entering a network is classified and marked in order to receive a specific quantitative or qualitative QoS[11].

The Diffserv architecture includes a number of functional elements known as per-hop behaviors, packet classifiers and traffic conditioners. They are implemented in several nodes (eg routers) along the network. A per-hop behavior (PHB) is the means by which a sequence of packets obtains some level of service. It may be seen as the differential treatment which a packet will receive. It may be defined in terms of network resources (i.e. buffer), traffic characteristics (e.g. delay, loss), etc.

A packet classifier starts by selecting the packets in a input traffic stream by using either the DS code point of the packet header or a combination of one or more header fields, such as IP destination address, IP source address, DS field, and IPv6 flow ID and/or other packet attributes. After that, it forwards them to an element of traffic conditioner for further processing. Thus, a classifier splits an input traffic stream into one or more output streams. A traffic conditioner is an entity which performs control functions intended to enforce traffic rules. It may contain meters, markers, shapers, and droppers. These components are described briefly as follows:

a) Meters: are used to monitor the arrival time of packets in order to verify that they are conforming to their traffic characteristics in the traffic characteristic agreement (i.e. traffic profile). The meter provides the resulting information to the other components of the traffic conditioner.

b) Markers: set the DS code point field in the IP packet to a particular value. For example, it may mark packets which have been classified by the classifier as a member of a particular flow. It also may re-mark previously marked packets which, for example, are not conforming to their traffic profile (see meters).

c) Shapers: delay packets in a traffic stream by using buffers, so the traffic conforms to its traffic profile.

d) Droppers: discard some or all the packets in a traffic stream so that the traffic stream conforms to its traffic profile.

The functional elements of the DiffServ architecture may be implemented in different nodes in a network. A node (e.g. a router) which is enabled to support differentiated services functions is called a DS node. A DiffServ specification classifies the nodes according to their location in a DiffServ region and the functions they perform. The following terminology applies to a DiffServ network. The DiffServ Working Group has defined several classes of services so far. They are defined in terms of PHBs and include Expedited Forwarding, Assured Forwarding, and Best-Effort Forwarding.

a) **Expedited Forwarding:** provides a virtual leased line end to-end service, which is characterized by low loss, low latency, low jitter, and assured bandwidth. It is also called "premium service". It may suit applications such as video broadcast, voice-over-IP, and virtual private networks.

b) **Assured Forwarding:** provides a service based on an "expected" usage profile. This profile indicates the level of performance (service assurance) uncertainty the user may tolerate (user expectation), more than a strict guarantee (like RSVP may provide). During periods of congestion some packets may still be dropped, but it may be acceptable for the user. The drop precedence values determine which packets are likely to be dropped during periods of congestion. In order to provide a level of forwarding assurance, a certain amount of resources (bandwidth and buffer space) are allocated for an assured forwarding class, and each IP packet must be marked with a drop precedence value.

c) **Best-Effort Forwarding:** is the default service given when there is no other agreement in place. It corresponds to the common best-effort service with no QoS guarantee

2.7 Network Data Traffic

"In 2003, a Wall Street financial company did an extensive study to identify and categorize the number of different applications on their networks. They found over 3000 discrete applications traversing their infrastructure. Further research has shown that this is not uncommon for larger enterprises."[7]. As described by Cisco, there are hundreds of thousands of data networking applications. Some of them are TCP and others are UDP; some of them are delay sensitive, others are not; some of them are burst in nature, others are steady; some are lightweight, others require high bandwidth, and so on. The variation of applications is not only from one another, but even the same application can vary significantly in one version to another[7, 37]. According to [7] the Cisco QoS Baseline identifies four main classes of data traffic, according to their general networking characteristics and requirements. These are best effort data, bulk data, transactional or interactive data, locally-defined mission-critical data.

2.7.1 Best Effort Data

The Best Effort class is the default class for all data traffic. An application will be removed from the default class only if it has been selected for preferential or deferential treatment [38]. When addressing the QoS needs of Best Effort data traffic, it recommended to allocate adequate bandwidth to it. Since the majority of applications will default to this class; it is recommended to reserve at least 25% the bandwidth for Best Effort traffic[7].

2.7.2 Bulk Data

This data class is intended for applications that are relatively not interactive and not drop sensitive and they typically span their operations over a long period of time as background occurrences. These applications are applications like FTP, E-mail, Backup operations, Database synchronizing or replicating operations etc. [38].

Bulk Data traffic should have a moderate bandwidth guarantee with constrained approach in order to prevent it from dominating a link. The advantage of provisioning moderate bandwidth

guarantees to Bulk Data applications rather than applying policers to them is that Bulk applications can dynamically take advantage of unused bandwidth and thus speed up their operations during non-peak periods. This in turn reduces the likelihood of their bleeding into busy periods and absorbing inordinate amounts of bandwidth for their time-insensitive operations[7].

2.7.3 Transactional or Interactive Data

The Transactional or Interactive Data class, is a combination to two similar types of applications: Transactional Data client-server applications and Interactive Messaging applications. The response time requirement of Transactional Data client-server applications separates it from generic client-server applications. If we consider Transactional Data client-server applications such as SAP, PeopleSoft, and Data Link Switching, the transaction is a foreground operation; the user waits for the operation to complete before proceeding. It is recommended that Transactional Data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations they support[7, 38].

2.7.4 Locally-Defined Mission-Critical Data

Our expression for the term "locally-defined" is used to stress the purpose of this class, which is to provide each enterprise with a premium class of service for a select subset of their Transactional Data applications that have the highest business priority for them. Mission-Critical Data traffic should have an adequate bandwidth guarantee for the interactive, foreground operations they support[7].

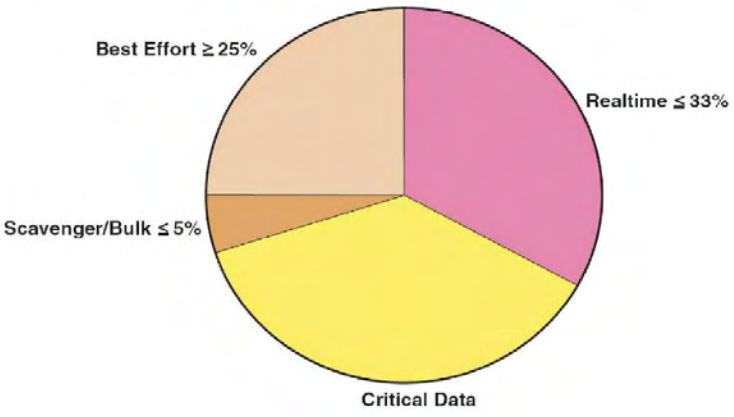

Figure 2-1 CISCO's Real-time, Best Effort and Scavenger Queuing Rules

If the platforms support a variety of queuing structures, network administrators should configure consistent queuing policies according to platform capabilities to ensure consistent PHBs (Per-Hop Behaviors). On a platform that only supports four queues with Class of Service (CoS) based admission (such as a Catalyst switch) a basic queuing policy could be as follows:

•Real-time (33%)

•Critical Data (42%)

•Best Effort Data (25%)

•Scavenger/Bulk (5%)

2.8 Review of Related work

For the purpose of getting a good understanding on the research problems and to have a good knowledge on how to conduct the research as well as to get a good methodology to use the researcher has reviewed several literatures of researches conducted on similar issues. In addition to that, to have a clear understanding of problem and its associated issues the researcher has improved his knowledge on the area by reading several books that give good knowledge on the field. To mention few of the reviewed literatures we have presented them as follows.

2.8.1 Local related work

Previously several researchers have conducted their study to improve the performance of a wide area network with different approaches.

Tsegaye Berhanu [10] has studied on Performance analysis for wide area network for bandwidth optimization by taking Addis Ababa University as his case study target. In his study he has used real time data to measure the performance of the wired AAU wide area network. By using a network performance measurement tool named Solarwind Performance Monitoring, he was able to study the performance of the network by the use four network performance measurement metrics (latency, packet loss, response time and availability). His study has mainly concentrated on the wired WAN of Addis Ababa University network. By the use of SNMP performance measurement and analysis of the data captured he has developed WAN optimization framework.

HAILAY WELDEGEBRIEL [14] has studied on developing dynamic bandwidth allocation prototype model for campus network based on network traffic analysis by taking Mekele University wide area network as his case study target. In his research he used offline weblog data to analyze the bandwidth utilization and to find the most frequent requests for the internet resource. For online traffic analysis the researcher used tools named ntop which is used for analyzing protocol distribution and MRTG which is used and Network traffic analysis at different links. The offline network traffic analysis is based on the server log file which is taken from Mekele University web server. This Server log files are records of web server activity. has analyzed the university campus user bandwidth utilization trends and has proposed conceptual framework for dynamic bandwidth allocator in intercampus network.

A study made by Yacob Gobena[9] on "Developing WAN Optimization Model to improve the Performance of Business Critical Applications: The Case of UNECA" has used SNMP performance measurement methods to find the resource utilization and he also has collected device configuration and network protocol configuration in UNCA network. He has used four steps in his experimentation process that enabled him to investigated the impact of various WAN factors. And has developed a WAN optimization model that can improve the performance of business critical applications over the WAN enterprise environment. The WAN optimization model developed was expected to give a solution for WAN enterprise environments. This model is developed based on the experimental result of the study and considering of the selected metrics. The researcher used UNECA wide area network for his case study. In his experimentation he conducted collection of application's traffic flow data by using a tool called Net flow analyzer as a first step. the second step he used was to simulate and experiment the collected traffic flow data by using a tool called OPNET modeler, the third step was analysis of the simulated traffic data to investigate the challenges of the critical applications over the WAN by using different optimization methods. Finally, he developed WAN Optimization model which is considered to identify the real-time bottleneck status based on the results of the analysis.

2.8.2 Global Related Work

When we look at global related works researchers named Moria E. Villapol, Eric A. Games and Nevdith Morales[11] have conducted a study on how to improve network throughput using QoS

building blocks approach at Central University of Venezuela. They hypostasized that, there is a problem of congestion between the university and the service provider(CANTV) during office hours on weekdays. To obtain data for the study, they used MTRG and measure the performance of their WAN links. And the result indicates that the outbound traffic is fare compared to the capacity of the link. To get more specific data about congestion they an active performance measurement method by downloading data from different sites using FTP. The result they obtained from this experiment shows that there is a packet lost between the computer and their ISP router but there is almost no packet lost between the computer and the university router which is the internal side of their network.

Their researchers clearly summarized and presented different QoS building blocks framework (shaping, classification, marking and scheduling), which allows QoS at the level of packets and let the engineers and designers to select which best suit the necessities of the UCV's network users. And also they propose to modify the current network topology, that is extended start to mesh topology which connects each device(IDFs) each other.

The researchers concluded that applying the proposed solution would bring a significant performance improvement than that of over-provisioning UCV's network. But characterizing the UCV's network traffic and simulating scenarios based on the proposed solution are left for future work.

A master's thesis done by as student at Lappeenranta university by a student named Maiju Kansanen [34] has studied his research taking Lappeenranta University of Technology network for his case study. His study has taken the protocols defined in the wide area network for critical applications can cause performance problems on the organizations network. For his research he has used the four performance metrics (bandwidth, latency and delay and throughput). The research shows as there are several factors that may be considered as a cause for performance issues for a wide area network. The focus of the study is evaluating the performance of commonly used application protocols such as TCP/IP in organization networks. After finding the performance problems in organization WANs, the study focusses on methods for improving WAN performance. Accordingly, WAN acceleration is proposed as a possible solution for improving WAN performance.

Chapter 3

3 Business Understanding

In order to have a clear understanding of the target organizations wide area network architecture, services and configuration of the devices deployed in the wide area network, the researcher has made frequent visit to University campuses and has interviewed the ICT staffs responsible for specific issues. From the under taken survey the researcher was able to find first hand information's and was able to collect the necessary data's from the devices and the ICT professionals.

In this section the University WAN environment is illustrated and finally critical points that are chosen to contribute for the research are presented.

3.1 Addis Ababa University Wide Area Network

As we have tried to highlight in chapter one, Addis Ababa University is one of the biggest universities in Africa. At present, the University has 15 campuses that are geographically at different locations and with a distance ranging from 1 to 120 kilometre from the main campus. The universities wide area network consists different size local area connections with different number of users.

The wide area network connectivity is composed of optical fibber connection and MPLS-VPN connection subscribed from the Ethio-telecom data line service to connect remote campuses local area networks including the main campuses local area network.

The connectivity types and subscribed connectivity bandwidth are depicted in the table below.

No	Acronym	Name	Internet BW (Mbps)	WAN link	VPN BW (Mbps)
1	MC	Main Campus (6killo)	600	Fiber	40
2	CBE	Collage of Business & Economics			
3	CNS	Collage of Natural Science (4 killo)			
4	YSM	Yared School of Music			
5	SFA	School of Fine Arts			
6	DS	Dental School			
7	AAiT	A.A. Institute of Technology (5 killo)	300 + Share with MC		

8	SCJ	School of Communication and Journalism	Share with MC	MPLS-VPN	4
12	SPH	Saint Paul Hospital		MPLS-VPN	4
10	SAHS	School of Agriculture and Health science (Fiche)	6	MPLS-VPN	4
9	CVMA	Collage of Veterinary Medicine	20	MPLS-VPN	4
11	SOC	School of Commerce	10	MPLS-VPN	20
13	CHS	Collage of Health Science	250	MPLS-VPN	4
14	EiABC	Ethiopian Institute of Architecture and Building Collage	100	MPLS-VPN	4
Total			**1286 Mbps**		**84 Mbps**

Table 3-1 AAU WAN link connectivity types bandwidth

3.1.1 Addis Ababa University Wired WAN

The university wide area network consists wired and MPLS-VPN connectivity for its remote campuses connection. The wired wide area network of the university consists of seven campuses including the main campus where the main datacentre and core of the entire wide area network of the university is located. These campuses are located in a 3killometer radius from the main datacentre where the fibber termination is made to. This WAN consists seven campuses of the university. Namely, Main campus (6killo), Addis Ababa Institute of Technology (5killo), Science campus (4killo), Collage of Business and Economics, Dental school, School of fine arts and Yared school. These seven campuses are connected with mesh topology for redundancy purpose.

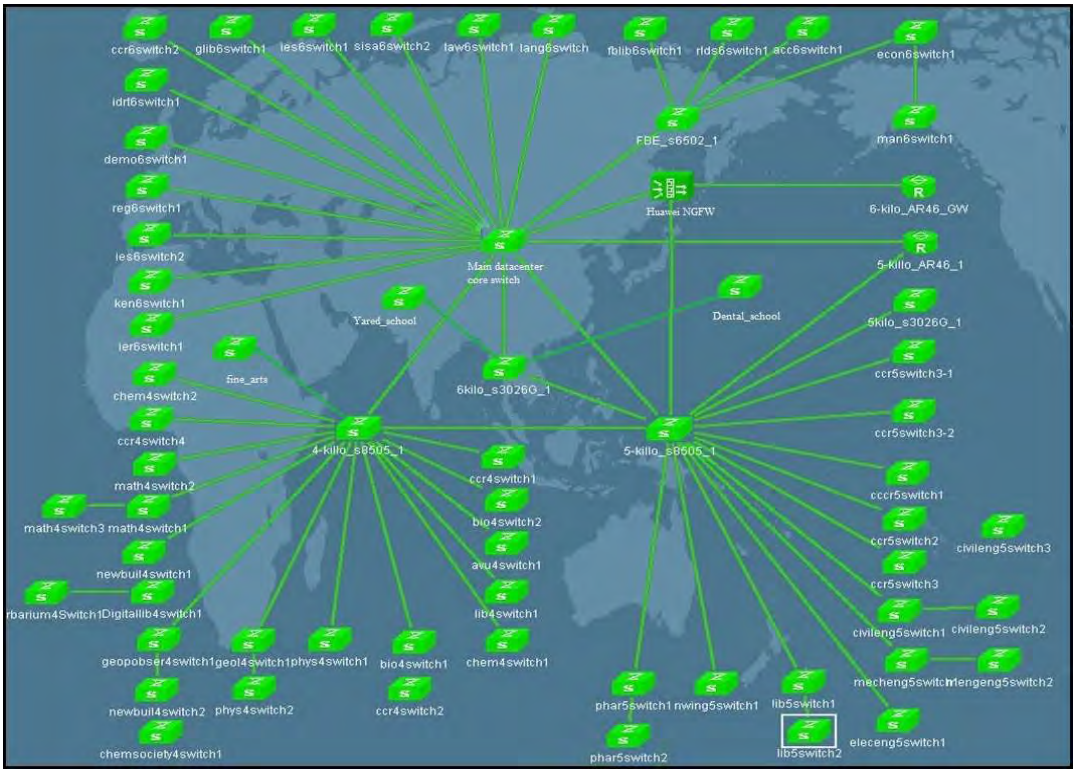

Figure 3-1 Wired AAU WAN

3.1.2 Addis Ababa University MPLS-VPN connected WAN

The remote campuses network architecture is designed to serve the campuses users need on its own. Basic services like DHCP, DNS have been deployed in their own server rooms. All the remote campuses of the university have their own datacentres and have devices that are capable of handling requests of services necessary for their local area network users. The network devices deployed in the local area networks of the remote campuses have a majority Huawei product and there are also few unmanaged hubs in their local area network.

The campuses network is divided with several virtual local area networks. As we have to highlight in chapter one, Addis Ababa University is one of the biggest universities in Africa. At present, the University has 15 campuses that are geographically at different locations and with a distance ranging from 1 to 120 kilometre from the main campus. The universities wide area network consists different size local area connections with different number of users. The campuses networks are sub divided in to different virtual local area networks to make ease of the network management and administration processes.

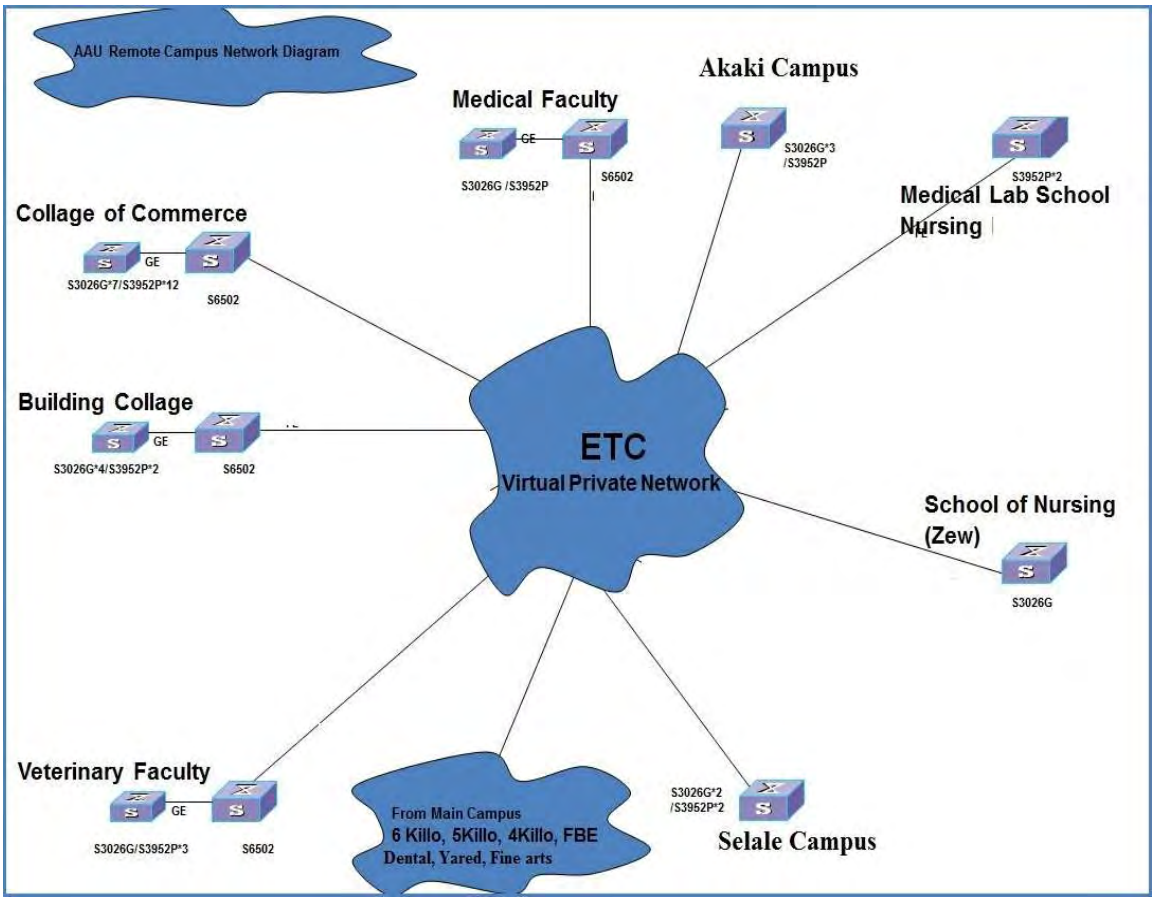

Figure 3-2 AAU MPLS-VPN WAN

3.2 Addis Ababa University Data Centers

The university has small to medium size datacentres in all its remote campuses. Regardless of their size and standards, the datacentres are structured to give all the requirements of services that is expected from the local area network users. The network infrastructure of the remote campuses is a three-layer architecture with merged approach. The core devices in the datacentre are majorly layer 3 devices which are configured to perform the inter-VLAN routing for the local area networks. Some of the campuses datacentres are designed to support disaster recovery and backup sites for the main datacentre. Because of their old structures and lack of concern, most of the remote campuses datacentres don't meet most of the criteria's of datacentre standards but they are serving in an acceptable condition for the connectivity purposes required.

3.3 Addis Ababa University Main Data Center

The main datacenter of the university is the core point where all the remote campuses and all of the university network users relay on. The main datacenter hosts almost all of the centralized services for the university community. It serves as an internet gateway for most of the remote campuses users.

In addition to that the main data center is the link for the seven campuses connected with fiber optics to the rest of the campuses that use VPN connectivity to connect to the university wide area network.

Figure 3-3 AAU Main Data Center

3.3.1 Main Data Center Network Infrastructure

The university datacenter architecture is designed as a three-layer architecture that consists of the core layer, distribution layer and access layer structure. The main datacenter network design has 3 main network sections. These are server farm network, Internal network and DMZ network. The server farm network holds services that run for the internal network users and the resources that are deployed in the server farm can only be accessed from AAU network only. Which means they are not accessible from any network outside AAU or the internet. The DMZ network holds services that the university made for the public access like the website of the university.

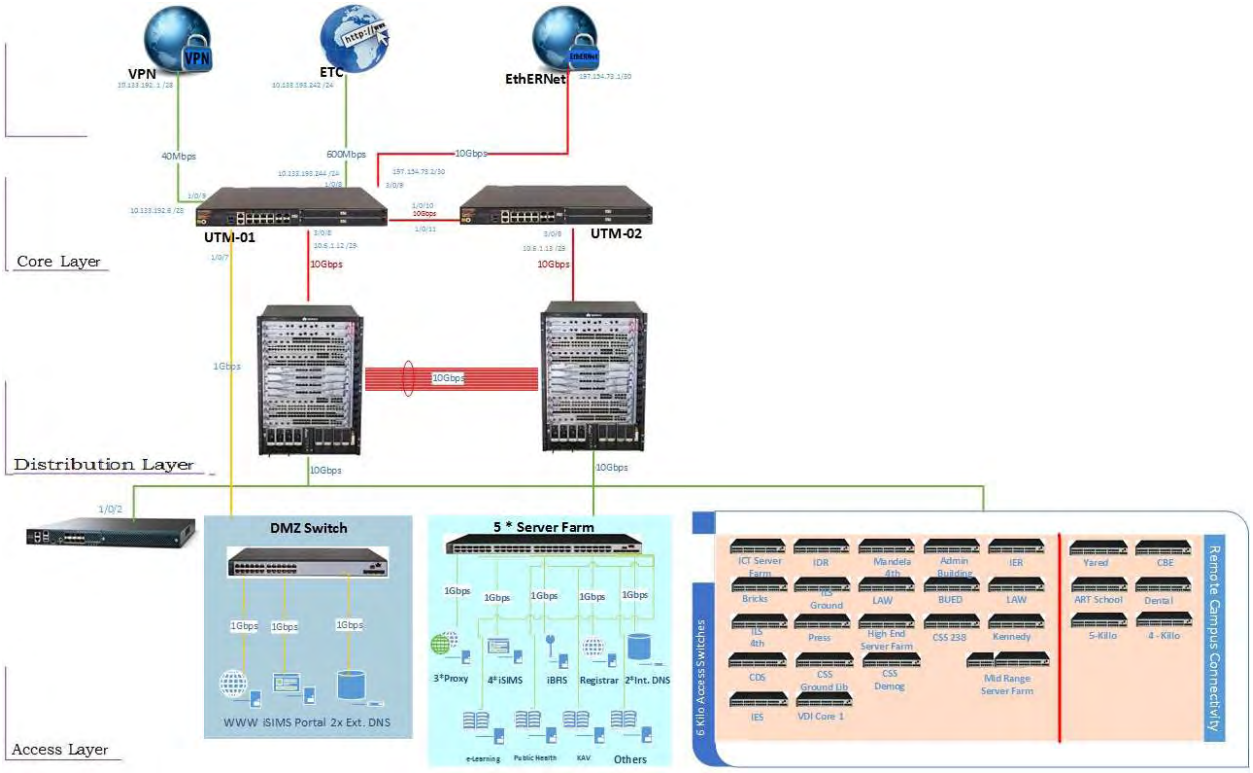

Figure 3-4 AAU Network Infrastructure

3.3.2 Main Data Center Core layer connectivity

It is most recommended that core switches and core routers be redundant. Core switches and core routers are placed in the core layer. This ensures layer 3 availability of the gateway IP address when one of the routers goes down. In most cases the routing and switching capability will be integrated into one device[39]. Addis Ababa Universities Core layer connectivity has 10Gps connectivity to the distribution. The excess amount connectivity bandwidth has provided the core

layer to be free from congestion and to have high reliability of data packet delivery. AAU network has redundant devices for the core layer connectivity. The core layer devices are configured with redundant links having 10Gbps bandwidth with each other. At the core layer of the network there are two types of devices. The core switch and the unified threat management device which also operates for path selection activity.

3.3.3 Main Data Center Distribution Layer Connectivity

This layer primarily provides layer 2 distribution through a switched network. This distribution can be solved by any layer 2 loop-free protocol. Among the most widely used solutions are the Spanning Tree Protocol, or some of the more developed protocols like RSTP and MSTP [39]. The Universities distribution layer connectivity has un uplink of 10Gps and a down link of 1Gbps. All the devices in the distribution layer are Huawei products. Although it is not uniform the spanning tree protocols are defined on the layer 2 devices.

3.3.4 Main Data Center Access Layer Connectivity

Access layer is used to connect end devices like servers, user's computers, printers, etc. In most cases servers are connected via two or more network adapters. For other devices which do not support resilient network connections we terminate the resilient topology on the last switch[39]. Addis Ababa University access layer connectivity has 1Gps uplink and 1Gps downlink connection. Though there are few layer two devices with 1Gps downlink.

3.4 WAN Services

The universities ICT infrastructure expected to deliver different services to the university community. In addition to the provision of connectivity there are technical services that it is required to provide user community. These are services such as Domain Name Service (DNS), E-mail services for facilitating managed internal communication without relying on the internet service, active directory services which enable users of the university community use resources anywhere in the network by roaming their profiles.

The university service is also expected to provide the necessary security for user data and should ensure the privacy related issues. The other set of service expected to provide direct support to

end-users in their normal day-to-day activities by providing access to resources like centralized storage for data sharing, automation of operational activities services e-learning, etc.

The university WAN service has almost all the basic technical services that are expected from one higher education institution. At the current state the university is changing its operational processes from manual to automated systems. The university has recently implemented Integrated Student Information Management System (ISIMS). In order to automate the finance information system, the university has deployed integrated finance management system that is controlled by the ministry of finance and economic development.

The finance system is deployed in the countries wereda-net WAN and the university has added an additional link to its WAN gateway. In addition to the newly added services different services and applications are deployed in the main datacenter housed in Sidist Kilo campus. This can show us that the university WAN is being used by several services and applications like Web Service using WordPress, e-mail Service using Gmail, e-Learning Service using Moodle, Doman Name System using BIND, Proxy Server using squid, Library System, eGranary/Internet in a BOX, Video Conferencing, DHCP Service, FTP Service, UCIS /University College Information System, Desktop cloud using Huawei solution, IP telephony service using Huawei solution.

3.4.1 Services Platform

Most of the services in AAU are Linux or Unix based services. Using the Unix and the Linux platform for the server side applications has enabled the ICT save cost of license and has made the systems not to be affected by viruses. Linux systems are well known for their ability to run for years without failure; in fact, the system administrators have mentioned as Linux systems have never seen a crashed. That's great for users of every kind, but it's especially for academic environments with a lot of concurrent connection which causes service crashes. Linux also handles a large number of processes running at once much better than other operating system platform does. In Linux systems there is no need for rebooting the system whenever there is a change in configuration. Whereas in Windows configuration changes typically require a reboot causing inevitable downtime. The other advantage is that Linux, which is based on Unix, was designed from the start to be a multiuser operating system. Only the administrator, or root user, has

administrative privileges, and fewer users and applications have permission to access the kernel or each other. That keeps everything modular and protected.

3.4.1.1 Proxy Servers

A Proxy Server is a server that acts as a gateway between a Local Area Network (LAN) and the Internet[40]. In the Addis Ababa University's network, the ICT office has a three proxy servers using Squid software to handle this task. There are three Proxy Servers in the AAU's network which are assigned to a domain name of cache.aau.edu.et. These are 10.90.10.70, 10.90.10.11**,** and 10.90.10.55**.**

Two of them, 10.90.10.70 and 10.90.10.55 are running on a Unix Operating System on Solaris 10 & 10.90.10.70 is running on a Linux Operating System on Ubuntu Server 12.04. The Proxy servers provide an increased performance and security, in some cases; the proxy servers log can be used to monitor employee's use of outside resources using SARG (Squid Analysis Report Generator) software and SAWMILL to generate the report of resource utilization. and in other cases the proxy servers can be used as a security mechanism using SquidGuard software. The Proxy Servers are servers that offer a computer network services to allow clients on the local network to make indirect network connection to other network services on the internet [41]. A client connects to the proxy servers, and then requests a connection, file, or other resources available on different servers.

How proxy servers work

The proxy server provides the resource either by connecting to the specified server from the Internet or by serving it from the cache, stored on the server[40]. A common proxy application is caching web proxy; this is what is implemented in AAU's network. This provides a nearby cache of web pages and files available on a remote web server/internet, allowing the local network clients to access them more quickly or reliably.

When it receives a request for the web resources (specified by a URL) a caching proxy looks for the resulting URL in its local cache. If the Proxy Servers found it locally it returns the document immediately from the cache. Otherwise Proxy Servers fetches it from the remote server/internet, return the results to the client on the local network and saves a copy in the cache/disk. The cache

usually uses an expiry algorithm to remove documents from the cache, according to their age, size and access history. If the cache stays long the Proxy Serves disk will be full, so this is mandatory to remove the cache after some time[14].

Analyzing traffic load by using proxy servers

The SAWMILL and SARG software's use the squid log files that are stored in the */usr/local/squid/log/* directory of the system and with the name of cache.log file to generate a detailed report. For the purpose of this study we have collected eight days' log files from all the three proxy servers of the university.

Benefits of using proxy

Squid is a proxy server and web cache daemon. It has a wide variety of uses, from speeding up a web server by caching repeated requests, to caching web, DNS and other computer network lookups for a group of people sharing network resources, to aiding security by filtering traffic. Although primarily used for HTTP and FTP, Squid includes limited support for several other protocols including TLS, SSL, Internet Gopher and HTTPS [14].

Caching is a way to store requested Internet objects (e.g. data like web pages) available via the HTTP, FTP, and Gopher protocols on a system closer to the requesting site. Web browsers can then use the local Squid cache as a proxy HTTP server, reducing access time as well as bandwidth consumption. This is often useful for Internet service providers to increase speed to their customers, and LANs that share an Internet connection. Because it is also a proxy (i.e. it behaves like a client on behalf of the real client), it can provide some anonymity and security. However, it also can introduce significant privacy concerns as it can log a lot of data including URLs requested, the exact date and time, the name and version of the requester's web browser and operating system, and the referrer [40].

A client program (e.g. browser) either has to specify explicitly the proxy server it wants to use (typical for ISP customers), or it could be using a proxy without any extra configuration: "transparent caching", in which case all outgoing HTTP requests are intercepted by Squid and all responses are cached. The latter is typically a corporate set-up (all clients are on the same LAN)

and often introduces the privacy concerns mentioned above. Squid is primarily designed to run on Unix-like systems but it also runs on Windows-based systems also.

3.4.1.2 Web Service

As per the need to announce and have an external and internal communication medium the university. is to advance the university's mission, brand and messages to core constituents which are prospective students, their families, trustees, alumni, organizations, the media and casual visitors. The medium is designed in a dynamic and user-friendly and can readily adapt to meet the diverse and changing needs of its constituents and, provides targeted and user-friendly interaction (e.g. prospective student inquiries, requests for information, online applications, online giving, etc.) with key constituents. The other purpose is to provide information to the campus community (i.e. faculty, staff, administration, students, trustees, alumni, parents). This will be done in a manner that is relevant, fresh and engaging (i.e. mission centric, current information and easily navigable).

3.4.1.3 e-Learning Service

As an academic institution the AAU uses the E-learning service for its many advantages, acting as an addition or support to regular classes. e-Learning Service is not intended to replace traditional teaching but to enhance and support certain aspects, improving the overall quality of university teaching. E-learning includes integration of media into teaching and uses a central platform for organizing communication processes. E-learning also helps for several purposes like provision of learning materials irrespective of time and location via a learning platform, activating learning in groups as well as self-learning, supports construction of knowledge and competence through communicative and collaborative assignments, active study methods instead of one-way presentation of material etc. For this service the university has deployed a Linux based server platform which runs Moodle E-learning software.

3.4.1.4 UCIS /University College Information System

This system is implemented for undertaking the process of automating registrar system. The university registrar uses this system for keeping the records of students' information. The system is developed as a web based system and uses MSSQL for its database.

3.5 AAU Network Management and Performance Measurement Tools

AAU uses different kinds of network performance measurement tools. Amon those tools Nagios, cacti, Solar wind are few of them. Majorly the ICT relies on the monitoring tools that are deployed with the network devices like Huawei e-sight solution to measure the performance and status of devices in the network. The eSight software has an end-to-end Network Management System (NMS) which can operate big networks like networks of educational institutions, and government agencies. Network devices, quality, and services are depicted visually, providing instant insights and guiding the user through workflows optimized for managing the IP network.

3.6 Performance Baseline

A network performance baseline has two things in it, which are network performance data and network metrics used to state the normal everyday working conditions of a network infrastructure. Several researches have given several procedures for network performance baseline[8]. In most books it is recommended to use manufacturers baseline as performance baselines. For the purpose of this research we have chosen to use manufacturers baselines for network performance data like interface utilization, CPU usage, device alert etc. But for network metrics we have used baselines that are approved by IETF[7].

3.7 Data Collection

The data collection stage of this work gathers the raw data from the network monitoring systems and proxy servers log files. This has been done by SNMP performance measurements, which is one of the performance measurement methods to collect data about the network status and devices performance like PRTG and eSight to conduct the measurements. The other data that we collection method is undertaken by passive data collection, collection of log files from proxy servers.

3.8 Network Traffic Analysis

In stage two, the raw data are processed in different ways to gather useful information about the measurements [42]. The network traffic analysis is done through analyzing network traffic from the proxy log files. At this step deep investigation has been done on the network traffic source, destination, file types, URL's, amount of data downloaded and number of requests per user.

Chapter Four

4 Result and Discussion

4.1 Introduction

In this section we have tried to present the result of analysis from two angles of our study. Traffic analysis from a proxy log files and the measurement of network performance metrics with respect of quality of service standard for an enterprise wide area network. We have also tried to present the proportion of bandwidth utilization trend in Addis Ababa University wide area network.

4.2 Traffic Analysis Using Proxy Log

The traffic analysis in AAU WAN environment is done by analyzing proxy server logs by using several parameters. Parameters like Page request, visits or number of users requesting, bandwidth or size of data traffic and page views have been used. Using those parameters, we have tried to analyze the traffic from traffic source which means source of the traffic (user who generated the traffic and the campus it belongs to) and its bandwidth utilization ratio, bandwidth abuse by streaming video, playing online games, identifying number of users from the traffic source etc. have been done.

By analyzing web traffic data, we were able to get a total number of users that were using the main datacenter infrastructure to be 13,305. In order to get a good statistical approach to the problem we have clustered the users based on campus and their utilization similarity.

4.2.1 Campus based Clustering

From a total number of 13,305 users registered in the 8 days' traffic data we have created 14 clusters of network users for our data analysis. By the use of identifying their source IP we have categorized those users in to their respective campus. Here follows the summarized figure of traffic with its parameters.

Source IP	Users	Requests	Page views	Sessions	Utilization in GB	Utilization Ratio
6Killo	3088	19203098	16590067	24815	1,901.92	19.22%
4Killo	2906	19292951	16536832	18506	1,937.10	19.58%
5Killo	2471	31190008	27872119	13140	3,384.43	34.20%
CBE	740	8442594	7687880	4684	606.81	6.13%
VDI	664	17171347	14064278	12445	1,413.01	14.28%
SRV farm	142	856434	719010	440	169.38	1.71%
Journalism	259	1119637	916480	1760	96.97	0.98%
Sefere Selam	151	747268	679467	713	41.00	0.41%
CHS	1949	3416760	2822500	6259	140.45	1.42%
EiABC	55	25348	20724	126	1.43	0.01%
CVMA	512	2294931	1927879	2072	96.84	0.98%
SoC	58	430766	360382	372	33.76	0.34%
Zewditu	141	393599	323353	584	22.33	0.23%
Paster	169	509422	406899	696	49.78	0.50%
Total	**13305**	**105094163**	**90927870**	**86612**	**9895.205485**	**100.00%**

Table 4-1 Statistics of traffic registered in 8 days

4.2.2 Bandwidth Resource Utilization Ratio

Bandwidth Resource Utilization Ratio (BWUR) is the percentage of traffic generated by each cluster from the total of the population. It can be calculated by dividing the amount of traffic that is generated from the cluster to the total sum of the traffic that is generated from all the clusters. To present it in the form of percentage we multiply the result by 100%.

$$BWUR = \frac{Traffic\ generated\ by\ the\ cluster}{Total\ Traffic\ generated}\ X\quad 100\%$$

To clearly show the resource utilization of the campuses we have chosen to show the result of our analysis by using a pie chart. The following figure shows the resource utilization ratio of each cluster or campus at the core layer of the main datacenter network infrastructure.

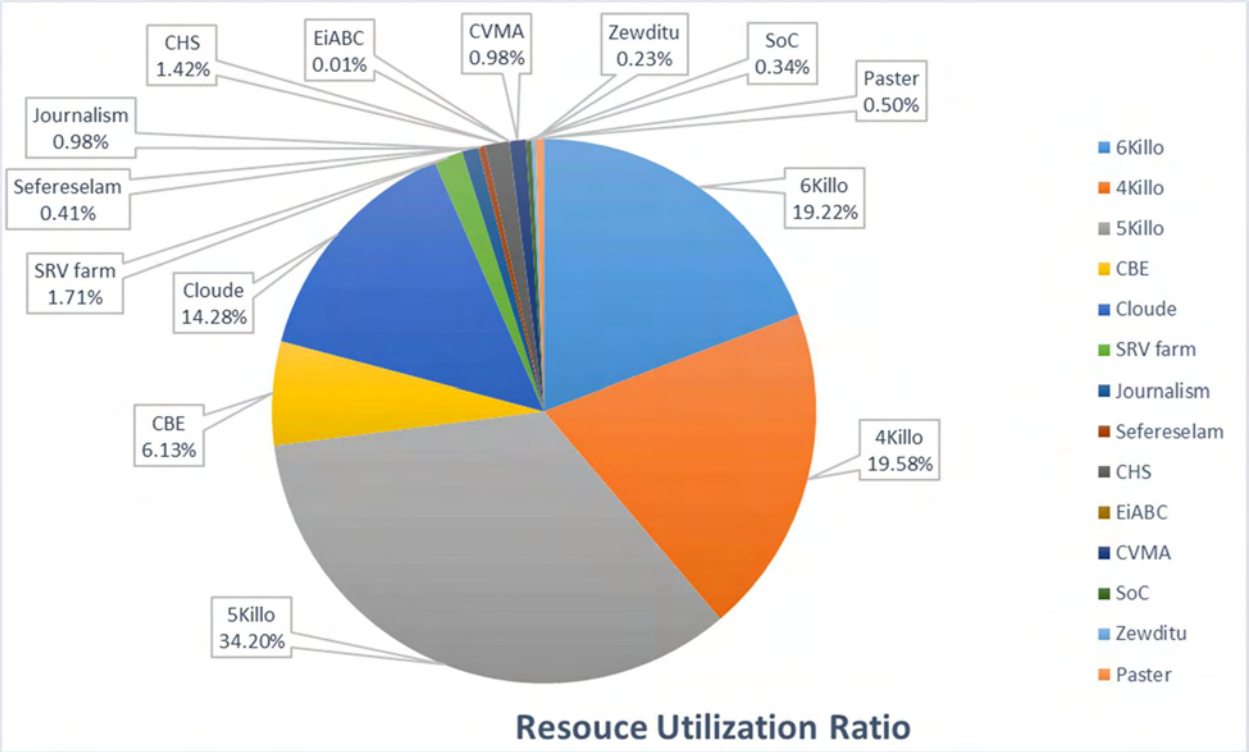

Figure 4-1 Bandwidth Resource Utilization Ratio

Bandwidth resource utilization ratio representation, 5Killo campus uses 34.20 percent of the total bandwidth resource. As we have tried to show the internet subscription of the university, 5Killo campus has 300MB bandwidth connectivity that is dedicated for the campus and in addition to that it shares 34.2% of the main campus internet resource. Which shows as the cluster uses 205.2Mbps (34.2% X 600Mbps) from the main campus internet bandwidth and 300Mbps from the internet line dedicated for itself. This shows as the campus consumes a total of 505.20Mbps.

Next to 5Killo campus 6Killo, 4Killo and the desktop cloud clusters utilize majority of the bandwidth resource. The desktop cloud is implemented in the 6Killo and 4Killo campuses students dormitory and library to be used by the students. Which means as 14.2% of the resource is utilized by the students who use the desktop cloud only.

In order to show the proportion of the number of users in each campus or cluster and the amount of data traffic that the users are generating, we have presented it in the following bar-chart.

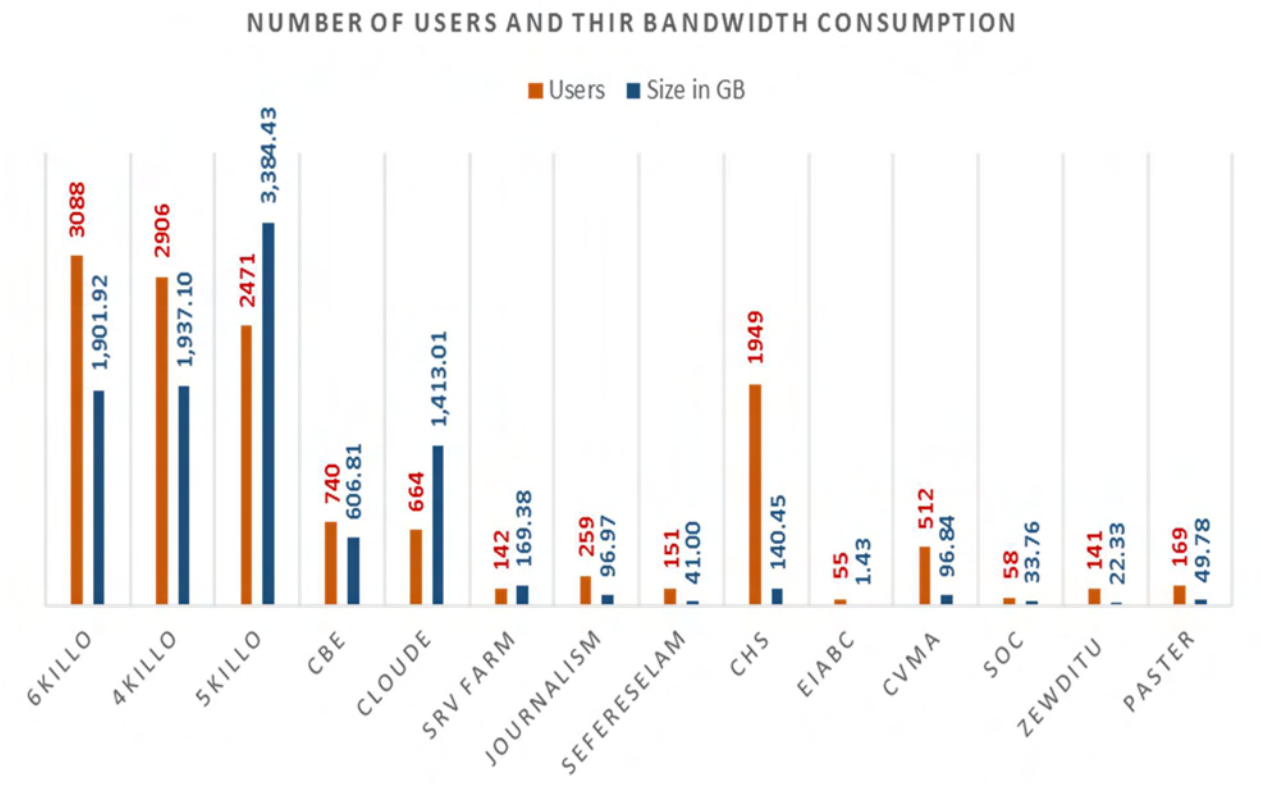

Figure 4-2 No of users and their bandwidth consumption

4.2.3 Resource Utilization Based Clustering

Resource utilization based clustering mainly focuses on the individual users and the cluster is made by ranking the users based on their bandwidth resource utilization amount. In order to clearly understand the resource utilization proportion, we have created five clusters of users based on their bandwidth consumption similarity. To create the five clusters first, we arranged the users based on their resource utilization rank in descending order. Next we clustered the first top 10% of the users as the supper bandwidth consumers. Next we created the preceding 20% of the users as high bandwidth consumers. We then clustered the next 40% of the users as average bandwidth users. The users that are below the average bandwidth consumers are clustered as low bandwidth consumers and they are 20% of the total user. The last cluster is the cluster which consists the bandwidth starved users which are the remaining 10% of the population.

Source IP	Cluster	Traffic in GB	Percentage of BW
First 10% users	**Supper Bandwidth consumers**	7,042.87	71.19%
11-30% users	**High bandwidth consumers**	2,301.75	23.27%
31-70% users	**Average Bandwidth consumers**	536.73	5.43%
71-90% users	**Low Bandwidth consumers**	11.99	0.12%
Last 10% users	**Bandwidth starved Users**	0.14	0.0014%
Total		**9,893.49**	**100.00%**

Table 4-2 Resource Utilization Based Clustering

4.2.4 Using Resources Responsibly

As an academic environment use of the university network and ICT resources should support the basic missions of the University in teaching, learning and research. Users of AAU network and computer resources are expected to properly use and protect information resources and to respect the rights of others. AAU provides a wide variety of computing and networking resources to all qualified members of the University community. Access to computers, computing systems, and networks owned by the University is a privilege which imposes certain responsibilities and obligations and which is granted subject to University rules and user ethics. All users of these resources must have to comply with some specific policies and guidelines that governs their use, and act responsibility while using shared computing and network resources, including wireless.

4.2.5 Bandwidth Utilization trend

In principle all users are expected to have the same amount of bandwidth sharing right for the same type of service for the same role in the organization regardless of their network cluster. As we have explained earlier, we have categorized users of the network in to five categories based on their utilization trend. The first 10% of the users are the most traffic generating users which utilize the majority of the bandwidth. The next 20% of the users are the users that are high bandwidth utilization categories of users next to the first top 10% of users. The next category of users are the average users which are 40% of the total users and consuming smaller amount of bandwidth than the first two classes of the network users. The other last 20% of network users are the category of users that are experiencing bandwidth shortage. The last 10% of users are users that are simply using very little of network bandwidth.

Users in rank	Requests	Page views	No. of users	Sessions	Traffic in GB	Percentage of BW
First 10%	58,173,085	51,464,877	1,331	18,026	7,042.87	71.19%
Second 20%	30,443,042	25,577,212	2,651	25,134	2,301.75	23.27%
Third 40%	15,265,058	12,769,700	5,333	31,336	536.73	5.43%
Fourth 20%	1,111,744	1,018,606	2,662	9,610	11.99	0.12%
Fifth 10%	88,690	87,389	1,328	2,496	0.14	0.0014%
Total	**105,081,619**	**90,917,784**	**13,305**	**86,602**	**9,893.49**	**100.00%**

Table 4-3 Resource utilization based on classification of users

Based on the result of our analysis from the proxy log files of the eight days, we can easily see that majority of the bandwidth is getting consumed by a very small number of users. The first 1,331users quantify 10% of the total network users that are using the main datacenter network for internet and other services. This 10% of users are utilizing 71.9% of the total bandwidth resource. Below Pie chart (Fig. 4.3) shows the ratio of network resource utilization ratio by bandwidth abusers.

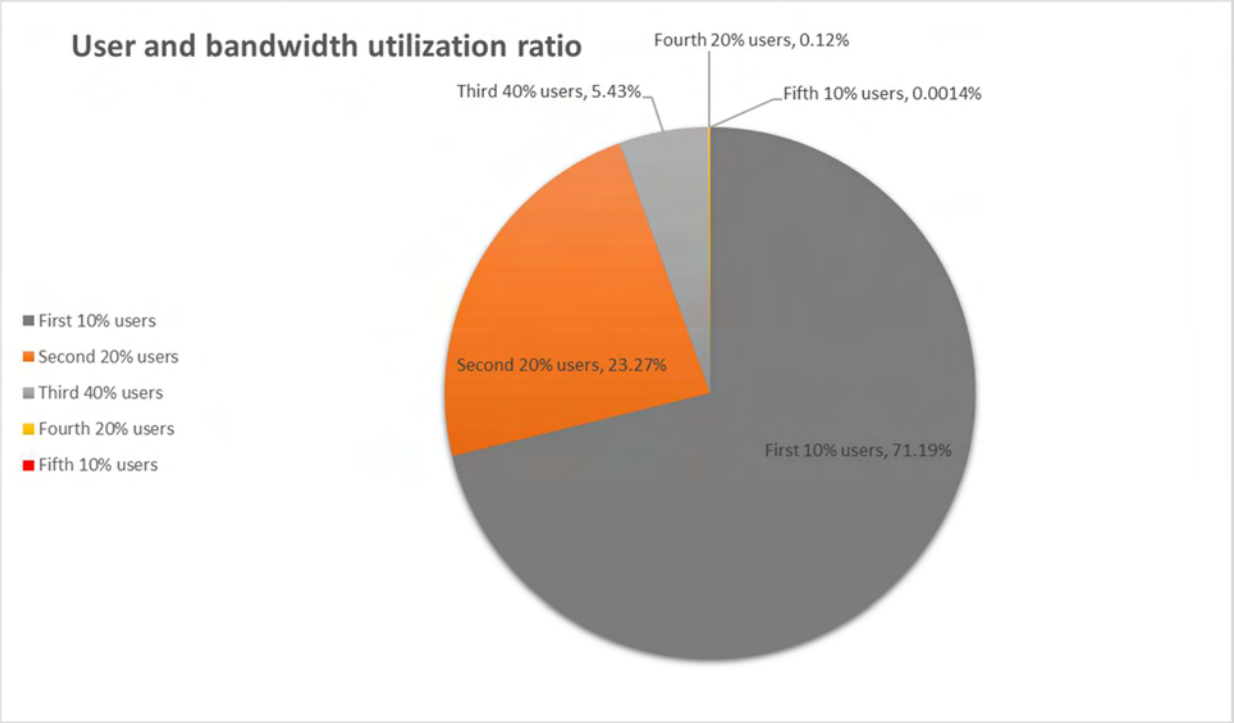

Figure 4-3 Users bandwidth utilization ratio

Since all users have different interest and different level of greediness in use over the available network resource, the bandwidth consumption of some group of users is quite different from the

other user groups. In the case of AAU network, the bandwidth resource utilization has a big difference with the number of users consuming the bandwidth resource. The group that has majority of the users is consuming a very small amount of the bandwidth resource. To show the relationship of the number of users and size of traffic generated by the users we have presented it with the following bar-chart.

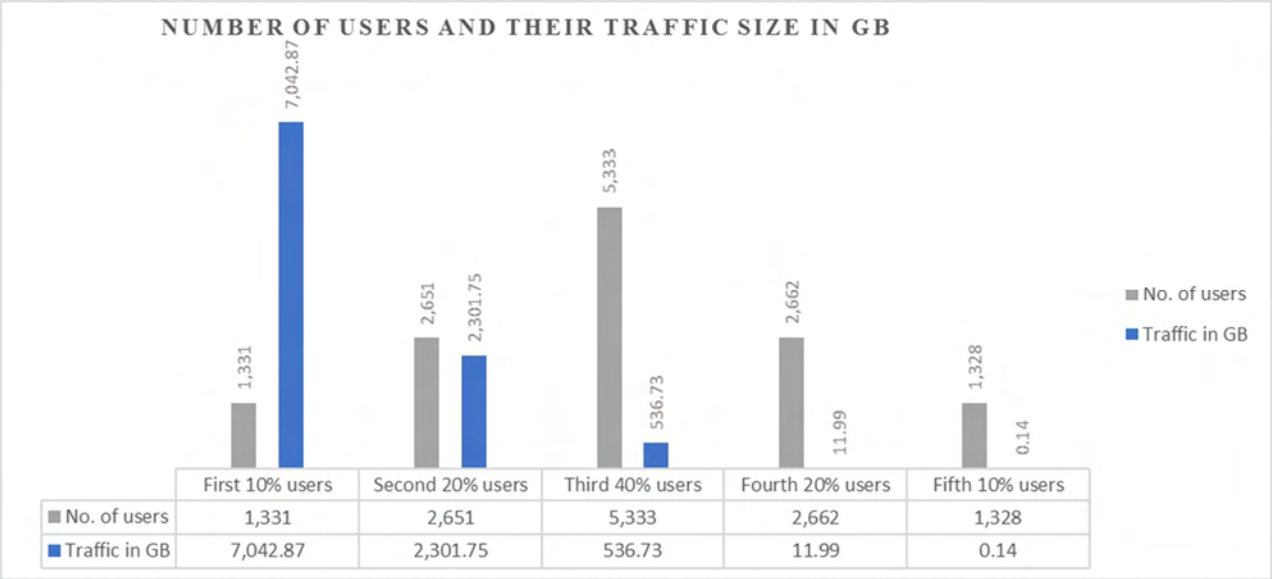

Figure 4-4 Number of users and the size of traffic they generate

4.3 Traffic based on Page domain categories

While analyzing the traffic types based on the traffic URL we were able to see that only few of the requested URL's domain are unrelated to educational purposes. In order to clearly understand the interest of the users in surfing the internet we have tried to categorize the most frequently accessed domains in to their respective groups. The groups are named based on Squid Guard domain membership and we the naming of the categories' is given by the researcher to create a good clarity on the subject matter. The categories' created are explained as follows.

Video: this category is containing list of URL's that are viewed by the user to download video files to directly stream video contents from the web. This category majorly has three domains (.googlevideo.com, .youtube.com and .xvideo.com).

Game: this category contains URL's that are categorized by the squid guard as a game site. This sites contain online games that are used to create a group game or play single player mode games.

Social: this category is for the sites that are categorized as a social media sites like Facebook, twitter, LinkedIn etc…

News: this category contains sites are used to broadcast news. In this category we have grouped the URL's sub-domains of sites that contain news content in it. For example, Yahoo site contains yahoo news and yahoo mail, so we have categorized the visit for the '. yahoo.news.com' as a news domain and 'mail.yahoo.com' as a Mail domain.

Mail: domains that are created for mail service like gmail.com are grouped in this category. This category contains sub domains of sites that dedicated for mail service.

Malware: in this category we grouped sites that are accessed for being redirected by malwares that embedded on the users' browser. This category mainly has sites that are categorized by the squid guard domain as a redirector and malware domain.

AAUMCA (Addis Ababa University Mission Critical Applications): this category has the universities main application URLs' traffics that are generated for the universities day to day operational activities. URL's that are requested to access the universities local applications like library system, eLearning systems, registrar system, etc… are categorized as AAUMCA.

IP Browsing: in this category we have sites that are not accessed by domain names. The sites that are accessed by port number 80 (http port) and do not have domain name are assumed to be accessed for specific purpose. The sites that are accessed by public IP addresses that is not subscribed by the university are groped in this category.

Tutorial: the sites domains that are grouped as a tutorial are the sites that contain educational contents. These domains mostly contain video and audio contents. So, in order to avoid redundant domain category, the researcher has taken an action of excluding domains that have been categorized in other domains groups from the search domain.

Other: The category created as 'other', contains all the rest of the URLs' that are not previously included in any of the above domains.

While categorizing URLs' in to their respective domains, we were able to find two most frequent domains request for video (. googlevideo.com and xvidio.com). Though their number of request

is too small, their utilization ratio is beyond it is expected. To show the bandwidth consumption of some domains we have presented it as follows.

Domain category	Requests	Page views	Users	Traffic in GB	Traffic ratio
Video	14,545,827	14,338,124	407,968	5,024.82	50.79%
Game	932,997	890,998	4,078	14.06	0.14%
Social	2,414,828	2,368,238	109,305	67.21	0.68%
News	166,744	80,471	10,815	8.33	0.08%
Mail	1,083,483	894,485	39,295	51.95	0.53%
Malware	606,154	606,154	16,468	56.79	0.57%
AAU MCA	829,519	745,393	2,834	18.26	0.18%
IP Browsing	1,438,869	1,396,242	22,167	172.45	1.74%
Software update	5,962,253	5,910,613	118,149	123.21	1.25%
Tutorial	1,403,763	1,319,436	1,376	4.68	0.05%
Other	75,697,182	62,367,630	4,590,898	4,351.85	43.99%
Total	**105,081,619**	**90,917,784**	**5,323,353**	**9,893.60**	**100.00%**

Table 4-42 Bandwidth utilization based on top visited URL's domain categories

From the above table we have seen as the number of request to access the video sites is much smaller than the number of request to access other domains. But the bandwidth consumption of video sites is more than the total sum of the bandwidth consumed all other domain categories. To clearly show the bandwidth resource consumption of the video streaming and download of the audio video files we have presented it with the following pie-chart.

$$BUR = \frac{Traffic\ generated\ to\ access\ the\ domain}{Total\ Traffic\ generated}\ X \quad 100\%$$

Where, BUR: Bandwidth Utilization Ratio

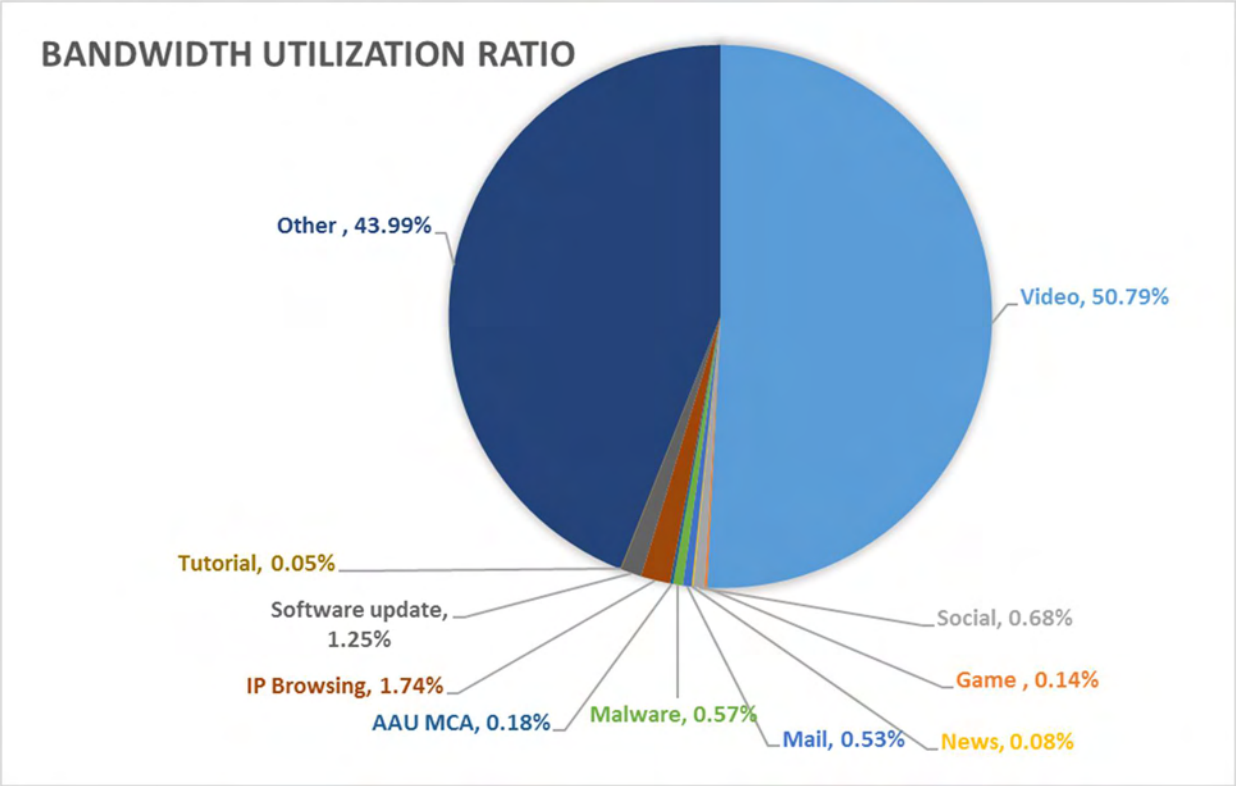

Figure 4-5 Bandwidth utilization ratio based on traffic categories

4.4 Request Ratio

In statistical calculation ratio means the division of the value of part of a group to the total sum of the values in the group or cluster in consideration. Request ratio is the amount of share that a specific domain to the total sum of the domains that are considered in the study. The request ratio figure shows the amount of requests that a specific domain is requested with respect to the total sum of all other domains. The request ratio to some extent signifies the interest of the users in accessing the resources.

$$RR = \frac{No.\ of\ request\ to\ the\ domain}{Total\ no.\ of\ requests}\ X \quad 100\%$$

Where, RR: Request ratio

From the result of our analysis we have seen as the ratio of request by the domains that are categorized in video domain is 13.84%, but the ratio of traffic utilization by the domain is 50.78%

of the total traffic size. This shows as the requests of video streaming and video files download requires high bandwidth.

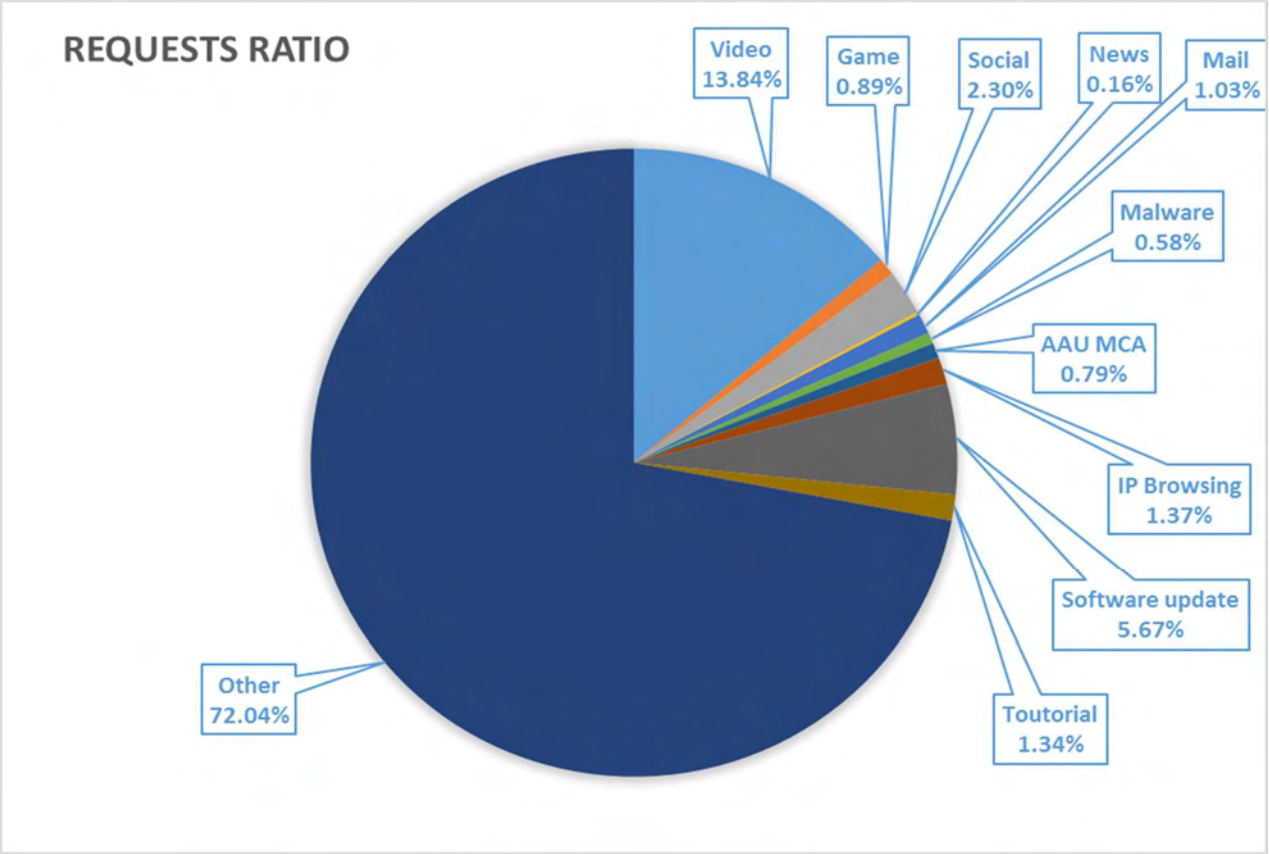

Figure 4-6 Request ratio

On the other hand, when we analyze the number of users who access the domains, we have seen as the number of users that are accessing the domains that consume very high bandwidth is very small.

4.5 Network Performance Analysis

In the network performance analysis section, we will present the results that we have discovered while analyzing the network performance based on the network performance metrics. Degrading or decreasing network performance in an already set up computer network system is the most

undesirable condition. As we have tried to show in chapter two of this work there are several factors which contribute to the decline of network performance, which their signs can be observed from quality changes in Quality of Service parameters measurement result [32]. The four main properties for a network connections or Quality of Service parameters are bandwidth, delay, jitter and throughput.

4.6 Network Performance factors

As we have seen on chapter two of this thesis, the cause of performance failure in a network, we have understood as resource shortage is one of the cause. So to clearly identify the performance failure reason we have started our study from analyzing the hardware resource utilization level of the core devices in the universities datacenters. Since the CPU usage and the Memory usage of the devices have a major impact on the network we have tried to see the amount of hardware resource utilization from the two sides.

4.6.1 CPU Utilization

CPU utilization refers to a devices usage of processing resources, or the amount of work handled by a CPU. Actual CPU utilization varies depending on the amount and type of managed routing and switching tasks. Certain tasks require heavy CPU time, while others require less because of non-CPU resource requirements. CPU utilization may be used to gauge system performance. For example, a heavy load with 100% CPU utilization may indicate insufficient CPU power support, or running at its maximum level of capacity [43].

CPU utilization indicates how much work load the device is performing relative to its total capacity. High percentage of CPU utilization mostly happen in networks with slow performance where services on the router fail to respond, for example, slow response in Telnet or unable to Telnet to the router, slow response on the console, slow or no response to ping, router does not send routing updates to other routers, high buffer failures[43].

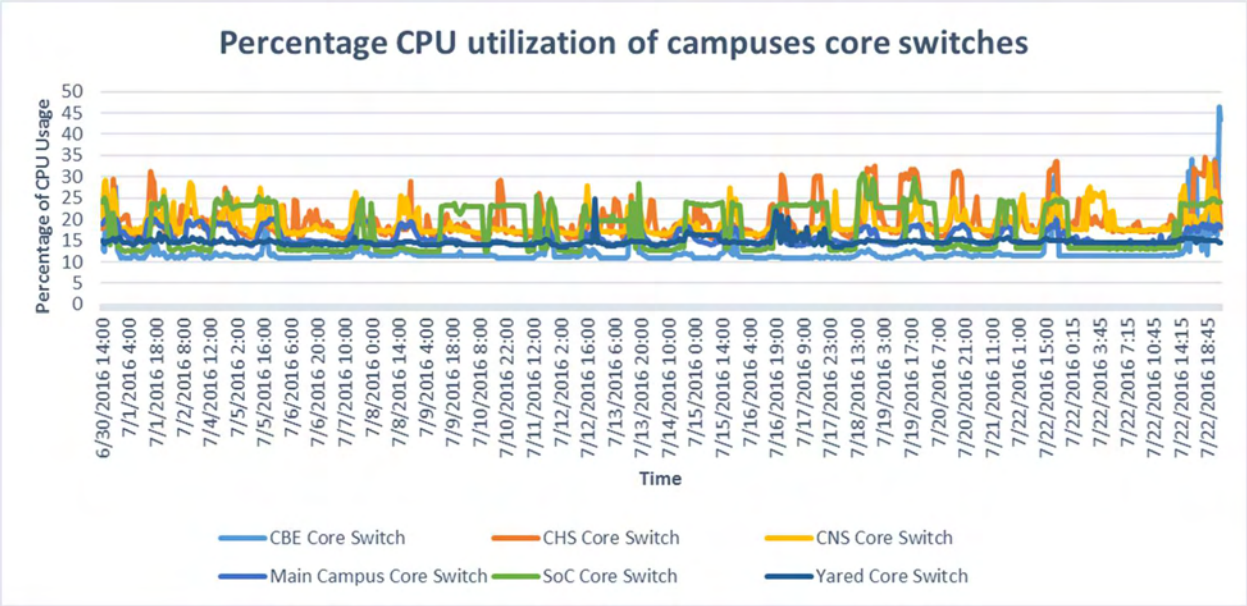

Figure 4-7 Percentage CPU utilization of campuses core switches

	CBE	CHS	CNS	Main Campus	SoC	Yared
Max. Value	46.33 %	34.5 %	33 %	20.67 %	30.67 %	24.67 %
Min. Value	10.84 %	15.38 %	16.4 %	13.89 %	12.33 %	13.5 %
Average	12.32 %	19.63 %	19.05 %	15.95 %	17.89 %	14.70 %

Table 4-5 Statistical values of 3 weeks data on CPU utilization of campuses core switches

As we have seen on Fig.4-7 above the CPU utilization of the core switch devices in all campuses shows as the devices processor is getting utilized below its average performance and it is a good indicator that the core network devices do not have problem associated with data processing. Table 4-5 Shows the statistical maximum value, minimum value and average of the CPU utilization of the respective campuses from a 3 weeks' data.

4.6.2 Memory Utilization

Like the CPU, memory is one of the resource that devices performance is affected. Network devices memory utilization ratio indicates the data load the device is having with respect to its total capacity. When there is a data that is received by the device to forward it to its destination, the device uses its memory as a temporary storage to process the task. Having a high figure in the memory utilization shows as there is either problem in processing the packets to send them to their

destination or there is a very high traffic load. From the 3 weeks' data we collected from the devices log and the eSight monitoring system the memory usage of the devices in the campuses show as the devices are operating in a good memory utilization percentage. Huawei devices resource utilization percentage threshold is 50% and none of the core switches in the campuses have reached the threshold level. Fig 4-8 shows the average memory usage of the devices from the 3 weeks' data collected from eSight monitoring system

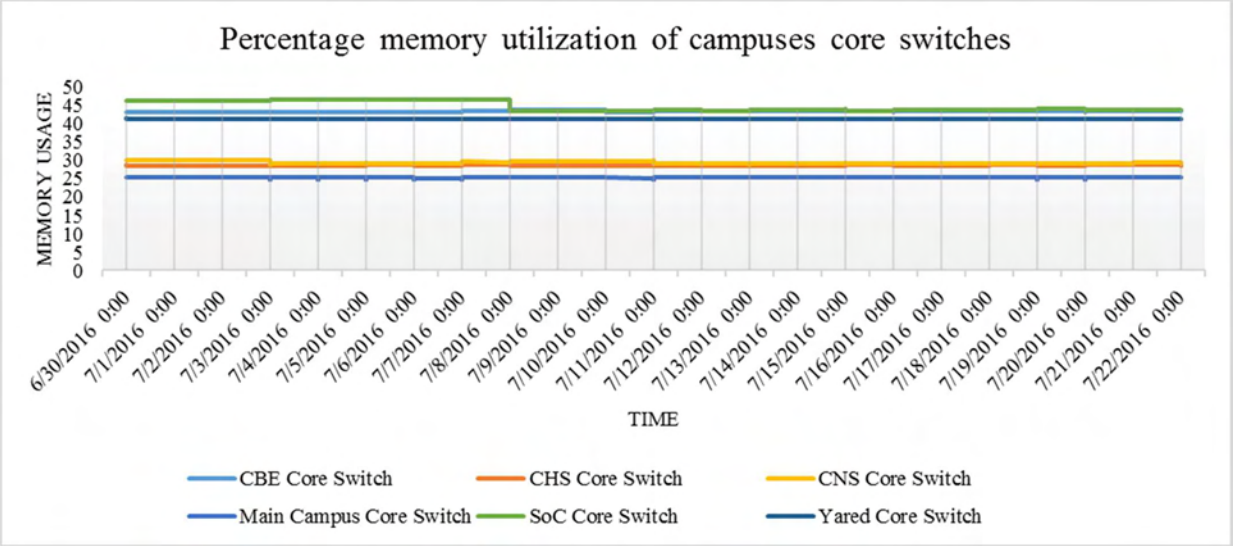

Figure 4-8 Percentage memory utilization of campuses core switches

From the above graph we can see that the memory utilization of core devices is also below average value. The memory utilization of the devices is different from one another which shows that the devices installed in the networks are with different workloads from one another. This condition also depends on the size of the network, number of users in the network and the traffic that the users generate to access resources from the wide area network or the internet.

	CBE	CHS	CNS	Main Campus	SoC	Yared
MAX	43.67 %	28.50 %	29.80 %	25.17 %	46.33 %	41.50 %
MIN	43.00 %	28.17 %	28.75 %	24.92 %	43.33 %	41.00 %
Average	43.24 %	28.31 %	29.12 %	25.12 %	44.52 %	41.00 %

Table 4-6 Statistical Max, Min and average value of campuses core switches memory usage

As we have seen in Table 4-6 the highest memory utilization figure is registered by the SoC (School of Commerce) core switch, which is 46.33% and CBE (Collage of Business and

Economics) campus core switch, which is 43.67%. Although the figure is higher than the other campuses core switches utilization ratio, it is still below the threshold value.

4.6.3 Bandwidth

As explained before bandwidth is a term used to tell the speed at which a traffic data travels in a network. Since the speed of bits in a media depends on limit of interface speed, we have also mentioned as it does not measure how fast bits of data move from one location to another. Data packets travel over electronic or fiber optic cables, the speed of each bit transferred is negligible. So, bandwidth measures how much data can flow through a specific connection at one time. The size of the bandwidth is limited by the interface speed. In the case of the WAN connectivity of AAU the link that the campuses are inter connected shows as the interface speed utilization ratio is very low. The interfaces speed of AAU WAN links for the wired media is in Gbps speed, which is too much more than the amount that the network users use. This condition has brought the core connectivity of the campuses to be fast enough.

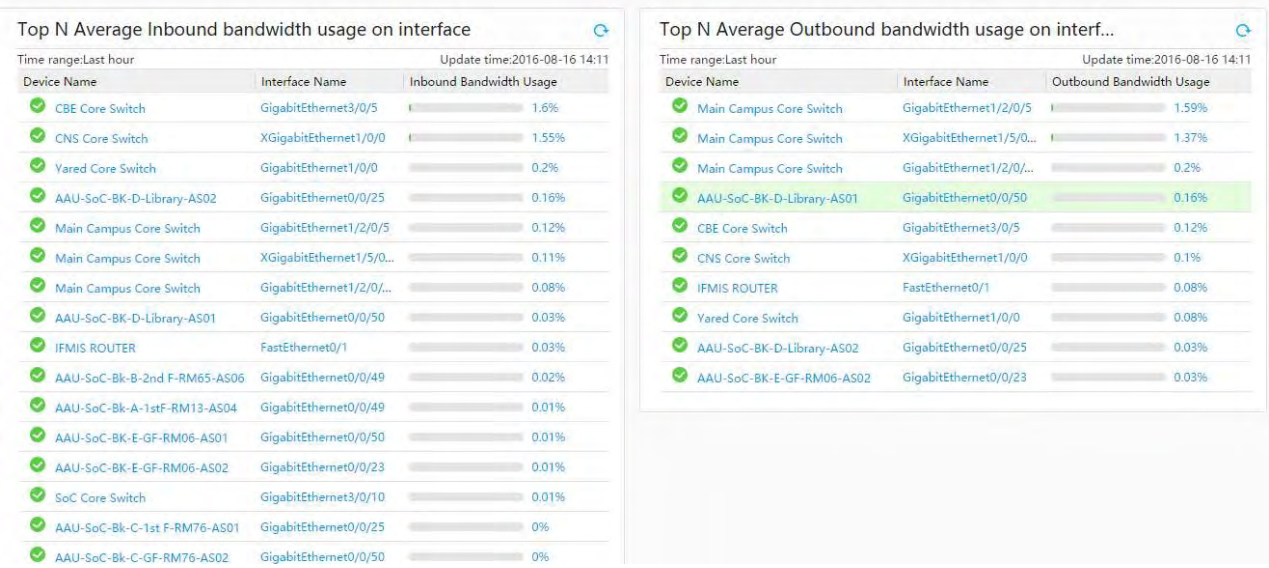

Figure 4-9 Real time interface utilization ratio captured from eSight

As we can see in fig. 4-1, the interface utilization ratio of CBE (Collage of Business and Economics) is 1.6 % inbound usage and 0.12% outbound, and the interface utilization ratio of CNS (Collage of Natural Science) is 1.55 % inbound usage and 0.1 % outbound usage. In the case of the non

4.6.4 Delay

Delay or a response time is the total time it takes after the client sends a packet till it gets a response from the server. Response time report represents the values of the server response time. Average and maximum response times are the most important characteristics of measuring network performance [10].

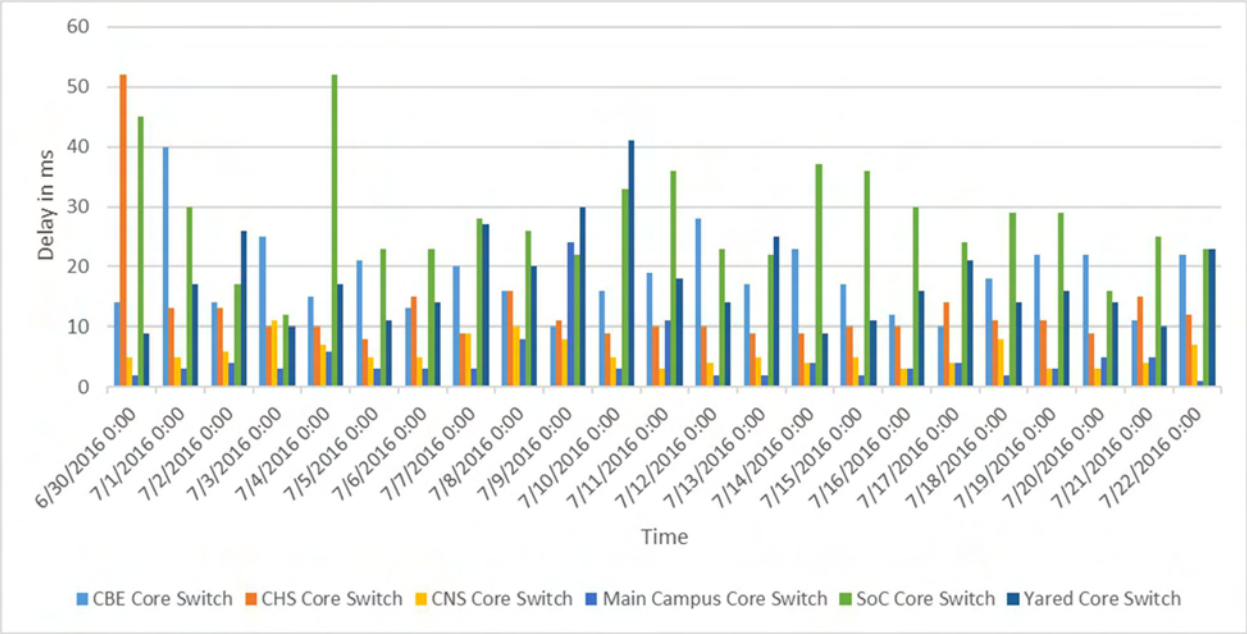

Figure 4-10 Average Delay in millisecond (ms) from the eSight to each of the core switches

	CBE	CHS	CNS	Main Campus	SoC	Yared Sschool
Max	*40ms*	*52ms*	*11ms*	*24ms*	*52ms*	*41ms*
Min	*1ms*	*1ms*	*0ms*	*0ms*	*2ms*	*1ms*
Average	*6.74ms*	*5.28ms*	*2.01ms*	*1.33ms*	*10.70ms*	*6.25ms*

Table 4-7 Average Delay in millisecond (ms) from the eSight to each of the core switches

4.7 Summary of Our Analysis

Higher education institutions network is different from other types of organizations networks in that it has a different types of users with different areas of expertise. The vast field of study in the campuses produces totally different types of need to utilize the network resource. Students who study computer related courses like programing use the network to test their codes sending some traffic through the network, to test applications that may affect the network operation or provide them more privilege than others while using resources. They may also try to attack the systems by developing their own codes. Students who study other courses like geography use applications like google earth which consumes high bandwidth. In such networks the need for having a good way of allocating resources to the proper purpose is mandatory. From our observation in the network traffic analysis by using proxy log analysis and network performance parameters analysis, we have observed as the network resource usage at the core layer of network does not have shortage of resources.

4.7.1 Proxy log analysis

From the proxy log analysis results we were able to see as more than 71% of the bandwidth is getting utilized by 10% of the users. In our campus based clustering we were able to see as the 34.2% of the bandwidth resource is getting used by 5killo campus uses. This campus has a 300Mbps internet line subscribed for use only by the 5killo campus users, but in addition to that amount of internet bandwidth, the campus users' bandwidth traffic is ranked first in sharing the main campus internet bandwidth. The number of users in 5killo campus which share the resource from the main campus are 2471 which is the third in the number of users when ranked by number of users next to 6killo and 4killo which have 3088 and 2906 users' respectively. But the resource utilization by 5killo users accounts more than double of the resource utilization by 4killo and 5killo campuses. The users that use a fiber media to connect to the main campus datacenter have a better bandwidth utilization figure than the users that are located in the campuses that use VPN media. This shows as the VPN bottleneck has affected the chance of users' resource sharing. This incident indicates as the network administrators have to create a technique that will help the users of the VPN media of connection to connect to the main campus to get prioritized service at the core layer.

In the second way of crusting, which is resource utilization based clustering, we have seen as the resource is getting utilized by only small number of users. The users that are categorized as supper bandwidth consumers consume 71.9% of the bandwidth resource. This amount of resource consumption has made the other users to be starved for bandwidth resource. Users that constitute 40% of the total users which are categorized as average bandwidth consumers are getting only 5.43% of the total bandwidth resource.

On the other hand, while analyzing the bandwidth resource consumption using domain categories we have seen as the bandwidth usage of the video site categories' have consumed more than half of the bandwidth resource which is 50.79% of the total bandwidth. Whereas the category that consists the mission critical applications of AAU consume only 0.18% of the total bandwidth resource. The category of domain that is named as other has 43.99% of utilization of bandwidth resource. This category of domain consists all types of sites that are not categorized in other domains.

When we look at the request ratio, we can see as the request for the domain that consumes the resource most, which is the video domain, consists only 13.84% of the number of requests. When we look at this ratio, the amount of request to access video sites is very high. A high amount of bandwidth may be consumed by passing a small video request. A single video download request may result in consuming high bandwidth to download a large amount of video file.

4.7.2 Network performance analysis

In order to analyze the network performance, we have started measuring the resource utilization of the devices at the core layer of AAU network. By network resources we mean the performance of the network devices that are operating at the core layer. This device list includes the core switches of the remote campuses and the core switch of the main campus which is performing switching and routing activities that arrive at the core later of the network from all campuses.

When we monitor the CPU utilization of the core devices we have observed as the devices are consuming only small amount of the central processing unit capacity that the devices have. The maximum amount of CPU utilization observed on the core switch of collage of business and economics campus which is 46.33%. This utilization amount indicates as the devices central processing unit is performing below its average capacity. The core switch which operates the

switching and routing of the traffics that come from all campuses, which is the 6killo campus core switch is using a maximum of 20.67% of its CPU capacity. The reason for having small CPU usage is because the core switches at the main campus has a better capacity than all the other core switch that are found in the other campuses.

The memory utilization rate of the devices shows as all the core switch devices are using below average capacity of the devices memory size. From the statistical analysis we have seen as the maximum memory utilization figure is observed on the school of commerce core switch which is 46.33%. This figure is the biggest figure that is ever registered in the campuses core switches memory utilization percentage, but this figure is still below average utilization level. From this we can see as the memory usage of the devices does not indicate any kind of memory shortage on the core devices.

The other factor that the core network performance is affected is, by the bandwidth that the devices interfaces are connected with each other. The percentage of utilization of the interfaces indicates as the devices interfaces are bottleneck to the core connectivity or not. From our data analysis on the core devices interfaces percentage of utilization we have observed as the interface capacity usage is below 2%. For example, the percentage of utilization of school of commerce core switch is 0.01%. This interface is the interface that the school of commerce network data traffic is forwarded through Ethio-telecom MPLS-VPN to be routed to the main campus. The MPLS-VPN bandwidth is 20Mbps. This amount of data size amounts 0.0195% of the interface capacity. From this figure we can see that the MPLS-VPN bandwidth is getting utilized at its highest level.

The response time analysis shows as some difference in the time that the network devices respond. The core switches that are connected with a fiber media are expected to have similar response time with respect to the other performance measurement analysis, which are memory utilization and CPU utilization ratio. While looking at the data collage of business and economics core switch has a higher response time figure. The average response time for the collage of business and economics core switch is 6.74ms this figure is looks high when we compare it with the other core switches that are connected with fiber media, which are 4killo and 6killo, with 2.01ms and 1.33ms average response time.

4.8 Overall summery

As we have physically observed on the visit of the main datacenters, we seen as AAU core network infrastructure has the most recent technology products and high capacity devices installed in the datacenters. AAU ICT has also interconnected the core layer and the distribution layer of the network with a fiber media which provides higher bandwidth for connection. The office has also subscribed high bandwidth for the internet service and relatively high MPLS-VPN connection bandwidth to connect campuses that are distant from the main campus. From this we can see as the office is trying to solve the issue by employing over provision technique. But from our analysis result we were able to see that the network resource is getting wasted or abused by online video streaming and downloading of video files which are not related to the institutes mission and vision. The other issue observed in the analysis was as the network resources are getting consumed by only few users that are found in different campuses. The traffic graph of the internet bandwidth shows as it is operating at its full level, but it doesn't mean that there is a bandwidth shortage for the university network users. The main thing that this incident is shows is as there is a lack of proper resource allocation for the network users.

4.9 PROPOSED SOLUTION

As we have seen on the summery part, over-provisioning has been the main mechanism used to reduce the network congestion problems of the AAU's network. But, over-provisioning presents several disadvantages. The main one is that the University is a public institution with a limited annual budget, so it cannot afford for more bandwidth due to its high cost in Ethiopian telecommunication bandwidth price. Also, overprovisioning neither ensures the necessary QoS; the lack of some QoS guaranties has led to poor user satisfaction while accessing resources from the internet. In this section, we propose a complementary solution to the network congestion problems.

According to [12] the best way to tackle network resource allocation is use of the building blocks of quality of service. When we consider other researches that are conducted to solve the congestion issues of wide area network bandwidths, most of the researches focus on solving such issues through applying quality of service techniques in several ways. One of the research that have a very good similarity with our scenario is the research conducted on academic institution network

[11] has brought a good solution to network congestion problem. This research is conducted for a campus with two internet gateways but in our scenario the AAU has multiple internet gateways on its different campuses.

The main objective of this study is to develop a suitable framework to provide QoS which is capable of determining the category of the application traffic and accordingly apply appropriate QoS measures which first identify the category of the application and then allocate the amount of bandwidth ratio it requires in order to provide fair bandwidth allocation for all the network users of AAU. The developed framework is based on the findings of experimental analysis of network performance traffic analysis conducted. The analysis result shows that the AAU WAN has a major bottleneck at the internet gateway and the MPLS-VPN lines because of high video and audio traffic is eating the bandwidth.

It is described following the *QoS building blocks framework* proposed in [7]. In this framework, the traffic and congestion control function such as, shaping, classification, marking, scheduling constitute buildings blocks for achieving QoS at the level packet. We initially classify the existing and coming applications of the institution and then describe the building blocks which may be considered in the design of the University's QoS network. After this process we will present our proposed the process and steps of our framework which is intended to solve the problem of the university network issue and finally, we present other issues which may be relevant in order to improve network performance and fault tolerance.

4.9.1 Services and Applications

Currently, the distributed applications of the university are treated as best effort. We propose to classify the existing and coming applications as section described in previous sections for traffic analysis part. For the web services we have considered the destination address as the criteria for classification. They have been categorized as Real time services, Critical services, Best effort services and Scavenger services.

Type of service	Application	Category based on Destination Address
Real time service		
	• *Video Conferencing* • *VoIP*	
Critical service	*Corporative applications (student management, Finance system, Registrar system, Library system, E-learning systems etc…)*	AAU MCA
Best effort service	*HTTP Browsing* *File transfer, etc.*	News
		Social
		Other
		Tutorial
		Mail
Scavenger service		IP Browsing Software update Game Video

Table 4-8 Service classification of AAU applications

4.9.2 Components of the framework

Our Framework is designed to operate on proxy servers with application of QoS configuration on gateway routers. The framework consists of four major building blocks to improve the network performance problem and reduce the packet loss of sensitive applications and services because of congestion over a WAN link where there is lack of enough bandwidth for the request by the Campus users. According to our study the main issue in the university network is bandwidth abuse than bandwidth shortage. Packets that are congesting the network and irrelevant to the campuses or the institutions mission should not be given equal priority to the mission critical applications. So in order to improve the performance, this study proposes a quality of service building blocks based framework that can solve these network performance issues predominantly at the WAN gateway congestion problem because of lack of sufficient network resources. The major components of the Framework are listed below:

- Traffic State Analyzer (TSA)
- Program Initiator (PI)
- Packet Classifier (PC)
- Traffic Conditioner (TC)
- Program Terminator (PT)

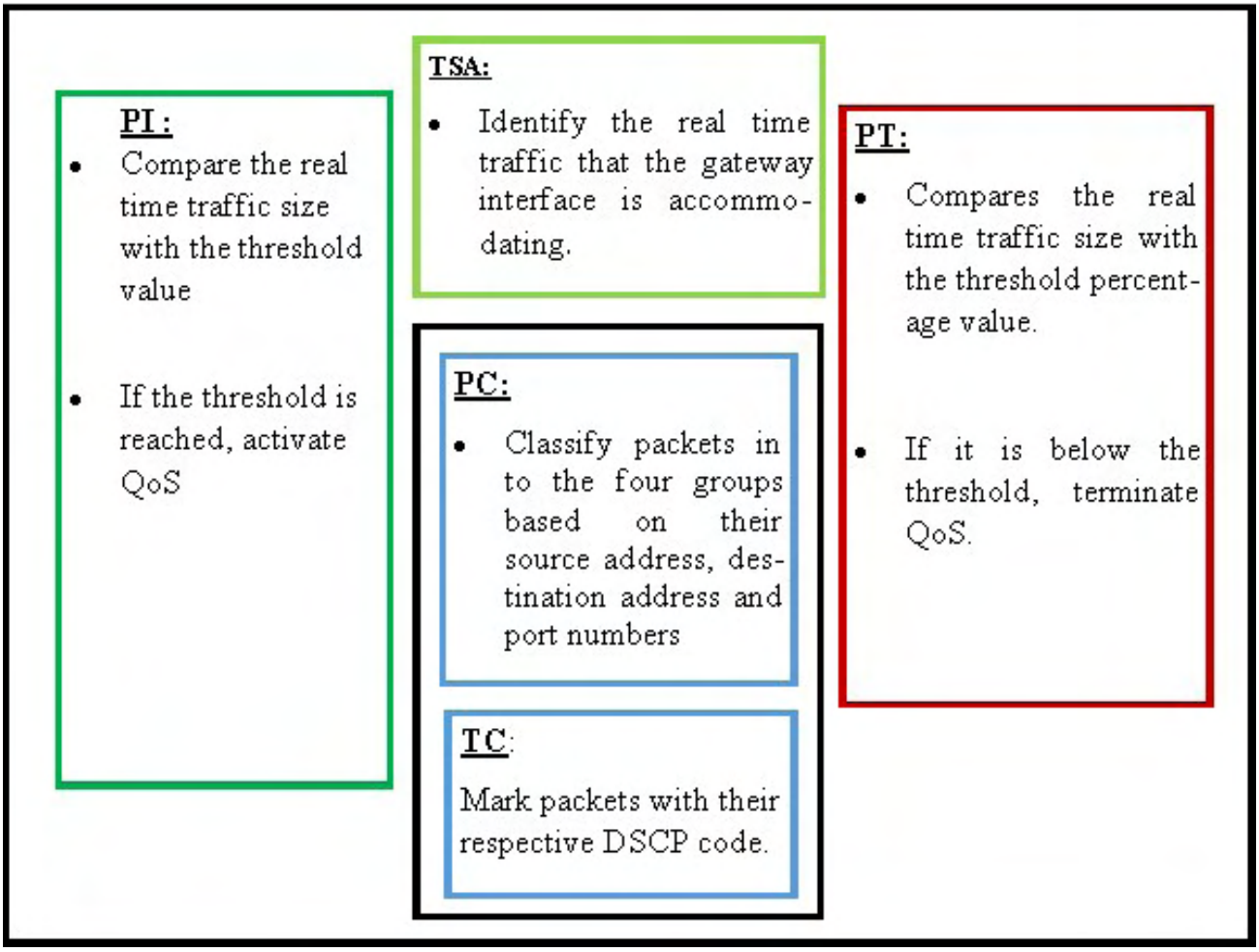

Figure 4-11 Proposed WAN optimization framework components

4.9.2.1 Traffic State Analyzer (TSA)

This component is the part of the framework that is responsible to frequently check the size of traffic that is exiting from the network. This component will run on the proxy server and it checks the amount of traffic that the proxy server is forwarding to the gateway router. Since the amount of resource after the gateway router external link will be limited the threshold value will be set as

the maximum amount of the subscribed bandwidth. This component first will convert the amount of bandwidth utilized in to percentage by using the total amount of bandwidth subscribed.

BWU = Amount of bandwidth utilized currently / Total bandwidth available *100%

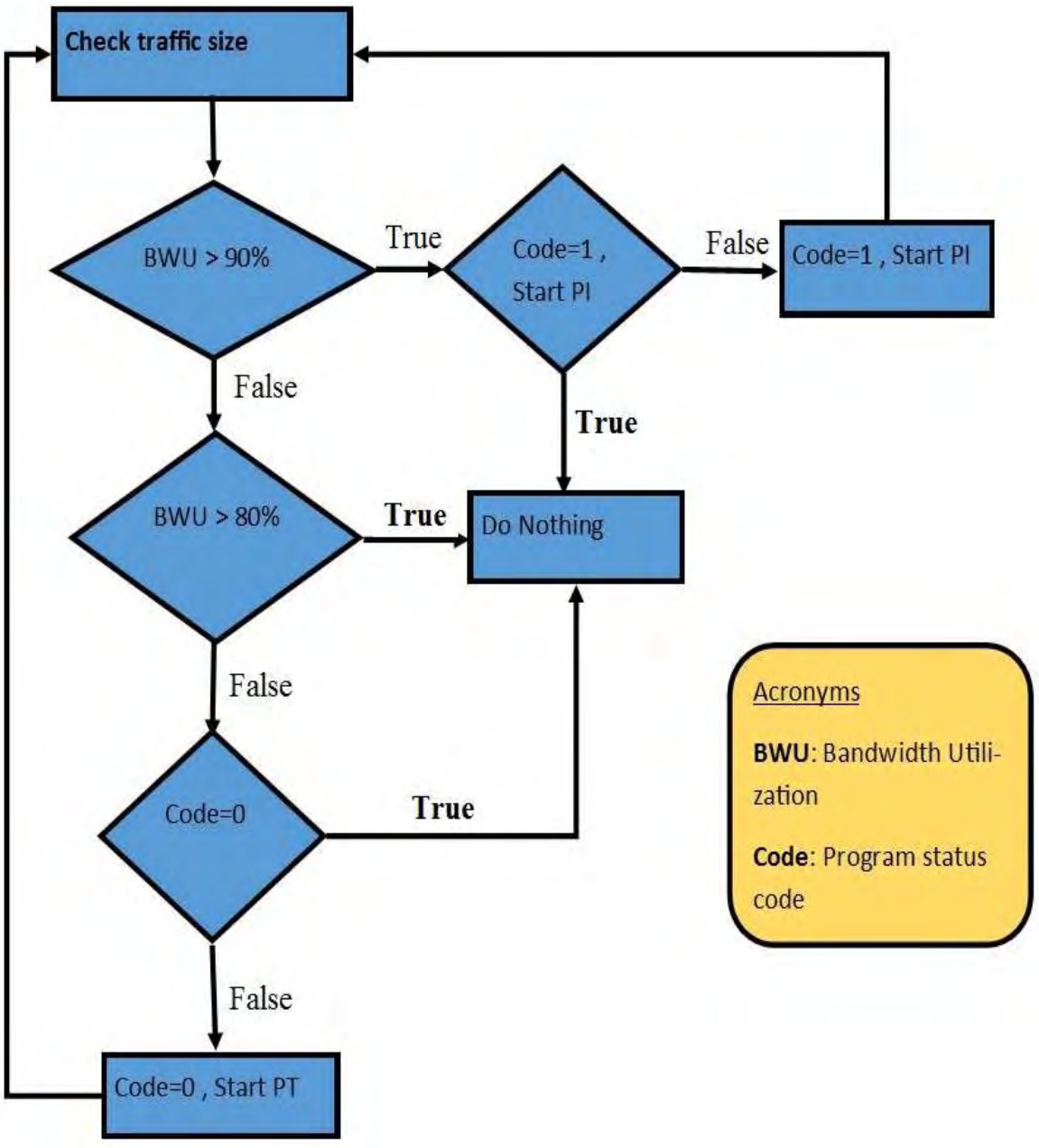

Figure 4-12 Traffic State Analyzer (TSA)

Traffic State Analyzer

```
Traffic_State_Analyser (Data type: D)
BEGIN
        READ Subsciribed_Bandwidth;
                READ traffic_size;
                        Bandwidth_Utilization = (traffic_size  / Subsciribed_Bandwidth) ;
IF (Bandwidth_Utilization > 0.9)
        {
                IF (Program_code == 1)
                        GO TO BEGIN;
                else
                        Program_Initiater ( );
                        Program_code =1;
                }
IF (Bandwidth_Utilization > 0.8)
        sleep (60000)
        GO TO BEGIN;
else
        IF (Program_code == 0)
                {
                        sleep (60000)
                        GO TO BEGIN;
        else
                Program_Terminator ( );
                Program_code = 0
        }
END IF;
GOTO BEGIN;
```

4.9.2.2 Program Initiator (PI)

Program initiated component uses a code of 0 and 1 to know whether the QoS program is active of not. When the program is activated the QoS value of the component will be assigned to 1 and the component will go to sleep until the program terminator (PT) component sends a code that tells it as the QoS value of the PT component has turned 1. In this case the PI component will be activated and turns its QoS value to 0. This component is responsible for triggering the application that is used to tag packets with a DSCP code in order to prepare them for routing. This component reads the threshold value and the real time traffic size from the first component and it takes action of initiating the QoS service application.

Program Initiater:

```
Program_initiater ( );
    {
        NET START Quality_of_Service application;
    End;
    }
```

4.9.2.3 Packet Classifier (PC)

The packet classifier component is the part that is responsible to classify the packets according to their classes of service. As it is mentioned above we have grouped classes of services in to four major classes. These are Real time services, Critical services, Best effort services and Scavenger services. Based on the conditions given by the network administrators the packet classifier component will classify the packets in to their respective class of service. In order to do that the component uses the packets source address, destination address or destination port number for classifying the packets.

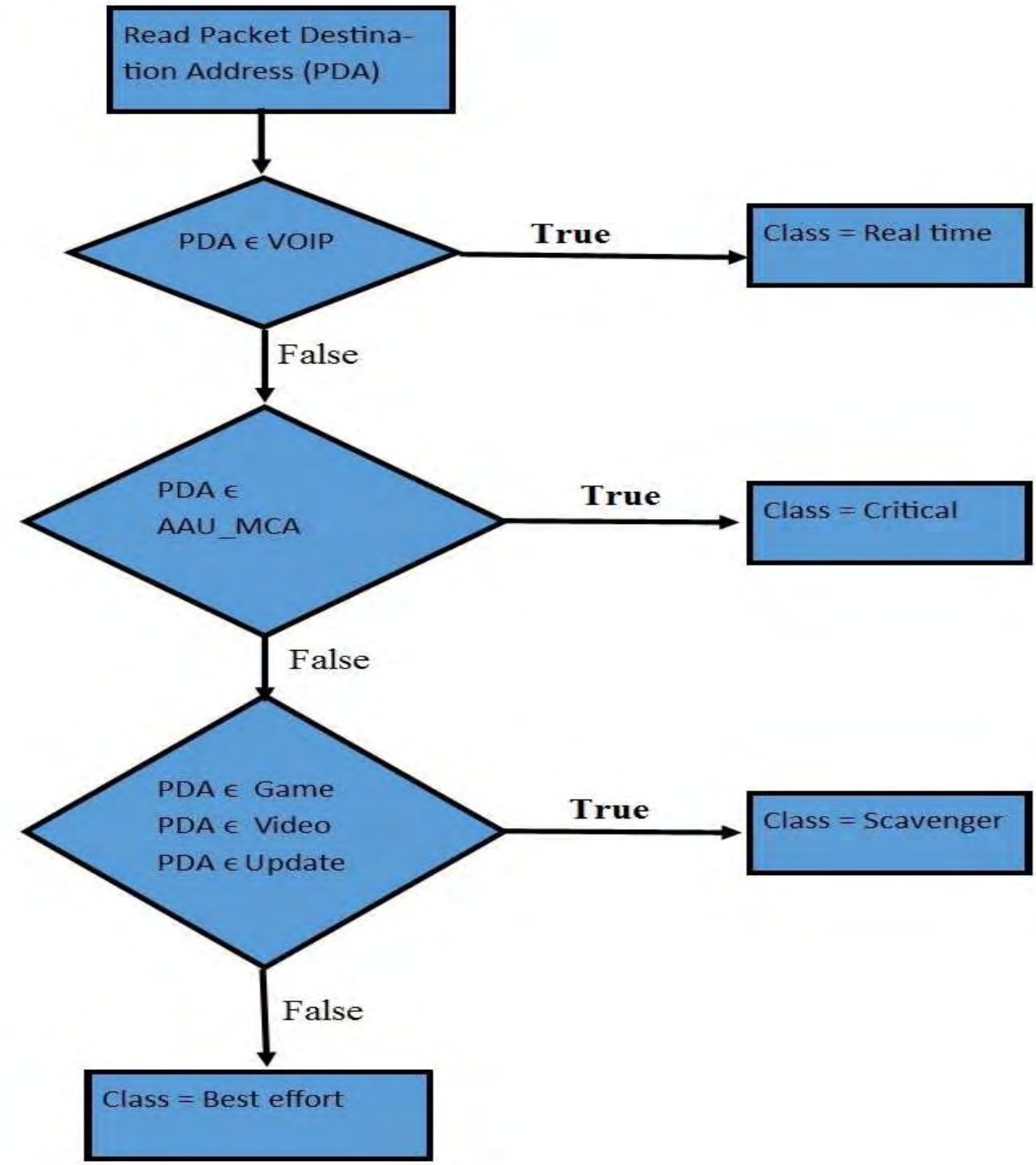

Figure 4-13 Packet classifier (PC)

Packet Classifier

```
Packet_classifier (Packet[id] )
        {
                X = GET Packet _Size
                READ Packet[id].Packet_Destination_Address;
                 For (i=0; I = x; i++)
                        If (Packet_Destination_Address == VOIP)
                                P.Class = Real-Time;
                        ELSEIF ( Packet_Destination_Address == AAU_MCA)
                        Packet[id].Class = Critical;
                        ELSEIF ( Packet_Destination_Address == games )
                        Packet[id].Class = Scavenger;
                ELSE
                        Packet[id].Class = Besteffort;
END
        }
```

4.9.2.4 Traffic Conditioner (TC)

Once the traffic is classified in to their respective class of service, the next step will be made by the traffic conditioner component. This component is responsible for tagging the packets with the DSCP code number in order for the router to forward them with the allocated amount of resource that is configured for them. This component uses four t DSCP marks for the four types of classes of services. Once the traffic is classified by the traffic classifier the next step will be to tag every packet to its respective DSCP code.

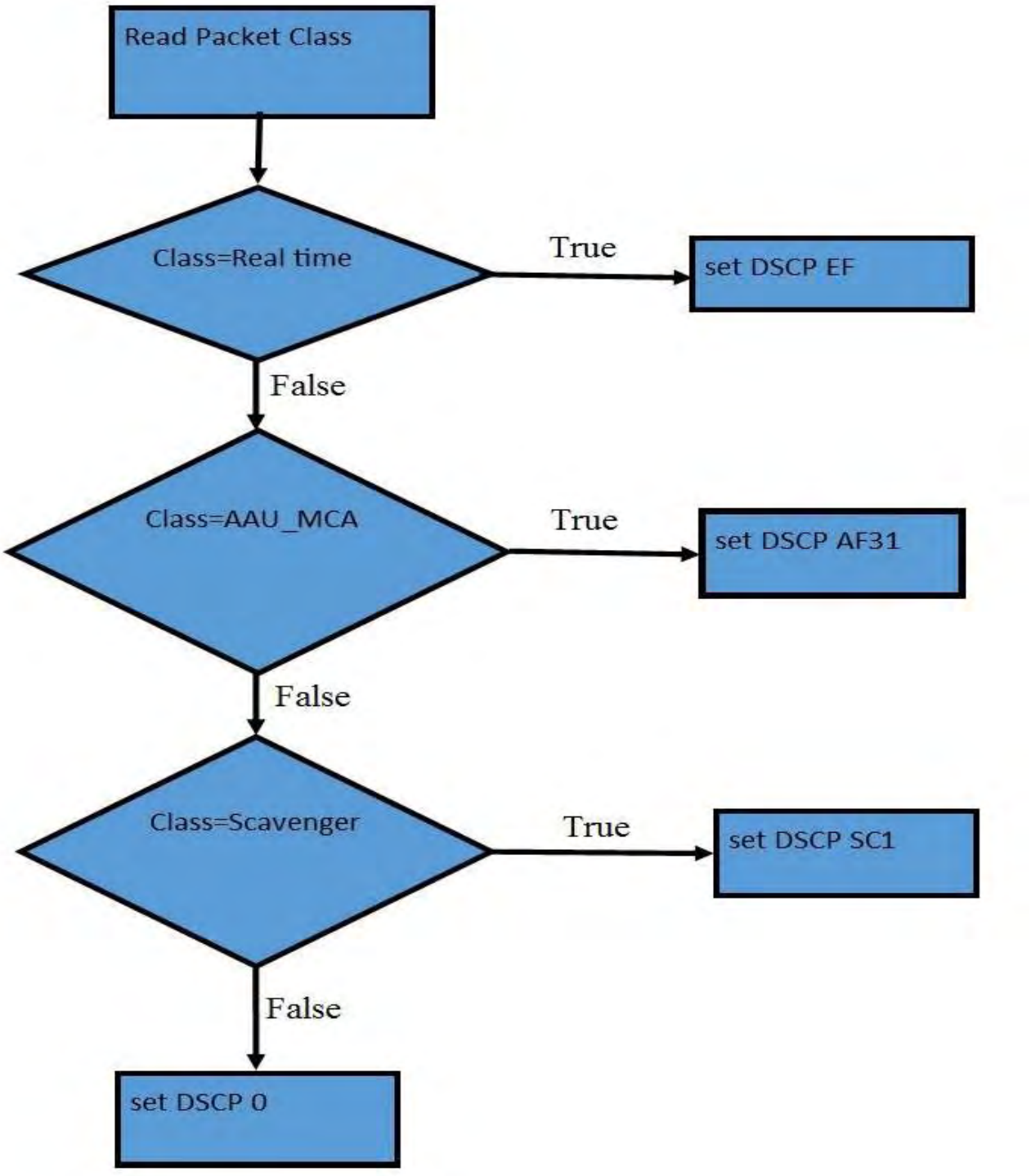

Figure 4-14Traffic Conditioner (TC)

Packet Classifier:

```
Packet_Classifier (Packet[id])
    {
        GET Packet[id].Class;
            IF (Packet[id].Class == Real_Time)
                SET DSCP = EF;
            IF (Packet[id].Class == Critical)
                SET DSCP = AF31;
            IF (Packet[id].Class == Scavenger)
                SET DSCP = CS1;
            ELSE
                SET DSCP = 0;
    }
```

4.9.2.5 Program Terminator (PT)

This component is initiated when the program initiator (PI) component sends a signal that indicates as there is congestion and as the component has activated the QoS program. This component is expected to terminate the program in order to save the resource consumption by the QoS application at times where there is no congestion in the network. The component reads the traffic size from the traffic analyzer and compares it with the threshold value, if the size of traffic is below the threshold then the program terminator component will terminate the QoS application and reports to the program activator (PA) to wait for traffic increase in order to activate the program whenever there is congestion.

Program Terminator:

```
Program_Terminator ( );
```

{

NET STOP Quality_of_Service application;

End;

}

4.10 Implementation

The above framework will be applied at the proxy servers of every campus which will directly be connected to the gateway router. The gateway routers will be configured in order to support quality of service. The approach of the framework is using differentiated quality of service framework with customized approach in order to make it suit with our countries scenario. The packet classification, packet conditioning process is taken care in the proxy servers and the scheduling work will be done on the router side. The difference of this frame work from other is that the QoS will be initiated only when there is traffic congestion. If there is no traffic congestion the network will operate as a best effort scenario.

4.11 Validation of the WAN Optimization Framework

Validation of the developed WAN optimization framework was undertaken in a real situation at AAU Network scenario. Domain experts of the university ICT office were selected from the senior ICT staffs and a brief presentation about the proposed framework was done by the researcher. The developed framework was also presented to the universities network administrator managers and experts. In order to avoid the variation of awareness among them about the framework a detailed discussion was done with the selected senior ICT network and system experts.

The evaluation process was done by filling a questioner that is prepared for the purpose of collecting evaluation of the proposed framework. While validating the developed framework, the application technique and feasibility of the developed framework has been evaluated by domain experts of the AAU ICT office. In this process three senior network administrator and two senior system administrators have actively participated in evaluating the proposed solution.

In order to evaluate the performance of the optimization framework we used the questioner depicted below:

Performance Value	**1**	**2**	**3**	**4**	**5**	NA
Description	Poor	Fair	Good	Very Good	Excellent	Not Applicable/ Not Known

Instruction: please put X mark on the appropriate value for the corresponding parameter of evaluation questions of the developed WAN optimization framework for AAU network environment.

Evaluation criteria	Performance value					
	1	2	3	4	5	NA
1. How do you rate your general understanding of the developed optimization framework?					5	
2. How do you rate the proposed implementation technique?				1	4	
3. Do the proposed framework solve the performance issue?					5	
4. How do you rate the applicability of the optimization framework?				2	3	
5. Are the components of the framework currently available in different ways					3	2
6. Do you think they can be integrated to provide the proposed framework?					3	2

Table 4-9 Optimization framework Validation questioner

As we have seen on the result of the questioner table above we are able to observe that 100% of the participants have clearly understood about the framework. This figure shows as the respondents' evaluation can be assumed as aware response. For the question prepared to evaluate the satisfaction of the target group on the framework 80% of the participants rated excellent and 20% of the participants rated very good.

For the question prepared for solution provision level of the framework the response of the participants was 100% excellent. For the question of feasibility measurement, the framework has been valued 60% excellent and 40% very good result. Our question for availability of the building blocks technique in separately we have received 60% highly available and 40% not know response. The last questing about integration of the available building blocks to develop the total framework was given the same result as the availability.

When we look at the respondents' data overall performance result from the domain experts indicate as our framework has a promising and encouraging applicability in improvement of the network performance problem in the institutions WAN environment.

Chapter five

5 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The objective with this master thesis was to provide solution for network performance problems in a wide area network with additional intention of providing network and system-administrators supportive solution in administration of wide area network. This was primary done by identifying the existing network resource capacity and users network resource utilization trend. For providing a remedy solution for the network identification of network resources, utilization of the existing resource (specially the bandwidth) and the properties that are important for the quality of the services.

The properties provide quality of service for the connection between the remote campuses networks. Secondary, some simple methods for analyzing and presenting the measured data was identified. These methods simplify the interpretation part of the administration of wide area networks. The three case studies were created to demonstrate the functionality for some of the tools used to measure the four properties in quality of service. In Case One, the objective was to make use of traffic measurement and analysis tools, to monitor the traffic on eight consecutive days at the gateway proxy servers. The real time traffic monitoring tool successfully measured the data at the core layer devices of the remote campuses. Both methods have provided enough information to create a good understanding of what is happening on the network for the last operational times of the network.

The network monitoring tool successfully measured the throughput at the gateway interfaces of all remote campuses and main datacenter nodes. In this case, the results have clearly indicated as there is a high traffic to the gateway than the available bandwidth, and this had probably something to do with traffic congestion at the gateway interface on the service providers side. This has also showed as there is no shortage of hardware resources at the nodes.

In Case Three, the objective was to make use of resource utilization trend analysis from the data collected, to benchmark the network resource utilization based on the IETF based framework for internet bandwidth utilization quota for applications and services. Accordingly, we were able to

classify the data traffic that are originated from inside network to the internet and to different WAN services implemented in the university. Because of the different need of users and variety of entertainment sites and services in the current times, users of the university network are mostly dependent on the internet and the wide area services for every activity. The centralization of services and applications and decentralization of network users in different university campuses, the WAN has become degraded in performance.

The three case studies together demonstrated all aspects of quality of service measurements, and together with the analysis and presentation methods described in this thesis, the network and system administrators should be able to administer the wide area network by using quality of service framework on the wide area network. In order to support that activity, we have developed an optimization framework that can support wide area network by using quality of service building blocks. The proposed solution was verified and approved by highly skilled senior experts on the area.

5.2 RECOMMENDATIONS

- To measure the network devices latency at every access links and evaluate the end devices performance issue with regard to the impact on the wide area network.
- Developing algorithm for the proposed framework and implementation of the proposed framework on active wide area network and evaluation of the developed framework.
- Another recommendation is evaluation of network application that are operating in the wide area network with respect to their interoperability with the operating system they are residing on.
- Study of the services that the wide area network is accommodating and their compatibility in operating with the existence of proxy servers at the core layer and their compatibility with several versions of web browsers.
- In this study it is not thought that full investigation is done on the institutions network because of several limitations to conduct the study like lack of time and budget. Hence future work is needed to investigate further analysis on the wide area network by using active monitoring systems.
- This study is conducted with the objective of evaluating the performance bottleneck of the AAU wired wide area network and developing a WAN optimization framework using

quality of service building blocks for institutions with similar network size and infrastructure.

References

[1] S. J. Fjeldbo, "ADMINISTRATION OF REMOTE COMPUTER NETWORKS," Master's Thesis, Department of Computer Science, OSLO UNIVERSITY COLLEGE, 2005.

[2] Okonigene R.E., Matthews V.O., Akinade B., J. S., and C. I.S., "Managing and Improving Upon Bandwidth Challenges in Computer Network," *Journal of Emerging Trends in Engineering and Applied Sciences (JETEAS),* pp. 482-486, 2011.

[3] P. k. Pankaj Rakheja, Anjali gupta, Aditi Sharma, "Performance Analysis of RIP, OSPF, IGRP and EIGRP Routing Protocols in a Network," *International Journal of Computer Applications*

vol. 48, pp. 0975 - 888, 2012.

[4] A. S. TANENBAUM, *computer networks*. Prentice Hall PTR: Pearson Education LTD, 2003.

[5] United States. Central Intelligence Agency., *The CIA world factbook 2014*. New York: Skyhorse Publishing, 2013.

[6] A. Webmaster. (2016). *About Addis Ababa University ICT Overview*. Available: http://www.aau.edu.et/services/ict/overview/

[7] C. H. Tim Szigeti, Robert Barton, Kenneth Briley, Jr., *End-to-End QoS Network Design: Quality of Service for Rich-Media & Cloud Networks* Cisco Press, 2013.

[8] D. J. Nassar. (2000). *Network performance baselining (1st ed.)*. Available: Online access for Columbia College Chicago users https://colum.idm.oclc.org/login?url=http://proquest.safaribooksonline.com/?uiCode=&xmlId=1578702402

[9] Y. Gobena, "Developing WAN Optimization Model to improve the Performance of Business Critical Applications: The Case of UNECA," Master's Thesis, School OF INFORMATION SCIENCE, ADDIS ABABA UNIVERSITY, 2013.

[10] T. B. NEGGA, "PERFORMANCE ANALYSIS FOR WIDE AREA NETWORK OPTIMIZATION THE CASE OF ADDIS ABABA UNIVERSITY," Master's Thesis, SCHOOL OF INFORMATION SCIENCE, ADDIS ABABA UNIVERSITY, 2014.

[11] E. A. G. María E. Villapol, Neudith Morales, "A Proposal to Improve Network Throughput Using a QoS Building Blocks Approach at Central University of Venezuela," 2005.

[12] S. Alvarez and Cisco Systems Inc., *QoS for IP/MPLS networks*. Indianapolis: Cisco Press, 2006.

[13] Y. G. SAHIN, S. BALTA, and T. ERCAN, "THE USE OF INTERNET RESOURCES BY UNIVERSITY STUDENTS DURING THEIR COURSE PROJECTS ELICITATION," *TOJET: - The Turkish Online Journal of Educational Technology,* vol. volume 9 p. 11, 2010.

[14] H. W. MARU, "DEVELOPING DYNAMIC BANDWIDTH ALLOCATION PROTOTYPE MODEL FOR CAMPUS NETWORK BASED ON NETWORK TRAFFIC ANALYSIS " Master's Thesis, DEPARTMENT OF INFORMATION SCIENCE, ADDIS ABABA UNIVERSITY, 2011.

[15] U. o. N. D. Australia, "The University of Notre Dame Australia internet usage policy," in *POLICY: EMAIL AND INTERNET USAGE*, ed. 16 November 2006: Vice Chancellery, 2006.

[16] U. a. Buffalo, "Computing and Network Use Policy: The State University of Newyork," in *Computing and Network Use Policy*, ed. University at Buffalo: Office of the Chief Information Officer, 2007.

[17] A. Webmaster. (2016). *Addis Ababa University at a glance* Available: http://www.aau.edu.et/about/aau-at-glance/

[18] A. Pitsillides, G. Stylianou, C. S. Pattichis, A. Sekercioglu, and A. Vasilakos, "Bandwidth allocation for virtual paths (BAVP): investigation of performance of classical constrained and genetic algorithm based optimisation techniques," vol. 3, pp. 1501-1510, 2000.

[19] L. Chitanana, "Bandwidth management in universities in Zimbabwe: Towards a responsible user base through effective policy implementation," *nternational Journal of Education and Development using Information and Communication Technology (IJEDICT),* vol. 8, pp. 62-76, 2012.
[20] C. g. M. Aras, J. F. Kurose, D. S. Reeves, and H. Schulzrinne, "Real time communication of tcp switched network," 2002.
[21] J. Mairs, *VPNs : a beginner's guide*. New York: McGraw-Hill, 2002.
[22] D. Fowler, *Virtual private networks : making the right connection*. San Francisco, Calif.: Morgan Kaufmann Publishers, 1999.
[23] J. M. Powell, "The Impact of Virtual Private Network (VPN) on a Companys Network," 2010.
[24] F. Halsall, *Data communications, computer networks, and open systems*, 4th ed. Wokingham, England ; Reading, Mass.: Addison-Wesley Pub. Co., 1996.
[25] M. Hassan and R. Jain, *High performance TCP/IP networking : concepts, issues, and solutions*. Upper Saddle River, NJ: Pearson/Prentice Hall, 2004.
[26] H. Hooper. (2012). *CCNP Security VPN 642-648 official cert guide (2nd ed.)*. Available: Online access for Columbia College Chicago users https://colum.idm.oclc.org/login?url=http://proquest.safaribooksonline.com/?uiCode=&xmlId=9780132966399
[27] G. Liang, *Network protocols*. Hauppauge, N.Y.: Nova Science Publishers, 2012.
[28] M.-S. K. a. J. W. H. Hyo-Jin Lee, Gil-Haeng Lee*, "QoS Parameters to Network Performance Metrics Mapping for SLA Monitoring.pdf," 2002.
[29] T. Bonald and M. Feuillet. (2011). *Network performance analysis*. Available: http://HZ9PJ6FE4T.search.serialssolutions.com/?V=1.0&L=HZ9PJ6FE4T&S=JCs&C=TC0000831963&T=marc&tab=BOOKS Available only to UIC users
[30] G. R. Ash and A. Farrel, *Network quality of service : know it all*. Amsterdam ; Boston: Elsevier/Morgan Kaufmann Publishers, 2009.
[31] S. Urushidani, K. Fukuda, M. Koibuchi, M. Nakamura, S. Abe, Y. Ji, *et al.*, "Dynamic Resource Allocation and QoS Control Capabilities of the Japanese Academic Backbone Network," *Future Internet,* vol. 2, pp. 295-307, 2010.
[32] J. E. I. WINARNO SUGENG, KHABIB MUSTOFA, AHMAD ASHARI, "The Impact of QoS Changes towards Network Performance," *International Journal of Computer Networks and Communications Security,* vol. VOL. 3, pp. 48-53, 2015.
[33] A. Kumar, M. Amarandei-Stavila, M. Robin, A. Siganporia, S. Stuart, A. Vahdat, *et al.*, "BwE," pp. 1-14, 2015.
[34] M. Kansanen, "Wide Area Network Acceleration in Corporate Networks," Master's Thesis, Department of Information Technology, Lappeenranta University of Technology, 2009.
[35] R. G. Cole and R. Ramaswamy. (2000). *Wide-area data network performance engineering*. Available: http://HZ9PJ6FE4T.search.serialssolutions.com/?V=1.0&L=HZ9PJ6FE4T&S=JCs&C=TC0000644623&T=marc&tab=BOOKS Available only to UIC users
[36] D. D. Kouvatsos. (2011). *Network performance engineering : a handbook on convergent multi-service networks and next generation Internet* [text].
[37] J. Downing, "Quality of Service (QoS) Networking," in *Internetworking Technology Overview*, ed, 1999.
[38] M. Welzl, *Network Congestion Control Managing Internet Traffic*. West Sussex PO19 8SQ, England: John Wiley & Sons Ltd, 2005.
[39] V. Z. Tomas Podermanski, "Recommended Resilient Campus Network Design," ed: CESNET led working group on Network monitoring, 2010.

[40] A. Sieminski, "The impact of Proxy caches on Browser," *International Journal of Computer Science & Applications*

vol. Vol. II, No. II, pp. 5 - 21, 2005.

[41] A. H. Mohammed Ali, D. A. Mustafa, and D. A. A. L. S. Abdalla, "Proxy Impact on Network Performance Review Study," *International Journal of Innovative Research in Science, Engineering and Technology,* vol. 03, pp. 17107-17113, 2014.

[42] R. J. Mahbub Hassan, *High Performance TCP/IP Networking*. Pearson Prentice Hall, 2004.

[43] techopedia.com. (2016, Aug. 25). *CPU Utilization*. Available: https://www.techopedia.com/definition/28291/cpu-utilization