

Improving multimedia transmission through enhanced multimedia devices

Ikenna Osuagwu

20905904

Dissertation submitted in partial fulfillment of the requirements for the degree
Master of Engineering at the Potchefstroom Campus of the North-West University

Supervisor: Prof PW. Stoker

November 2008.

ACKNOWLEDGEMENT

My sincere gratitude goes to the Almighty God, my creator and my savior; for His tender mercies, love and favour even in the face of tough times. He continues to be my Lord and I dedicate this work to Him. I am also indebted in no small way to my beloved wife, Ogechi (God's time), and my beautiful daughter, Onyinyechi (God's gift) who just turned one. Their love, support and encouragement has been second to none. They are the reason why I continue to strive to be the best I can be. My world revolves around them and I dedicate this research to them.

My immense gratitude goes to my instructor, supervisor and mentor, Prof. Piet Stoker; who has taught me the way to realizing my dreams. He has been there to show me the direction when I needed one; guide me when I derail, and his mentorship has encouraged me to advance beyond my expectations. I am also grateful to all my colleagues in South Africa. We started this program not having the slightest clue on how it will all end. Their collective support, encouragement and constructive competition have made all the difference. My heart goes out to all of them.

ABSTRACT

Multimedia transmission is the effective way of transmitting multimedia elements (comprising voice, audio, video, data etc) from one place to the other via internet enabled protocols and other means. The term 'effective' is used because multimedia transmission is a nightmare if the conveyance is not smooth, seamless and efficient. Over time, the world has seen tremendous improvement that started from the era of the first generation of multimedia generation to the point of multimedia transmission. Much has been said and done in this area and the world has become a connected enterprise because of the transmission of multimedia.

In spite of these successes that have been recorded in these areas, there are still many challenges facing multimedia transmission. What determines the progress of technology globally is the trends of evolution that multimedia transmission has gone through. An important challenge facing multimedia transmission is one that has been neglected for a long time. There has been deep neglect of the devices that are used in transmission while much emphasis has been on the protocols and the software that are being developed for multimedia transmission. Devices play a very important role in the realization of seamless transmission of media.

Lately, the world seems to have realized the fact that devices that do the actual transmission needs more attention. These devices are the ones that do the distribution and the transmission of the multimedia streams or signals. This has been highlighted in a recent research study that was referenced in the Cambridge Handbook of Multimedia Learning that showed that the expectations of meeting the world's target on multimedia has been reduced by half because of problems arising from the inefficiency of multimedia devices and not really from the protocol perspective as earlier perceived.

It is inline with the above that this research was titled "improving multimedia transmission through enhanced multimedia devices". Multimedia devices are the end to end units that are used in multimedia transmission. This research investigated the current devices that are being used, their deficiencies and the reasons that make them unstable for multimedia transmission. It focused on the real time multimedia transmission over the internet protocol (IP) through enhancing limited capabilities of the current multimedia devices. This will make way for new studies into newer devices that are better designed for the efficient multimedia flow. It will assure better quality end-to-end solutions in the area of multimedia distribution and transmission.

This research is broad enough to cover most of the major areas of multimedia transmission and cut across several industries and technologies. These might include industries that specialize in internet telephony; design and manufacture multimedia devices and multimedia technologies. Despite the huge

number of fields that this research cut across, the focus remained unchanged in highlighting this challenge and proffering a solution through enhanced end to end multimedia elements.

It is hoped that this research work will contribute to the solution of this area of challenge and bring to fore the work that should be done in this regard. The advantages of improved multimedia transmission cannot be over emphasized as there will be a tremendous reduction in the cost of long distance communication globally and smoother media transmission which makes use of the readily available internet protocols.

Key words

Bandwidth

Dichotomy between protocols and devices

Dissertation

Enhanced multimedia devices

Improving multimedia transmission

Jitters

Latency

Multimedia devices

Multimedia technologies

Packet Loss

Problems of multimedia transmission

Protocols

Recommended solutions to multimedia transmission

Technology trend

TABLE OF CONTENTS

<u>Page Content</u>	<u>Page Number</u>
Cover Page -----	i
Acknowledgement -----	ii
Abstract/Keywords -----	iii
Table of Contents -----	v
List of Tables -----	vii
List of Figures -----	viii
List of Abbreviations -----	ix
Chapter 1: Introduction -----	1
1.2 Problem Statement of Research Work -----	2
1.3 Research Objectives -----	5
1.4 Outline of Research -----	5
Chapter 2: Literature Review -----	7
2.1 Introduction to Multimedia Transmission -----	7
2.1.1 Multimedia Classification -----	8
2.1.1.1 Text -----	10
2.1.1.2 Audio -----	12
2.1.1.3 Graphics & Animation -----	13
2.1.1.4 Video -----	14
2.1.2 Multimedia Expectations from a Transmission Network -----	15
2.1.2.1 Real Time Characteristics (Limits on Delay and Jitter) -----	16
2.1.2.2 Need for Higher Bandwidth -----	16
2.1.3 Multicasting Support -----	17
2.1.4 Session Management -----	17
2.1.5 Security -----	18
2.1.6 Mobility Support -----	18
2.2 Overview of Multimedia Technologies -----	19
2.2.1 Voice over Internet Protocol (VoIP) -----	19
2.2.2 Video over IP -----	23
2.2.3 IP Multimedia System & Multimedia Messaging System (IMS & MMS) --	26
2.3 Introduction of Problems Based On Incapable Multimedia Devices -----	28
Chapter 3: Empirical Investigation -----	32
3.1 Latency & Transmission Delays -----	32
3.2 Jitters -----	37

3.3 Bandwidth -----	40
3.4 Packet Losses -----	42
Chapter 4: Evaluation of Solutions and Improvements -----	46
4.1 Solution to Latency & Transmission Delay Problems -----	46
4.2 Solution to Problems Associated with Jitters -----	49
4.2.1 Implementing Jitter Buffer in a Multimedia Device -----	50
4.3 Solution to Problems Associated with Bandwidth -----	52
4.3.1 Compression -----	52
4.3.1.1 Lossless Multimedia Compression -----	53
4.3.1.2 Lossy Multimedia Compression -----	53
4.3.2 Caching -----	54
4.3.3 Mirroring -----	55
4.4 Solution to Problems Associated with Packet Losses -----	56
4.4.1 Forward Error Correction -----	56
4.4.2 Packet Retransmission -----	57
4.4.3 Packet Loss Concealment – PLC -----	59
Chapter 5: Discussion on the Architecture of Multimedia Transmission -----	61
5.1 Architecture of a Standard Multimedia Transmission Platform -----	61
5.1.1 Multimedia Database -----	61
5.1.2 Multimedia Server -----	62
5.1.3 Multimedia Client -----	63
5.1.4 Multimedia Network or Communication Path -----	64
5.1.5 Multimedia Protocols -----	66
5.1.6 Software Components -----	68
Chapter 6: Conclusion & Recommendation -----	
6.1 Conclusion -----	69
6.2 Recommendations -----	71
References -----	73

LIST OF TABLES

	Page
Table 1: Text Compression Schemes -----	11
Table 2: Audio Compression Schemes -----	12
Table 3: Image Compression Schemes -----	14
Table 4: Video Compression Schemes -----	15
Table 5: Multimedia Types & Bandwidths -----	17
Table 6: Qualitative comparison of Voice over PSTN and VoIP -----	21

LIST OF FIGURES

	Page
Fig 2.1 Diversity of Multimedia data compositions -----	8
Fig 2.2 Classifications of Multimedia Types -----	10
Fig 2.3 The different forms of VoIP -----	22
Fig 2.4 The architecture of IMS -----	28
Fig 3.1 Measurement of latency -----	33
Fig 3.2 Measuring latency in roundtrip times -----	34
Fig 3.3 Roundtrip and Processing latencies -----	35
Fig 3.4 Communication path of a Multimedia transmission -----	35
Fig 3.5 Captured Trace of Network analysis using Ethereal Network Analyzer ----	36
Fig 3.6 Illustration of Parameters for Jitter Calculation -----	38
Fig. 3.7 How Steady Stream of Packet is Handled within a Device -----	39
Fig. 3.8 How Jitters are handled -----	40
Fig 4.1 Variation of Packet Arrival Time (Jitter) -----	49
Fig. 4.2 Block Diagram of a Jitter Buffer within a Multimedia Device -----	51
Fig 4.3 Architecture of a Cache & Disk Memory in a multimedia device -----	55
Fig 4.4 Different methods of solving packet loss problems -----	56
Fig 4.5 Packet Retransmission -----	58
Fig 5.1 The Evolution of the different types of Multimedia Networks -----	64
Fig 5.2 A typical Multimedia Application Environment -----	65
Fig 5.3 Multimedia Protocols & the IP network layers -----	67

LIST OF ABBREVIATIONS

ACM – Association for Computing Machinery
ARQ - Automatic Retransmission request
ASCII – American Standard Code for Information Interchange
CD – Compact Disc
FCC Federal Communications Commission
FEC – Forward Error Correction
FTP – File Transfer Protocol
GIF – Graphics Interchange Format
HTML – Hyper Text Mark-up Language
HTTP – Hyper Text Transfer Protocol
HTTPS – Hyper Text Transfer Protocol Secured
IP – Internet Protocol
IPv4 – Internet Protocols version 4
IPv6 – Internet Protocol version 6
ISO – International Standard Organization
JPEG - Joint Photographic Expert Group
LZW – Lempel -Ziv –Welch
MPEG – Moving Pictures Expert Group
PNG – Portable Network Graphics
POTS – Plain Old Telephone Service
PSTN – Public Switched Telephone Network
SMTP – Simple Mail Transfer Protocol
TCP - Transport Control Protocol
UDP – User Datagram Protocol
VCR – Video Cassette Recorder
VoIP – Voice over IP (Internet Protocol)

CHAPTER 1
INTRODUCTION

In our world today, the nations and their people have experienced the impact of multimedia transmission in development and civilization. The conveyance of information from one place to another is no more a problem and the excuse of not being able to relate properly over long distances is no longer viable. From the developed and technologically-advanced nations of Europe and North America, to the developing nations of Africa and Asia, digital divides are being bridged as a result of the evolution in multimedia transmission and their improvements.

“There is no barrier to what can be transmitted over transmission media and the internet has been mobilized effectively to provide seamless transmission of multimedia information across board. This is a reality that the world has come to face in contemporary times; and this has greatly impacted world development and information dissemination” (Powell, Eller & Shockley, 2003).

Every single day, volumes of information are transmitted across the globe; between individuals, businesses and families. Through the internet, packets of multimedia data are streamed worldwide using the latest technologies that evolve at the same rate with the advancement of civilization.

In a world that is largely dominated by globalization especially in our 21st century, the transmission of multimedia has come to be a major driver of globalization. There can't be any form of foreign trade without the transmission of information through catalogs, quotations, invoices and money transfers between sellers and their customers separated by geographical divides. There can't be any form of long distance learning if the instructors and students separated by thousands of miles are not able to use the facilities made possible by multimedia transmission. They use this medium to send out study materials, assignments and other class instructions back and forth. Students are able to form study groups and are able to share their experiences even from far distances.

“The advent of social networking, which has come to stay couldn't have been possible without multimedia transmission. Today, and everywhere around us, we see people belonging to social network groups like Facebook, hi5, MySpace, twitter etc. We see a lot of blog sites, wikis and online forums, where people go and share their thoughts and experiences on certain issues from far distances, without physical contacts. Also, the sharing of media types has become possible. People sign up to media sharing facilities provided by certain groups where videos, images and audio are released, transmitted and shared almost immediately because of multimedia transmission” (Ayesi, Saka & Ikwuemesi, 2002).

Multimedia transmission has been most important in the dissemination of information through news and podcasts across the globe. Events that make news are heard almost as soon as they happen. This has

been helpful, especially in the quick and prompt response to areas hit by deadly disasters. Also people travelling to warring areas are easily notified of the danger they may face in the areas they are visiting. This has given the news media a big boost as a result of multimedia transmission. Before now, the transmission of news and information was only text-based through newspapers and magazines with still images; then it moved to the audio-based news dissemination using radios. Today, it is possible to view on televisions, personal computers and handheld mobile devices real-time 'live' events as they happen. All these have been made possible by the developments witnessed in multimedia transmission. Distances and differences in race, religion and culture have been eroded as a result of these new technologies in multimedia transmission. The once staunch beliefs in remote areas have given way to civilization; driven largely by multimedia transmission.

Life in the present times has become so easy and interesting as a result of the advancements in multimedia transmission. This has been able to save people a lot of 'costs' attributed to transportation, which includes both financial cost and human life (Wikipedia, 2008).

This dissertation shall focus on this multimedia transmission, which has made all the above possible and is still doing more to improve lives, businesses and the global economy. The emphasis in this work shall be on improving the transmission of multimedia by the use of enhanced multimedia devices.

1.2 PROBLEM STATEMENT OF RESEARCH WORK

In recent years, wireless technologies have been growing rapidly. Not only are new wireless networks built, such as Wireless LAN and the third generation wireless networks (3G), but also more powerful wireless terminals are developed such as smart mobile phones and Personal Digital Assistants (PDA). With the improvement of wireless technologies, more people begin to enjoy wireless applications and services. However, multimedia applications and services via multimedia wireless networks are still limited and not as attractive as those on Internet. Nowadays, millions of people enjoy searching songs, movies, games and other kinds of multimedia resources and sharing/exchanging them with others.

Besides, communication with friends and strangers, multimedia applications are also very attractive and useful to many people. Emails, instant messages and other applications and services have become part of many people's daily life. Entertainment applications such as online movies, live TV stream, network radio and audio-visual conversation are catching up with this trend.

"Despite these important features of wireless multimedia technology, it is still very hard to provide seamless multimedia transmission. The complexity of multimedia technology has even made the problems facing multimedia transmission more enormous" (Ji Shen et al, 2004).

Improving Multimedia Transmission through enhanced Multimedia devices

Since the advent of multimedia technologies, a lot of attention has been focused on multimedia protocols as a way improving transmission. The existence of these problems today points to the fact that the problems facing multimedia transmission is not just protocol, but stem from other complexities.

Advanced multimedia protocols have been put forward in a bid to solve these problems, but some of the basic problems are still persistent. Big multimedia companies and regulatory bodies globally have invested heavily in protocol development in order to eradicate these problems, but it is obvious that a lot remains to be done.

A simple illustration of some the problems is what is experienced during an internet telephony session, which involves multimedia transmission. In most cases, it takes the called party several microseconds to hear the voice of the caller during a call session, even within short distances. This is as a result of the problems associated with latency and jitters. Most times, researchers and developers have always looked the way of the real-time protocols as a way of solving these problems. Even with these protocols, these problems are still there. Unfortunately, this is an issue associated with the multimedia devices. There are also other unpleasant experiences that users of multimedia applications experience.

An example with the video application is the case where the video at the receiving end experiences problems and frustrates the user during streaming. The transmission is slow because there are not enough buffers to handle the video streaming jitters at the receiver end of the transmission. In most cases, the packets transmitting the multimedia information are lost or dropped on the transmission path, causing the streaming to break intermittently or stop instantly. The jitter buffers that can be used to solve this problem are associated with multimedia devices and not with protocols.

Numerous examples abound with multimedia transmissions that are clearly linked to multimedia devices and not protocols.

Two schools of thought exist on the issue of multimedia devices and protocols; and their significance to the problems of multimedia transmission. They are two opposing views on this matter.

Some have argued that the best way to improving the transmission is by focusing more on the protocols, and investing more resources into more advanced protocol development.

A journal on software and protocols, *Journal of Systems and Software*, Volume 79 of 2006 stated that; "Even though there are problems with multimedia transmission, protocols remain the most viable option of improving multimedia transmission. Improvements in the protocols are the reason for some advancement in the latest multimedia technologies. Given the time, good software applications and protocols will meet the need of efficient multimedia applications and their transmission".

Others have disagreed and tend to believe that the problem lies in some other factors, but not just the protocols. They argue that some attention should be diverted to multimedia devices.

Improving Multimedia Transmission through enhanced Multimedia devices

The later argument on multimedia devices is reflective of the opinion in most scientific articles, recent research and related journals on multimedia and its transmission.

“Even though that it is true that the world has been fortunate to have all these advancements at this time and the opportunity to use them, it is still faced with a lot of challenges of providing seamless transmission of these multimedia contents” (Ji Shen et al, 2004).

“The problem with efficient multimedia transmission is because the terminals that provide the end to end transmission have been neglected and much emphasis has been on the software and protocols” (Ifi UiO, 2001).

“We have discovered that we are faced with a peculiar problem, and this is because we have diverted our focus from the main issue of efficient point to point multimedia streaming using the devices. This has improved the protocols over time and affected the development of better multimedia devices that make transmission better” (Godred & Reine, 2006).

“As the use of wireless local area networks spreads beyond simple data transfer to bandwidth-intense multimedia applications, addressing Quality of Service (QoS) issues is very necessary, therefore for an efficient wireless multimedia streaming solution, an in-depth understanding and comparative evaluation of these strategies are necessary to effectively assess and enable the possible tradeoffs in the performance of multimedia quality and implementation complexity that are provided by the various OSI layers” (Mihaela & Shankar, 2007).

“A recent study in a 2006 edition of Cambridge Handbook of Multimedia Learning showed that the expectations of meeting the world’s target on multimedia has been reduced by half because of problems from the multimedia devices and not from the protocols” (Richard Mayer, 2006).

“These problems of inefficient multimedia streams are having effects on the ability to deliver on the multimedia dependencies. Long distance transmissions have become almost impossible and is crippling the positive usage of the internet in offering long distance education; which is based on multimedia transfer of education materials (data, voice and video) from instructors to students over the globe” (Ian Rogers, 2006: 12).

The essence of this research is to look at the dichotomy that exists between protocols and devices and to determine in which direction the industry should place emphasis going forward.

The references above, all point to the need for an emphasized focus shift from the development of just software and protocols, to the development of more enhanced multimedia devices.

To understand the place of multimedia devices, it is important to understand that in multimedia transmission, different component parts come into play in transmitting multimedia information. There are the protocols, the networks and the multimedia devices. The focus on the protocols has largely diverted attention from one component that remains a significant factor in multimedia transmission; the multimedia devices. This research therefore, tends to investigate the trends in multimedia technology and highlight the effects of the increasing focus on protocols over multimedia devices. This will also

highlight the problems that multimedia transmission faces as a result of the neglect of multimedia devices. From the device perspective, it will recommend solutions that can be implemented to minimize the problems.

1.3 RESEARCH OBJECTIVES

The objectives of this research work are centred at realizing the following:

- To highlight the current trends in multimedia technologies and how they could be harnessed to improving multimedia transmission
- To introduce the concept of multimedia transmission as a means of investigating the problems affecting multimedia transmission
- To highlight the dichotomy that exists between the protocols and multimedia devices in enhancing a more seamless transmission process
- To investigate the need for a technical focus shift from protocols to multimedia devices in solving multimedia transmission problems
- Recommend potential solutions to the problems of latency, jitter, packet losses and other problems through enhanced multimedia devices.

1.4 OUTLINE OF RESEARCH WORK

Chapter 1 is the introduction of the research, the problem statement and the objectives of the research. It ends with the overall organization of the research work by chapter.

Chapter 2 is a review of the previous work done in multimedia transmission over wireless platforms and devices. These are works that have reviewed the technologies involved in multimedia transmission, and the problems they face. It describes the overview of technologies that have been used/being used to support multimedia transmission.

Chapter 3 focuses on the main problems of multimedia transmission from a general perspective. It gives a detailed background of these problems and their impact on multimedia transmission. It continued with the critical analysis of the components of these problems.

Chapter 4 is the evaluation of the solutions and methods to minimize the effect of the problems. These are more practical and scientific steps to improving multimedia transmission using enhanced devices.

Chapter 5 focuses on the architecture of a standard multimedia platform. It gives a general description of a typical architecture of multimedia transmission. It presents a general spectrum of the components of multimedia transmission. An abridged discussion on multimedia transmission is presented in this chapter.

Chapter 6, the last chapter, is the conclusion and recommendation. The conclusion puts forward the solution statement to the problems described in the research problem. It features an abridged description of the problems and their potential solutions.

It also includes the recommendations, which built around the solutions discussed in the conclusion.

They are recommendations and industry best practices to the multimedia industry and stakeholders.

The rest of the dissertation, post chapter pages, is the list of the references and the abbreviations used in this work.

CHAPTER 2

LITERATURE REVIEW

This chapter will present the previous works that are related to the theme of this dissertation. The study is on how to improve the transmission of multimedia signals through enhanced multimedia devices. This chapter is divided into several sections. An introduction of the concept of multimedia and its transmission; the overview of the multimedia technologies; and identification of problems associated with incapable multimedia devices.

2.1 INTRODUCTION TO MULTIMEDIA TRANSMISSION

The term 'multimedia' refers to diverse classes of media employed to represent information. Multimedia traffic refers to the transmission of data representing diverse media over communication networks (Khanvilkar, Bashir, Schonfeld, Khokhar, 2004).

Multimedia transmission is the technical concept for defining multimedia traffic.

The figure below (Figure1) is used to describe the diversity of the media classified into three major groups: (i) Text, (ii) Visuals, and (iii) Sound. As illustrated in the figure below, the symbolic textual material may include not only the traditional unformatted plain text, but also formatted text with numerous control characters, mathematical expressions, phonetic transcriptions of speech, music scores, and other symbolic representations such as hypertext. (Khanvilkar, et.al, 2004)

The visual content of multimedia may include line drawings, maps, gray-scale or coloured images and photographs, videos as well as animations, simulations, virtual reality objects, video- and tele-conferencing.

Equally, the sound content may include telephone/broadcast-quality speech to represent voice, wideband audio for music reproduction, and recordings of sounds such as electrocardiograms or other biomedical signals. The entire above are considered to be a part of multimedia. It is a broad range of material content. All other perceptory senses such as touch and smell, which can very well be considered as part of multimedia, are considered out of the scope of this work (Khanvilkar, et al. 2004)

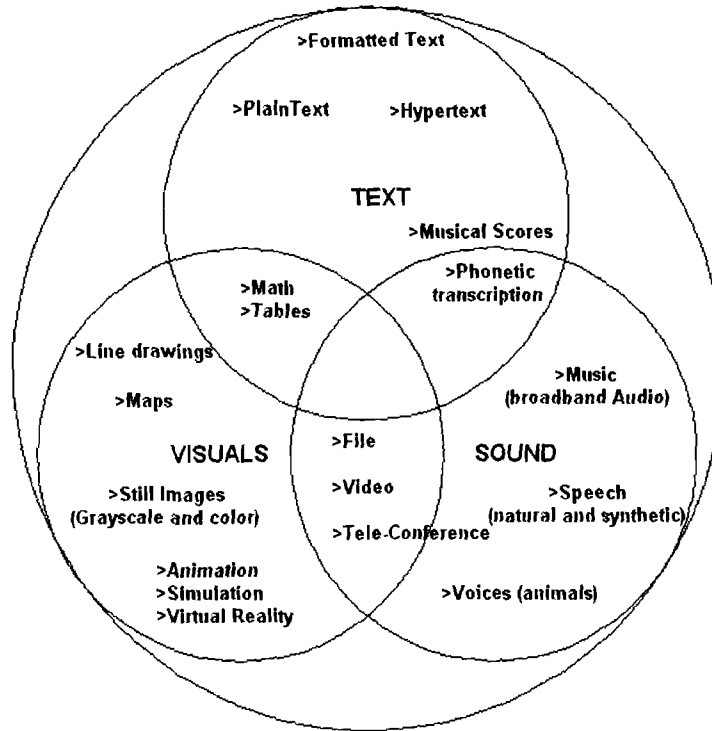


Figure 2.1 Diversity of Multimedia data compositions

Text is digitalized media type, while the other media types like sound and visual can be analog. These analog media types are expected to be converted into digital form using appropriate analog to digital conversion techniques before they can be transmitted. For the sake of this research, it is assumed that all the multimedia types that are to be mentioned in this work have been suitably digitized. The focus will be directed on the characteristics of the multimedia types, their transmission. Most importantly, the focus will also be on their enhancement using newer and better multimedia devices and end-to-end elements. In this regard, multimedia transmission deals with the generation of multimedia signals and the eventual transfer of these multimedia signals (which can be in the form of voice, video or audio) through networks that can handle multiple media types with ease and deliver scalable performance (Bertsekas & Gallager, 1987)

2.1.1 Multimedia Classification

Multimedia as known is a combination of multiple media types. The name **multimedia** was derived from a combination of **multiple** + **media**. The medium for the transmission of multimedia is through networks, which could be local metropolitan or wide. It could also be through Internet protocol (IP), through Bluetooth or through any other means, but the emphasis here is that multimedia can only be transmitted through a network (Avramovic, 2006).

Form a networking perspective, all media types can be classified as either Real-Time (RT) or Non Real-Time (NRT) as will be described below in Figure 1.2. The major difference between these two classifications is that RT media types like audio and video require either hard or soft bounds on the end-

to-end packet delay/jitter, while NRT media types like text or image files, do not have any strict delay constraints, but may be rigid constraints on error. Usually, there are basically two approaches to error control (Khanvilkar, et al. 2004). They are:

(i) Error detection followed by Automatic Retransmission request (ARQ) – request's retransmission of damaged or lost packets. This approach is used by TCP (Transport Control Protocol), a transport layer protocol in the TCP/IP protocol stack, to provide reliable connection-oriented service. Applications that require an error free delivery of NRT media typically use the TCP for transport.

(ii) Forward Error Correction (FEC) which provides sufficient redundancy in packets so that errors can be corrected without the need for re-transmissions. This approach can be used by UDP (User Datagram Protocol), another transport layer protocol in the TCP/IP protocol stack that provides connectionless unreliable service. Applications that exchange error-tolerant media types (both RT and NRT) typically use UDP for transport as it eliminates time lost in re-transmissions.

The RT media types are further classified as Discrete media (DM) or Continuous Media (CM), depending on whether the media traffic is transmitted in discrete quantum as a file or message, or continuously as a stream of messages with inter-message dependency. The real time discrete types of media has recently gained high popularity because of ubiquitous applications like MSN/Yahoo messengers (which are error intolerant) and instant messaging services like the stock quote updates (which are error tolerant). The RT continuous media types can be classified as delay tolerant or delay intolerant. Delay tolerant is usually used to signify that such media types can tolerate higher amounts of delay than their delay intolerant counterpart, without significant performance degradation.

Examples of RT, continuous, and delay-intolerant media are audio and video streams used in audio and video conferencing systems, and remote desktop applications. Streaming audio/video media, used in applications like internet webcast, are examples of delay-tolerant media types. Their delay-dependency is significantly diminished by having an adaptive buffer at the receiver that downloads and stores a certain portion of the media stream before starting playback. The entire classification of multimedia types is carefully illustrated in Figure 1.2 below. Based on these classifications, we can discuss some of the common media types and their defining characteristics in terms of bandwidth usage, error requirements, and real-time nature (Khanvilkar, et al. 2004).

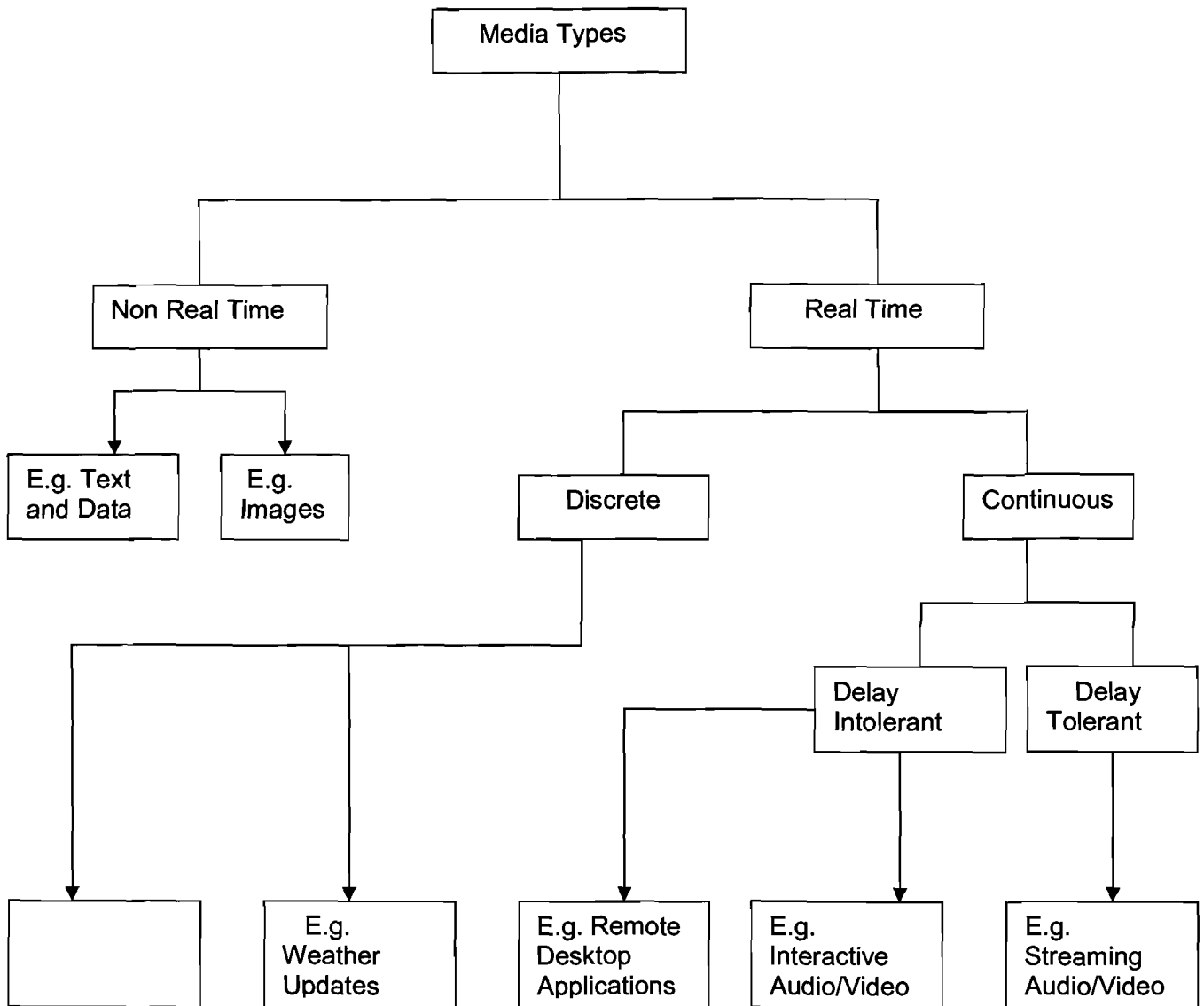


Fig 2.2 Classifications of Multimedia Types

2.1.1.1 Text

Text is the most popular of all the media types. It is distributed over the Internet in many forms including files or messages using the different transfer protocols such as FTP (File Transfer Protocol: used to transfer binary and ASCII or 'notepad text' files over the internet), HTTP (Hyper Text Transfer Protocol: used to transmit HTML pages) or SMTP (Simple Mail Transfer Protocol: used for exchanging e-mails (Cavusoglu, Schonfeld & Ansari, 2003).

Text is represented in binary as 7-bit US-ASCII, 8-bit ISO-8859 Unicode or 32-bit ISO-10646 character sets, depending on the language of choice and the country of origin. Bandwidth requirements of text

media mainly depends on its size, which can be easily reduced using common compression schemes as will be illustrated in Table 1 below. The error characteristics of text media depends largely on the application under considerations. Some text applications, such as file transfer, require text communication to be completely loss/error free and therefore use TCP for transport. Other text applications such as instant messaging may tolerate some errors as well as losses and therefore can use the UDP for transport (Cavusoglu, et al. 2003).

Applications that use text as primary media, e.g. web browsing or e-mail does not have any real-time constraints, such as bounded delay or jitter. These applications are called *Elastic Application*. However, applications like instant messaging (IM) do require some guarantees on the experienced delay. Overall, the text media has been around since the birth of the internet and can be considered as the primary means of information exchange (Cavusoglu, et al. 2003).

Compression Scheme	Comments
Shannon – Fano Coding	Use variable length code words, i.e., symbols with higher probability of occurrences are represented by smaller codes-words.
Huffman Coding	Same as Shannon-Fano coding
Lempel-Ziv-Welch (LZW)	LZW compression replaces strings and characters with single codes. It does not do any analysis of the incoming text. Instead, it just adds new string of characters it sees to a table of strings. Compression occurs when a single code is outputted instead of a string of characters
Unix Compress	This uses LZW with growing dictionary. Initially, the dictionary contains 512 entries, and is subsequently doubled till it reaches the maximum value set by the user.

Table 1: Text Compression Schemes

2.1.1.2 Audio

Audio media is sound/speech converted into digital form using sampling and quantization. Usually, sound and speech are considered to be in their analog forms and when these analog signals are converted to digital forms, the new signals are considered to be audio. Digitized audio media is transmitted as a stream of discrete packets over the network. The bandwidth requirements of digitized audio depend on its dynamic range and/or spectrum. For example, telephone-grade voice uses dynamic range reduction, using the logarithmic A-law (Europe) or μ -law (North America) capable of reducing the linear range of 12 bits to nonlinear range of 8 bits only. This reduces the throughput from 96 kbps to 64 kbps. A number of compression schemes along with their bit rates as illustrated in Table 2 below, are commonly used for audio media types (Braden, Zhang, Berson, Herzog & Jamin, 1997)

Voice/Audio Code	Used for:	Bit Rates(Kbps)
Pulse Code Modulation (G.711)	Narrowband speech (300 – 3300Hz)	64
GSM	Narrowband speech (300-3300Hz)	13
CS-ACELP (G.729)	Narrowband speech (300 – 3300Hz)	8
G.723.3	Narrowband speech (300 – 3300Hz)	6.4 and 5.3
Adaptive differential PCM(G.726)	Narrowband speech (300 – 3300Hz)	32
SBC (G.722)	Wideband speech (50 – 7000Hz)	48/56/64
MPEG layer III (MP3)	CD-quality music Wideband Audio(10-22Khz)	128 – 112 Kbps

Table 2: Audio Compression Schemes

The audio media type has loose requirements on packet loss/errors (or loss/error tolerant), in the sense that it can tolerate up to 1 to 2 percent packet loss/error without much degradation. Today, most multimedia applications that use audio, have inbuilt mechanisms to deal with the lost packets using advanced interpolation techniques.

The real time requirements of audio strictly depend on the expected interactivity between the involved parties. Some applications like Internet-Telephony (VoIP), which involves two-way communication, are highly interactive and require shorter response times. The audio media, in this case, requires strong bounds on end-to-end packet delay/jitter to be acceptable/decipherable quality. This jitter is what causes some delay in the voice transmission when making calls over the internet when the bandwidth is limited or making a circuit switched call over a very long distance. (Braden, et al. 1997).

Applications that use this media type are called *Real-Time Intolerant (RTI)* applications. In most RTI applications, the end-to-end delay may be limited to approximately 200 msec to get an acceptable performance. Other applications like Internet webcast, which involves one-way communication, have

relatively low interactivity. Interactivity, in this case is limited to commands that allow the user to change radio channels, which can tolerate higher response times. Consequently, this requires weaker bounds on delay/jitter and the applications that use such kind of media are termed *Real-Time Tolerant (RTT)* applications. They can be generally referred to as "*Streaming Audio*". (Braden, et al. 1997).

2.1.1.3 Graphics and Animation

This includes static media types like digital images and dynamic media types like flash presentations. An uncompressed, digitally encoded image consists of an array of pixels, with each pixel encoded in a number of bits to represent luminance and colour. Compared to text or digital audio, digital images tend to be a larger size. For example, a typical 4" x 6" digital image, with a spatial resolution of 480 x 640 pixels and a colour resolution of 24 bits, requires approximately 1Mbytes. To transmit this image on a 56.6 Kbps line will take at least 2 minutes. If the image is compressed at the modest 10:1 compression ratio, the storage is reduced to approximately 100KB and the transmission time drops to approximately 14 seconds. Thus some form of compression schemes are always used that cash on the property of high spatial redundancy in digital images. Some popular compression schemes are illustrated below in Table 3 (Chapman & Chapman, 2000).

Most modern image compression schemes are progressive, which have important implications to transmission over the communication networks. When such an image is received and decompressed, the receiver can display the image in a low-quality format and then improve the display as subsequent image information is received and decompressed. A user watching the image display on the screen can recognize most of the image features after only 5-10% of the information has been decompressed. Progressive compression can be achieved by the following means:

- (a.) Encoding spatial frequency data progressively
- (b.) Using vector quantization that starts with a gray image and later adds colours to it, and
- (c.) Using 'pyramid coding' which encodes images into layers, where early layers are of low resolution and the later layers progressively increase the resolution.

According to (Chapman & Chapman, 2000), images are error-tolerant and can sustain packet loss, provided the application used to render them knows how to handle lost packets. Moreover images, like text files, do not have real time constraints.

Compression Schemes	Comments
Graphic Interchange Format (GIF)	Supports a maximum of 256 colours and is best used on images with sharply defined edges and large, flat areas of colour like Text and line based drawings. GIF uses LZW (Lempel-Ziv-Welch) compression to make files small. This is a lossless compression scheme.
Portable Network Graphics (PNG)	Supports any number of colours and works best with almost any type of image. PNG uses the zlib compression scheme, compressing data in blocks dependent on the filter of choice. This is a lossless compression scheme and doesn't support animation.
Joint Photographic Experts Group (JPEG)	Best suited for images with subtle and smooth colour transitions such as photographs, grayscale/coloured images. This compression standard is based on the Huffman and Run-Length encoding. JPEG is a 'lossy' compression.
JPEG2000	Suitable for a wide range of images ranging from those produced by portable digital cameras through to advanced pre-press, medical imaging. JPEG2000 is a new image coding system that uses state-of-the-art compression techniques based on wavelet technology that stores its information in a data stream, instead of blocks as in JPEG. This is a scalable lossy compression scheme
JPEG-LS	Suitable for continuous-tone images. The standard is based on the LOCO-I algorithm (Low Complexity LOSSless COmpression for Images) developed by HP. This is a lossless compression standard.
Joint Bi-level Image Experts Group (JBIG)	Suitable for compressing black and white monochromatic images. Uses multiple arithmetic coding schemes to compress the image. This is a lossless type of compression.

Table 3: Image Compression Schemes

2.1.1.4 Video

Video is a sequence of images/frames displayed at a constant rate, e.g. 24 or 30 frames/second. Digitized video, like digitized audio, is also transmitted as a stream of discrete packets over the network. The bandwidth requirements for digitized video depend on the spatial redundancy present within every frame, as well as the temporal redundancy present in consecutive frames. Both these redundancies can be exploited to achieve efficient compression of video data (Braden, et al. 1997).

In Table 4 below, an illustration of some common compression schemes that are used for video is shown. The error and real-time requirements of video media are similar to the audio media type and so will not be discussed again.

Compression Scheme	Comment
MPEG-1	Used to produce VCR NTSC (352 x 240) quality video compression to be stored on CD-ROM using a data rate of 1.2Mbps. Uses heavy down-sampling of images as well as limits image rate to 24-30Hz to achieve this goal
MPEG-2	More generic standard for a variety of audio-visual coding applications and supports error-resilience for broadcasting. Supports broadcast quality video compression like Digital Video Broadcasting (DVB) and High Definition Television (HDTV). MPEG-2 supports four resolution levels: low (352 x 240), main (720 x 480), high-1440 (1440 x 1152) and high (1920x 1080). The MPEG-2 compressed video data rates are in the range of 3-100Mbps.
MPEG4(MP4)	Support low bandwidth video compression at the data rate of 64Kbps that can be transmitted over a single N-ISDN B channel. MPEG-4 is a genuine multimedia compression standard that supports audio and video as well as synthetic and animated images, text, graphics, texture, and speech synthesis.
H.261	Supports video communication over ISDN at data rates of px64 Kbps. It relies on intra and inter-frame coding where integer-pixel accuracy motion estimation is required for inter mode coding
H.263	The H.263 standard is aimed at video communication over POTS and wireless networks at very low data rates (as low as 18-64 Kbps). Improvements in this standard are due to the incorporation of several features such as half-pixel motion estimation, overlapping and variable block sizes, bidirectional temporal prediction, and improved variable-length coding options.

Table 4: Video Compression Scheme

2.1.2 Multimedia Expectations from a Transmission Network

This sub-section shall identify and analyze the requirements that a distributed multimedia application may enforce on the transmission network. Due to the vastness of this field, the list is going to be exhaustive; however, efforts will be made to include all the important aspects (from a general view point) that have significantly impacted the enhancements to the basic Internet architecture and its associated protocols.

Here, these requirements are divided into two categories, namely *Traffic Requirements* and *Functional Requirements*. The traffic requirements include limits on real-time parameters – such as delay and jitter, bandwidth and reliability; while functional requirements include support for multimedia services such as multicasting, security, mobility and session management. The traffic requirements can be met only by

enhancements to the basic Internet Architecture and improved multimedia equipments. The functional requirements can be met by introducing newer protocols over the TCP/IP networking stack whilst still considering the improvements in multimedia devices. The major task of this research work is to improve these traffic and functional requirements through enhanced multimedia devices or equipments.

According to (Ji Shen et al, 2004), multimedia device is a common factor in enhancing both traffic and functional requirements. There has to be a shift in focus from the common things that are generally defined as problems to multimedia expectations like the improvements in protocols and in architecture. This research will emphasize on the improvements in multimedia expectations using improved and enhanced multimedia devices embracing the latest technologies. In order to fully grasp the role of improved multimedia devices in enhancing multimedia transmission, this research will carefully observe the characteristics and the requirements of multimedia types.

2.1.2.1 Real Time Characteristics (Limits on Delay and Jitter)

As discussed earlier, media types such as audio and video have real-time traffic requirements and the transmission network must honour these requirements. For example, audio and video data must be played back continuously at the rate at which they are sampled. If the data does not arrive in time, the play back process will stop and human ears and eyes can easily pick up the artifact.

In Internet telephony, which is a major aspect of multimedia transmission involving multimedia devices, human beings can tolerate a latency of approximately 200msec (Braden, et al. 1997). If the latency exceeds this limit, the voice will sound like a call routed over a long satellite link, which amounts to degradation in the quality of the call. This is a major case in very long distance calls over satellite. This real-time traffic enforces strict bounds on end-to-end packet delay-time taken by the packet to travel from the source to the destination. Then there is the jitter, which is variability in the inter-packet delay at the receiver. The performance of distributed multimedia applications improves with decrease in both of these quantities (B. Furht, 1996).

2.1.2.2 Need for Higher Bandwidth

Multimedia applications require significantly higher bandwidths than conventional textual applications of the past. Moreover, media streams are transmitted using the UDP that does not have any mechanism to control congestion. The transmission network must be able to handle such high bandwidth requirements (Handley & Jacobson, 1998).

Below, in table 5 is a summary of bandwidth requirements of some common audio, image and video media types. Earlier in previous sections, compression schemes for the different media types have been discussed, it is worthy of mention here that these compression schemes are broadly divided into two: lossy and lossless. The lossy compression techniques eliminate redundant information from the data and

subsequently introduce distortion or noise in the original data. The lossless compression techniques do not lose any information and the data received at the other end of the multimedia device is exactly identical to the original data (Handley & Jacobson, 1998).

Multimedia Source	Sampling Rate	Bits/Sample	Bit Rate
Telephone Grade Voice (up to 3.4 KHz)	8000 samples/sec	12	96Kbps
Wideband Speech (up to 7KHz)	1600 samples/sec	14	224 Kbps
Wideband Audio Two Channels (up to 20KHz)	44.1 Ksamples/sec	16 per channel	1.412 Mbps
Image Source	Pixels	Bits/Pixels	Bit rate
Color Image	512 x 512	24	6.3 Mbps
CCIR TV	720 x 576 x 30	24	300 Mbps
HDTV	1280 x 720 x 60	24	1.327 Gbps

Table 5: Multimedia Types & Bandwidths

2.1.3 Multicasting Support

Multicasting refers to single source of communication with simultaneous multiple receivers. Most popular distributed multimedia applications require multicasting. For example, the multi-party audio/video conferencing is one of the most widely used services in Internet telephony, and this is a one application that multicasting support is employed. Multicasting is relatively easier to achieve for one-way communication than for two-way communication. This is the case in the Internet Radio, where multicasting is used to create a spanning tree consisting of the sender at the root who sends out packets and the receiver at the leaves who receives the packets. However, in the case of two-way communication like Internet telephony among multiple parties, there would be a need to have some form of audio mixing functionality that will mix the audios from all participants and only relay the correct information. Without this audio mixing, a two-way communication channel will need to be established between each participant in an all-to-all mesh function, which will amount to waste of bandwidth. (B. Furht, 1994)

2.1.4 Session Management

The session management functionality includes:

(i) *Multimedia Description*: This enables a distributed multimedia application to distribute session information such as multimedia type (audio, video, image or data) used in the session, media encoding schemes (PCM, MPEG II, etc), session start time, session stop time, IP addresses of the involved hosts, etc. It's often essential to describe the session before establishment, so that the entire participants in the session will be aware of the kind of multimedia to be transmitted (Handley, Perkin & Whelan, 2000).

(ii) *Session Announcement*: This allows participants to announce future sessions. For example, there are thousands of Internet radio stations over the internet, each web-casting different channels. Session announcement allows such radio stations to distribute information regarding their scheduled shows, so a user finds it easier to tune-in to the preferred show (Handley, et al. 2000).

(iii) *Session Identification*: A multimedia session often consists of multiple media streams (including continuous media like audio and video; and discrete media like text and images) that need to be separately identified. For example, the sender might choose to send the audio and video as two separate streams over the same network connection, which the receiver needs to decode synchronously. Another example is that the sender might send the audio and video streams together, but divide the quality of the multimedia into a base layer and some enhancement layers. This will enable the low-bandwidth receivers to receive only the base layer (audio), while high-bandwidth receivers receive the enhancement layers (video). In this way, the receiver separates the streams, which were sent out in a single stream (Handley, et al. 2000).

(iv) *Session Control*: As described above, a multimedia session involves multiple media streams. The information contained in these data streams is often inter-related, and the multimedia communication network must guarantee to maintain such relationships as the streams are transmitted and presented to the user.

This is called Multimedia Synchronization and can be achieved by putting timestamps in every media packet. This functionality is the situation in a VCR, VCD and DVD, while watching a video or a CD player, when listening to a CD (Handley, et al. 2000).

2.1.5 Security

Security is one issue that has been neglected most times in multimedia transmissions.

With the increasing usage of online services to offer multimedia transmissions, it is now apparent that security issues are quite significant.

Security provides the following to multimedia data: Integrity (to ensure that data cannot be changed in mid-flight), Authenticity (to ensure that data comes from the right source) and Encryption (to ensure that data cannot be deciphered by any third party). Therefore the Integrity, Authenticity and the Encryption of Data are important features for a multimedia device to have in order to assure security for the transmitted data (Johnston, 2000)

2.1.6 Mobility Support

The advent of wireless and cellular networks has also enhanced multimedia transmissions through mobility. The introduction of IEEE 802.11x wireless LAN, which can operate at speeds exceeding 54Mbps, is one major area where mobility has been added to multimedia transmission. Another emerging network is the IEEE 802.16 WiMAX technology, which takes care of the smaller area coverage

and limited mobility of WLAN. But the advantage of the WLAN is that they don't require licensing, thus eliminating significant investments into license purchase; easy to set up and they are relatively available.

Mobility support has added another dimension to the complexity of multimedia networks. The enhanced multimedia devices that are to be discussed in this study are expected to have the robustness of seamless connections to wireless networks so as to support mobility (Kent & Atkinson, 1998).

2.2 Overview of Multimedia Technologies

This section shall be looking at some of the multimedia technologies that are in use at the moment. Like mentioned in the problem statement, a lot of research has taken place in the protocols that enhance multimedia transmission. What this means is that there are a lot of new protocols and recent technologies in multimedia transmission, but the problem remains in the area of improving the transmission using enhanced multimedia devices. This section shall focus mainly on the general multimedia technologies. It will be good to mention here that there are so many technologies involved with multimedia. However, this work shall highlight those technologies that have more relevance to multimedia devices.

In the next chapter, the problems that these technologies have failed to solve in improving multimedia transmission as a result of incapable multimedia devices will be discussed.

2.2.1 Voice over Internet Protocol (VoIP)

VoIP is an application carrying voice over the packet switched Internet Protocol (IP) network, and is today a trend rapidly emerging in telecommunication. It is often used abstractly to refer to the actual transmission of voice (rather than the protocol) implementing it). This later concept makes it to be described more as IP telephony, Internet telephony, Voice over broadband, or broadband telephony.

VoIP is unique because it is a revolution in technology that saves cost as it reduces the cost factor prevalent in common circuit switched networks. The global evolution of the Internet and the wide spread growth of networks have been the driving force behind the evolution of VoIP. The technology has become a potential alternative to and supplement of the traditional telephony systems over PSTN (Public Switched Telephone Network), by providing a versatile, flexible, and cost-effective solution to speech communications. Now, the vast resource of the internet is used to transmit voice or audio signals through the internet network (Tariq & Malkagiri, 2007).

Voice-over-IP systems carry telephony signals as digital audio, typically reduced in data rate using speech data compression techniques, encapsulated in data-packet stream over IP. Since the inception of the first computer network, VoIP has been a subject of interest. By 1973, voice was being transmitted over the early internet. This followed a progressive trend until this moment, where VoIP provides value added services like voicemail and caller ID, which were more traditional with the PSTN technologies. It is

Improving Multimedia Transmission through enhanced Multimedia devices

expected that the total VoIP industry in the United States is set to grow by 24.3% in 2008 to \$3.2 billion (Leon-Garcia & Widjaja, 2000).

There are numerous reasons that make VoIP a widely accepted technology for multimedia and transmission. This is because VoIP can facilitate tasks and provide services that may be difficult to implement or more expensive using the common PSTN networks. Examples of this functional values include the following:

a). The ability of VoIP to transmit more than one telephone call over the same broadband connection. This can make VoIP a simple way to add an extra telephone line to a home or office. This is quite unlike the traditional PSTN which is not as robust in accommodating newer lines when needed. Because VoIP applications are sent through the internet, the issues with adding more lines are solved.

b). It is now possible to integrate conference calling, call forwarding, automatic redial, and caller ID and other zero or near-zero cost features in VoIP. This is a rapid shift from the traditional telecommunication companies (Telco), who charge higher extra charges for these value added services or features.

c). With the VoIP internet technology, it is now possible to make secure calls using standardized protocols (such as secure Real-time Transport Protocol). Most of the difficulties of creating a secure phone connection that has been difficult to implement with traditional phone lines, like digitizing and encryption, are already in place with the VoIP. With security concerns all over the world today, it has become necessary to encrypt and authenticate the existing data stream.

d). There has been a new advantage of location independence with VoIP. This is because one needs just an internet connection to get a connection to a VoIP provider through the VoIP device. This was not so with the PSTN networks, where one must be within the coverage area to have access to their networks. This makes it possible that one can be anywhere and make calls as long as there is a sufficiently fast and stable internet connection.

e). Integrating the VoIP technology with other technologies is very easy compared with what happens in the Telco, where it takes the signing of new contracts and supply of new equipments to integrate services. This has made VoIP as attractive as it can be, because VoIP can be integrated easily with other services like video conversation, message, or data file exchange in parallel with conversation, audio conferencing, managing address books, and passing information about whether others are available to interested parties. Recently, the integration of VoIP to location mapping, which helps in location identification, has become possible.

f). With the advent of VoIP, advanced telephony features such as call routing, screen pops, and IVR implementations are easier and cheaper to implement and integrate. The fact that the phone call is on

Improving Multimedia Transmission through enhanced Multimedia devices

the same data network as a user's personal computer opens a new door to a lot of new possibilities because while on a call, the user can also use the internet on the PC for several other things (Liu & Mouchtaris, 2000).

In order to highlight the basic differences between the VoIP and the traditional PSTN networks, this research will do a qualitative analysis of voice over PSTN and over IP. Below is the table:

Concept	Voice over PSTN	Voice over IP
Switching	Circuit switched	Packet switched
Bit Rate	64kbps pr 32kbps	14kbps with overhead
Latency	<100ms	200 – 700ms
Bandwidth	Dedicated	Dynamically allocated
Cost of access/billing	Business customers, monthly charge for line, per minute/second billing	Business customers, cost of IP infra-structure, monthly charge for line, cost of computer, monthly charge to ISP
Equipment	Dumb terminal, intelligence in network	Integrated smart programmable terminal, intelligence not in network
Additional features and services	Requires reprogramming or changes in network design. Not flexible	Easy enough to add new features without major changes, due to flexible protocol support
Quality of Service (QoS)	High (extremely low loss)	Low and variable, depends on IP bandwidth and internet stability
Authorization & Authentication	Only once when service starts	Potentially required per call
Regulations	Many at federal and state levels	Few regulations yet
Network availability	99.999% up time	Level of reliability still sketchy
Electrical power failure at customers' premises	Not a problem; powered by separate source from phone company	Will have problems as equipment may be down. Power from other sources is not easy to obtain
Security	High level of security because one line is dedicated to one call	Possible eavesdropping at routers, unless secured
Standards/status	Mature	Emerging problems in interworking

Table 6: Qualitative comparison of Voice over PSTN and VoIP

Source: Communications of the ACM, January 2002/ Vol. 45, No. 1

2.2.1.1 Different Forms of VoIP

The one interesting thing about VoIP is that there is not just one way to place a call through VoIP. This is where the issue of the multimedia devices used in transmitting multimedia through VoIP technology comes into play. There are three different “flavours” of VoIP service in common use today (Tariq & Malkagiri, 2007). They include:

- **ATA:** This is the simplest and most common way which can be done with the use of a device called an ATA (Analog Telephone adaptor). The ATA allows you to connect a standard phone to your computer or your internet connection for use with VoIP. The ATA is an analog-to-digital (A-D) converter. It takes the analog signals from the traditional phone and converts it into digital data for transmission over the Internet.
- **IP Phones:** These specialized phones look just like normal phones with a handset, cradle and buttons. But instead of having the standard RJ-11 phone connectors, IP phones have an RJ-45 Ethernet connector. IP phones connect directly to the router and have all the necessary hardware and software right onboard to handle the IP call. Recently, there are Wi-Fi IP phones, which allow subscribing callers to make VoIP calls from any Wi-Fi hot spot.
- **Computer-to-Computer:** This is perhaps the simplest way to use the VoIP. Here, one doesn't have to pay for long-distance calls. All the person needs is the software, installed on the PC, a microphone on PC, speakers, a sound card and an internet connection. This is the application for most commonly known VoIP like Skype, msn messenger, Yahoo messenger etc.



Fig 2.3 The different forms of VoIP

Source: Federal Communication Commission, <http://www.fcc.gov/cgb/>

2.2.1.2 Devices used in VoIP technology

The architecture of internet telephony is similar to the traditional telephone networks in many ways, but it also has some significant differences. This research is focusing more on the multimedia devices used for multimedia transmission. It also focuses on ways these devices can be enhanced to improve the way multimedia is being transmitted. Taking a look at these devices provides the clue of the limitations that multimedia devices face as a result of long neglect. The most common devices used in Internet telephony are end systems, gateways and signaling servers.

- **End Systems:** These are electronic devices with which client or users place and receive calls. These end systems respond and initiate to signaling, and receive and transmit media. These devices also maintain the track of calls and their status. The end systems are the physical interfaces that the users see and identify. They include the IP phones, computers, the connectors, the speakers, the microphones that are used directly by the user.
- **Gateways:** These are devices that allow calls to be placed to and from other telephone networks. This is the interconnect point between the network and other networks (PSTN or internet networks). The gateways are responsible for adapting protocols; and recognizing the route to the outside networks. Common applications of gateways are the Voice gateways being manufactured by Cisco, Huawei, Alcatel, Nortel and other equipment providers.
- **Signaling servers:** These devices handle the application-level control of the routing of signaling messages. They are typically used to perform user location services; a signaling server can maintain information about where a user can be currently be found and forward or redirect call setup requests to the appropriate current location. Signaling servers are the devices which, from the point of view of feature-creation, are most similar in functionality to service control or switching points in the circuit-switched network; they can programmatically direct, block, or alter call signaling messages based on their own internal logic (Kurose & Ross, 2001).

Subsequent sections shall look at these devices and understand the limitations of each group of devices mentioned above. This will help recognize ways to improve on these limitations and enhance a more effective multimedia transmission.

2.2.2 Video over IP

Video over IP is a video-compression technology which involves transmitting video, audio and data signals as a packetized data over an IP network. The digital video (which is hitherto transmitted through cable and satellite) is reduced to a bit stream (MPEG-2/MPEG-4), and then transmitted through the internet.

The field of video transport is constantly evolving. As new technologies emerge, they are put to work in transporting video signals. This has been true for radio, coaxial cable, microwave, satellite, and optical fiber, and will also hold true for Internet Protocol, or IP.

Several recent technology trends have combined to make video transport over IP networks useful of applications today. They include the following:

- **Transition to Digital Video:** Video production has almost completely migrated from analog to digital technology during the past 20 years, and today even reaches into the homes with digital camcorders, digital television displays, and digital broadcasts from terrestrial, satellite, and cable television providers. One result of this change is that video signals no longer need to be transported on specialized analog networks but can now take the advantage of the wide range of digital technology that are available.
- **Advances in Video Compression:** Compression technology has rapidly evolved from the first MPEG-1 standard in 1991 to today's Advanced Video Codec for MPEG-4. Compression means that acceptable video signals can be sent over networks with limited capacity, including those that serve many households and most businesses around the world.
- **Growth in IP Network Capacity:** The rapid growth of Internet traffic (by factor of 10,000 in the decade up to 2003) and the wide-spread adoption of broadband access lines, mean that the IP network capacity needed for video transport has reached critical mass. (Wes Simpson, 2005)

Video Over IP technology, categories and devices

Video signals are based on analog technology, just like voice. They are carried via expensive transmission circuits, but with the advent of Video over IP, these video signals can now be captured in the analog format, digitized, streamed and managed over IP networks (Marcel Dekker, 2001)

The first step is the capturing of the video content. This can be accomplished via several means. The content is processed, compressed, stored and edited on a video server (a major multimedia device). The content can either be "live" (that is captured and processed in real-time) or pre-recorded and stored. These transmissions can be sent via the IP network to either one or several recipients.

Video transmission and presentation over IP services can be looked at in these different forms, and they include:

- a. Video broadcast {IPTV (IP television)}
- b. Video-on-demand
- c. Video conferencing

Of all these forms, video conferencing is full duplex (involving a two-way communication); all the others are one way transmissions.

- **Video broadcast over IP** is a network-based one way transmission of video file content. The endpoint is merely a passive viewer with no control over the session. Video broadcast can be

either unicast or Multicast from the server. In a Unicast configuration, the transmission is replicated by the server for each endpoint viewer. In a Multicast configuration, the same signal is sent over the network as one transmission, but to multiple endpoints, or simply a group of users.

This technology is implemented in corporate environments as a means to distribute training, presentations, meeting minutes and speeches. It is also utilized by universities, broadcasters, webcast providers to transmit video information. As regards the multimedia devices involved in this kind of technology, the main factors to consider here are; the number of users, their bandwidth to the server, and the length of presentation or video length. The multimedia device that will be best employed for this technology should be one that takes into account the bandwidth capabilities of the factors mentioned above.

- Video on demand over IP allows a user to request a streamed video stored on a server. This technology differs from broadcast video in that the user has the options to stop, start, fast-forward or rewind the video as the service is interactive. Video on demand also has another feature in that it is generally accompanied by usage data allowing viewing and billing of video services or video time. Video on demand can also be used for real-time viewing, even though it is generally used for stored video files, This technology is used for e-learning, training, marketing, entertainment, broadcasting, and other areas where the end user has needs to view the files based on their schedule and not the schedule of the video supplier. The multimedia devices involved in the video on demand technology include the following:
 - a. The video server (which is the store house and could be archive server or cluster of servers)
 - b. The application control server, which initiates the transmission and controls the presentation
 - c. An endpoint with a converter to submit the viewing request and control playback
 - d. A PC or network-based device to record or convert the video files.

The seamless implementation of this kind of technology involves the enhancement of the above mentioned devices, and not the real-time protocol or other protocols. Enhancing these devices in the area of interoperability, compatibility and bandwidth requirement will improve multimedia transmission of video on demand.

- Videoconferencing over IP is a combination of full duplex audio and video transmissions which allows people in two different locations to see and hear each other as if participating in a face-to-face conversation. The camera is utilized at both end-points to capture and transmit video signals. Microphones are used at each endpoint to capture and transmit speech, which is then played through speakers. The communication is real-time and generally not stored. This technology can be used for applications including corporate communications, telemedicine, tele-health, training, e-learning, telecommuting and customer service. Videoconferencing can be point-to-point (one user to one user), or multipoint (multiple users participating in the same session). The major multimedia devices involved with videoconferencing include:
The MCU (Multipoint Conference Unit), which is maintained at a central location. This unit allows multiple video feeds to be viewed simultaneously. It provides key functions for multicast.

Improving Multimedia Transmission through enhanced Multimedia devices

The Gatekeeper, which is a box that controls the bandwidth, addressing, identification and security measures for the conferences.

The video gateway, which provide the services, and the most common being the translation between endpoint terminals and other terminal types.

The terminals (or endpoints), which are the videoconferencing systems on a Wireless or wired LAN, which supports video, voice and data.

The codec, which accomplishes the video processing in the video terminal. This is the short for coder/decoder. This codes and encodes the multimedia signals, by compressing and decompressing the signals as they enter and leave the videoconferencing system.

There is also the acoustic echo canceller, which helps to cancel out echo generated as a result of feedback of sound signals from the audio output back to the audio input of the same system (Marcel Dekker, 2001).

Delivering high quality video over IP requires offering the same or better quality of experience than what broadcast and cable TV has offered. This is the major challenge facing the Video over IP technology. This cannot be solved through protocols, but through enhanced multimedia devices that offers seamless video transmission. This is heightened by the fact that video requires extremely high throughput compared to voice. What impact the quality of video traffic are the packet losses, which are witnessed as pixilation, jitter, and freeze-frame. Prioritizing network traffic puts time-sensitive packets ahead of data packet transmissions. A strong infrastructure in the multimedia devices, which will have plenty of headroom, bandwidth and capacity, will be the single greatest factor to enhance the multimedia transmission of the video media. This is the basic fact that this research work tends to solve and provide a situation where these packet losses are solved by the use of enhanced multimedia devices that eliminates these losses and makes for improved multimedia transmission (Schonfeld, 2000)

2.2.3 IP Multimedia System & Multimedia Messaging System (IMS & MMS)

This research has discussed the Voice and Video over IP, but there is one concept that tends to bind all multimedia elements and technologies. This offers the architectural framework for delivering internet protocol multimedia to mobile users. That concept is IP Multimedia System (IMS). According to the 3GPP (3rd Generation Partnership Project), IMS is not intended to standardize applications itself but it is meant to aid the access of multimedia and voice applications over wireless and wireline terminals. This is the architecture that enables the convergence of voice, video, data and mobile network technology over an IP-based infrastructure. It fills the gap between the two most successful communication paradigms, cellular and Internet technology. The vision of the IMS is to provide cellular access to all services that the Internet provides.

Architecture of IP Multimedia System

The architecture of IP multimedia systems supports a wide range of services that are enabled based on the protocols that best support the service. Also, for each of the services supported, there is a wide range of devices that are meant for each specific service under the IMS platforms.

From the diagram above in Fig 2.4 below, IMS architecture delivers multimedia services that can be accessed by a user from various devices via an IP network or traditional telephony system. The underlying network architecture can be divided into three layers (Device Layer, Transport Layer, and Control Layer), plus the service & application layer (Marcel Dekker, 2001).

- **Device Layer:** The IMS architecture provides a variety of choices for users to choose end-point devices. The IMS devices such as computers, mobile phone, PDAs, and other digital access devices are able to connect to the IMS infrastructure via the network. Other types of devices however, try to connect to the IP network via the PSTN gateway that links the traditional networks to the IP networks. This device layer defines the properties of each of the devices as it relates to its connection to the IMS network. The interest of this research work is on this layer by trying to enhance these IMS devices in such a way that they can improve the multimedia transmission.
- **Transport Layer:** This layer is responsible for initiating and terminating the protocol sessions and providing the conversion of data transmitted between analog/digital formats and the IP packet format. It forms the gateway for interconnectivity between the traditional and IP networks.
- **Control Layer:** This layer offers the Call Session Control Function. It handles the registration of the devices and processes the signal messaging of the appropriate application server in the service layer. Here, there is a database, called the Home Subscriber Server (HSS) that stores the information and features of each subscriber to this service. This layer is also referred to as the IMS layer (refer to diagram below).
- **Service & Application Layer:** On top of the IMS layer, there is the service and application layer. The services and the applications that execute them are all run by the service layer. There are a number of services that run on this layer and are controlled by the servers which include the Instant Messaging server, the Presence server, the Group-list management server and others.

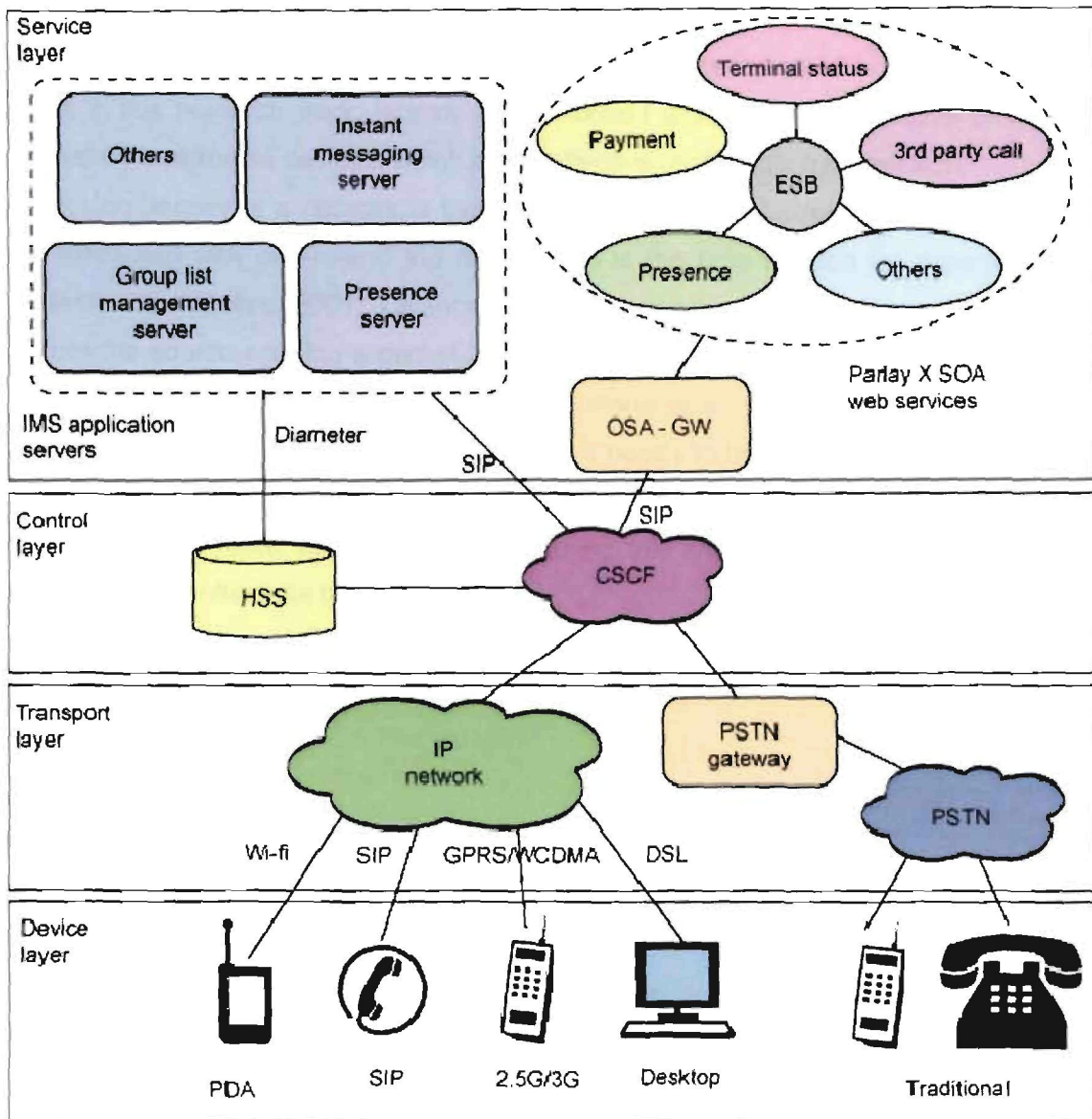


Fig 2.4 The architecture of IMS

Source: IBM Developers Forum

There are various services that are offered under the umbrella of the IP Multimedia services. Some of the common IMS services include: Multimedia conference, Multimedia Messaging service (MMS), Device location, Presence Management or instant messaging, Geocoding and Mapping etc. These services have become ubiquitous today and as a result, the increasing problems of multimedia transmission remain largely noticeable. These problems are discussed in the subsequent section, which seeks to address the problems that face the technologies that have been discussed in this chapter.

2.3 INTRODUCTION OF PROBLEMS BASED ON INCAPABLE MULTIMEDIA DEVICES

The previous sections have been highlighting the scholarly work and research done by different people at different times on the subject of multimedia transmission, and how this can be improved through enhanced multimedia devices. In this section, the emphasis will be on discussing the problems of multimedia transmission based on incapable multimedia devices. It shall highlight some of the effects of

the neglect in this aspect. "In the multimedia industry, more than 70% of the focus is on releasing newer versions of the protocols and draft recommendations of protocols. This leaves just about the rest for the multimedia devices. It has become common to believe that the multimedia devices will run any software installed or protocols. This is a myth. This has caused a long-time neglect of multimedia devices. It should be noted that multimedia devices are as important to multimedia transmission as are the protocols" (Godred & Reine, 2006).

Multimedia devices are involved at every stage of multimedia transmission. The transmitting, receiving ends and the intermediaries in multimedia transmission all make use of multimedia devices. The multimedia transmitters, the multimedia networking devices and the multimedia terminals are all part of the device spectrum. Even though these devices are all used at the moment to transmit multimedia data, their capabilities have not been boosted to take care of these problems. Improving multimedia transmission involves the recognition of these problems, finding solutions to them, and implementing these solutions in the multimedia devices in the form of enhancements.

Some of the problems facing multimedia transmission are end-to-end transmission problems. Problems like packet losses, jitters, latency and transmission delays have hampered the effectiveness of multimedia transmission. Others include problems with limited bandwidth and the fact that the multimedia devices use up a large chunk of bandwidth; network security and power requirements in multimedia devices. Solving end-to-end transmission problems start with enhancing the multimedia devices to meet these challenges.

In multimedia, these end-to-end transmissions are via the multimedia devices within wireless networks. Little work has been done to enhance these end-to-end multimedia transmissions.

Compared with wired networks, there are several problems that make end-to-end wireless multimedia transmission difficult to achieve. This has largely constituted the problems affecting efficient multimedia transmission. It is obvious that these problems exist with the multimedia devices. They include the following:

a. Low Capability and Limited Resource of Terminals

Although with the development of Wi-Fi, more personal computers can join the wireless networks, most wireless terminals are still PDAs and mobile phones. These wireless terminals typically have a small liquid crystal display supporting quite few colors and low resolution, a slow processor and a very small size of memory space etc. However, multimedia applications usually require high graphic processing capability, large size of memory space and also a big screen to display pictures and videos. The requirements of these multimedia applications are higher than these multimedia devices can handle. This results to problems with these applications when they are being transmitted with these devices

b. Diversity Of Wireless Terminals

In the Internet, although there are many types of multimedia data, they are in different data formats. Computers are still able to handle most of multimedia data formats. However, wireless terminals often support only a limited set of data formats due to their low capability and limited resource. The main reason is that manufacturers of wireless devices often focus on their portability and mobility, and their multimedia processing capabilities are sacrificed. So this result in non-compatibility of communication between two different multimedia terminals; because their supported data formats may be not compatible, the multimedia transmission is affected. In some cases, data from one terminal could be converted into the format supported by the other; such conversions are often expensive, not feasible and are not acceptable to resource-limited multimedia terminals.

c. Low Bandwidth Of Wireless Networks

Today, there are three most popular multimedia wireless networks, the GPRS, the Wi-Fi and the 3G network. GPRS can provide a dedicated bandwidth to each subscriber up to 56Kbps. Wireless LAN (Wi-Fi) can provide a shared bandwidth of up to 11Mbps, which may be quite lower when many people use it simultaneously. The 3G provides an adaptive bandwidth which can rise to 2Mbps when the user is static, and falls to 128Kbps if the user is moving.

With the facts above, it is therefore no doubt that these problems are real and needs to be minimized or solved entirely. This can only be done if research, investments and technologies are shifted from protocols to multimedia devices; where the problems reside. In recent past, multimedia transmission over wireless and IP networks have been studied and researched on, but little progress has been made in this area. This lack of result emanate from the fact that the focus has been heavy on protocols. Also in most multimedia research, multimedia devices are not taken into consideration.

It has seemed rather unfortunate that in spite of all the importance of multimedia devices to multimedia transmission, this importance is always neglected for other issues that are not so pressing. The importance of multimedia devices to the effective transmission of multimedia forms can be looked at from this basic outlook, and this is the main reason why there should much emphasis on the development of enhanced multimedia devices.

From an end-user perspective, the multimedia terminal is the most important piece for enjoying a multimedia service. It greatly determines the Quality of Experience (QoE) of any service. The reason for this is because the user does not know of the interconnecting elements between the multimedia device and the service, but the fact that the device is not functional enough to his taste makes the service inefficient.

Improving Multimedia Transmission through enhanced Multimedia devices

While the infrastructure standard are being developed rapidly, there is no consensus on how to create the “user experience” for multimedia devices, nor the ways in which the multimedia applications interact with each other within the multimedia device.

All the multimedia technologies discussed above experience some underlying problems. The ultimate goal for every multimedia transmission is to preserve both the mission-critical data in the presence of the multimedia forms, and also to preserve the quality of the multimedia being transmitted. To see the impact to which multimedia devices contribute to the problem of multimedia transmission, consider the impact of multimedia devices to the quality of transmitted media on the following: latency or delay, (which is the amount of time it takes the packet to transverse the network); the jitter, (which is the variation in delay from packet to packet); the bandwidth, (which is the data rate that can be supported on the network); the packet loss, (which is the percent of packet that do not make it to their destination). Other factors include the large multimedia storage space and the fact that the multimedia devices are not built to accommodate this multimedia capacity. All these factors are all affected as a result of the configurations in the multimedia device being used for transmission. Improvement to these factors can only be through effective research on the multimedia devices and not on the protocols.

When transmitting over networks that support multimedia transmission, there is bound to be losses. The loss of packets is a persistent problem, particularly given the increasing popularity of the multimedia. These packet losses are experienced more because of the inability of the multimedia device to accommodate these losses. Bandwidth reservation in the multimedia devices and the move towards an integrated service management in the devices can be possible ways of solving this problem.

Transmission delay is a major problem that multimedia devices contribute to inefficient multimedia transmission. The research in the field of multimedia communication has indicated that humans can tolerate end-to-end delays between 150 to 300 ms in a two way conversation. This end-to-end delay or latency is the difference in time between the time the multimedia signal is produced by the sending device and the time it is received by the receiving device. Since a sending multimedia device has to collect enough multimedia data to fill a packet before sending it, larger packets add more delay that shorter ones do. These large packets and the delays as a result of the sending and receiving multimedia devices make multimedia transmission a nightmare (Martin Isenburg, 1996).

Almost all the problems highlighted above as affecting multimedia transmission are results of the neglect on multimedia devices. If these problems can be resolved on the multimedia devices by enhancing them, then most of the limitations to a seamless and effective multimedia transmission will be put to rest. An in-depth investigation of these problems is the subject of the next chapter, which is the critical analysis of the investigated problems.

CHAPTER 3

EMPIRICAL INVESTIGATION / CRITICAL PROBLEM ANALYSIS

In the previous chapter, the discussion was based on the research that has been carried out by scholars and researchers on multimedia transmission and how to enhance multimedia transmission. The chapter is centred on identifying the common problems facing multimedia transmission.

“In the world of multimedia technology, management has realized that understanding the problems with multimedia is a sure key to solving the problems. Because multimedia has shown up in every facet of globalization and management processes, the problems abound. A careful illustration of the impacts of these problems leads the way to realizing feasible solutions to problems with multimedia technology” (Wachol, 2005).

The identified problems are associated with the following:

transmission latency or delays, jitters, bandwidth problems, packet losses, storage space requirements and plus other minor problems that stare make multimedia transmission a nightmare as a result of incoherent multimedia devices.

In this chapter, these problems will be analyzed more closely to determine their impact on multimedia transmission and how the neglect on multimedia devices has contributed largely to these problems. According to (Minoli & Minoli, 1998), “it is a fact that majority of the problems facing multimedia transmission is as a result of inadequate provisions in multimedia devices”. In this chapter, these problems, which have compounded over the years because of the much attention on protocols, will be analyzed.

The problems highlighted above have arisen because of the increase in the number of multimedia applications in the IP based networks. The increase in multimedia technologies has triggered an increase in the number of “wasteful” multimedia transmissions (Minoli & Minoli, 1998). The real-time multimedia applications usually encode video and voice in a format that handles loss of full packets. Less tolerant real-time multimedia applications such as video conferencing or video streaming prefer UDP (User Datagram Protocol) to avoid unacceptable delay introduced by packet retransmissions. UDP is considered ‘selfish’ and ‘ill-behaving’ because TCP (Transmission Control Protocol) throttles its transmission rate against the network congestion whereas UDP does not have such control mechanisms.

3.1 LATENCY & TRANSMISSION DELAYS

With an increasing interest in multimedia usage and transmission, there is the rising need to understand the causes and effect of latency in multimedia transmission as it affects the multimedia devices. Latency is the time delay between the moment something is initiated, and the moment one of its effects begins or become detectable. The word thrives from the fact that during the period of latency, the effects of an

action are latent, meaning “potential” or “not yet observed”. Latency could be as a result of a variety of reasons, but in this research work, latency will be looked at from the perspective of the transmission delays caused by multimedia devices, which affect effective multimedia transmission.

The transmission latency of a network is the delay between the occurrences of two events. Latency is commonly measured with end-to-end trip time, which is the time needed for a packet to travel from source to destination (Dallas, 2001). Latency in a packet-switched network is measured either one-way (the time from the source sending a packet to the destination receiving it), or a round trip (the one way latency from source to destination plus the one-way latency from the destination back to the source. A late packet in a multimedia stream can miss the time it needs to be played and is viewed as a loss. It is this phenomenon that makes the latency a nightmare during multimedia transmission. Because of the delay that occurs as a result of the late packet during this time, the transmission is broken and this affects the seamless multimedia transmission.

Since latency is a time delay along a path, it can be affected by:

- Interconnecting devices
- Link speeds
- Atmospheric conditions
- General faults

We are interested in checking how the interconnecting multimedia devices affect latency. The latency within a multimedia device is usually measured across the processing points (Hassan & Nayandoro, 2001). This can be expressed in the figure below.

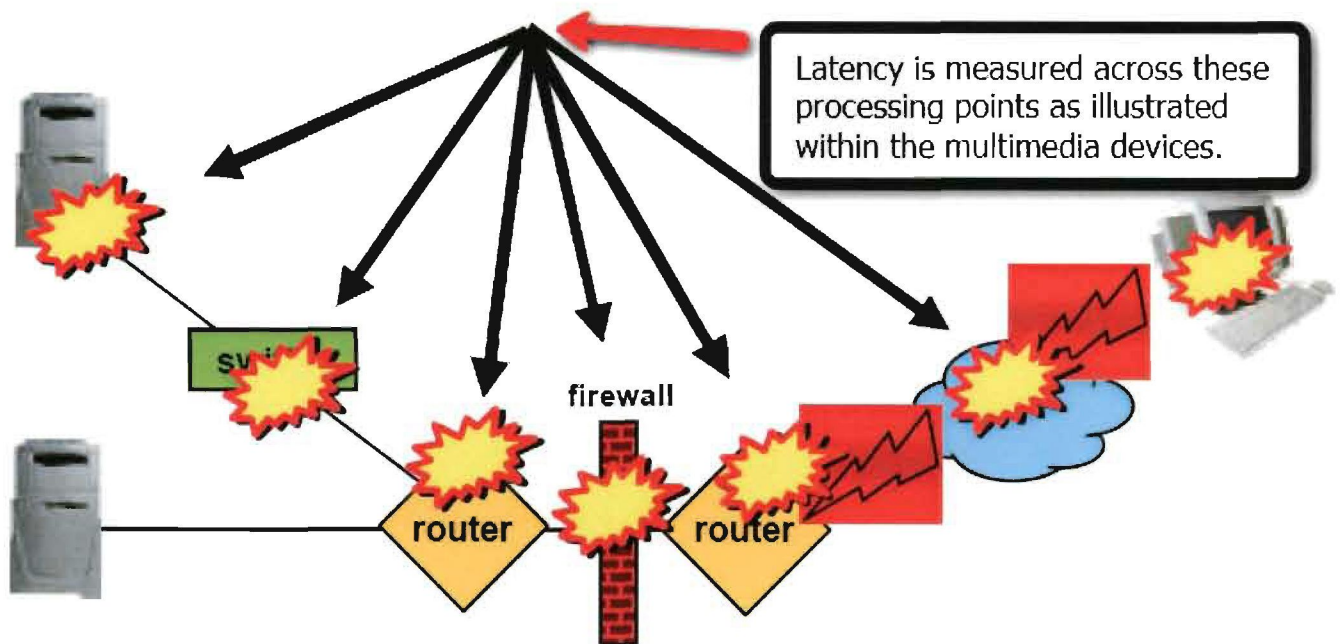


Fig 3.1 Measurement of latency

Latency can also be measured in roundtrip times. This is for simplicity in order not to take into account if the outbound or inbound transfer is faster or slower. The measurement of latency in roundtrip times is illustrated below as shown in Figure 3.2.

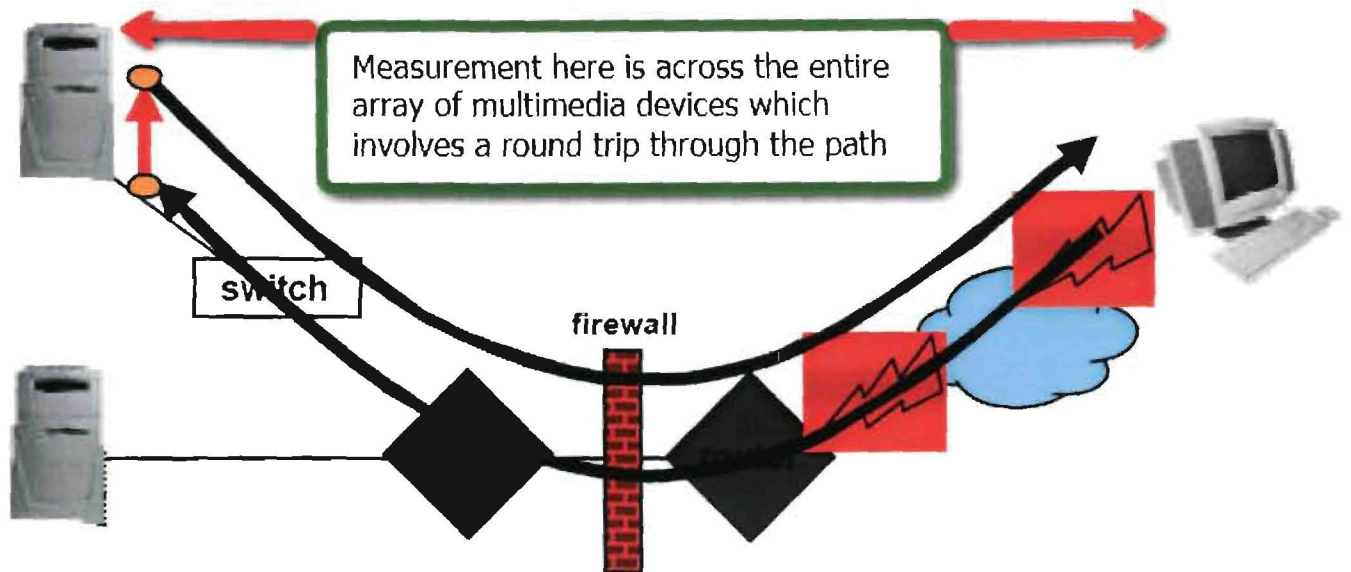


Fig 3.2 Measuring latency in roundtrip times

The issue of latency is experienced when one wants to transfer a large file over a modem. This usually takes several seconds, or even minutes. The less the data that is sent, the less time it takes, but there is a limit. No matter how small the amount of data, for any particular network device, there is always a minimum time that one can never beat. This is what is called the latency of the device. For a typical Ethernet connection, the latency is usually about 0.3ms (milliseconds – thousandths of a second). For a typical modem link, the latency is usually about 100ms, which is about 300 times worse than Ethernet. Now if one wanted to send ten characters over a 33kbit/sec modem link, one might think that the total transmission time would be: $80 \text{ bits} / 33000 \text{ bits per second} = 2.4\text{ms}$.

But unfortunately, this does not translate to 2.4ms. It takes 102.4ms because of the 100ms latency introduced by the modems at each end of the link. If you want to send a large amount of data, say 100Kb, then that takes 25 seconds, and the 100ms latency isn't very noticeable, but if you want to send a smaller amount of data, say 100bytes, then the latency is more than the transmission time.

To express the above calculation and understand how latency is calculated generally, it will be nice to look at latency from the two types of latency:

- a. Wire latency – This is the time to cross a link
- b. Processing latency – This is the time to process a request

Calculating latency is an addition of these two latencies;

Expressing this mathematically,

Wire Latency (roundtrip latency) = A

Processing Latency = B

Total latency = C = (A + B) i.e. Roundtrip latency + Processing latency

Calculating both latencies from one side of a communication path is expressed as:

$$C \text{ (processing + roundtrip)} - A \text{ (roundtrip)} = B$$

To illustrate this mathematical statement, the figure below is drawn for more visibility:

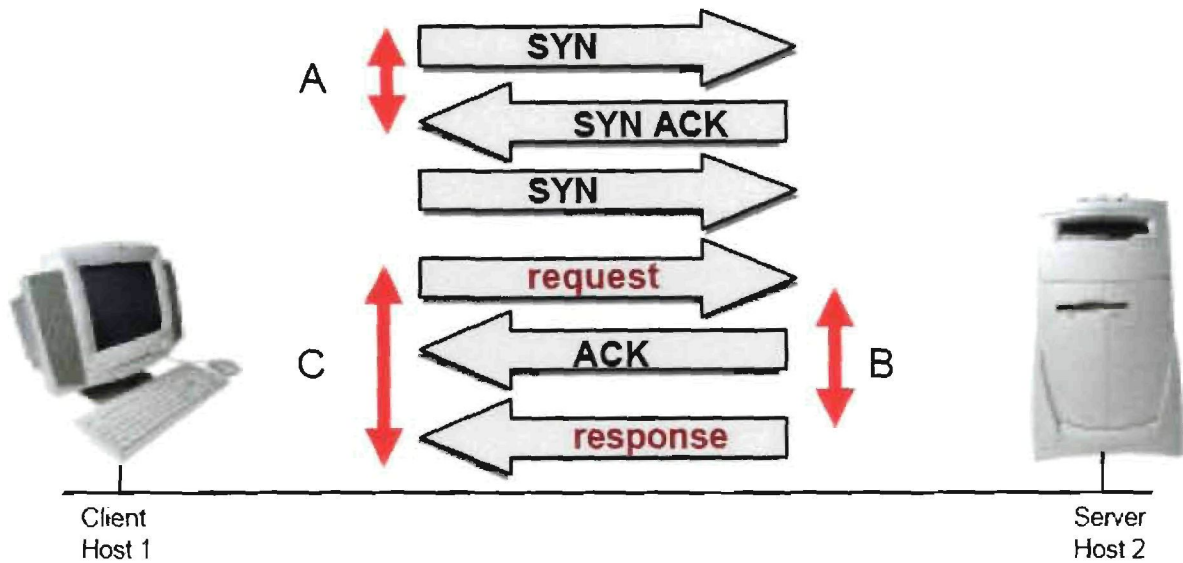


Fig 3.3 Roundtrip and Processing latencies

Measuring latency can be aided with the use of protocols and network analyzers like ethereal, aero peek etc. These tools can be used to measure the latency along a communication path. The higher the latencies, the more distorted the transmission of the multimedia information will be.

This can be illustrated using an example of information that involves a multimedia device sandwiched between a client and a server.

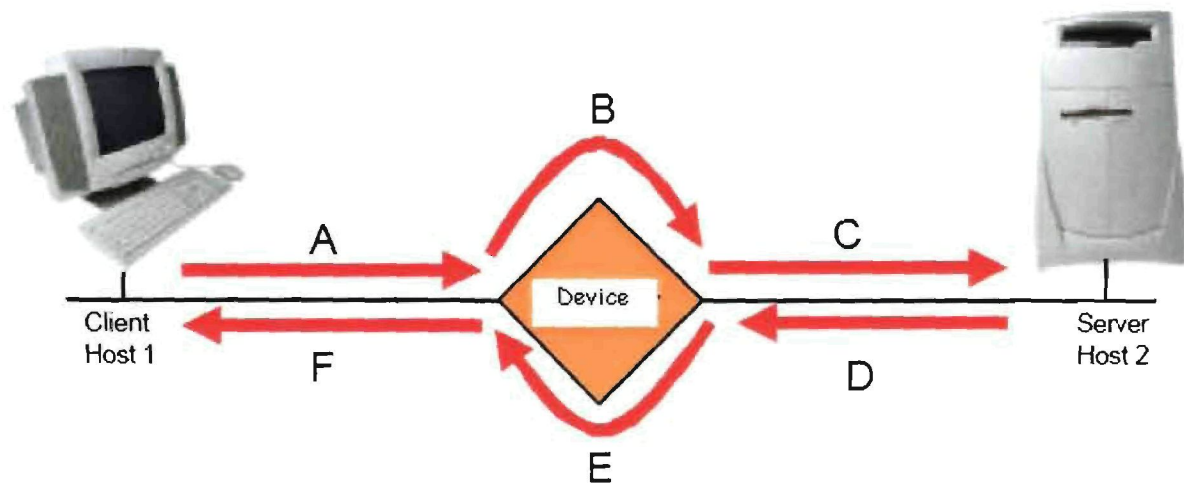


Fig 3.4 Communication path of a Multimedia transmission

Using the ethereal tool for measuring latencies, the time definitions for the latencies can be visualized in the tool. With this time that has been captured, the latencies can be calculated using the formula stated above.

The diagram below in Fig. 3.5 is a captured trace of a multimedia transmission involving a multimedia device between a client and a server as drawn above. This captured trace was done with the ethereal tool.

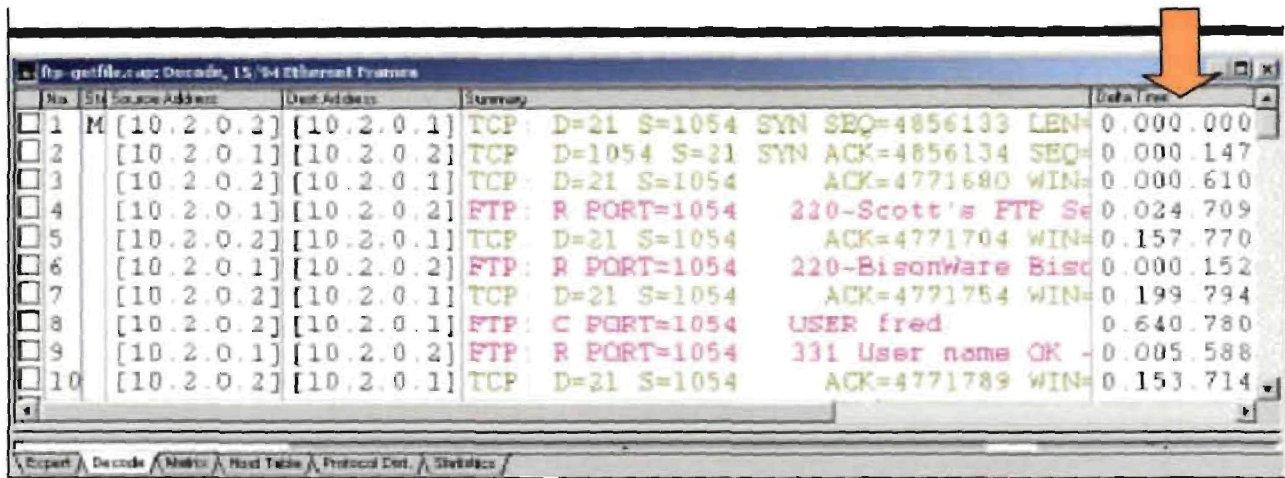


Fig 3.5 Captured Trace of Network analysis using Ethereal Network Analyzer.

From Fig. 3.5, the latencies can now be calculated thus:

1. From SYN (#1) to SYN ACK (#2) = 147 microseconds (μ s) --- A {147 μ s = 0.147ms}
2. From Request (#8) to Response (#9) = 5.588milliseconds (ms) --- C

Using the formula: C (processing + roundtrip) – A (roundtrip) = B, the latency for a one-way communication path (A – B – C or D - E – F) is given as: 5.588ms – 0.147ms = 5.441ms of latency

Like mentioned above, latency is measured in seconds, or milliseconds (ms). What this means is that if a multimedia device has a latency rate of 50ms, this means that the device will take just 50ms to send a multimedia message across. For a typical multimedia device in a non-congested network, the typical latency is 50ms, but when the network is congested, the latency could be as high as 1000ms. When this happens, the “playback” applications in the multimedia device are most affected.

As the name implies, a playback application is intended for playing back what is occurring at the sender side as similar as possible to the receiver. Typical examples of playback applications are the multimedia signals like voice, video and data. Unfortunately, these multimedia signals are most affected a lot by latency. However, different multimedia signals are affected in differing percentages. The video is most affected by latency, while data is less vulnerable to latency.

The negative implication of latency is that it affects the concept of a seamless multimedia transmission and reduces the quality of the information or data at the other end of the transmission. Another implication of latency is that it leads to a timeout of signals. This occurs when an application running in a multimedia device is waiting for some information and this information is unavailable for sometime (timeout time). The information becomes unusable and having this information after the timeout becomes useless. This is more dominant in multimedia devices that transport their data through the Internet.

Solving the problem with latency in multimedia devices has to do with solving the problem of packet queues in the multimedia devices. This is necessary because when some burst of packets reach a multimedia device and the device cannot dispatch these packets as it's supposed to, it deletes the

packets (by dropping them). The information will be lost in the process, or it can try to en-queue the packets, trying to wait a little. The formation of these queues in the multimedia devices is the main reason why the latency phenomenon appears.

3.2 JITTERS

Jitter is one of the devastating problems that face multimedia devices and their effectiveness in multimedia transmission. Jitter is an unwanted variation of one or more characteristics of a periodic signal in multimedia transmission. Commonly, 'jitter' has two meanings; the first meaning is the variation of a signal with respect to some clock signal, where the arrival time of the signal is expected to coincide with the arrival of the clock signal. This meaning is used with reference to synchronous signals and might be used to measure the quality of circuit emulation. However, the second meaning has to do with variation of a metric (e.g., delay) with respect to some reference metric (e.g., average delay or minimum delay). This is the meaning that will be applied in this research to discuss the problem of jitter. Jitter can apply to a number of signal qualities (e.g. amplitude, phase, pulse width or pulse position), and can be quantified in the same terms as all time-varying signals. There are two types of jitters: random jitter and deterministic jitter (IETF journal, 2002).

Coming home to the interest of this research work, jitter can be referred to the variation in the delay of packets. The concept of jitter is closely related to latency. In relation to latency, jitter is the variation in the latency of packets at the destination. It is the change in latency from packet to packet. If the jitter value is high, the performance in some time-sensitive applications, such as Voice over IP, Video over IP and other multimedia applications get affected.

RFC 4689 (IETF) defines jitter as the absolute value of the difference between the Forwarding Delay of two consecutive received packets belonging to the same stream.

For example, when two packets (packets A and B) are sent through a network:

Packet A takes 15ms to transverse the network

Packet B takes 18ms to transverse the network

The difference in latency between the two packets in the pair is 3ms

Jitter = $| 15 - 8 | = 3\text{ms}$

Taking into account RFC 4689, calculations of jitter requires the measurement of four parameters:

- Transit time of the first packet in the pair
- Receive time of the first packet in the pair
- Transit time of the second packet in the pair
- Receive time of the second packet in the pair

If A is the first packet and B is the second packet, then the jitter can be expressed as:

$$| (RxA - TxA) - (RxB - TxB) |$$

Jitter is always expressed as a positive number, so the absolute value of the difference is used. Average jitter is defines as the average value of the jitter of consecutive packet pairs. If the latency is constant

and each packet experiences the same delay, then the jitter would be 0 (since the difference is latency from packet to packet would not change)

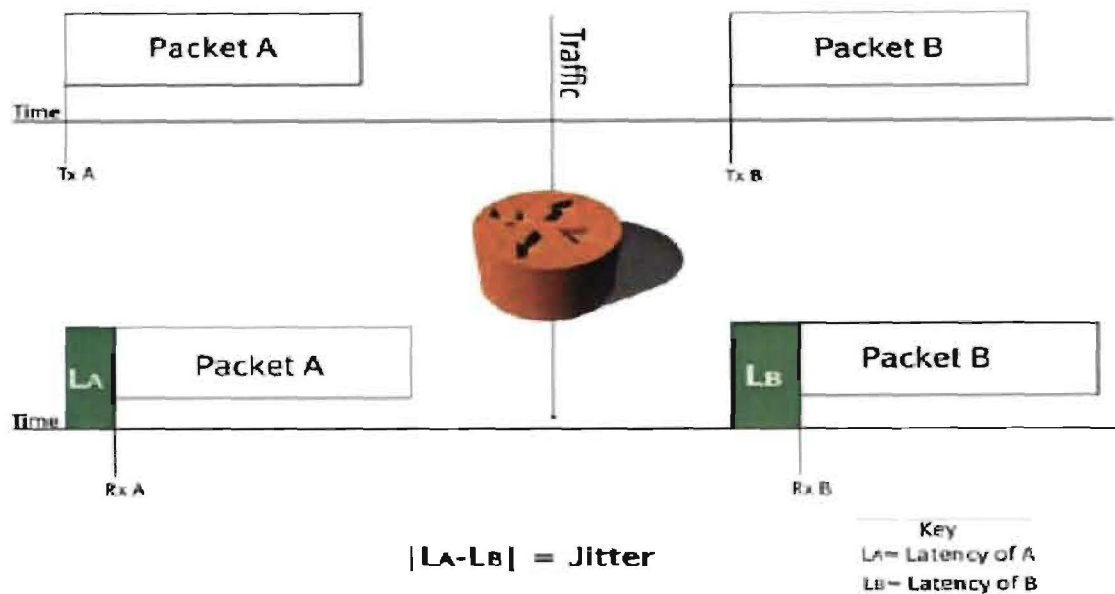


Fig. 3.6 Illustration of Parameter for Jitter Calculation

Jitter can be also calculated to get an actual value of jitter within a communication path in the case that you are considering the jitter of just one packet. In this case, this process is used:

The symbols for calculation include:

s – signifies the source

r – signifies the receiver

rr – signifies the receiver report

This formula is applied in calculating jitters:

$$r \text{ Jitter } J = sJ + [1/16 * (d - sJ)]$$

where d = transit – previous transit

sJ = Previous or source Jitter

transit = Arrival (Current time) – Timestamp from the incoming packet

To understand this concept, this example is used: if the arrival time for a packet is 1000ms and the timestamp from the incoming packet is 700ms; what is the Jitter value?

Solving this example using the formula above:

Initially, the source transit = 0 and the source or previous jitter = 0 (because this is the first packet)

For the first packet:

$$\text{transit } d = 1000 - 700 = 300\text{ms}$$

$$d = 300 - 0 \text{ (because the previous transit = 0 at source)} = 300$$

$$\text{Therefore, Jitter} = sJ + [1/16 * (d - sJ)] = 0 + [1/16 * (300 - 0)] = 18.75\text{ms}$$

Assuming a second packet is considered with respect to the first packet, and the arrival time for the second packet is 1800ms, with a timestamp of 1700ms, then:

$$\text{transit} = 1800 - 1700 = 100$$

$$d = 300 \text{ (from first packet)} - 100 \text{ (from source of second packet)} = 200$$

It should be remembered that the previous Jitter when considering the second packet is the jitter for first packet = 18.75

$$\text{Therefore, the Jitter for second packet} = sJ + [1/16 * (d-sJ)] = 18.75 + [1/16 * (200 - 18.75)] = 30.01\text{ms}$$

So the Jitter at the receiver end of the multimedia transmission for the first and the second packets are 18.75ms and 30.01ms respectively.

Jitter is a typical problem of the connectionless networks or packet switched networks. This happens because the information that is contained in the packets is divided into packets and each packet can travel by a different path from the sender and the receiver multimedia device.

The recommended value for the jitter between the starting and final points of communication through the multimedia devices should be less than 100ms.

The diagram (Fig. 3.7) below shows how a steady stream of packets is handled within a multimedia device. But due to network congestion, improper queuing, or configuration errors, this steady stream can become lumpy, or the delay between each packet can vary instead of remaining constant. The next diagram below (Fig. 3.8) is used to illustrate what happens to the packets and how the jitters are handled. The real solution and that of other multimedia transmission problems will be discussed in the next chapter, where the solutions and the improvements proposed will be discussed in greater details.

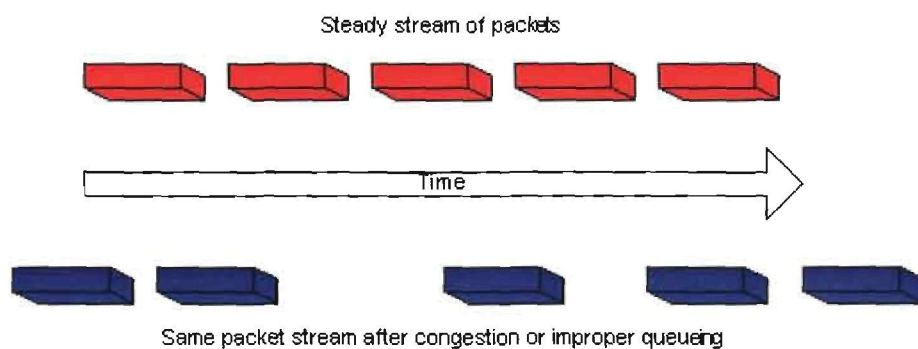


Fig. 3.7 How Steady Stream of Packet is handled within a Device

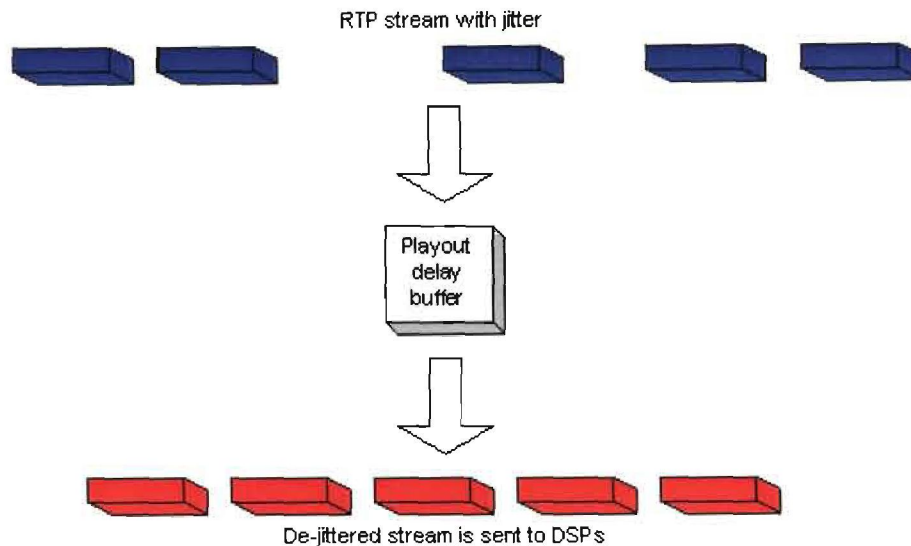


Fig. 3.8 How Jitters are handled

The issue of playout delay buffer or jitter buffer is introduced here, but this and other measures used in solving the multimedia transmission problems will be the main subjects of the next chapter.

3.3 BANDWIDTH

Bandwidth is the amount of data that can be carried from one point to another in a given time period (usually in seconds). It is usually expressed in bits (or data) per second (bps). Occasionally, it is expressed as bytes per second (Bps). To understand this better, a modem that works at 57,500bps has twice the bandwidth of a modem that works at 28,800 bps. In general, a link with a high bandwidth is one that may be able to carry enough information to sustain the succession of images in a video presentation.

It is important that bits are not confused with bytes. Bandwidths are normally specified in bits per second, and so a 56K modem has a theoretical top speed of 56kilobits per second, and a 100BaseT Ethernet has a bandwidth of 100 megabits per second. There are a thousand bits in a kilobit, and a million bits in a megabit. Further more, there are eight bits in a byte, but there are 1024 bytes in a kilobyte etc.

When discussing bandwidth, upstream and downstream bandwidth are considered. Upstream bandwidth is the bandwidth that is used during 'uploads", that is going from the multimedia client to the multimedia server; while the downstream bandwidth is the bandwidth used during "downloads", that is from the server to the client. There is always a difference between these values and that is the origin of the problem with bandwidths in multimedia transmission (Guojun, 2007).

The bandwidth requirements of multimedia servers and clients for varying multimedia applications increases linearly with the client arrival rate since the multimedia clients arrive and presumably stay in the system for some reasonable time. This increase in the rate precludes the use of the particular application and hence the need for a proper bandwidth sharing mechanism to solve the problem with

bandwidth. Most multimedia streams are in serious need for stable bandwidth. To meet these bandwidth requirements, there has to be scalability in the rate of allocation of bits. But here, there are still problems that have to do with estimating the point in time to change the bit rate of the transmitted bit stream. Estimating the available network bandwidth for each multimedia device is very challenging and crucial for multimedia streaming services.

To calculate the bandwidth of a network, it is important to distinguish between bandwidth and throughput. The bandwidth is the speed a network element (or multimedia device in this case) can forward traffic. Both physical and available bandwidths are independent of the hosts and protocol type. The throughput is the amount of data that is successfully sent from a host to the destination host, including hardware and software. Throughput also has two characteristics – achievable and maximum. There have been confusions between the available bandwidth and achievable throughput, and between capacity and maximum throughput. It is confused that to maximize the throughput via some tuning technique will make the throughput equal to the available bandwidth (Guojun, 2007).

Bandwidths are calculated based on the following parameters:

Capacity C, which is the maximum number of bits per second a network element (or multimedia device) can transfer. The capacity of an end-to-end path is determined by the slowest network element along the path.

Utilization U, is the percentage of the capacity on a link or path currently being consumed by aggregated traffic. Therefore, $U = \text{Traffic} / C$

Then the *Available Bandwidth A*, is the capacity minus cross traffic utilization over a given time interval. This is stated as:

$$\begin{aligned} \text{Available bandwidth } A &= \text{Capacity} - \text{Traffic} \\ &= C \times (1 - U) \end{aligned}$$

Today, bandwidth has become a major issue in delivering multimedia transmission effectively from the context of multimedia devices. Multimedia applications impose significant packet resource requirements in the bandwidth constrained networks. Under these conditions, it is very difficult to provide effective multimedia transmission with good quality of service guarantees. In particular, the delay constraints associated with real-time multimedia pose the greatest challenge. Real-time multimedia is particularly sensitive to delay, as multimedia packets require a strict bounded end-to-end delay, i.e. every multimedia packet must arrive at the client device before its playout time with enough time to decode and display the contents of the packet. The flow of these real-time applications do not typically back off when they encounter congestion, thus they are called unresponsive or aggressive flows. As a result, they aggressively use up more bandwidth than TCP flows. Because these multimedia applications run on the UDP, they tend to use a huge chunk of bandwidth within the multimedia device. Getting the multimedia applications to run on TCP is one way of solving this problem of bandwidth and make multimedia

transmission more efficient. So, this is the problem of bandwidth and how it has become a problem for multimedia transmission. The situation can be improved by enhancing multimedia devices to run their multimedia applications effectively by reducing their bandwidth consumption. This is what happens with a multimedia device with bandwidth constraints (Natarajan, 2003).

3.4 PACKET LOSSES

Packet loss occurs when one or more packets of data travelling within a multimedia device fail to reach their destination. This is usually caused by a number of factors, including signal degradation over the network medium, oversaturated network links, corrupted packets rejected in-transit, faulty multimedia device or networking hardware, maligned system drivers or multimedia applications. Other causes could include inadequate signal at the destination, natural or human-made interference, excessive system noise, software corruption and other causes.

The effect of this is that packets are lost and this impacts on the transmission of the multimedia packet by causing noticeable performance issues with streaming technologies, voice over IP, video over IP, online gaming, videoconferencing and other multimedia applications.

Specifically, for the multimedia types, in text and data, the packet loss produces errors; in videoconferencing environments, it can cause jitter; in pure audio communications, such as VoIP, it can cause jitter and frequent gaps in received speech. In the worst cases, packet loss can cause severe mutilation of received data, broken-up images, unintelligible speech or even complete absence of a received signal. For this reason, it is important that enough packets reach the destination for the multimedia applications to be recognizable; however, it is not a given that 100% of all packets must reach the intended target. In actuality, a multimedia application can lose up to 5% of its packets and still be recognizable. A higher drop-rate beyond this 5% will result in an incomplete and possible indecipherable message (Bow, 2006).

Driving it home, these losses mainly occur when congested multimedia transmission devices are forced to drop packets of information, given the impossibility to forward or to dispatch them. Some protocols, as TCP, deal with this problem asking to the sender to resend the lost packets again, previously slowing the rate of transmission. But some other protocols, less intelligent than TCP, can not deal with this situation, affecting the correct behaviour of applications that use them as transport mechanism. Those protocols are not intelligent than TCP, but because their designers are not interested to improve them, but for keeping them as light and fast as possible to be used to transport real-time multimedia applications. These applications demand a constant bandwidth over short periods of time, and this requirement can't be satisfied by TCP congestion control response.

For example, when transmitting voice, video and other multimedia traffic, most designers use UDP protocol as the transporting medium. They encapsulate multimedia frames over RTP (real-time protocol),

and TRP over UDP. They select UDP because it's a light protocol, very fast and adequate for low latency transmissions. Using RTP, they give sequence support to multimedia frame transportation, not serviced by UDP itself (multimedia frames must be counted, frame lost must be as low as possible, and must be ordered in its original sequence before being play backed at destination application).

But UDP, being simple and fast as is required for transporting real-time data, has a problem when comparing with TCP. It's an unresponsive protocol; this means that when congestion begins, and packets' dropping is a notorious situation, it does not respond as TCP adjusting its flow of transmission for adopting itself to the receiver capacity, avoiding this way more losses, but instead it maintains its flow rate, worsening the situation. Nevertheless, this behaviour is a "must" for multimedia applications. Given this explanation, this work concludes that congestion problems affect a lot more real-time traffic such as voice, video and multimedia, than other type of traffic that travel over TCP. Actually, when more and more applications are real-time applications, the problem is becoming more and more complicated. This becomes necessary that network engineers pay special attention to the "Quality of Service" premise, and try to implement QoS tools to deal with these situations (Bow, 2006).

Apart from the problems that have been discussed above, there are various challenges facing the current multimedia technologies. These have contributed to the problems experienced with efficient multimedia transmission. These challenges or deficiencies have been the major focus of most world bodies and organizations that are stakeholders in multimedia technology standards. These views from major bodies are enumerated below to further buttress the issues raised in this chapter.

The Internet Engineering Task Force (IETF) journal for the 2nd quarter of 2007 stated that efficient multimedia systems are not just simple extensions of conventional computing systems. The delivery of time-sensitive information, such as voice and video, and handling of large volumes of data require special considerations to produce successful wide spread applications. There has been a major failure in enhancing multimedia devices to meet the present multimedia challenges. This has widened as a result of the less emphasis in this regard.

The current multimedia technologies have failed in addressing some of these needs in the areas of the following important issues:

- Quality of Service QoS management (Real time delivery and adaptability to the environment)
 - *Reducing the Jitter experienced in multimedia communication and transmission*
 - *Decreasing the latency involved in such transmission to minimize transmission delays*
 - *Minimizing the extent of packet loss in multimedia transmission*
 - *Making available physical bandwidth in the delivery path to users*
- Information Management
 - *Indexing*
 - *Retrieval*

Improving Multimedia Transmission through enhanced Multimedia devices

- User satisfaction
- Security concerns, especially management of content rights
- Storage organization and management (IETF Journal, 2007).

An IBM journal on Research and Development in Multimedia gave an outline of the components of a multimedia infrastructure. It stated that a generalized environment supporting multimedia applications represents a user having access to an application or service through an access network. A multimedia server delivers services (i.e. multimedia data streams) over a series of possibly heterogeneous networks to the clients.

The need to provide flawless service to users places significant requirements on a large number of components, which have to work well together. This is the area where problems are experienced with the current multimedia technologies and the devices that run these applications. In order to understand the deficiency these several heterogeneous systems involved in these current multimedia technologies, one can systematically analyze the components. This will help to evaluate various options and design challenges being faced with the technologies.

- Server subsystems need to store, manage, and retrieve the multimedia data streams upon a user request.
- Network subsystem needs to transport, deliver, adapt (and perhaps transform) the data streams isochronously (i.e., at a specified rate, without observable delay) to the clients
- Client subsystems need to receive and/or pre-fetch the data streams and manage the presentation of multimedia data.
- There is the need for multimedia application programs that deal with relationships among data frames and media segments; and manage user navigation and retrieval of these multimedia data.

A good multimedia technology has to address all these components highlighted above in coordination, since each component has evolved individually up to now (IBM Journal of Research and Development, 2006).

Real time challenges are also part of the major challenges facing today's current multimedia technologies as they try to handle multimedia transmission from the device point of view. In multimedia networking, some of the deficiencies in multimedia technologies include:

- Compared with traditional textual applications, multimedia applications usually require much higher bandwidth. Current multimedia devices have not fared well in handling this requirement.
- Most multimedia applications require the real-time traffic. Audio and video data must be played back continuously at the rate they are sampled. If the data does not arrive in time, delays arise which in effect leads to network congestions as a result of undelivered data clogging the path.
- Multimedia data stream is usually bursty. For most multimedia applications, the receiver has a limited buffer. There are limited buffers in today's multimedia devices and this has led to an

overflow or underflow of the application buffer. This has resulted in the clutter of the multimedia data stream within the transmission network (IEEE Xplore 2.0, 2006).

Also, of growing concern and a major challenge to current technologies is the problem of energy management. A multimedia device consists mainly of four functional blocks; display (optional), processing module, communication module, and power management block. The first three are power consumers of the device, while the last one is the energy source. In line with a world that preaches energy conservation, designing multimedia devices that is energy efficient has been a deficiency for the current technologies. With the demands for a seamless multimedia transmission that is in-sync with the current protocols and standards, conserving energy in these multimedia devices to meet these standards is a major challenge (IEEE Xplore 2.0, 2006).

With the increasing diversification in multimedia devices and software platforms, it has become increasingly difficult to generate interoperability of multimedia devices and services on all levels. A good example of this is the Internet enabled television sets. Even though TV internet browsing is possible today, it is not yet a popular service, since rendering and navigation of arbitrary content on TV sets. Still worse, there is no common application platform today. Even though the some organizations like IEEE, ITU, IETF and other bodies, have tried to create standards, interoperability of multimedia devices is still a major challenge to multimedia transmission in the present age (IEEE Xplore 2.0, 2006).

Confronting these problems and the challenges is the theme of the next chapter, which shall discuss the solutions and possible improvements to the problems and challenges that have been discussed in this chapter. Also, the next chapter shall feature the branded solution to the challenges highlighted by the IBM Journal of Research and Development.

CHAPTER 4

EVALUATION OF SOLUTIONS AND IMPROVEMENTS

In the previous chapter, the problems with multimedia transmission within the context of inefficient multimedia devices were discussed. This introduced to terms like transmission delay, jitters, latency, packet losses and bandwidth. These concepts were examined to understand how they affect multimedia devices in transmitting multimedia applications properly.

In the chapter, the solutions to these problems will be looked at in deeper details. This shall include all the known solutions and anticipated or suggested solutions to these problems raised in the previous chapter. Also, possible improvements on the present situation in multimedia transmission and multimedia devices will also be discussed. At the end of this chapter, the ways of solving the problems highlighted in the previous sections of this research work will be discoursed. The solutions will be arranged based on the arrangement of the problems in chapter three.

4.1 SOLUTION TO LATENCY AND TRANSMISSION DELAY PROBLEMS

Latency in section 3.1 was explained as the expression of how much time it takes for a packet of data to get from one designated point to another. In some usages, latency is measured by sending a packet that is returned to the sender and the round-trip time is considered as the latency.

Latency issues is usually a nightmare in multimedia transmission as there results in noticeable delays between packets of data. This delays leads to distortion in the transmission as delayed packets are discarded as a result of the delay.

Real time applications like Video and voice require that all packets arrive in a timely manner. If the packets are lost, then the synchronization between the encoder and the decoder within the multimedia device is broken (disturbed), and errors are propagated through the rendered multimedia application. If packets are excessively delayed, then they become useless to the decoder and are treated as lost (Davidson & Peter, 2005).

To solve this problem with latency and to ensure that multimedia packets arrive and leave the multimedia device as the designated time that will avoid delays or losses, a proposal is made. This will involve the management of the protocols that are involved with multimedia transmission. A little manipulation of these protocols will achieve the desired effect.

The solution lies in the use of the User Datagram Protocol (UDP) for the transmission of the actual multimedia data while the TCP is used for the transmission of control information. UDP is often used for transmission of multimedia applications since error correction and retransmission convey unacceptable time delays in the multimedia sequence (Davidson & Peter, 2005).

Layered above the IP layer are these two unique transport protocols, USP and TCP. UDP is a lightweight extension of IP, implementing an unreliable packet transport protocol. UDP is used as the transport platform for real time multimedia applications. UDP traffic typically forms some 8% of the total volume of traffic on a public Internet network, and the majority of this traffic is attributable to the real time

applications. UDP traffic is non-adaptive, so that a UDP protocol stack will not attempt to adjust its sending rate to fit within available network capacity.

On other hand, latency most affects TCP. TCP is a reliable data transfer protocol. The requirements for reliability in the data transfer implies that the protocol will detect any form of data corruption on the part of the network and retransmit until the data is transferred without errors and successfully. This 'stop and retransmit' implies that there is no fixed rate for data transfer, nor will any implicit timing of packets be preserved by TCP. TCP is not a real-time protocol, and hence tries to maximize its data transfer rate through dynamic rate adjustments. The way TCP achieves this is to continually test the network to see if a higher data transfer rate can be supported. When TCP encounters packet loss, it assumes that the loss is due to network congestion, and the protocol immediately reduces its data transfer rate.

The TCP uses a 'sliding window control' to send data. By this, it is meant that the sender sends a sequence of packets (a 'window'), and then holds a copy of these packets while awaiting an acknowledgement (ACK) from the receiver that the packets have arrived. Each time an ACK for new data is received, the window is advanced by one packet, allowing the sender to send the next packet into the network (Davidson & Peter, 2005).

The way in which TCP adjusts its rate is by increasing the window each time an ACK for new data is received, and reducing the window size when the sender believes that a packet has been discarded. Typically, in a steady state, the sender sends a burst of packets, and then waits for a corresponding burst of ACKs before sending any further data. When starting a TCP session, a control method called 'slow start' is used, where the window is increased by one packet each time an ACK is received. Across high latency paths, such as satellite paths, thus function can result in packet bursts, where the sender sends a string of packets and then sits idle, waiting for the corresponding string of packets.

With TCP, these packet bursts are injected into the network at twice the rate that is available on the data path. The network must perform rate adaptation by using queues to adapt the sending rate to the bottleneck rate. The larger the round-trip delay time (RTT), or in other words, the greater the latency, the greater this burstiness of the data (McPhillips, 1999).

Now a sender within the multimedia device cannot continuously increase its sending rate without limit. At some stage, the receiver within a multimedia device will signal that its receiving buffer is saturated, or the sender will exhaust its sending buffer, or a network queue resource will become saturated. In the last case, this network queue saturation will result in a packet loss. When TCP experiences packet loss, the TCP sender will immediately halve its sending rate and then enter "congestion avoidance" mode. In the congestion avoidance mode, the TCP sender will increase its sending rate by one packet every RTT interval. Paradoxically, this is somewhat slower than the "slow start" rate (McPhillips, 1999).

Improving Multimedia Transmission through enhanced Multimedia devices

For TCP, the critical network characteristic is the latency. The longer the latency, the more insensitive TCP becomes in its efforts to adapt to the network state. As the latency increases, TCP's rate increase becomes slower, and the traffic pattern becomes burstier in nature. These two factors combine to reduce the efficiency of the protocol and hence the reliability of the network and the multimedia device to carry data.

This is how the TCP becomes a major problem with the latency. Like proposed above, using the TCP for the control information and using the UDP for the real multimedia data solves the problem of latency. Because of the reliability of TCP as explained above, it is assured that the information that controls the transmission of the multimedia is taken care of. The UDP, which is a less reliable protocol, now carries the multimedia data. Because it's not so reliable and doesn't take errors into account, it is assured that the data multimedia data is transmitted through a medium that will not want to retransmit as a result of 'errors'. This way, the expected information is received on time and within the limits of expectation, and the latency that used to be a problem is reduced.

The above is one major way of solving latency in multimedia transmission; by managing the transport protocols to transmit different aspects of the information. Using the TCP for the control information for the multimedia application and using the UDP for the actual multimedia application will reduce latency. The question is how to integrate this into the present day multimedia devices. This is a proposal to the hardware and software vendor companies that are involved with the manufacture of multimedia devices. This has to do with adjusting the design of the multimedia devices to favour the transmission of two blocs of information; the control information and the real multimedia application.

This adjustment in design will also include that the transport protocols are redesigned such that the initial control information is conveyed through the network through the TCP; while the second and actual multimedia data is conveyed through the UDP. Both the sending and reception of the multimedia data to and from the multimedia device will use the same protocol on both sides of the link. What this means is that the TCP sends and is also used to receive the control information; while the UDP sends and is used to receive the multimedia data.

Another minor method of solving the latency problem is to use very large buffers in the multimedia devices. With these large buffers, the present configuration of multimedia devices where TCP is used to transmit multimedia data and the control information will be effective. What this means is that there is enough space in the buffer to take care of the delays that are encountered with TCPs handling of the 'connection errors'. Tuning a TCP stack to support window scaling, selective acknowledgements and large TCP buffers is another way of limiting the effects of latency. Setting up large buffers on the multimedia devices that connect to the network also will have a positive effect on the TCP.

The only reservation this research work has with this method is the cost of implementing this and the fact that it's not a permanent solution. This is because, if there are too many 'errors' encountered by the TCP, a point is reached where these buffers will get saturated again, and the latency problem is introduced to the multimedia transmission.

4.2 SOLUTION TO PROBLEMS ASSOCIATED WITH JITTERS

Jitter has been defined in section 3.2 above as an unwanted variation of one or more characteristics of a periodic signal in electronics or telecommunications. It is the variation in the delay of packets, so it is closely related to latency. The devastating effect here is that while latency can be handled by the steps outlined above, having a variation in the delay tends to prove somewhat difficult to handle. This is why jitters have affected the efficient transmission of multimedia data.

Jitter is one issue that exists in packet-based networks. While in a typical multimedia environment, the sender is expected to reliably transmit multimedia streams at a regular interval. These multimedia streams can be delayed throughout the network and not arrive at the same regular interval at the receiving station.

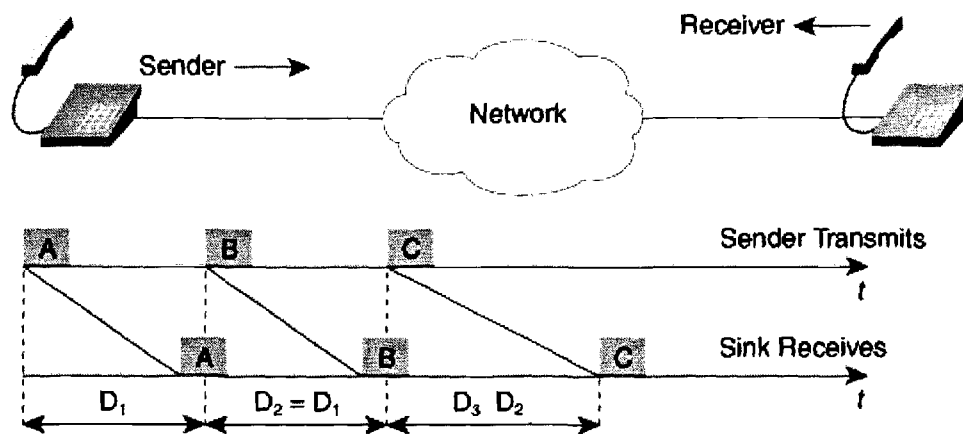


Fig 4.1 Variation of Packet Arrival Time (Jitter)

In the figure above, it can be observed that the amount of time it takes for packets A and B to send and receive is equal (i.e. $D_1 = D_2$). Packet C encounters delay in the network, however, and is received after it is expected. The packets in this case represent the multimedia streams in a multimedia application.

To solve this problem involved with jitter, the concept of *Jitter Buffer* is introduced. The jitter buffer is a solution to the problems that arise in multimedia transmission as a result of jitter. In the jitter buffer, the inter-arrival packet delay variation is concealed. Multimedia streams in multimedia applications have highly variable packet-interarrival intervals. A buffer is something that acts as a temporary storage of data, used to compensate for differences in the rate of flow of data between components.

The idea with the jitter buffer is to create temporary storage for these multimedia streams. It happens in such a way that the number of streams that arrive late are counted. Then the ratio of these late streams to the number of successfully processed streams is created. This ratio can then be used to adjust the jitter buffer to target a pre-determined and allowable late-packet ratio.

The implication of this is that the amount of jitter experienced should determine the size of the jitter buffer that needs to compensate for the late multimedia streams (Cisco IP Journal, 2008).

4.2.1 IMPLEMENTING JITTER BUFFER IN A MULTIMEDIA DEVICE

In multimedia, a jitter buffer is a shared data area where multimedia packets can be collected, stored, and sent to the processor in evenly spaced intervals. This is because variations in packet arrival time can occur because of network congestion, timing etc. The jitter buffer, which is located at the receiving end of the multimedia connection, intentionally delays the arriving packets so that the end user experiences a clear connection. With this, there is very little distortion in the streams.

The whole process of jitter buffer is to implement a solution which requires collecting packets and holding them long enough to allow the slowest packets to arrive in time to be played in the correct sequence. Buffer space is implemented in a network element for these purposes. The amount of buffer space required is typically set at 1.5 times the delay variation time. Jitter Compensation Delay varies from network to network, but even for a well designed network, it can contribute 10 to 20 ms to the overall delay (Cisco IP Journal, 2008).

There are two kinds of jitter buffers:

- Static jitter buffers
- Dynamic jitter buffers

The static jitter buffer is hardware-based and is usually configured by the manufacturer of the multimedia device. That is why it is static, because it is embedded into the hardware configuration of the multimedia device. The dynamic jitter buffer is software-based and can be configured by the network administrator to adapt to changes in the network's delay (Cisco IP Journal, 2008).

The jitter buffer is a buffer that stores frames (multimedia packets) as they are received from the network, and outputs them at a constant output rate, thus absorbing the packet delay variation (jitters). As long as the average rate of reception of the packets is equal to the constant output rate, the jitter buffer allows the packets to be outputted at the constant output rate. This happens even though they are not necessarily received at a constant rate.

The jitter buffer is able to do this by introducing *delay* into the communication path. This means that there is a delay while the packet travels through the jitter buffer until it is processed, or 'played out' at the receiver. This delay between the receipt of the packet and the payout of the packet within the jitter buffer is called "play out time offset". (Ouelletter et al, 2008).

The jitter buffer is designed in such a way that it is large enough to store a sufficient number of multimedia packets. This is to ensure that the slowest data in a multimedia stream has sufficient time to arrive at the device on the receiver end before playback.

If the jitter buffer is too small in size or capacity, the packet loss will be much, which produces distortion in the stream. However, if the buffer is too big, this increases the playout time offset (defined above), and this results in packet stream overlap in the received signal.

There are several methods of implementing jitter buffer. One approach is to measure the variation of packet level in the jitter buffer, and the buffer size is adapted to match the calculated jitter. This means that the first process is to calculate the jitter within a multimedia devices involved in the network. This approach will work better in a system with a consistent packet performance such as networks that run on the Asynchronous Transfer Mode (ATM). It is not so useful in systems with highly variable packet inter-arrival times.

The second approach for implementing the jitter buffer is to count the number of packets that arrive late, and creating a ratio of these packets to the number of packets processed successfully. This ratio is then used to adjust the jitter buffer to target a pre-determined allowable packet ratio. This approach works best with multimedia devices with high packet arrival variability like IP networks (Cisco IP Journal, 2008). This research work recommends this method because of its level of accommodation for variability in packet inter-arrival.

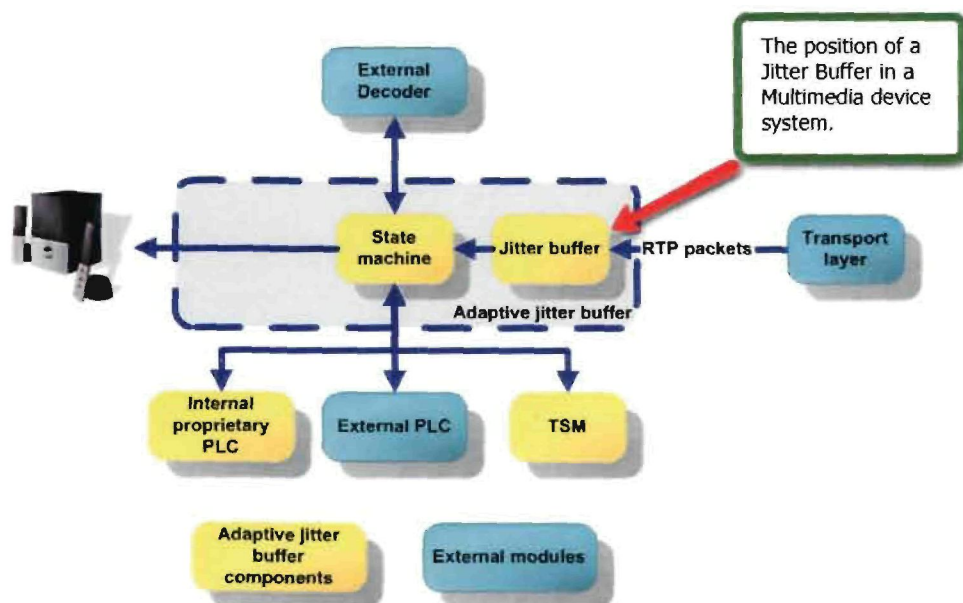


Fig. 4.2 Block Diagram of a Jitter Buffer within a Multimedia Device

The figure above illustrates the place of a jitter buffer in a typical multimedia device. The one represented above is a voice decoder (vocoder) in a Voice over IP application.

To illustrate a jitter buffer from a vendor perspective, Cisco is one vendor that readily calls to mind. Cisco has a Cisco IOS (Internet Operating System), where Real-Time Protocol (RTP) timestamps are used to determine what level of jitter that exists within the network. The jitter buffer found within the Cisco IOS software is considered as a dynamic queue. This is a dynamic jitter buffer because it can be re-configured by the administrator and it is software based. This queue can grow or shrink exponentially depending on the interarrival time of the RTP packets.

Although many vendors choose to use static jitter buffers, Cisco found that a well-engineered dynamic jitter buffer is the best mechanism to use packet based multimedia applications. Static jitter buffers force the jitter buffer to be either too large or too small. This results in a drop in the quality of multimedia transmission, due to either lost packets or excessive delay. The Cisco jitter buffer dynamically increases or decreases based upon the interarrival delay variation of the last few packets (Cisco IP Journal, 2008).

4.3 SOLUTION TO PROBLEMS ASSOCIATED WITH BANDWIDTH

The issue of bandwidth and its effects on multimedia transmission has been a recurring issue for quite a long time. The inability to have enough bandwidth for use has made the seamless transmission of multimedia and its components harder. This is evident in the fact that network resources are limited and the making these resources adequate is still a long way coming (Natarajan, 2003).

Simply selecting a networking technology and topology is not enough. Maximizing the effective bandwidth is essential to ensure that multimedia applications are transmitted effectively. Most network transfers and multimedia transmissions use TCP/IP, a protocol that supplies a logical client/server communication structure. This protocol consists of a network layer (IP) and a transport layer (TCP). These layers have unique characteristics and limitations especially with bandwidth management. The transport layer, which is the core of the multimedia transmission, limits the TCP packet size to the maximum network interface. It does this to avoid fragmentation (Natarajan, 2003).

Summarily, the problems that affect multimedia transmission from the bandwidth perspective include bandwidth consumption, congestion and traffic being discarded. These are the issues with bandwidth that needs to be solved for an efficient multimedia transmission.

There are many techniques that can be used to solve the problems that are directly related to bandwidth. Here, these solutions will be discussed based on their applications in solving these bandwidth problems.

4.3.1 Compression

Compression is one of the solutions to the problems associated with limited bandwidth. There is multimedia data-specific compression. This include still pictures (images) compressed with JPEG, PNG

or GIF; moving pictures (video) compressed with MPEG, Motion JPEG, and other video codecs; audio (voice) compressed with MP3, WAV, AAC, A-law, μ -law and other audio compression techniques.

The techniques involved in multimedia compression can be viewed from these perspectives:

- *Lossless compression*: This means the recovery of the original multimedia representation
- *Lossy compression*: This means recovery of a representation similar to the original
- *Hybrid compression*: A combination of the two above. E.g. include JPEG, MPEG, which combine several approaches to achieve compression (Furht, Smoliar & Zang, 1996).

4.3.1.1 Lossless Multimedia Compression

This is a class of compression algorithms that allows the exact original multimedia data to be reconstructed from the compressed data. The term *lossless* is in contrast to *lossy*, which only allows an approximation of the original data to be reconstructed. Lossless multimedia compression is used in many multimedia applications. The technique for lossless multimedia compression is that it takes neighbouring pairs of data points, stores their difference and sum, and continues with the sums on a higher level with lower resolutions. A lossless compression technique uses algorithms for compression. Examples of lossless compression include: PNG, GIF, TIFF, and JPEG 2000 (for images); WavPack, WMA Lossless etc (for audio); Huffvuv, JPEG 2000, CorePNG (for video). The general purpose lossless compression algorithms include Run-length encoding (RLE), Lempel-Ziv-Welch (LZW), Huffman coding, Shannon-Fano algorithm and Arithmetic coding (Wikipedia, 2008).

4.3.1.2 Lossy Multimedia Compression

A lossy multimedia compression is a multimedia compression technique where some amount of the original data is lost. Lossy compression is most commonly used to compress multimedia data (audio, video, and still images), especially in multimedia applications, like streaming media and Internet telephony. By contrast, lossless compression is usually used for text and data files, such as text articles. Lossy compression formats suffer from generation loss (Webopedia, 2007).

There are two basic lossy compression schemes; lossy transform codecs and lossy predictive codecs. In lossy transform codecs, samples of the multimedia data are taken, chopped into small segments, transformed into a new basic space, then quantized and coded. In lossy predictive codecs, previous and/or subsequent decoded data is used to predict the current multimedia sample. The error between the predicted data and the real data, together with any extra information is quantized and coded.

Examples of lossy compression methods include the following:

JPEG (image), MPEG – 1, 2, 4 (video), AAC, ATRAC, MP3, WMA (audio), G.711, G.726, AMR (speech). Their advantage over the lossless compression is that they produce smaller compressed files, while still meeting the requirements of the application (Wikipedia, 2008).

The essence of these compressions is to reduce the size of multimedia stream that passes through the communication network, whilst still maintaining the integrity and quality of the stream at the receiving end. By doing this, the amount of multimedia stream that passes through the multimedia device during a multimedia transmission session is reduced. This means that fewer resources are to be used for better results. This reduces the rate of occupation of network resources and this way, the available bandwidth is conserved. A conserved bandwidth helps in ensuring a seamless multimedia transmission.

4.3.2 Caching

Caching is one of the methods of solving bandwidth problems. It is the strategy of keeping a copy of multimedia files in a local location within the multimedia devices. The essence of this is to make it less “expensive” to view or use the same multimedia data again. This is called local caching and it is one way of reducing the extensive use of bandwidth and ensuring that the network is not “clogged”. This reduces the inefficiency in multimedia transmission (Dias, Cope & Wijayarathne, 1996).

The implication of caching is that it involves a cache memory within the multimedia device where “cached” data is stored. It involves storing multimedia data temporarily in the memory cache so that the data can be requested easily by the multimedia applications.

Caching techniques used in multimedia can be divided into two groups; frequency and interval caching. Frequency caching techniques try to cache the multimedia data that is requested more frequently. These techniques require that the file access pattern was registered in the past in order to select the data to be cached. However, the most useful caching techniques are the ones based on intervals.

Intervals caching algorithms benefit from the sequential access pattern of multimedia files. A multimedia file that is being played does not repeat any of its data during the play time. With intervals caching, when another client requests the same file to be played, then some data portions of the file are reused in the system (Fernandez, et. al, 2007).

This data portion, d can be calculated using this equation:

$$d = t * b$$

where d = data portion

t = Interval time, which is the time between when client 1 and client 2 requested the file

b = File Bandwidth, which is the bandwidth needed to play the file.

The diagram (Fig. 4.3) below illustrates the architecture of the cache memory within a multimedia device. Inside a multimedia device, there are multimedia applications. Each application has a client module, which the devices use to connect to the server in order to fetch multimedia data.

The client module runs the client request for one client. When a client (or multimedia application) needs a block of data, it makes a request asking for it in the cache memory.

The cache memory receives all the requests from the client modules, and checks if the requested block is cached already. If it is cached, the cache memory uses an algorithm to send the data block back to the client. If the data is not cached, the cache memory selects a “victim” block to leave the cache memory (if the cache memory is full). Then it uses some algorithm to send a request to the disk memory asking for the data. The new data from the disk memory will now be cached after it sends the data to the client.

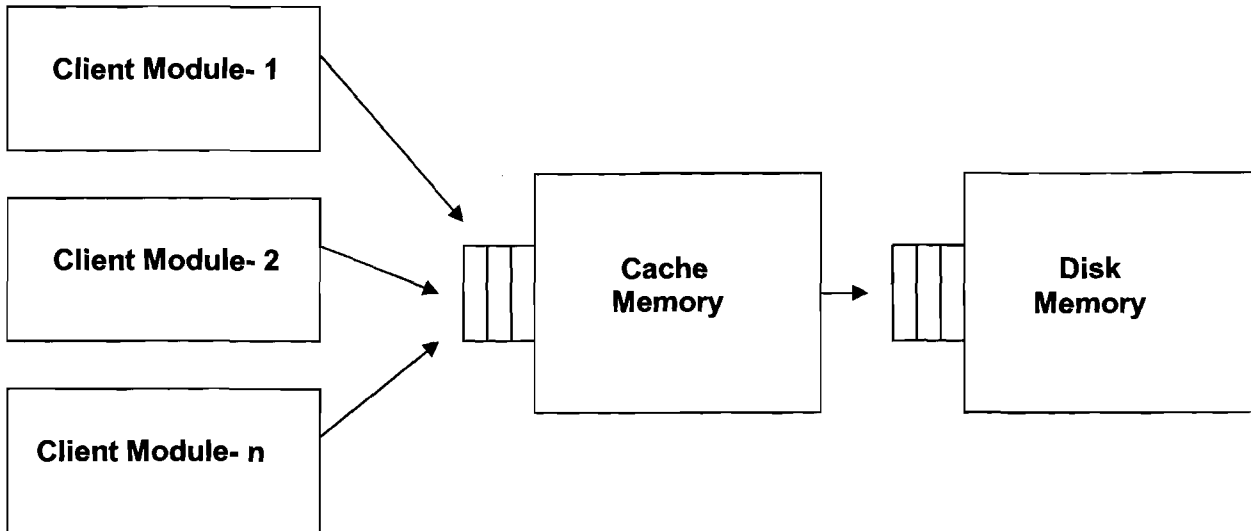


Fig 4.3 Architecture of a Cache & Disk Memory in a multimedia device

This process of caching, where the cache memory stores the multimedia data reduces the bandwidth problems that affect the multimedia transmission. It requires a lot of bandwidth between the cache memory and the disk memory. Because the cache memory is not always accessing the disk memory to get data, the available bandwidth is conserved.

The recommendation of this research work is that during the design of a multimedia device, the cache memory should be large enough to take more data. If the cache memory is large enough, then it can store more multimedia data and shed less of the “victim” blocks, which the client could access later (Fernandez, et. al, 2007).

4.3.3 Mirroring

This is one of the solutions proposed by this research as one way of solving the problems of bandwidth in multimedia transmission. Mirroring is the process of placing copies of large multimedia files that are frequently accessed closer to the end-users (client device). This involves the replication of multimedia servers based on network usage and traffic. The essence of this is to reduce bandwidth usage by providing local access to these files to the client device. The name *mirroring* is derived from the fact that each of the multimedia servers is a mirror of the others (Dias, Cope & Wijayarathne, 1996).

Mirroring is widely used to replicate ftp sites, and in more sophisticated forms, to distribute web content. This research proposes this solution for multimedia platforms. This solution will free the multimedia

network of frequent accesses to the central multimedia server. This frequent request from the many clients to a central resource for the multimedia information engages the bandwidth of the network. This leads to reduced time of responses between the multimedia devices (client and server), and affects seamless multimedia transmission.

The only implication of this solution is the cost of implementation. This will involve that based on distance, frequency of use and traffic; the multimedia servers are mirrored / replicated and placed in locations close to the clients.

The only visible demerit of this solution is that the clients are forced to request multimedia data from their closest servers. If there is a problem with the mirroring or the link is broken, the multimedia device, then the multimedia transmission is affected. However, once these shortcomings are fixed, mirroring provides a feasible way of managing bandwidth situations in multimedia

4.4 SOLUTION TO PROBLEMS ASSOCIATED WITH PACKET LOSSES

In section 3.3, packet loss was defined as a situation where one or more packets being transmitted within a multimedia network fail to reach their destination. There are basically two types of packets losses; namely, “gaps” and “bursts”.

“Gaps” are periods of sporadic drop of single packets, while “burst” is the period of dropping large packets of data. Packet losses are caused by a variety of factors which has already been highlighted in the section 3.3. In this section, the solutions to packet loss will be examined.

There are so many ways of solving the problem of packet loss. These methods can be viewed from the perspective of the sender device and the receiver end of the transmission. The figure below (Fig. 4.4) shows the different methods of solving packet loss problems from the sender and receiver ends.

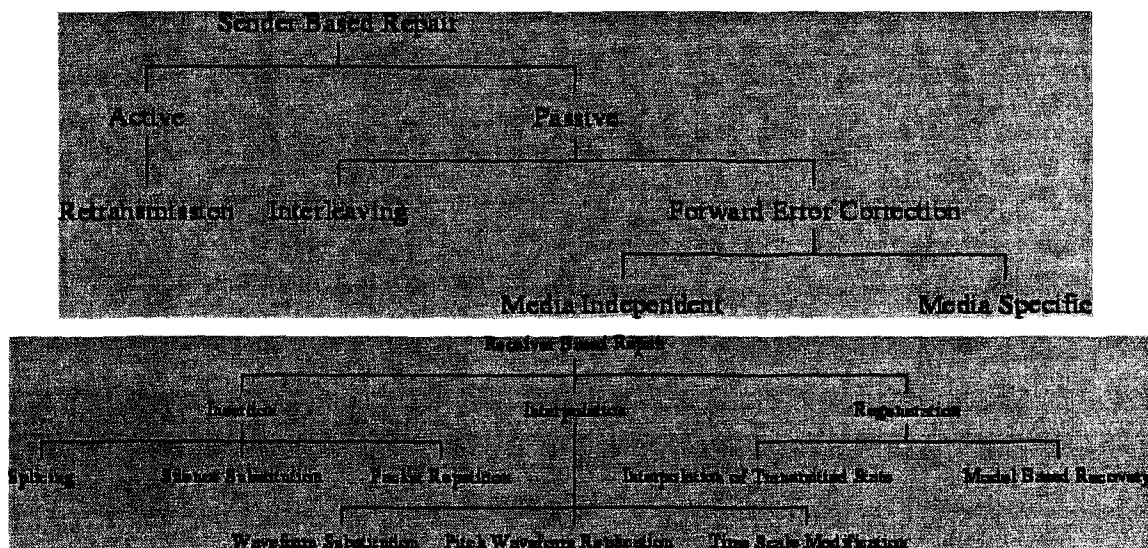


Fig 4.4 Different methods of solving packet loss problems

Source: Tel Aviv University, Israel.

For this research, some of the solutions which are most viable and realistic in solving packet losses in multimedia transmissions will be discussed.

4.4.1 Forward Error Correction - FEC

Forward error correction (FEC) in multimedia transmission is a system of error control for data transmission where the sender adds redundant data to the multimedia information. This redundant information is also known as error correction code. The problem with packet loss is that packets are lost or discarded while being transmitted from sender to receiver.

Therefore, FEC allows the receiver to detect and correct errors without the need to ask the sender for additional data. It's able to do this because of the presence of redundant data (also known as *parity* bits), which takes the place of similar blocks of information that are lost initially.

The advantage of forward error correction is that a back-channel is not required and retransmission of data can be avoided. Retransmission of data leads to higher cost because of higher bandwidths requirements for data retransmission.

The FEC method of solving packet loss becomes much expedient where retransmission is so costly or almost impossible.

To understand the forward error correction, it can be expressed in the form of the probability. Here, the calculation is the probability of a successful frame (or packet) transmission.

For example, in a multimedia transmission involving an MPEG frame made up K original packets, there is the need to calculate the probability that there will be no packet loss.

FEC adds redundant (or parity bits) to the original packets. FEC coding adds $N-K$ redundant packets to the frame. If K or more of the N packets are received;

The probability of a successful frame transmission is expressed mathematically as:

$$q(N, K, p) = \sum_{i=K}^N \left[\binom{N}{i} (1-p)^i \times p^{N-i} \right]$$

FEC is implemented by adding redundancy to the transmitted information using a predetermined algorithm. Each redundant bit is invariably a complex function of many original information bits. The original information may or may not appear in the encoded output. The codes that include the unmodified input in the output are *systematic*, while those that do not are *nonsystematic* (Wikipedia, 2008).

Therefore, the solution lies with the parity (or redundant bits) replacing the lost packets in the block of data. This reduces the amount of packets lost and makes for a more efficient multimedia transmission.

4.4.2 Packet Retransmission

Despite the cost on bandwidth of implementing this solution, packet retransmission is one practical way of solving problems of packet losses. Just as the name sounds, it really makes sense to *retransmit* a *lost* packet.

In many packet-based communication systems such as TCP/IP based systems, packets are communicated over a channel, and if the packet is lost along the way, they are retransmitted. These retransmissions consume significant resources and bandwidth, but where the resources are plentiful, retransmission is viable in solving eliminating packet loss.

It is worthy of note that packet retransmission occurs with the TCP protocol and is not possible with the UDP protocol. This is because the TCP protocol is more reliable because of the retransmission.

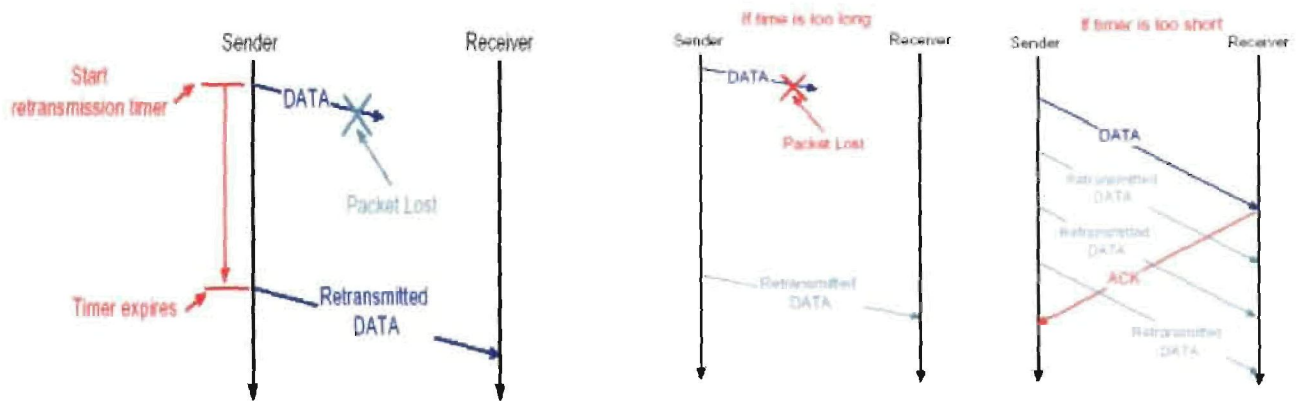


Fig 4.5 Packet Retransmission

Fig. 4.5 above illustrates what happens when a packet is lost between a sender and a receiver within a multimedia communication path. It explains the principle of packet retransmission in multimedia transmission to avoid packet losses.

The sender starts a retransmission timer after a packet is sent out from the sender. After the packet has been sent, the sender starts the timer while it waits for an acknowledgement (ACK) from the receiver. A time is set for the timer (called retransmission timeout RTO). If this timer expires (that is if this retransmission timeout time is reached) before the ACK arrives, the sender assumes that the previous packet was lost in transit. This assumption prompts the sender to resend/retransmit the packet (Hsu, 2003).

One implication that is obvious is with the retransmit timer. If the timer is too long, then the performance of the transmission is affected or degraded. With long timer values, the packets are not retransmitted promptly to cover the effect of the previously lost packets. This leads to latency or delay in packet arrival. Alternatively, if the packet is too short, then there will be unnecessary retransmission of packets even in situations where there are no packet losses. With shorter timer values, the sender resends the packets almost immediately if it doesn't receive an acknowledgement. This leads to huge abuse of bandwidth resources and also affects the quality of the multimedia information transmitted.

In order to avoid the complexities mentioned above, the retransmit timer has to be set to manage the multimedia transmission/retransmission properly. The set of equations of necessary to calculate this time:

The retransmission time out (*RTO*) is calculated from the *RTT* (Round Trip Time);

Where *RTT* is the time between the packet transmission and the receipt of acknowledgement ACK, i.e.

$$RTT = t(ACK) - t(SENT)$$

The *RTT* fluctuates and varied heavily for in different calculations. To have a more steady value, a new term *SRTT* (Smoothed Round Trip Timer) is used. *SRTT* is the average of the *RTT* over time, and it is used as the estimate for the *RTT*.

$$Err = RTT - SRTT$$

$SRTT = RTT + (g * Err)$ - where *g* is recommended to be set to 0.125 (1/2³)

$D_s = D + [h * (|Err| - D)]$ - where *h* is recommended to be set to 0.25 (1/2²)

where *D* is the mean deviation and *D_s* is the Smoothed mean deviation

The mean deviation is a good approximation of standard deviation. It is used here to avoid computing square roots (which is the case for standard deviations).

Finally, the retransmission time out time is calculated using the above parameters as:

$$RTO = SRTT + (n * D_s) - \text{where } n \text{ is recommended to be set to } 4(2^2)$$

This retransmission time out time is measured in microseconds (μ s), milliseconds (ms) and seconds(s).

This depends on how large the network is and the speed of the bandwidth (Hsu, 2003).

From the research conducted for this work, it is recommended that this calculation should be embedded in an algorithm within the software application running in the multimedia device. During the transmission, the multimedia device 'reads' the scripts of the software code for each application. This way, the device is able to 'read' the algorithm and use the calculation above to determine the *RTO* for retransmission of lost packets. This is one of the solutions to packet loss problems in multimedia devices. The only disadvantage of this solution is the bandwidth consumption, its implementation and its restriction to multimedia applications and devices that uses only TCP protocols as its mode of transport.

4.4.3 Packet Loss Concealment - PLC

Packet loss concealment PLC is a technology designed to minimize the practical effect of lost packets in multimedia transmissions. It is one of the different solutions to packet loss problems. This is a solution that is based on the receiver side of the transmission path. The packet loss concealment mitigates the effects of packet loss, which is the failure of one or more transmitted packets to arrive at their destination. Here, packets delayed for a brief time can be dealt with at a given destination by means of a *buffer*. This buffer stores incoming packets for a certain interval and assembles them after that interval has elapsed. However, a more serious problem arises for packets that are dropped or long-delayed.

There are different techniques of implementing PLC. They include the following:

- Regeneration
- Insertion

- Interpolation

In waveform substitution or regeneration, the missing gap (missing packets) is reconstructed by repeating a portion of already received multimedia data. This is called the artificial regeneration of the packets prior to the lost one. Here, the lost packets are replaced with the regenerated packets of the previous stream.

In the insertion (or zero insertion) technique, the lost packets are replaced with a substitute packet (fill-in packet). Here, the lost bits within the packets are replaced with zero bits to fill the bytes that are lost. This is the simplest form of packet loss concealment, and it's much easier to implement.

In the interpolation technique, different types of pattern matching and interpolation/extrapolation are used in order to find a suitable replacement to the packets that are lost. This technique is much harder to implement and more expensive, however, it is more reliable (Mahfuz, 2001).

These methods and techniques are the methods that this research work recommends for solving the problems that face multimedia transmission. These are based on the current challenges that have been described in chapter 3 of this work. There are also other minor challenges that need to be improved on for a seamless multimedia transmission in this age. They are minor because they don't pose as much threat as the ones consider major. It is the hope of this research work that these solutions and recommendations will be implemented. This will provide for a seamless multimedia session, where transmission will be smooth and efficient.

These solutions lead to the next chapter, which shall discuss on the general requirements of a multimedia transmission platform. It shall be discussions on the expectations of a platform that runs a seamless multimedia transmission. The discussion shall be based on the components of multimedia transmission architecture. The discussion shall isolate the composite parts of a multimedia network and discuss on the expectations and requirements for each of these parts.

This shall give vendors of multimedia devices a framework or blueprint of a multimedia platform, where most of these problems and challenges are non-existent or minimal.

CHAPTER 5

DISCUSSIONS ON THE ARCHITECTURE OF A MULTIMEDIA TRANSMISSION PLATFORM

In the chapters preceding this one, the problems, the solutions and improvements to multimedia transmission have been discussed. The research looked at the major and some of the minor issues that mitigate efficient multimedia transmission.

Here in this chapter, the research shall be focusing on the expectancies (or otherwise outlooks) of a multimedia transmission platform. The expectancies here could also mean the general requirements of the architecture of a multimedia transmission platform. The essence of this chapter is to make the beneficiaries of this research work to understand the overall expectations and architecture of a multimedia transmission. It's not enough to mention just the problems and the solutions to multimedia transmission. A careful look at the expectations of a multimedia transmission platform is needed to reveal the picture of a standard multimedia transmission platform.

5.1 ARCHITECTURE OF A STANDARD MULTIMEDIA TRANSMISSION PLATFORM

The word "architecture" here means the structure or organization of a standard multimedia transmission profile. A standard multimedia transmission platform comprises of the following components:

1. Multimedia Database
2. Multimedia Server
3. Multimedia Client
4. Network or Communication path
5. Protocols
6. Software - (applications that runs within the above listed hardware)

Numbers 2 and 3 join to form the multimedia devices, which has been the emphasis of this research work.

5.1.1 Multimedia Database

A multimedia database is a database that hosts one or more primary media file types such as .txt (documents), .jpg (images), .swf (videos), .mp3 (audio), etc.

They can be classified into three main categories:

- a. Authentication Multimedia Database: also known as the Verification Multimedia database. This is used to authenticate the multimedia client applications before they can access the database.
- b. Identification Multimedia database: this is a database that compares one-to-many.
- c. Biometric Multimedia Database: this is a newly emerged database and specializes in automatic human verification like the fingerprints and other human identities based on algorithms (Tech FAQ, 2008).

The summary is that the multimedia database is the location where the multimedia data is located or stored. It is from this storage that the servers retrieve the multimedia data for the clients. The multimedia databases provide almost all the basic functions that a traditional database provides. It also provides new and improved functionalities and features. They are required to provide a holistic framework that stores, processes, retrieves, transmits and presents a list of media data types in a wide variety of formats.

5.1.2 Multimedia Server

Multimedia servers store, manage multimedia objects and deliver data streams in real-time, in response to requests from the clients or the users. In addition, multimedia servers process the stored information before delivery to users. The contents of a multimedia server usually comprises of a whole range of multimedia applications. This could range from huge amounts of multimedia deposits (like video and audio) to small amounts (like text and images).

Irrespective of the capacity of a multimedia server, the main aim of having a multimedia server is to provide the platform with maximum number of clients. These are the number of clients that can simultaneously access the applications, with minimal latency, little jitter and minimal bandwidth consumption.

Most times, the multimedia servers perform the roles of multimedia databases. The servers are not so different from the network file servers; this is because they also support the storage and the retrieval of multimedia data. In situations where the multimedia database is embedded in the multimedia server, frequent resource usage between database and server are minimized.

The multimedia servers are positioned so that the clients are able to retrieve the multimedia objects from the servers for real-time 'playback'. Making the multimedia servers available at close distances to the servers help in reducing bandwidth consumption (mirroring in chapter 4).

For a robust multimedia server, the retrieval is also expected to be interactive. This means that the multimedia clients can pause, stop, resume and even record the real-time media application.

A typical multimedia server can be centralized or distributed. A centralized multimedia server is good for small scale multimedia transmission and sessions. For a huge multimedia transmission profile, a distributed multimedia server system makes for conservation of network resources.

A multimedia server is expected to have the following component parts:

- *Multimedia data retrieval scheduling and placement*: performs the function of balancing the load, efficient data retrieval and scheduling of system resources.
- *Buffer Management*: for taking care of the variances in data retrieval to minimize latency and jitter effects. This also implements memory-based techniques such as sorting.

- *Admission control and quality of service*: for maintaining the performance criteria of the server. It guarantees end-to-end performance by preventing stream overload.

Some of the characteristics of a multimedia server include the following; multimedia servers:

- Receive requests from multimedia clients
- Dispatches responses to client requests
- Executes database retrievals, because of its direct connection the multimedia database
- Performs and provides the back-end infrastructure to the multimedia architecture
- Manages multimedia data integrity

5.1.3 Multimedia Client

A multimedia client is a multimedia application or system that accesses another service (remotely or in the same location), on another system known as the multimedia server. Here, the multimedia client is used to describe both the software applications and the hardware clients or multimedia devices that host these applications. A typical example of a multimedia client is the web browser that runs on a local computer or workstation. Both the software applications and the computer that runs the software for the web browser constitute the multimedia client.

The multimedia client is a very essential component of a multimedia architecture. This is because the “friendly” multimedia applications are part of the multimedia client. This is the part of the multimedia architecture that is considered as the “front-end”, because the multimedia users interact with it. The multimedia servers, the network, the multimedia database form the “back-end” of the architecture.

The multimedia client comprises of two types:

a. Multimedia client sender; b. Multimedia client receiver.

The multimedia client can function in both ways as a sender and as a receiver. When the client is the one sending out multimedia streams to the multimedia server, it acts as a sender. When the client is receiving multimedia streams from the multimedia server, it acts as a receiver.

Some of the characteristics of a multimedia client include the following; the multimedia client:

- Sends out and receives multimedia applications from the server
- Waits for and receives replies and requests
- Is the front-end of the application that the user sees and interacts with
- Contains solution-specific logic and provides the multimedia interface
- Connects to a small number of server at once
- Is the end-user part using graphical user interface GUI to appeal to end-users

5.1.4 Multimedia Network or Communication Path

The communication path or the network is the path through which the multimedia data or streams are transmitted. This path comprises of the network connections between the different nodes of the multimedia architecture. The network provides the link between the database, the server and the clients. The speed and the capacity of the network depend on the available bandwidth and the size of the multimedia data being transmitted. The network is important in a multimedia platform because of the inherent characteristics of a multimedia stream. Most multimedia data are transmitted on real time, hence the importance of a communication path through which the data can be transmitted.

There are different types of communication paths or networks in a typical multimedia architecture. These different types of communication paths have developed over the years. The diagram (Fig. 5.1) below illustrates the evolution of communication networks in a multimedia platform.

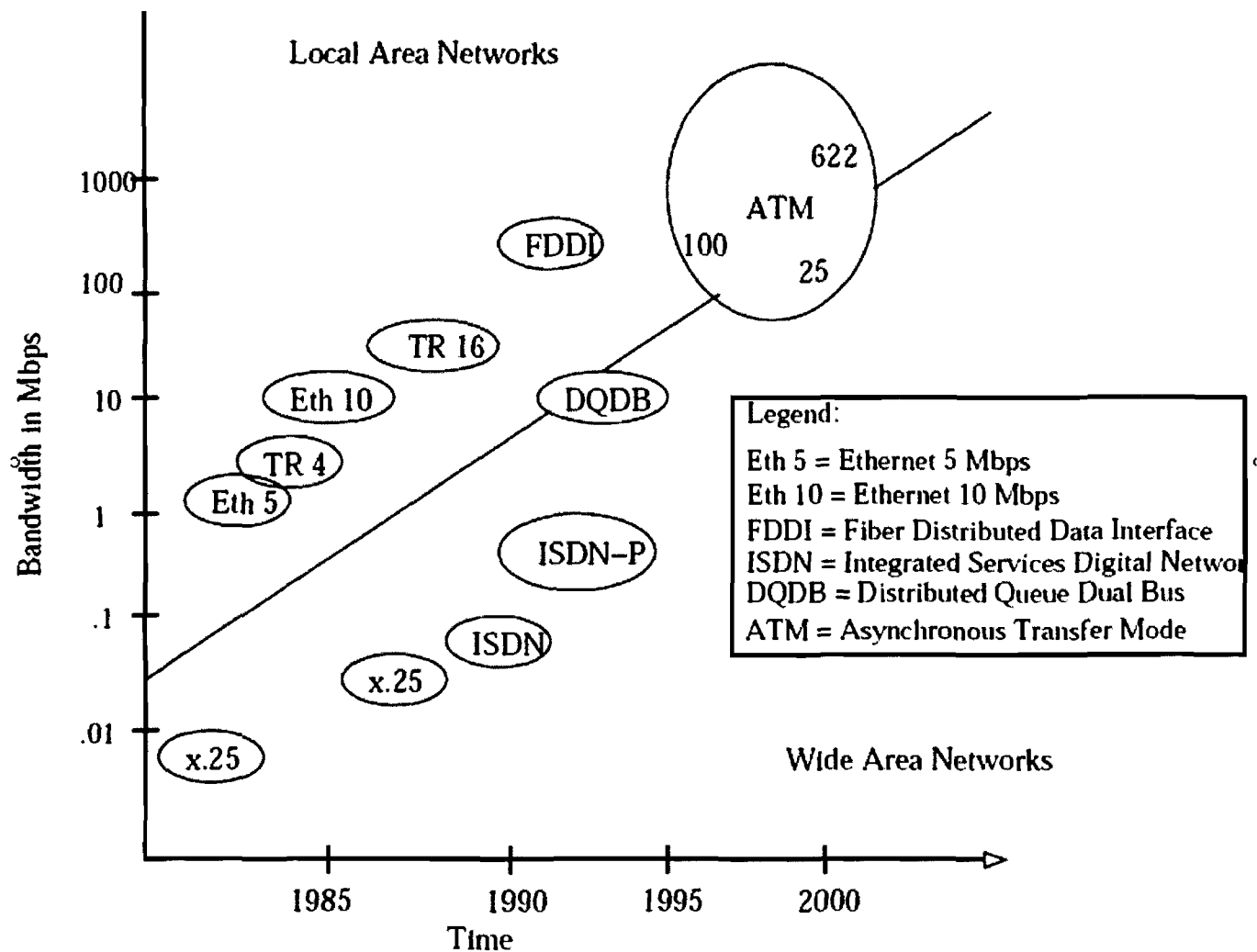


Fig 5.1 The Evolution of the different types of Multimedia Networks

Source: India Institute of Science

The diagram illustrates the different types of multimedia networks that have evolved. With the recent advancements in technology, the multimedia network has been expanded for improved bandwidth.

The multimedia networks that drive today's multimedia transmission include the following:

- GPRS – General Radio Packet Service (for the 2.5G networks – circuit switched technology)
- 3rd Generation Networks (3G network, UMTS, IMS, EVDO – hybrid of circuit and packet switched technology)
- 4th Generation Networks (4G networks – Long Term Evolution LTE)
- Local Area Network (LAN), Wide Area Network (WAN)
Wireless LAN – Wi-Fi(IEEE 802.11a/b/g/n standard)
- WiMAX (IEEE 802.16 standard)
- Internet, (which forms the backbone of the entire multimedia networks).

These network standards have evolved as a result of the need to improve the multimedia network and transmission through enhanced network throughputs.

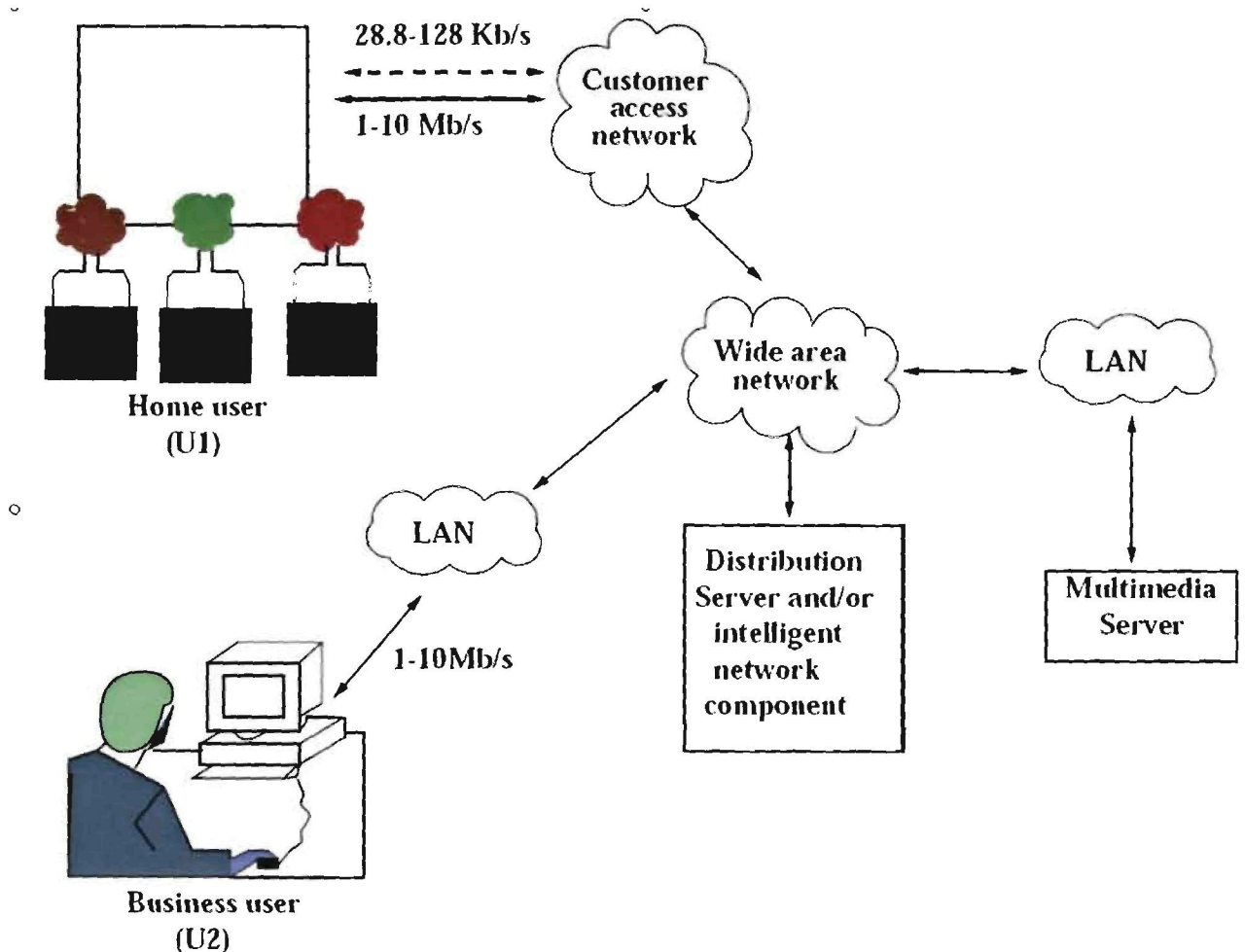


Fig 5.2 A typical Multimedia Application Environment

Source: India Institute of Science

5.1.5 Multimedia Protocols

Multimedia protocols are standards that control and enable the connection, communication, and data transfer between two multimedia devices or end-points. They define the general rules that guide the syntax, synchronization, and the semantics of communication.

There are so many protocols in the different areas of technology; however, the emphasis here is on multimedia protocols. The place of protocols in a multimedia transmission suite is very huge. Without the protocols, there won't be any guide for the definition of multimedia transmission.

The problem has always been the persistent focus on multimedia protocols as a way of improving multimedia transmission.

Multimedia protocols are specific to the needs of multimedia and its transmission. This is because most of the multimedia data are transmitted in real-time. This makes the real-time protocols the major protocols for multimedia transmission. There can be different classifications of the multimedia protocols. They can be grouped into *Multimedia protocols over TCP* and *protocols under UDP*. They can also be grouped into *Real-time* and *Internet Telephony* protocols. The popular multimedia protocols in use today include:

a. Real-Time Transport Protocol (RTP): This is standardized application layer protocol for delivering multimedia information (audio and video) over the multimedia network. It provides the support for the transmission of real-time data streams. The RTP usually needs support from lower layers, and it is supported by UDP. This is because the UDP does not need retransmission and is connectionless. So delay is reduced in the UDP and therefore, it makes it suitable for a RTP, where delay is a liability.

b. Real-Time Control Protocol (RTCP): This is an application layer protocol that works in conjunction with the RTP to monitor the quality of service of the multimedia stream.

c. Resource Reservation Protocol (RSVP): This is a network control protocol that allows the receiver to request a special end-to-end quality of service for data flows. It is a one way reservation (from receiver to sender).

d. Real-Time Streaming Protocol (RTSP): This is a client-server presentation protocol that enables the controlled delivery of multimedia streams over the IP network. This is the protocol that allows the media player (or client) to control the transmission of the media stream (like Pause/Resume, Playback, and Fast-forward/Rewind etc). It also allows conference recording, retrieval of a media from a server etc.

Protocols (a–d) above are collectively referred to as Real-Time Protocols. They are usually transmitted through the User-Datagram Protocol (UDP), because of the delay reductions in UDP transmission.

e. Session Initiation Protocol (SIP): This is a signaling protocol that is used for initiating, managing, and terminating multimedia sessions (like data, voice, and video streams). SIP comes into existence

because of the need to have a two way transmission. It manages the session, by negotiating changes in the capabilities or requirements of the end points (multimedia clients) during communication. It also “tears down” or terminates the communication, when not in use. This is done to efficiently manage bandwidth.

f. Session Description Protocol & Announcement Protocol (SDP & SAP): Apart from the SIP that manages the signaling in multimedia sessions, announcements are handled by two protocols. They are the Session Description Protocol, SDP and Session Announcement Protocol, SAP. They display the multimedia session names within the traffic. They function on top of both the TCP and UDP protocols.

g. H.323 Protocol: This is sometimes called the ITU-T “Umbrella” documentation. It defines the protocols to provide audio-visual sessions on packet networks. This is used more effectively by networks that employ voice and video conferencing. It functions with different networks like IP, PSTN etc. Under the IP network, the H.323 protocol uses the TCP protocols for transmission.

Protocols (e-g) above are collectively referred to as Internet Telephony protocols.

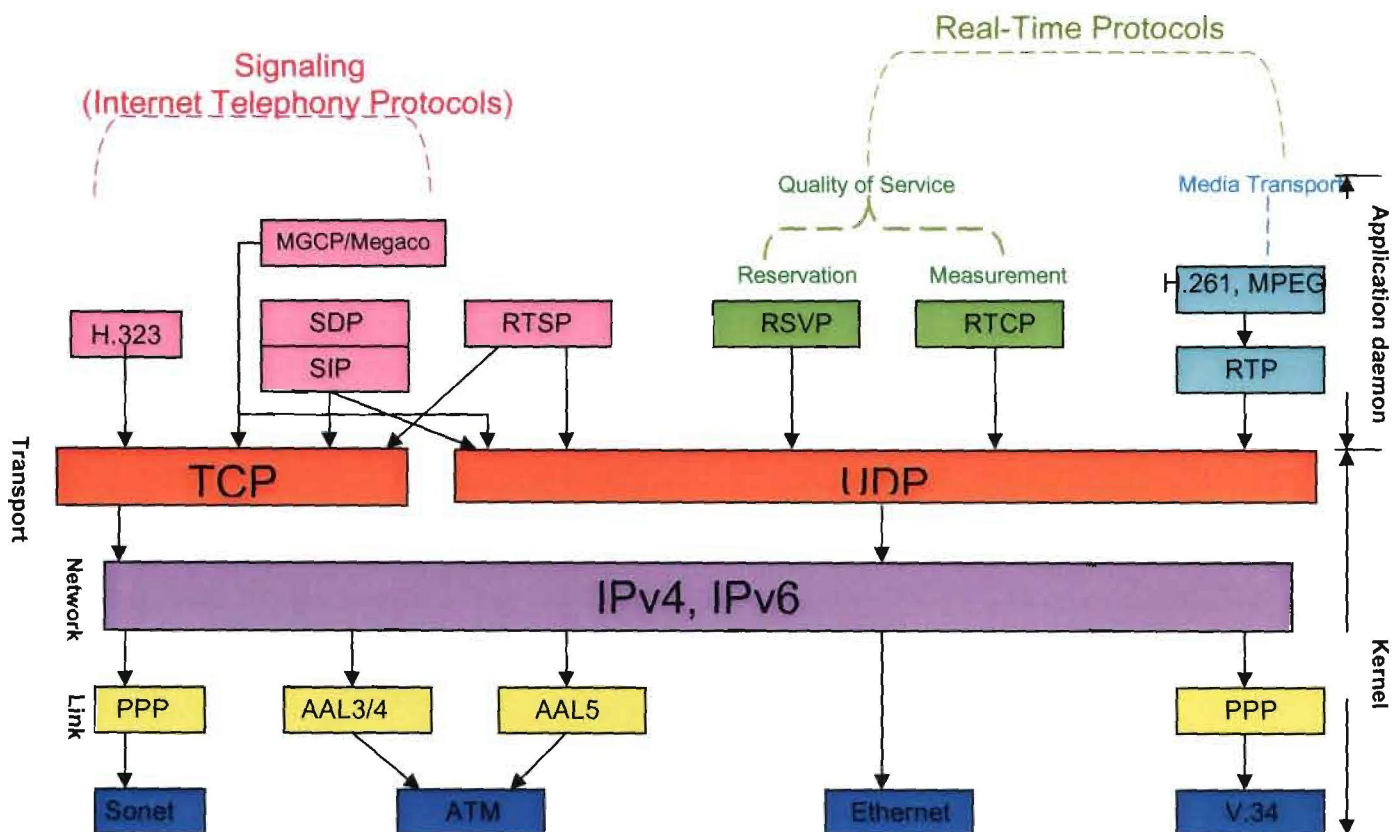


Fig 5.3 Multimedia Protocols & the IP network layers

The diagram above (Fig. 5.3) gives an outline of the different multimedia protocols that are used presently. It also shows the different IP protocol stack and includes the different layers of an IP network (physical, link, network, transport and application). It shows the inter-relationship of the protocols and

their dependencies. A careful look at the protocols outlines the explanation above (under the protocol sub-section).

5.1.6 Software Components

One integral part of multimedia platforms is the software. Just like in every other field of technology, software dominates all the aspects of multimedia. This includes in the propagation, encoding, transmission and decoding of the multimedia information. There are software components in all the layers of the multimedia platform.

A typical platform consists of the hardware stack (servers, databases, hardware devices) and the software stack (database applications, server applications, client programs, protocols etc.).

This implies that the entire concept of multimedia revolves around the software components. The development and the rapid advancement of multimedia have been made possible by advancements in software development.

The number of multimedia software are running into numbers and still counting. Some of the common programming languages that are used in developing these software programs include Java, C, C++ and others.

Discussed above is the basic structure of typical multimedia transmission architecture. There are so many other minor components that make up a multimedia platform, but the above are the fundamentals of a multimedia transmission system. In this research work, more emphasis has been on improving multimedia transmission in the multimedia clients and multimedia servers (or multimedia devices).

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 Conclusion

Since the inception of multimedia transmission, there has been the need to make it better and more efficient. Understanding the trend of the focus on multimedia is the key to improving the transmission of multimedia information. While some have argued that protocols remain the best option to resolving the problems facing multimedia transmission; others have argued against it by saying that more focus should be on multimedia devices.

The Journal of Systems and Software, Volume 79 of 2006 supports more emphasis on protocol. It stated that “even though there are problems with multimedia transmission, protocols remain the most viable option of improving multimedia transmission. Improvements in the protocols are the reason for some advancement in the latest multimedia technologies. Given the time, good software applications and protocols will solve every problem associated with multimedia applications and their transmission”.

Ifi UiO, 2001 in an article on end-to-end transmission, negates this idea and supports more emphasis on multimedia devices. “The problem with efficient multimedia transmission is because the terminals that provide the end to end transmission have been neglected and much emphasis has been on the software and protocols”.

Godred & Reine, 2006 also supports the idea of a shift from protocols by saying that “we have discovered that we are faced with a peculiar problem, and this is because we have diverted our focus from the main issue of efficient point to point multimedia streaming using the devices. This has improved the protocols over time and affected the development of better multimedia devices that make transmission better”.

It is evident from the contents of this research that there is the need for an urgent re-direction towards enhancing multimedia devices to meet the challenges of multimedia transmission.

The trend in multimedia technology has transcended beyond focusing all the attention and resources only on protocols and software, while ignoring devices, which is a very significant component.

The essence of this research work has been to highlight the trend in multimedia technology and highlighting the dichotomy that exists between the focus on protocols and on multimedia devices. This has led to the discovery of numerous underlying problems, because of the lack of focus on multimedia devices. The entire chapters of this work have been dedicated to stating the fact that multimedia devices play a major role in multimedia transmission. Not realizing this role has lead to inefficiencies in providing seamless multimedia transmission.

Improving Multimedia Transmission through enhanced Multimedia devices

This research is of the opinion that multimedia devices, not just protocols, should be enhanced to improve multimedia transmission. This is judging from the fact that improvements in protocols lately has not lead to significant improvements in multimedia transmission. It has to be recognized that that protocols cannot solve these problems independently. Enhancing the capabilities of these multimedia devices is one of the sure ways of improving the quality of multimedia transmission. The reality is that enhanced multimedia devices will work in consonance with these protocols to improving multimedia transmission.

Therefore, the multimedia industry should recognize this technology trend and focus more on multimedia devices. The manufacturers of multimedia devices should take up the challenges by emphasizing more on the research and development of these devices to enhance them. A more aggressive approach to multimedia devices is needed in this situation. Since the protocols are not independent components of multimedia transmission, it is meant to be compatible with these enhanced multimedia devices. This means that devoting more resources on multimedia devices should also involve the integration of these devices to the protocols.

As protocols alone cannot solve the problem of multimedia transmission, the same applies to enhanced multimedia device alone. Hence, there is the need for a structured integration of these enhanced multimedia devices and protocols in the future.

Based on the work in the previous chapters, this research has also proposed a set of potential methods of enhancing multimedia devices. This it did by isolating the problems that face multimedia transmission and recognizing that they are device specific, not protocols. It is evident therefore that the solution falls within the multimedia devices and not the protocols. It has taken steps to propose these solutions. These are conceptual solutions for each of the problems that have been identified. The solutions are to be implemented in the core design of multimedia devices.

It is difficult and sometimes impossible to implement a solution on a finished product. Based on this, enhancing these multimedia devices will involve having these solutions implemented at the beginning stage of the lifecycle of the multimedia devices. From the design to the manufacturing of the multimedia devices, the emphasis should be on how to make these devices achieve the desired result of actualizing seamless multimedia transmission. The enhanced devices should be tested with the other components of multimedia transmission to ensure that the devices are compatible with the existing architecture.

For problems associated with latency and jitters, there should be an implementation of packet delay buffers. These buffers can only be embedded in the design of the multimedia devices, to be able to counter the delays. This implementation is not feasible in protocols, hence the need for a shift in focus to multimedia devices. Solving packet losses require packet acknowledgement timers to be built in the hardware design of the multimedia devices. This will be able to trigger re-transmission of packets in the

sender devices when packets are lost. Actualizing this will be in the multimedia devices and not in the protocol stack as is the case presently. Therefore, a shift of emphasis to multimedia device is once again evident, in improving multimedia transmission.

It follows from the above that the trend in multimedia technology supports a shift in strategy from protocols towards multimedia devices. This shift is recommended in order to solve real life problems facing multimedia transmission. With the fast development and delivery of software and protocols, these problems have not gone away. Therefore, implementing these set of solutions to newly produced multimedia devices will enhance them largely.

However, a more holistic approach is required. This involves focusing more on multimedia devices, whilst still recognizing the need for proper integration and interoperability with the protocols. The combination of these enhanced multimedia devices with the developed protocols will see a huge improvement in multimedia transmission.

6.2 Recommendations

Proffering solutions to the problems that have long affected the way people use multimedia has been the major focus of this dissertation. This exercise has made it possible for this research to make recommendations that could enhance a more seamless transmission of multimedia.

It is important for these recommendations to be taken in and effected to improve the way multimedia is transmitted.

To enhance a smoother transmission of multimedia information, the problems that have been highlighted must be taken into consideration. This is the first process. The next process is to understand that the trend in technology at the moment supports a more focused approach towards enhancing multimedia devices. Then, a workable solution should be implemented to solve the problems. Most of the times, the stakeholders in the multimedia industry are ignorant of this trend. The aim of this research is to highlight this trend, so that the industry can work towards it.

It is hoped that this research work will be resourceful to all stakeholders in the multimedia industry and its applications. A collective understanding of the trend, the problems, the focus and the solutions, will increase the Quality of service delivered in multimedia transmission. Focusing on a single component of multimedia transmission architecture (in the case of protocols) does not solve the problem of multimedia transmission. This research has kick-started the process of focusing more on multimedia device enhancements.

The standard organizations that have authority over multimedia contents should make regulations for the multimedia industry to bother more on the devices. There seems to be no standard way of manufacturing multimedia devices. This has been the major cause of the crisis. This chaotic circumstance; where there

Improving Multimedia Transmission through enhanced Multimedia devices

are no standards over multimedia devices has led to problems of interoperability and compatibility within multimedia. There must be standards and recommendations for multimedia devices. Organizations like IETF, ITU-T, ANSI, IEEE, IMA (Interactive Media Association), WiMedia Alliance, MPEG should take up the responsibility to releasing standards for multimedia devices. It is not surprising though, because these organizations were constituted as protocol-standard bodies.

This is the challenge and only a reversal of trend can improve the transmission of multimedia from the device perspective.

Looking ahead, it looks like the world is far away from completely actualizing this focus shift and solving these problems. However, the relief lies in understanding this trend and working towards reducing the frustration that materializes as a result of these problems. It is the desire of this research that most of the solutions and improvements recommended in this work will be implemented to enhance subsequent multimedia devices in improving multimedia transmission. This shift of focus to devices will help jumpstart a new era of multimedia device enhancements.

REFERENCES

1. Z. Avramovic, "Policy based routing in the Defense Information System Network", IEEE Military Communications Conference, vol: 3, pp: 1210-1214. 2004
2. D. Bertsekas and R. Gallager, *Data Networks*, Prentice Hall, 1987.
3. S. Blake, D. Black, M. Carlson, E. Davies, et. al., "An Architecture for Differentiated Services", RFC2475, December 1998.
4. R. Braden, L. Zhang, S. Berson, S. Herzog and S. Jamin, "Resource Reservation Protocol (RSVP): Version 1 Functional Specification", RFC2205, September 1997.
5. B. Cavusoglu, D. Schonfeld, and R. Ansari, "Real-time adaptive forward error correction for MPEG-2 video communications over RTP networks", Proceedings of the IEEE International Conference on Multimedia and Expo, Baltimore, Maryland, 2003.
6. N. Chapman and J. Chapman, *Digital Multimedia*, John Wiley and Sons Ltd, 2000.
7. B. Furht, *Multimedia Tools and Applications*, Kluwer Academic Publishers, 1996.
8. B. Furht, "Real-time issues in distributed multimedia systems", Proceedings of the Second Workshop on Parallel and Distributed Real-Time Systems, April 1994, pp: 88-97.
9. M. Handley and V. Jacobson, "SDP: Session Description Protocol", RFC2327, April 1998.
10. A.B. Johnston, *Understanding the Session Initiation Protocol*, Artech House, 2000.
11. S. Kent and R. Atkinson, "Security Architecture for the Internet Protocol", RFC2401, November 1998.
12. A. Leon-Garcia and I. Widjaja, *Communication Networks: Fundamental Concepts and Key Architectures*, McGraw-Hill, 2000.
13. H. Liu and P. Mouchtaris, "Voice over IP Signaling: H.323 and Beyond", IEEE Communications Magazine, October 2000.
14. J. Kurose and K. Ross, *Computer Networking: A top-down approach featuring the Internet*, Addison Wesley, 2001.
15. Web Simpson, *Video Over IP: A Practical Guide to Technology and Applications*, 2005.
16. Martin Isenburg, *Transmission of multimedia data over lossy networks*, McGraw-Hill, 1996.

17. Marcel Dekker, *Compressed Video over Networks*, Sun & Reibman (eds.), 2001.
18. D. Schonfeld, "Image and Video Communication Networks", (Invited Chapter). Handbook of Image and Video Processing. A. Bovik (ed.), Academic Press: San Diego, California, Chapter 9.3, pp. 717-732, 2000.
19. Minoli, D., Minoli, E., "Delivering Voice over IP Networks", John Wiley & Sons, 1998.
20. Hassan, M., Nayandoro, A., "Internet Telephony: Services, Technical Challenges, and Products", *IEEE Communication Magazine*, 2000.
21. Jin Guojin, *Algorithms and Requirements for Measuring Network Bandwidth*, Distributed Systems Department, Berkeley California, 2007.
22. Jonathan Davidson, James Peter, *Voice Over IP Fundamentals*, Cisco Press, 2005.
23. E. McPhillips, *The Factors Affecting the Growth of VoIP*, HP Laboratories Bristol, 1999.
24. Ouellette Michel, et al, *Adaptive Jitter Buffer Control*, IBM Journal for Research & Development, 2006.
25. M. Natarajan, "Multimedia and Data Transfer Technology: The Challenges and Delivery", DESIDOC Bulletin of Information Technology, Vol. 23, No.4, July 2003.
26. B. Furht, S.W. Smoliar, H-J. Zang, "Video and Image Processing in Multimedia Systems", Kluwer Academic Pub., 1996.
27. Gihan V. Dias, Graham Cope and Ravi Wijayarathne "A Smart Internet Caching System," Proc. INET '96, Internet Society, 1996.
28. Javier Fernandez, Jesus Carretero, Felix Garcia, Jose Perez, A. Calderon, "Enhancing Multimedia Caching Algorithms Performance Through New Interval Definition Strategies", IEEE Computer Society, 2007.
29. Sam Hsu, *Computer Network Programming; TCP Overview*, Florida Atlantic University, Florida, 2003.
30. Ejaz Mahfuz, *Packet Loss Concealment for Voice Transmission over IP Networks*, McGill University, Montreal Canada, 2001.
31. <http://www.patentstorm.us/patents/7359324/description.html>, 2008.
32. http://www.webopedia.com/TERM/L/lossy_compression.html, 2008.
33. <http://www.cisco.com>, Cisco IP Journal, 2008
34. <http://www.wikipedia.com>, Multimedia_Literacy, 2008.

Improving Multimedia Transmission through enhanced Multimedia devices

35. Jerzy Wachol, *Multimedia in Management facing Globalization processes and Cognitivisticis*, Annales UMCS Informatica AI 4, 2006.
36. O.O. Ayeni, D.N. Saka & G. Ikwuemesi, *Developing a Multimedia GIS Database for Tourism Industry in Nigeria*, Faculty of Engineering, University of Lagos, 2002.
37. Tamara Powell, Susan Eller & Daniel Shockley, *Preparing Students for the 21st Century: Using Multimedia Technology to Enhance Writing Instruction*, Louisiana Tech University, Louisiana, 2005.
38. Ifi UiO, *Enhanced Speech and Audio Coding Technologies Enable Innovative Mobile Multimedia Services*, University of Oslo, Norway, 2001.