

Investigating The Delivery of High-Quality Videoconferencing over Low-Bandwidth Networks using Multicasting and Quality of Service (QoS)

Ugochukwu Ogbonna

Abstract—Videoconferencing is a technology which enables organizations overcome challenges involving cost, logistics and safety of participants when organizing meetings and conferences. However, a major problem faced with this method is the availability of bandwidth, which is a concern for most developing countries, as video data consume huge bandwidth. The paper seeks to investigate the realization of a cost-effective network prototype capable of delivering high-quality videoconferencing, which can be implemented on low and medium-level networks and links. Preliminary research on multicast techniques and architectures and QoS strategies were undertaken in order to achieve this objective. Quantitative and qualitative tests were carried out to determine if there were any improvements in video quality before and after deployment of QoS. Analysis of results showed a significant improvement in the quality of video feeds after application of selected QoS design.

Index Terms— Multicasting, Protocol Independent Management, Quality of Service, Telepresence, Videoconferencing

1 INTRODUCTION

ORGANIZATIONS and corporate bodies in the world are expressing increasing worries over travel safety and security in attending meetings. In addition to these worries are concerns of travel costs and other expenditure incurred in organizing these meetings. More recently however, corporations are beginning to adopt the use of videoconferencing as a solution to this problem. Its ease of affordability and facilitating meetings and teamwork interaction has led organizations to utilize as a solution and so far, it has proved valuable in simplifying task execution with fewer resources.

Then again, most available videoconference solutions are quite capital intensive and are targeted at 'big' companies. Telepresence, a variation of videoconference capable of delivering HD quality videoconferencing, deployed by businesses lately, is quite expensive (Bielski, (2008)) entailing massive bandwidth and costly hardware and design requirements (Lazar, (2007)). Generally, telepresence kit marketed by vendors require certain high-grade network devices and bandwidth requirement, as well as specific visual and acoustic prerequisite designs which are somewhat financially demanding, especially to small and medium scale enterprises (SMEs). This makes videoconferencing quite difficult, if not impossible for SMEs to adopt in their infrastructure.

The aim of this paper is to investigate the possibility of realizing high-quality videoconferencing for SMEs. This is based on the

assumption that SMEs have average or basic network design and devices and running on a basic WAN connection (in this case, a 2Mbps WAN connectivity is used). This research will investigate the possibility of delivering high quality videoconferencing over 2Mbps WAN connection across 3 sites.

This paper is structured in the following ways: section 2 reviews the background technology considered in this research and the experiments undertaken in realizing and testing the suggested design. Section 3 discusses the approaches carried out in realising the system design. Section 4 examines the results of the tests, discussing the outcomes. Section 5 then concludes the paper, recommending areas for further research.

2 MATERIALS AND METHODS

As stated previously, the objective of this paper is targeted at researching high-quality videoconferencing for SMEs. As part of our definition of SMEs, we assume that these organizations have "medium-scale", readily available network equipment and link accessibility. In other words, we make the following assumptions for our design:

Availability of typical Cisco integrated services routers (e.g. 28xx series) and switches (e.g. 29xx series) for designing and managing the core layer of small business networks.

1. The existence of a 2Mb/s WAN link for interconnection of main/regional sites. This assumption was made based on the average deliverable internet speeds from top 10 ISPs in Nigeria.

2. This research is designed across 3 sites (represented as Lagos, Abuja and Enugu sites) to represent the presence of this small business across the 3 main geopolitical zones of the country.

In the subsequent sections, the two main important technologies leveraged on in realizing the intended design (Multicasting and Quality of Service) are examined.

2.1 Multicasting

There are 3 defined means of data transmission – unicasting, broadcasting and multicasting. Unicasting defines a one – to – one communication whereas broadcasting involves one – to – all communication. However, multicasting involves data transmission from one to a selected group of nodes. Multicasting offers management of bandwidth and available networks resources as well as support for seamless, multipoint distribution of multimedia data. This attributes make multicasting a favourable technique for videoconferencing.

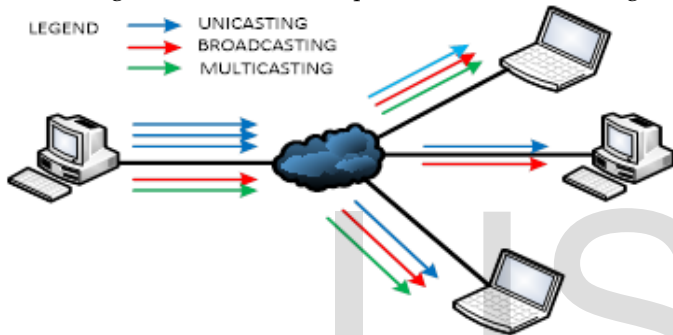


Fig. 1. Illustration of unicast, broadcast and multicast.

In deploying multicast, 3 architectures are involved – IP Multicast (IPM), Overlay Multicast (OM) and Application Layer Multicasting (ALM). IPM implements multicasting through simultaneous delivery of duplicate packets sent from the source host, to destination nodes who would have indicated their desire to receive these packets by joining a specific multicast group tied to transmission of the said packets. However, OM is poorly suited over low links, causing quality-related issues. ALM involves additional control overhead added to compatibility issue with quality of service. IPM however, presents best efficiency for delivering multimedia data while conserving bandwidth and processing overhead. It is for this reason that it is selected for this design.

2.1.1 IP Multicast Protocols

According to Cisco, there are 4 main multicast protocols supported on Cisco devices;

- Distance Vector Multicast Routing Protocol (DVMRP)
- Cisco Group Management Protocol (CGMP)
- Protocol Independent Management (PIM), and
- Internet Group Management Protocol (IGMP)

DVMRP is a distance-vector based algorithm protocol designed for computing routes through a network and is usually deployed on the internet backbone. CGMP is a Cisco proprietary protocol deployed on their catalyst switches. For the purpose of the work, only the last 2 of the above (PIM and

IGMP) list will be considered as they are industry-standard, independent multicast protocols.

2.1.1.1 PIM

For IPM operations, 4 delivery modes exist:

- PIM Dense Mode (PIM-DM),
- PIM Sparse Mode (PIM-SM),
- PIM Sparse Dense Mode (PIM-SDM), and
- Bidirectional PIM (Bidr-PIM),

PIM-DM is prone to congestion as traffic flooding is involved and is not recommended for IP multicast design [12]. PIM-SM and Bidr-PIM pose optimal path issues as the metric from receiver to RP might not be better than from receiver to source. PIM-SDM, which combines DM and SM functionality was selected for this research because it ensures scalability and efficient network resource management by ensuring optimal path selection in data delivery. Although it uses DM initially for RP location, it immediately switches to SM upon RP discovery thereby conserving network resources.

2.1.1.2 IGMP

IGMP is a standard protocol for managing group membership for multicast delivery in networks and restraining traffic. It information about multicast data in networks are provided to routers, controlling multicast traffic flow across the network through multicast queries. It employs a query-reply mechanism, regularly updating group membership and specifies how hosts register with routers in order to receive requested traffic.

2.2 Quality of Service (QoS)

In the previous section, multicasting was identified as a technique for ensuring efficient multimedia traffic delivery. However, data traversing a network are by default, treated equally. In a unified network, this can prove disastrous as multimedia data can consume most of the bandwidth. A strategy needs to be defined for managing continuous UDP (multimedia) and TCP traffic (ordinary data).

QoS is a method that ensures reliable and consistent traffic delivery with least possible delay while measuring service availability and transmission quality across converged networks allowing various traffic types contend unevenly for available network resources. This section examines the available QoS models and mechanisms utilized in realizing our intended videoconference design.

2.2.1 QoS Architectures

IETF has standardized 2 QoS frameworks for designing QoS strategies – Integrated Services (IntServ) and Differentiated Services (DiffServ) models. IntServ QoS provides end-to-end guaranteed QoS through resource and pre-request policy admission control. However, in the bid to sustain this guarantee, IntServ nodes need to maintain policing, scheduling, queuing and flow-state updates. Reference [4] indicates that this makes IntServ non-scalable, leading to increase in overhead with increased flow as each flow requires continuous signalling. Hence, its application across the Internet, where real-time applications are unable to work well due to variable queuing delays.

DiffServ provides QoS by assigning service levels to packets on a hop-by-hop basis. Different packets are arranged in classes using Differentiated Service Code Point (DSCP) – a 6-bit IP-header marking – for defining priority levels for each packet

type. Although DiffServ cannot deliver end-to-end QoS guarantee on its own, using it with queuing techniques, can guarantee end-to-end QoS. Moreover, it is scalable and does not require additional overhead like IntServ, thereby achieving optimized time and resource management. Moreover, DSCP allows traffic to be classified and marked appropriately, allowing routers to assign priority levels to traversing packets based on the DSCP mark given to it. Hence, its selection for this research.

2.2.2 Traffic Classification and Marking

As stated in the previous section, DSCP values will be used to classify and mark packets. Classification/Marking allows packets to be grouped into same behavioural aggregates (BAs) and then treated with the desired service priority level. According to [11], classification and marking can be done either in layer-2 (Data-link Layer) or layer-3 (Network Layer).

Layer-2 classification/marking involves analysing the 802.1Q field of the Ethernet frame. Within the frame is a 2-byte tag control identifier (TCI) field. This field contains a priority field made up of 3-bits defined by IEEE802.1p for defining classes of service (CoS) used for tagging frames as they traverse the network. However, CoS tags usually don't cross routers and hence, are not maintained end-to-end making network administrators adopt a different tagging method altogether.

Layer-3 tagging involves the use of 8-bit type of service (ToS) field in the IP header. The first 6-bits define DSCP while the last 2-bits are reserved for flow control. The values of the DSCP bits define the per-hop-behavioural (PHB) treatment given to it. Figure 2 shows the classes of PHB defined by the IETF used in traffic classification:

Default - describes setting the first 3-bits of the DS field to 0 (000) resulting in conventional packet delivery service using best effort (BE).

Assured Forwarding (AF) - this class has 12 subdivisions and is realized by setting the first 3 DS-field bits to 001, 010, 011 and 100 (AF1, AF2, AF3 and AF4 respectively). Each class is then subdivided into 3 subclasses defining drop precedence (DP). Classes of higher category and DP receive higher priority.

Expedited Forwarding (EF) - involves setting the DSCP field to 101110 (decimal 46), and is the standardized service model adopted by the IETF for delivering assured bandwidth with minimal jitter, packet loss and delay.

for maximum link usage efficiency

2.2.3 Congestion Management

Congestion usually arises as a result of data speed mismatch, and aggregation at confluence nodes or points. This is mostly caused by packets arriving at a node faster than they leave that same node. Congestion can result in jitters and delays in networks if not properly managed.

References [5] and [8] identified the main strategies for managing congestion as;

- First-In-First-Out (FIFO)
- Priority Queuing (PQ)
- Weighted Fair Queuing (WFQ)
- Class-Based Weighted Fair Queuing (CBWFQ), and
- Low Latency Queuing (LQ)

Table 1: Various subdivisions of the AF class

Drop Precedence (DP)	Class #1	Class #2	Class #3	Class #4
Low DP	AF11 (001010)	AF21 (010010)	AF31 (011010)	AF41 (100010)
Medium DP	AF12 (001100)	AF22 (010100)	AF32 (011100)	AF42 (100100)
High DP	AF13 (001110)	AF23 (010110)	AF33 (011110)	AF43 (100110)

CBWFQ offers more bandwidth management, as well as traffic transport and queue management prioritization, compared to its prior alternatives. Where excess bandwidth is available, it allocates the bandwidth to traffic based on predefined criteria of classifying traffic priority. However, the absence of queuing methods makes this technique unfit for real-time traffic. LLQ, on the other hand, combines high priority queuing with CBWFQ, guaranteeing bandwidth to high priority (HP) traffic while allowing residual bandwidth to be shared across other traffic types - the idea being to guarantee delivery HP traffic type with insignificant delay. However, for the purposes of this work, CBWFQ will be combined with LLQ. This is to ensure that high priority queuing is present for real time traffic, while CBWFQ guarantees the appropriate bandwidth needed for transport of other mission-critical data.

3 SYSTEM NETWORK DESIGN AND TESTING

As previously stated in the introduction, this study is based on a 3-site design. In order to simulate this arrangement, 3 routers and switches were used in the design to represent the 3 sites, with an additional switch used to represent the WAN connecting the 3 sites. Suitable IP addresses were then assigned and unicast communication between the devices was realized using Cisco's enhanced interior gateway routing protocol (EIGRP).

3.1 Multicast Design

Arranging a multicast strategy for the intended network requires selection of multicast addresses for communication of multicast data. The Internet Assigned Numbers Authority (IANA) has specified reserved globally scoped address range of 224.0.1.0-238.255.255.255 for multicast data transmission between organizations and across the Internet, from which the

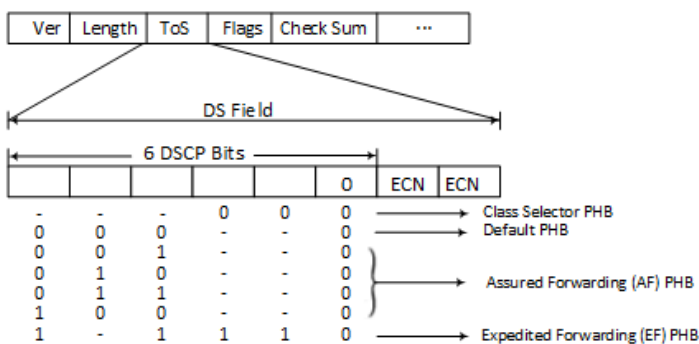


Fig. 2. IP Header showing DS field and DSCP bit values

According to [3] and [11], EF specifications are industrially recommended for voice/VOIP delivery. Video streaming and conferencing is usually recommended to be designed with AF41

following addresses were selected for this design;

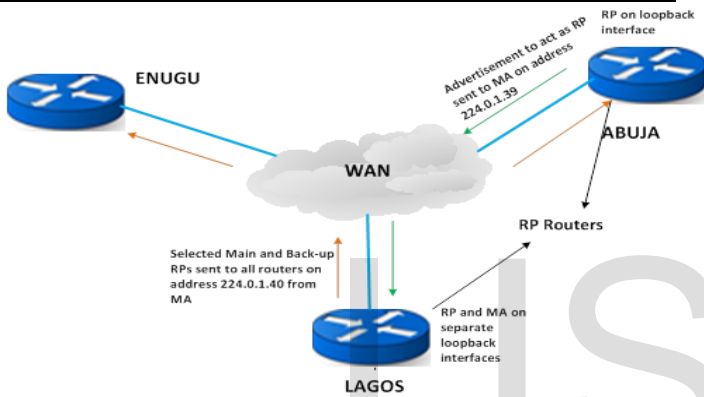
Table 2: Selected Multicast Addresses

SITE	MULTICAST ADDRESS
LAGOS	225.1.1.1
ABUJA	225.2.2.2
ENUGU	225.3.3.3

The rendezvous points (RP) and mapping agents (MA) necessary for the operation of PIM-SDM were distributed across the each router's loopback addresses as follows:

Table 3: RP and MA selected IP addresses

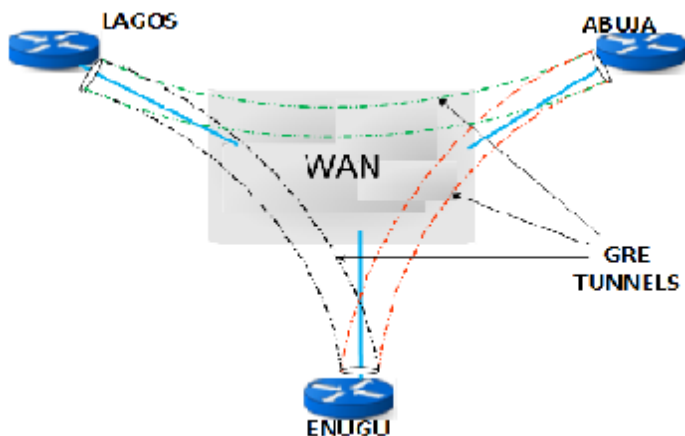
Type (Location)	Interface	IP Address
RP ₁ (Lagos)	Loopback Interface 0 (lo0)	10.1.2.0/32
RP ₂ (Abuja)	Loopback Interface 0 (lo0)	10.2.2.0/32



MA (Lagos)	Loopback Interface 1 (lo1)	10.1.3.0/32
------------	----------------------------	-------------

Fig. 3. RP Design Illustration

In order to facilitate the operation of the multicast traffic, Generic Routing Protocol (GRE) tunnels will be configured and used for transmitting multicast traffic so as to ensure that they are kept separate from other traffic. GRE tunnels create logical interfaces and encapsulate traffic whilst simulating separate point-to-point connections for the encapsulated traffic within the existing network. In this project, the tunnels are configured such that they originate and terminate at loopback interfaces configured on the routers. This is to ensure all multicast traffic



are encapsulated as

Fig. 4. Illustration of the GRE tunnel arrangement

GRE traffic before reaching the physical interfaces. All real time voice and video traffic are passed into the logical tunnel so that they appear as GRE traffic when they reach the physical interface. The resultant GRE traffic and other forms of traffic are then sent down the link. This enables us treat both sets of traffic

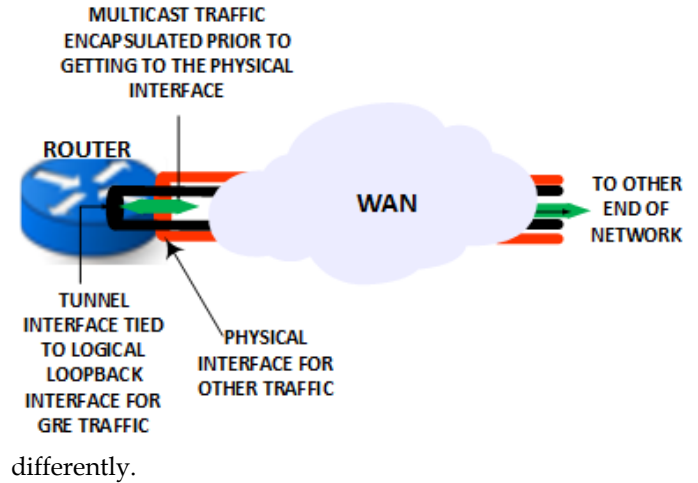


Fig.5. GRE encapsulation of multicast traffic within the router

3.2 System Design

In order to implement the selected techniques for this research, a unicast network has to be deployed first. The unicast network built to realise the needed 3-site connectivity is shown in figure 6. Here, we have chosen a router and a switch to represent a site, with the switch labelled WAN representing the ISP. The bandwidth connectivity between each site and the WAN were limited to 2Mb/s, which is the approximate round-off of a typical T1 line. 3 laptops with VLC software were used at each 'site'. On each laptop, 2 different VLC screens representing the video feed from the other 2 sites, were opened. A simple illustration for this design is shown in figure 3. As such, each site will receive two multicast video feeds while transmitting one.

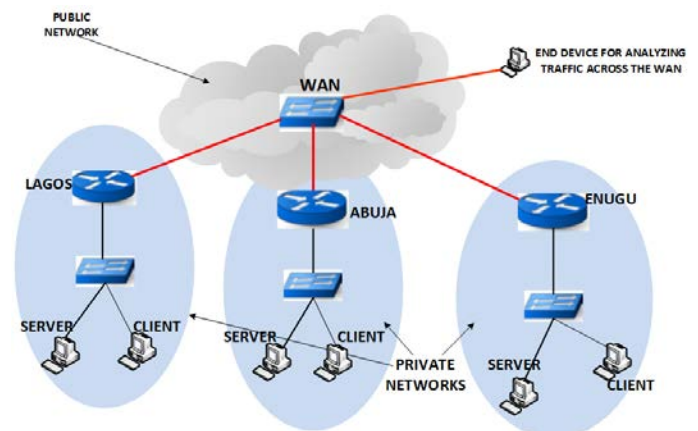


Fig. 6. Device arrangement for simulating the network model

3.3 QoS Design

In making QoS considerations, the recommended guidelines for video delivery were implemented. In [11], the authors stated that the recommended guidelines for video/voice delivery is as tabulated in table 4.

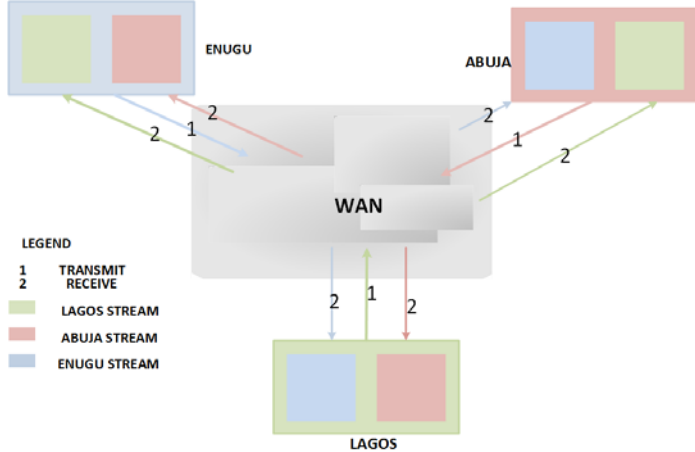


Fig. 7. Multicast Traffic Transmission/Reception Design

Table 4: Recommended Video Delivery Guidelines

QoS Element	Interactive Video	Streaming Video
Jitter	No more than 30ms	No specific requirements
Latency	No more than 150ms (one way)	4-5 seconds
Packet Loss	1%	5%
DSCP Value	AF41	CS4

As stated earlier, LLC and CBWFQ will be combined to achieve traffic classification and marking. Although the 11-class model which is recommended by Cisco is normally used on most networks, a 4-class model is used in this work. This model is modified from the suggested 11-class model.

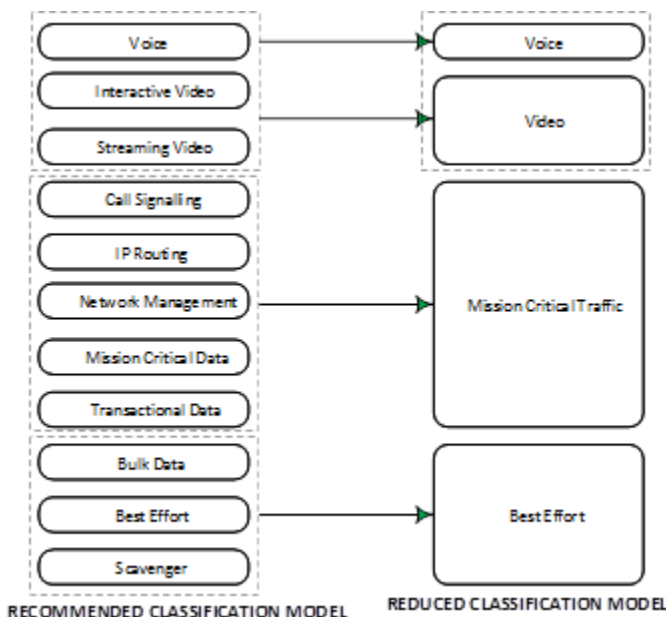


Fig. 8. Classification Model based on recommended Guidelines

In the selected model, interactive and streaming video are grouped together since video delivery is being realised through multicast video streaming. Other necessary network protocols and data necessary to operate are grouped as Mission Critical traffic. For the purposes of this research, TCP traffic will be generated and used to represent Mission Critical Traffic. All other traffic not belonging to any of the above mentioned groups are classed as best effort. In applying the recommended QoS baseline, the following DSCP values have been chosen for each traffic class for the reduced model.

Table 5: Selected QoS Classification/Marking values

Traffic Classes	DSCP Value	IPP Value
Video	AF41/Decimal (34)/Binary (100010)	Decimal (4)/ Binary (100)
Mission Critical	AF31/Decimal (26)/Binary (011010)	Decimal (3)/ Binary (010)
Best Effort	0/Decimal (0)/ Binary (000000)	Decimal (0)/ Binary (000)

In order to calculate bandwidth allocation for each class, the percentage allocation for voice, interactive and streaming video will be grouped together as real-time traffic since voice calls are not considered here. Best effort, bulk and scavenger allocation will be grouped together as Best Effort traffic, while the remaining recommended traffic will be classified as mission critical traffic. This is illustrated in the following calculations.

$$\text{Real Time} = 33\% (\text{real-time traffic}) + \text{approx. } 12\% (\text{video portion of critical traffic}) = 45\%$$

$$\text{Mission Critical} = 37\% - 12\% + 5\% = 30\%$$

The remaining bandwidth will be left for best effort traffic. With available WAN link bandwidth of 2Mbps, the required guaranteed bandwidth will be;

$$BW_G = 45\% \text{ of } 2Mbps = 0.45 \times 2Mbps \approx 900kbps \text{ approx. downwards}$$

To obtain the required compression rate for each link with 3 video streams we refer to the formula:

$$BW_G = 900kbps = 3\{CB + (20\% \times CB)\}; CB = 250kbps$$

Hence, each video stream will be compressed at not more than 250kbps using H.264 codec to cater for the overhead in order to stay within the reserved percentage. The cumulative size for each video stream plus overhead will require:

$$(250 + 0.2 \times 250)kbps = 300kbps \text{ bandwidth}$$

Using the same method above, we deduce the bandwidth requirement for Mission-Critical traffic to be:

$$BW_{MC} = 30\% \text{ of } 2Mbps = 0.3 \times 2Mbps \approx 600kbps$$

Best effort traffic will be configured as First-In-First-Out allowing it to use whatever bandwidth is left. The percentage selection were based on Cisco's recommended bandwidth percentage allocation.

3.4 Testing

A combination of various software were used for quantitative and qualitative testing so as to ensure accurate and unbiased results. Faultline and Colasoft Capsa 7 Enterprise softwares were used to conduct quantitative tests while ITU-T's recommended 5-point Mean Opinion Score (MOS) was used for qualitative testing. Five participants were selected to assess the video feed quality for the qualitative test. The average MOS scores gotten before and after QoS deployment were used to substantiate the quantitative results.

Table 6: ITU-T recommended 5-point MOS opinion scale

Opinion	Score
Excellent	5
Very Good	4
Good	3
Poor	2
Very Poor	1

JPerf software was used to generate excessive traffic so as to cause network congestion. Causing network congestion ensures that the QoS design adopted in our project kicks in. The cumulative traffic across each site was captured using Colasoft.

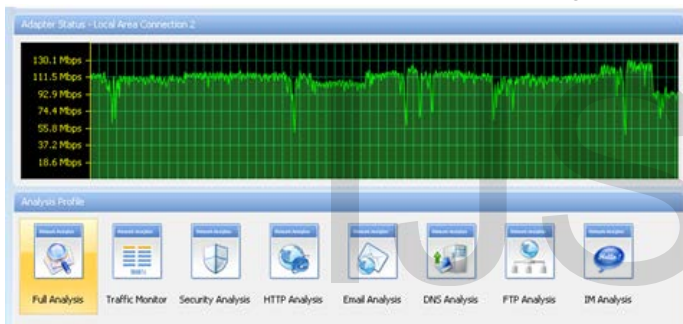


Fig. 9. Colasoft extract with traffic across site capped at 111.5Mbps

The testing phase was divided into 2 namely:

Quantitative and qualitative test before QoS implementation – Here the network is congested with TCP traffic generated using JPerf. With the QoS policies deactivated on each site's router, the video quality at each site is examined using Faultline (for quantitative) and human perception (qualitative).

Quantitative and qualitative test after QoS implementation – Here, the network is congested with TCP traffic, with the QoS policies activated on each site's router. The video quality at each site is also examined using Faultline and human perception.

4 RESULTS AND DISCUSSIONS

4.1 Before QoS Implementation

4.1.1 Results

The system was initiated without QoS policies in place and video feeds were requested from the various sites. The output from Faultline was captured as indicated in figure 10. The MOS results from the selected participants are given in the table below;

Table 6: MOS scores from participants before QoS deployment

Participant	P1	P2	P3	P4	P5
Rating	3	3	5	3	4

Fig. 10. Faultline extract showing Overview of video statistics

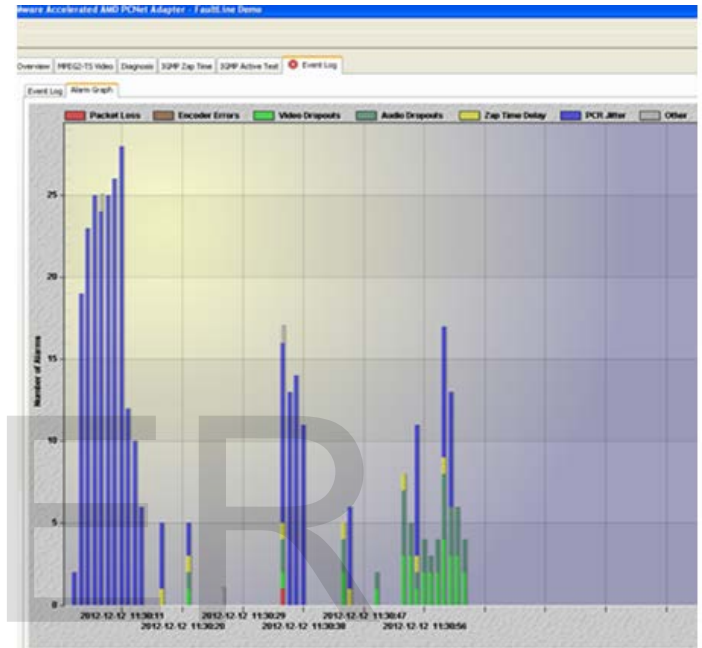


Fig. 11. Faultline extract showing graphical analysis of video statistics

4.1.2 Discussion of Results

As seen from Fig.10, the network registered video and audio dropouts during transmission. This is confirmed in the graphical output in Fig.11, which further shows series of jitters and delays resulting from packet losses. However, though the quantitative analysis seems to show reduced video quality, a look at the average MOS result of 3.6 seems to indicate "averagely good" video feed for the participants.

4.2 After QoS Implementation

4.2.1 Results

The system was again initiated, however this time with QoS policies in place and video feeds were requested from the various sites. The output from Faultline was captured as indicated in figures 12 and 13. The qualitative scores from the participants after QoS application are given in the table below:

Table 7: MOS scores after deployment of QoS policies

Participant	P 1	P 2	P 3	P 4	P 5
Rating	3	3	5	5	4

Fig. 12. Faultline extract showing overview of video statistics with QoS policies deployed

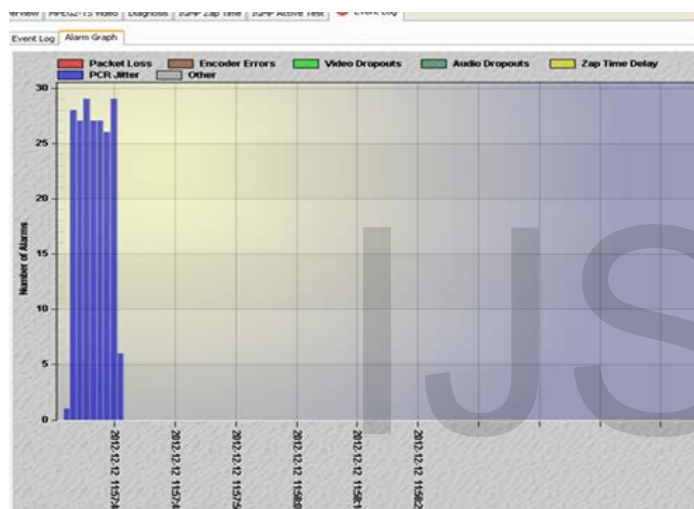


established a strategy for combining multicasting and QoS to realize a promising videoconferencing solution for SMEs. However, this research has not considered the implications and interoperability of other areas such as security and IPv6 interoperability-areas which will require further additional research.

REFERENCES

- [1] L. Bielski, "Telepresence: costly but very cool," *ABA Banking Journal* vol.100, pp. 50-52, April 2008.
- [2] Network Protocols Configuration Guide Part 1, Cisco Systems Inc., San Jose, CA, 2007.
- [3] Enterprise QoS Solution Reference Network Design Guide, Cisco Systems Inc., San Jose, CA, 2008.
- [4] B. E. Fgee, J. D. Kennedy, W. J. Philips, W. Robertson and S. Sivakumar, "Comparison of QoS performance between IPv6 QoS management model and IntServ and DiffServ QoS models," in *Proc. 3rd Annual CNSR*, 2005, pp. 287-292.
- [5] M. J. Fischer, D. M. Masi and J. F. Shortle, "Approximating Low Latency Queueing Buffer Latency," in *AICT 2008*, pp. 188-194.
- [6] P. Flavius and P. Ferdi, "Internet Service Delivery Models: Evolution and Current Issues," in *CyberC 2011*, pp. 146-153.
- [7] *IP Multicast Explained*, Metaswitch Networks, London, UK, 2009.
- [8] M. Z. Islam, M. S. Islam, A. K. Haque and M. Ahmed, "A comparative analysis of different real time applications over various queuing techniques," in *Proc. ICIEV*, 2012, pp. 1118-1123.
- [9] E. B. Kelly, "Quality of Service in Internet Protocol (IP) Networks." Presented at International Communications Industries Association, for Infocomm. Las Vegas, USA, June 12-14, 2002.
- [10] I. Lazar, "Telepresence: Looking beyond the Hype," *Business Comm. Review*, vol.37 pp. 14-15.
- [11] M. S. Lipu, M. N. Haque, T. F. Karim, M. L. Rahman and M. H. Morshed, "Quality of service model selection criteria and traffic controlling algorithm of a native IP Network for telecommunication services," in *ICEIT 2010*, vol. 3, pp. V3-144 – V3-148.
- [12] U. Shaukat, A. Qayyum and S. Hussain, "Performance evaluation of multicast protocols for multimedia traffic," in *ICET 2009*, pp. 383-388.
- [13] Techloy. (2012, March) Top 10 ISPs in Nigeria: Which Do you Use? [Online] Available: <http://techloy.com/2012/03/12/top-10-isps-nigeria/>.
- [14] L. Xiaohua and C. Kai, "The Research and Application of IP Multicast in Enterprise Network," in *ICICIS 2011*, pp. 191-194.

Fig. 13. Faultline graphical extract showing video statistics with QoS policies active



4.2.2 Discussion of Results

As observed from the new MOS results, there was a significant improvement in the average video feed quality as perceived by the participants (denoted by the average score of 4.0). However, the quantitative analysis results shows a considerable improvement after QoS policy application. This is illustrated in Fig. 12. Fig.13 shows the absence of video and audio dropouts. It can also be seen from Fig.11b that the delays and losses been removed. Also, we notice that jitters, though registered initially by Faultline, cleared off completely after a while. Generally, there has been a significant improvement in the quality of the video feeds on the links after the activation of the designed QoS.

5 CONCLUSION AND RECOMMENDATION

This research was undertaken with the intent of investigating the techniques and strategies which can be used to implement quality videoconferencing over relatively low links. With novel videoconferencing technologies like telepresence more or less out of reach for SMEs, this research was aimed at devising an alternative but effective means of achieving satisfactory videoconferencing with manageable cost. This research has