A DUAL BUFFER SCHEME FOR COMBACTING FRAME ERASURE PROBLEMS IN VoIP COMMUNICATION

Olufade, F.W.Onifade

Department of Computer Science,   University of Ibadan

Abstract

This paper refers to the transmission of speech across data-style networks. This form of transmission is conceptually superior to conventional circuit switched communication in many ways. Quality of service (QoS) is fundamental to the operation of a VoIP network that meets users' quality expectations. However, the implementation of various security measures can cause a marked deterioration in QoS. These complications range from firewalls delaying or blocking call setups to encryption-produced latency and delay variation (jitter). Because of the time-critical nature of VoIP, and its low tolerance for disruption and packet loss, many security measures implemented in traditional data networks are simply not applicable to VoIP in their current form' firewalls, intrusion detection systems, and other components must be specialized for VoIP.

To address the QoS issue of packet loss, this research work develops a new model making use of adative buffering technique. Our model implements dual buffers, strategically placed to manage and ameliorate frame erasure problem amongst other QoS factors. With this implementation, the new model achieved a better reduction in the amount of packet loss.

Keywords: Delay, Buffer, Packet Loss, QoS, VoIP, CS/GR, Frame erasure

1 INTRODUCTION

Data networks were initially designed – and are still widely used – to interconnect computers, their servers and the management platforms between them. The applications which use data networks are generally computer application for the exchange of data between machines. The use of data networks experienced a boom following the widespread introduction of computers in companies and universities in the wake of the progress made in the field of information technology (mini-computers, the microcomputers, including portable). Data networks display a certain "neutrality" vis-à-vis the applications they support, being generally characterized by the transfer of data in "packets", each containing all or part of the data to be transferred between two computers and the address of the destination computer. The network is essentially responsible for transferring the packets to the destination computer with a quality of service defined by the user, understood here as the application on the transmitting computer, without any prior knowledge of the nature of that application. The advantages here is that of a multiservice network whose architecture is not subordinated to a specific application, it even possible to use the network at a later stage for applications which had not even come into being at the time of its original deployment. This flexibility had led, for instance, to the use of data networks (especially IP ones) for new types of "human-related" communication applications, like voice and video transmission.

Reliable, high-quality voice-quality voice services, the kind that users expect from the PSTN is the ultimate objective of VoIP. The quality issues of VoIP are very complicated and challenging. It is hard to achieve same level of Quality of Service as in PSTV.

Here are the main parameters that affect voice quality in VoIP and what can be done to maximize quality.

- Bandwidth: Available rate for internet connection always tops the list of factors affecting voice quality in VoIP conversations. The bandwidth for VoIP is the key for voice quality. For instance, in a dual-up connection, the provided quality cannot be compared to other medium like broadband. A broadband connection will work right, as long as it is not spotty, and shared with too many other communication applications.
- Equipment: The VoIP hardware equipment employed can greatly impair the resultant quality. Poor quality equipments are normally the cheapest ones (but not always!). It is therefore always good to have as much information as possible on a router or IP phone before investing on it and starting to use it. Possible information can be sought from reviews and discussions in forums. It might also be that the hardware chosen is the best in the world, but the still the problems persist– in this case, suitability of purpose comes to play.
- Phone frequencies: The frequency of the IP phone may cause interference with other VoIP equipment. There are cases where people using 5.8 GHz phones were having voice quality problems. When all troubleshooting tricks failed, changing the phone to one with a lower frequency (e.g. 2.4 GHz) solved the problem.

QoS is one of the major issues concerning VoIP. There must be certain level of QoS in VoIP if it is expected to displace PSTNs in the future. Quality of Service (QoS) is fundamental to the operation of a VoIP network. Despite all the money VoIP can save users and the network elegance it provides, once it fails to deliver at least the same quality of call setup, voice relay functionality and voice quality comparable to the traditional telephone network, then it will provide little or no added value. The implementation of various security measures can degrade QoS. These complications range from delaying or blocking of call setups by firewalls to encryption-produced latency and delay variation (jitter). QoS issues are central to VoIP security. If QoS was assured, then most of the same security measures currently implemented in today's data networks could be used in VoIP networks. But because of the time-critical nature of VoIP, and its low tolerance for disruption and packet loss, many security measures implemented in traditional data networks just aren't applicable to VoIP in their current form. The main QoS issues associated with VoIP that security affects are presented here:
- Delay/ latency
- Jitter
- Echo
- Packet Loss

2 REVIEW OF QUALITY ISSUES IN VoIP
2.1 DELAY/ LATENCY

Latency is the time between the moment a voice packet is transmitted and the moment it reaches its destination. VoIP delay or latency is characterized as the amount of time it takes for speech to exit the speaker's mouth and reach the listener's ear. The time it takes a packet to get across the packet switched network to its destination. There are three types of delay:
- Propagation Delay: The physical distance between the two ends of the data network determines how long it takes to propagate a signal between them. This delay is proportional to the speed of light, that is, the time needed by the physical signal as it passes through copper, optical, or wireless media. Light travels through a vacuum at a speed of 186,000 miles per second, and electrons travel through copper or fiber at approximately 125,000 miles per second. A fiber network stretching halfway around the world (13,000 miles) induces a one-way delay of about 70 milliseconds (70ms). Although this delay is almost imperceptible to the human ear, propagation delays in conjunction with handling delays can cause noticeable speech degradation.

- Processing Delay: Also known as Handling delay, processing delay defines many different causes of delay (actual packetization, compression, and packet switching) and is caused by devices that forward the frame through network.
- Serialization delay: Serialization delay is the amount of time it takes to actually place a bit or byte onto an interface.

Delay is caused by slow network links. This is what leads to echo. There are two ways latency is measured: one direction and round trip. One direction latency is the time taken for the packet to travel one way from the source to the destination. Round-trip latency is the time taken for the packet to travel to and from the destination, back to the source. In fact, it is not the same packet that travels back, but an acknowledgement.

Latency is measured in milliseconds (ms) – thousands of seconds. A latency of 150ms is barely noticeable so is acceptable. Higher than that, quality starts to suffer. When it gets higher than 300ms, it becomes unacceptable. Below are the effects of latency over voice quality:

- It slows down phone conversations
- Untimeliness can result in overlapping noises, with one speaker interrupting the other
- Causes echo
- Disturbs synchronization between voice and other data types, especially during video conferencing.

2.2 JITTER

Jitter refers to non-uniform packet delays. It is often caused by low bandwidth situations in VoIP and can be exceptionally detrimental to the overall QoS. Variations in delays can be more detrimental to QoS than the actual delays themselves. Jitter is the variation of packet inter-arrival time. Jitter quantifies the effects of network delays on packet arrivals at the receiver. Packets transmitted at equal intervals from the left gateway arrive at the right gateway at irregular intervals as shown in Figure 1.
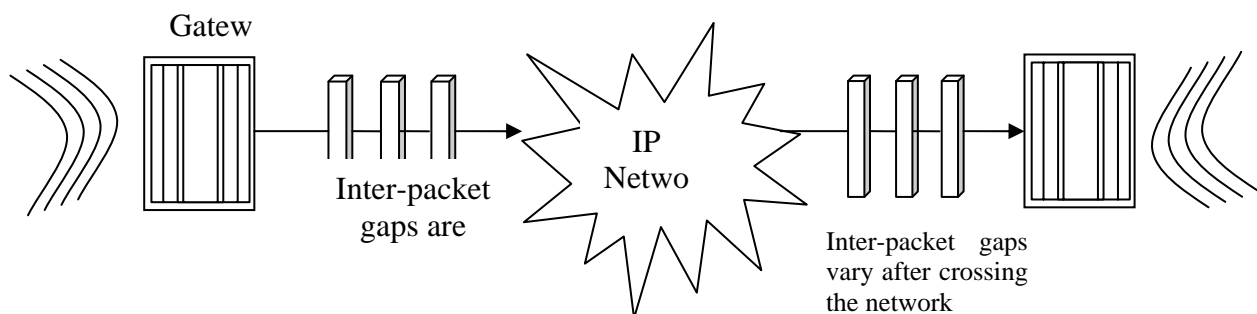

Figure 1: Demonstration of Jitter variation in VoIP

Unreasonable jitter makes speech cracked and difficult to understand. Jitter is calculated based on the inter-arrival time of successive packets. For high-quality voice, the average transmitter and the standard deviation should be low. Jitter is one issue that exists only in packet-based networks. While in a packet voice environment that sender is expected to reliably transmit voice packets at a regular interval. These voice packets can be delayed throughout the packet network and not arrive at the same regular interval at the receiving station.

Store-and-forward and queuing congestion in switches and routers along the way can lead to further packet spacing unpredictability, and thus jitter. Generally speaking, the more hops a packet has to travel, the worse the jitter. For example, VoIP packets that are sent at 20 millisecond (ms) intervals may arrive at 20, 45, 10, 15, 25 ms intervals. Low levels of jitter are easily handled by the jitter buffer at the receiver's end. Unfortunately, a jitter buffer adds additional delay to voice

reaching the ear piece in case other packets need to catch up. The packet stream is simply delayed for say 40ms in order to release packets at a steadier pace compared to their arrival time.

2.3 PACKET LOSS

According to [1], VoIP is exceptionally intolerant of packet loss. Packet loss can result from excess latency, where a group of packets arrive late and must be discarded in favour of newer ones. It can also be the result of jitter, that is, when a packet arrives after its surrounding packets have been flushed from the buffer, making the received packet useless. VoIP-specific packet loss issues exist in addition to the packet loss issues already associated with data networks; these are the cases where a packet is not delivered at all. Compounding the packet loss problem is VoIP's reliance on real time protocol (RTP), which uses the unreliable UDP for transport, and thus does not guarantee packet delivery. However, the time constraints do not allow for a reliable protocol such as TCP to be used to deliver media. By the time a packet could be reported missing, retransmitted, and received, the time constraints for QoS would be well exceeded.

VoIP frames have to traverse an IP network, which is unreliable. Packets may be dropped as a result of network congestion or data corruption. Furthermore, for real-time traffic like voice, retransmission of lost packets at the transport layer is not practical because of the additional delays. Hence, voice terminals have to deal with missing voice samples also referred to as frame erasures. The effect of packet loss on voice quality depends on how the terminals handle frame erasures. [4]

The loss of a packet results in missing information when the audio signal is received. Depending on the number of packets lost, the sound quality at the receiving end may suffer. In the IP approach, packet loss forms an integral part of the concept routers and are obliged (with the random early detection algorithm) to destroy packets in order to avoid possible congestion. In [6], there are four main causes of packet loss:

- Exhausted lifetime (TTL = 0)
- Receiving end delay greater than the jitter buffer
- Destruction by a congested module
- Packet invalidity due to transmission faults

Cisco Systems developed many qualities of service (QoS) tools that enable administrators to classify and manage traffic through a data network. If a data network is well engineered, you can keep packet loss to a minimum. However, with dynamic emergence of new technologies, QoS requirements and delivery continue to call for serious attention.

2.4 PACKET LOSS CONCEALMENT

According to [4], Cisco systems' VoIP implementation enables the voice router to respond to periodic packet loss. If a voice packet is not received when expected (the expected time is variable), it is assumed to be lost and the last packet received is replayed, as shown in Figure 2. Because the packet lost is only 20ms of speech, the average listener does not notice the difference in voice quality.
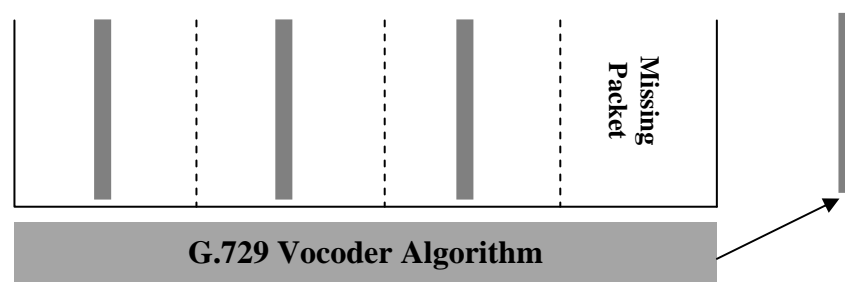

Figure 2: Packet Loss with G.729

Using Cisco's G.729 implementation for VoIP, assuming that each of the lines in the figure represents a packet, packet 1, 2, and 3 reach the destination, but packet 4 is lost somewhere in transmission. The receiving station waits for a period of time (per its jitter buffer) and then runs a concealment strategy.

This concealment strategy replays the last packet received (in this case, packet 3), so the listener does not hear gaps of silence. Because the lost speech is only 20ms, the listener most likely does not hear the difference. You can accomplish this concealment strategy only if one packet is lost. If multiple consecutive packets are lost, the concealment strategy is run only once until another packet is received. [4]

Because of the concealment strategy of G.729, as a rule of thumb G.729 is tolerant to about five percent packet loss average across an entire call. [4]

The proposed model has been designed to match the needs the QoS issues in VoIP, most especially, the packet loss. It attempts to address the issue of packet loss and also improve the rate of latency of packets while communication occurs.

3 DESIGN ISSUES & MODEL DISCUSSION
3.1 COMPRESSION AND DECOMPRESSION (CODEC)

VoIP converts the sound of your voice into packets of data, sends them across the network being used as the transport, and then reconstructs them into sound at the receiver's end of the network. The sound information, often called the payload, is put together with overhead information identifying where the packet should go to create the final packet transmitted over the network.

Compression breaks down into three parts:

- Frame delay: Unlike signal digitization, which is carried out in a continuous manner, compression relates to a specific data length. Waiting for that information can involve an appreciable processing time.
- Coding delay: This delay, which involves compression by synthesis based on prediction, is required by the encoder in order to know, while it is functioning, how the signal is evolving.
- Processing delay: Time taken by the algorithm to compress a frame. This depends on the processor and on the algorithm used.

Present within the scope of a VoIP terminal are audio codec and video codec.

Audio Codec: They enable the signal from the audio equipment for the transmission and decode the incoming audio code. Required functions include encoding and decoding of G.711 speech and transmitting and receiving a-law and μ-law formats. Optionally, G.722, G.723, G.728, and G.729 can be supported.

Video Codec: It is optional, but if provided, must be capable of encoding and decoding video signals according to H.261 Quarter Comment Intermediate Format (QCIF).

3.2 BUFFER

Several methods of buffer management schemes have been provided with each of them having their merits, demerits and range of applicability. These are presented in a taxonomical order. A detail discussion can be found in (ref). The design of these buffer schemes fall into either space priority or time priority.

Buffers are a major contribution to this research work and are aimed at showcasing the effectiveness of its combination with a choice scheduling algorithm to properly manage the rate of packet loss in VoIP network. The asynchronous nature of VoIP has come with the advantage of severally multiplexed traffics on the same route but with the great cost of delay and loss of packets, thereby necessitating the need for buffering. As earlier mentioned, traffic of this nature needs more than just a buffer, but an adaptive/dynamic buffer. The buffering technique employed is known as Complete Sharing with Gradual Release (CS/GR) a dynamic and adaptive buffering type.

3.3 TERMINALS

These are LAN client endpoints that provide real-time, two-way communication. H.323 terminals must have a system control unit, media transmission, audio codec, and packet-based network interface. Optional requirement include a video codec and user data applications.

The following functions and capabilities are within the scope of the terminal as shown in Figure 3.4

- System Control Unit – Provides H.225 and H.245 call control, capability exchange, messaging, and signaling of commands for proper operation of the terminal.
- Media Transmission – Formats the transmitted audio, video, data, control streams, and messages onto network interface. Media transmission also receives the audio, video, data, control streams, and messages from the network interface.
- Audio Codec – Encodes the signal from the audio equipment for transmission and decodes the incoming audio code. Required functions include encoding and decoding G.711 speech and transmitting and receiving a-law and μ-law formats. Optionally, G.722, G.723, G.728, G.729 encoding and decoding can be supported.
- Network Interface – A packet-based interface capable of end-to-end Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) unicast and multicast services.
- Video Codec – optional, but if provided, must be capable of encoding and decoding video according to H.261 Quarter Comment Intermediate Format (QCIF).
- Data Channel – Supports applications such as database access, file transfer, and audiographics conferencing (the capability to modify a common image over multiple users' computers simultaneously), as specified in Recommendation T.120

3.4 NETWORK

A telecommunication network may be described as the set of infrastructure which enables the transfer of information from one point of connection to the network to another. The notion of network arises from the need to share infrastructure in order to optimize their costs, it being possible for a single transmission line to be utilized by different users for different periods. To these internal lines we must add distribution or access networks in order to reach all potential users of the network. Telephony networks have in the course of time undergone major evolutionary changes, driven essentially by technological progress in various fields (switching, transmission, access and maintenance). The latest of such change is digitization of their transport technology, which has had a considerable integrating influence.

The technology used for the purpose of transporting voice within telephony networks is known as "circuit switching". It is based on the principle that a resource (circuit) must be reserved for a call from the time of its setting-up to its conclusion. The size of this resource – expressed as a bit rate since the digitization of telephone networks – is 64 Kbits/s. The limit was chosen at the time because it made for the efficient digitization of human voices samples with the spectrum lying between 300 and 3400 Hz6. More recent voice coding techniques allow for a considerable reduction in the 64 Kbits/s defined for a circuit; however, since this is the rate that is used in most of the active and transmission components throughout the global telephony network, it would be difficult to change it without incurring excessive cost and without jeopardizing one of the main qualities of that network: the universal service. As a case in point, one can mention that although modem wireless GSM network use an encoding that consumes as low as 8 Kbits/s on the radio part, this is transformed to a 64 Kbits/s encoding when voice reaches the mobile switches.

The network to be simulated consists of one or more parameters. All parameters may be specified by user at the time of component creation and may be modified. For the simulation of the new model in this research work, the following parameters to be used are discussed below.

- Exhaustive Lifetime of packets in buffer
- Buffer Content Destruction Rate.

- Scalability of buffer.

The crux of this research is to determine the importance of buffering and its placement. Consequently, the buffering technique adopted is termed complete sharing with gradual release (CS/GR), a detail discussion can be found in Aderounmu, et al., (2004). The placement of our buffer is as shown in figure 4 as opposed to the exiting model shown in figure 3.
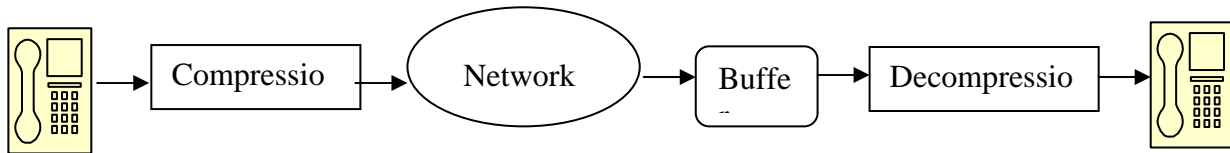
Transmitting Terminal
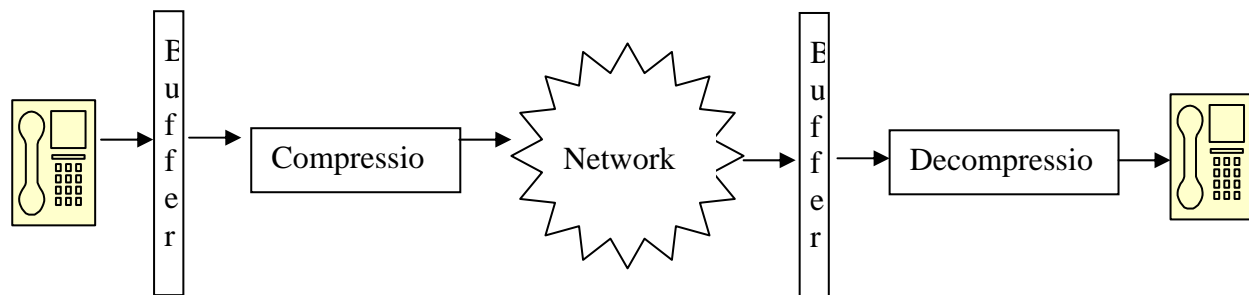
Figure 3: Existing VoIP buffering Model

Figure 4: The Proposed VoIP Buffering Model

In figure 4, we presented our proposed model in which dual buffer was employed rather than the single type. In the first stage before compression module of voice packets, an adaptive/dynamic buffer is placed immediately after the packets is being transmitted, while in the second phase, when the compressed packets have been sent over the channel, and before the decompression of the packets have been sent over the channel, the received packets must pass through another buffer similar to existing design at the receiving end to facilitate synchronization between both end. To address the issue of packet loss, the following were considered:

- Type of queuing technique.
- Terminals being used.
- Compression and Decompression.

4 PERFORMANCE ANALYSIS CRITERIA

To perform evaluation of the research work a platform is provided for pointing out the inherent problems in the existing single buffer model which implements the static buffer scheme with the manner by which the new model intends to remedy the "injustice" in the method of placing the buffer in order to guarantee the negotiated QoS parameters as it affects the quality of transmission expected from the severally sent voice packets without absolute favoritism towards a particular class of services.

The objective of this research work is to support a set of QoS parameters and classes for VoIP services and minimize network and end system complexity, packet loss and delay while

guaranteeing maximum network utilization. To this end, performance analysis criteria will be determined based on the support for the following:

- Exhausted lifetime: evaluations include the time it takes for the voice packet moving in the network to be processed at the nodes (from the transmitting terminal to the receiving end).
- Destruction by a congested module: evaluates the rate at which voice packets are destroyed when it encounters an overcrowded buffer.
- Scalability: Since it is expected that delay associated with traffic busyness be adequately controlled and packet loss minimized if not completely stopped, the following variables were considered
  - o Volume of traffic.
  - o Number of channels.
  - o Number of buffers.

4.1 RESULT ANALYSIS

Simulation results aim to quantify the previous enhancement and enable comparisons between the use of a single buffer as before in VoIP communication and using two buffers as in the proposed scheme. In the implementation, voice packets are sent over a single exhaustive lifetime of voice packets in buffer, the rate of destruction of buffer content in a congested module, and the scalability of the buffer.

4.1.1 PERFORMANCE EVALUATION

The exhaustive lifetime of voice packets in buffer, rate of destruction of buffer content in a congested module, and scalability of the buffer is computed and varied for both the single buffering scheme and the dual buffering scheme and are tabulated. Change in the volume of traffic is static, (i.e. the difference between the first and the next one is the same), for this experiment, the difference between the first and the next one is 10,000 and the maximum volume of traffic used in this experiment is hundred. The figures below shows the result for this experiment and it shows exhaustive lifetime of voice packets in buffer, rate of destruction of buffer content in a congested module, and scalability of the buffer for the different number of channels.

Figure 5 shows the volume of traffic taken in thousand to exhaust the lifetime of packets for various number of channel using both the single buffering model and the dual buffering model.
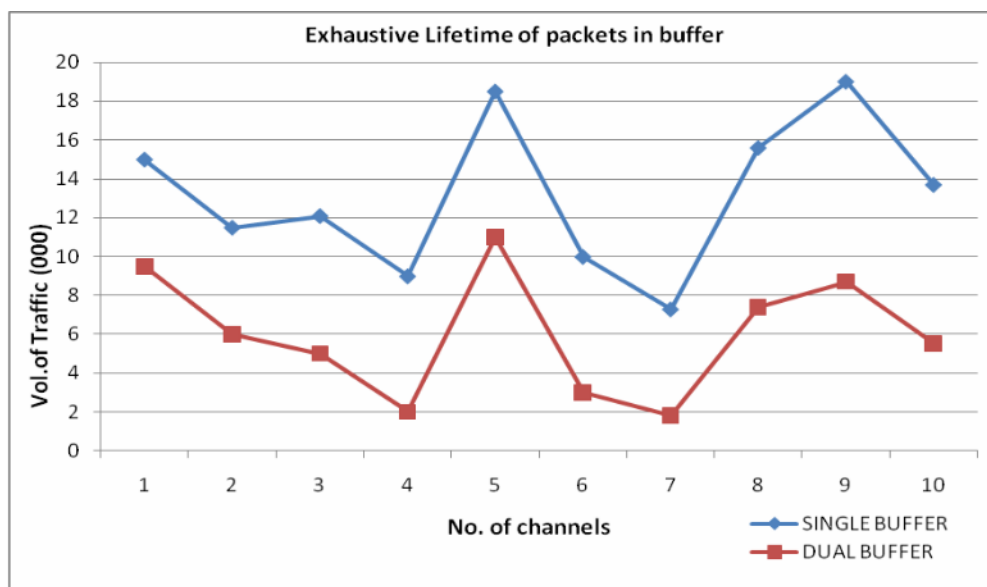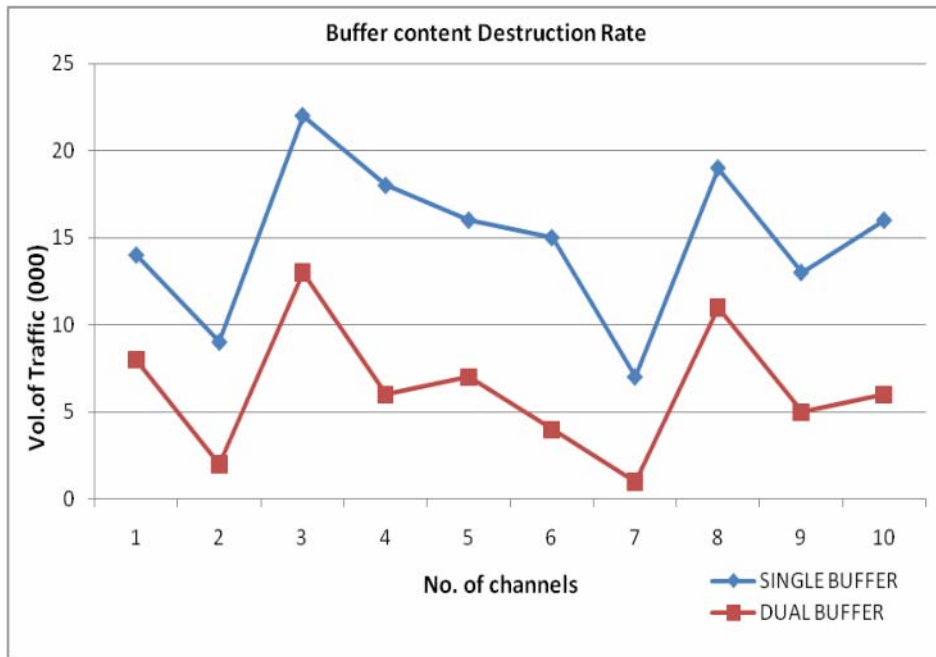

Figure 5: Exhaustive lifetime of packets in buffer

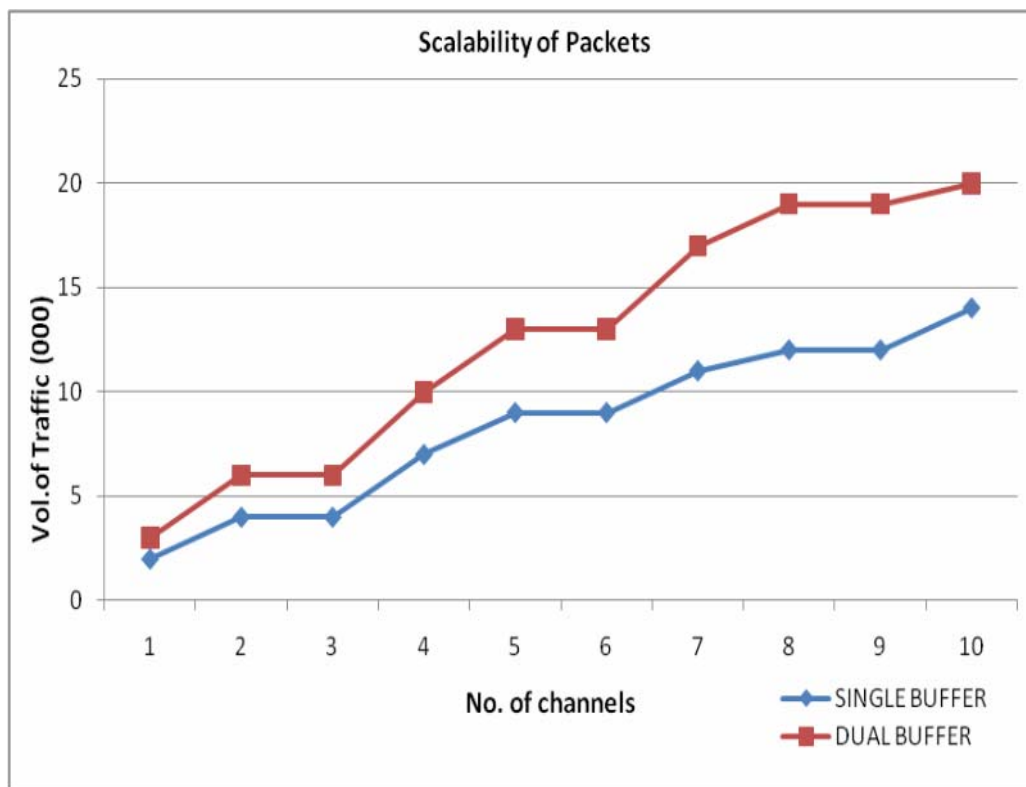Figure 6: Performance evaluation of buffer content Destruction Rate

Figure 7: Performance evaluation of Scalability of Packets

4.1.2 RESULT DISCUSSION

This research work has presented the combination of two buffers as the important factor on the design of this new model for managing delay and packet loss. According to figure 4, the total time spent in a model with a single buffer is spent in a model is greater than the total time spent in

the dual buffer model. The large time spent by the voice packets in this single buffered system causes delay which if it's too much makes communications unattractive to users and if the delay varies inconsistently, packets get lost. With the dual buffer model, if the same voice packet is sent, its lifetime is much more reduced and so delay is minimized and the packet loss is also reduced.

According to figure 5, we can see that the rate at which voice packets in a single buffer are destroyed in a congested module is higher than that of the dual buffer model. This is caused by the fact that when a voice packet arrives at the module and the buffer is overcrowded; the voice packet gets destroyed while waiting, this destruction has led to packet loss. With the dual buffer model, when the same packets are transferred, the destruction rate is reduced because not in all cases will the two buffers be overcrowded so hardly will a packet get destroyed while waiting.

According to figure 6, a dual model is more scalable than a single buffered model. A model with implementing a two buffered model is able to adapt to changes easier than a single buffered implementation.

REFERENCES

[1] Debashish Mitra (2001), Network Convergence and Voice over IP, Technology review#2001-2.

[2] Vickers B. J., Kim J. B, Suda T. & Hong D. P. (1993): "Congestion Control and Resource Management in diverse ATM Environments", Trans. Inst. Electronics, Communication engineers of Japan (IECE J). Journal, J76-B-I (11), pp 759 – 774.

[3] Aderounmu G. A., Oyeniyi S. J., Adagunodo E. R. & Akinde A. D. (2000): "A New Buffer Management Scheme for Multi-QoS Traffic Over ATM Switching Systems". International Journal of Information and Computing Science. Vol. 3, No. 2, pp 57

[4] Angus Ma (2001): Voice over IP (VoIP). pp 3 – 4

[5] Wildpackets, Inc (2006): VoIP: Changing the dynamics of Network Analysis, March 2006. White Paper, pp 5 – 7

[6] IP Telephony / ITU-D (2003): "The Essential Report of IP Telephony." ITU E-Strategies Unit. pp 91.

[7] Aderounmu G. A., Ogwu F. J. & Onifade O. F. (2004): "A Dynamic Traffic Shaping Technique for a Scalable QoS in ATM Networks." International Conference on Computing, Communications and Control Technologies. August 14-17 – Austin, Texas, USA, pp 332 – 337.

[8] Marko Leppänen, Voice over IP, Helsinki University of Technology, Department of Computer Science.

[9] Onifade O.F.W & Ojesanmi O.A. (2005): "Agent-Based Stock Monitoring and Security Systems (ASTOM)", Journal of Research in Physical Science, Volume 1, Number 1, 2005. pp 56 – 60

[10] Francis Ogwu, Mohammad Talib, Ganiyu Aderounmu, & Onifade, O.F.W., (2008):"A Dynamic Traffic Shaping Technique for a Scalable QoS in ATM Networks". The International Arab Journal of Information Technology, Vol. 5, No. 4, pp 93 – 100, October 2008