Application Level Performance of Multimedia Services

Parag Pruthi, Dragos Ilie and Adrian Popescu

Department of Telecommunications and Signal Processing
University of Karlskrona/Ronneby
371 79 Karlskrona, Sweden

ABSTRACT

Quality of Service (QoS) is a difficult term to define for multimedia applications. The main reason is that both audio and video quality are subjective and difficult to quantify. Much work has been done in the past to map the subjective quality of video and audio into measurable quantities. Unfortunately, when it comes to IP environments, not much experience and mathematical work exists that can be used to define robust metrics for measurement of QoS.

In this paper, we report on measurements of multimedia QoS and try to map subjective criteria to discrete measurables in terms of packet loss rates, packet delays, and other quantities. We report the results of measurements done at the application level and show how network characteristics affect the perceived quality of multimedia applications. In particular, we analyze the application traffic generated by MBone clients in a distributed network education scenario.

In order to measure the traffic, we have implemented software on a non-intrusive probe developed by NIKSUN INC.* that can accurately monitor all traffic from a variety of networks. We have developed MBone aware software modules which can not only play back the recorded streams but also provide the essential statistics in real-time. We report in detail the results of our study of a particular end-to-end MBone session.

Keywords: Quality of Service, MBone, stream identification, data extraction, statistical analysis, non-intrusive monitoring, application monitoring

1. INTRODUCTION

Multimedia distributed systems are required to generate, process, store, retrieve, communicate and present different data objects comprised of mixed data types, which may include video, images, audio, voice and data. For most actual and future multimedia applications, data sets are created and fetched from many and dispersed sources, and transmitted to different users at interactive terminals. The integration process requires assembling the data sets based on both spatial and temporal constraints. Furthermore, the picture is even more complicated because of the fact that, besides an efficient integration of isochronous and non-isochronous services, temporal integration of isochronous communications with different temporal granularities (periodicities) must be also supported. Unique sets of requirements are imposed on the communication component of a multimedia system, due to the size and characteristics of different multimedia objects. Also, extra requirements in terms of coexistence, integration and interaction of different media by mutual synchronization heavily complicates the issue of multimedia communication.

Early trials on delivering high-quality media over "long-haul" or wide area network (WAN) links have shown that delivering QoS for multimedia applications on IP networks is not just a matter of interconnecting all the "high-tech" pieces. In fact, each of the pieces must be individually optimized, not individually as isolated components, but as an integral part of the whole. Furthermore, since multimedia application sources (e.g., audio, video) are of type "streaming", they put additional requirements in terms of rather hard inter as well as intra stream synchronization in order to be adequately rendered at the user front ends. The situation gets finally even more complicated because of many specific problems related to the specific application, e.g., in the case of distanced education there

Further author information: (Send correspondence to A.P.)
P.P.: E-mail: parag@niksun.com
D.I.: E-mail: dil@itm.hk-r.se; Telephone: +46 455 385661; Fax: +46 455 385667
A.P.: E-mail: adrian@itm.hk-r.se; Telephone: +46 455 385659; Fax: +46 455 385667

*http://www.niksun.com

are problems related to "live" classrooms, large virtual class sizes, need for the ability of "continuous" feedback between the lecturer and the virtual class rooms, and need for off-line access.

The focus of this paper is on real measurements from a network being used for distance networked education. In particular, large sets of MBone multimedia traffic generated by the actual network distributed education system between the USA and Sweden are analyzed. The need for real measurements at the application level measurements is motivated. The intricate relationship that exists among application level parameters, network characteristics and performance of multimedia applications is demonstrated.

With increases in network bandwidth and the push for QoS, MBone multimedia traffic is a good representation of the kind of network traffic expected in the very near future on long-haul links. A rather broad view today seems to be that increasing capacity alone will solve the needs of future applications like multimedia, but the equation is not so simple. Most Internet service providers today are striving to optimize their network as they find that revenues generated from the sale of bandwidth alone are not sufficient for their long-term survival. Also, todays networks are not properly engineered to handle large quantities of multimedia traffic while providing good QoS. Network upgrades and engineering are necessary if value added services like multimedia will form a viable revenue generating service. Therefore, the proper engineer of both applications as well as network resources takes on a critical role in cost-effective delivery of these services in the future. In a joint project between the University of Karlskrona/Ronneby (Sweden) and NIKSUN INC. (USA), we are investigating the effective dimensioning of different parts of a networked education system to deliver high-quality real-time and interactive multimedia content.

In this paper, the measurement infrastructure and the results of measurement studies of a set of actual multimedia conferences are reported. The capability for link to application level measurements is described and the relationships among application level QoS, application parameters, packet delays and losses are presented. Quality of Service (QoS) itself is a term difficult to define for certain types of traffic such as video over IP. There are many factors that may affect the QoS and it is often the agreggated effect of these factors that has a noticeable impact on the quality the user experiences. Since video and audio are subjectively better analyzed, we used test subjects to subjectively mark those intervals during which the multimedia audio and video appeared to be unsatisfactory to the test subject. Using the subjective evaluation as a basis for multimedia quality, we developed some measurable metrics that are good indicators of the subjective loss in quality. These metrics are described and their practical utility is presented.

The paper is organized as follows. The first part (section 2) is devoted to stream identification, playback and data extraction in MBone sessions. The second part (section 3) is focused on QoS measurements and description of metrics for estimating QoS in MBone-based multimedia services. In particular, various results are reported based on our methodology to discover MBone sessions, playback MBone sessions, extract specific data statistics, perform real-time statistical analysis and QoS metric identification and evaluation. Finally, results are presented for end-to-end performance (inter-arrival delay and packet loss) for characteristic audio and video MBone sessions.

2. STREAM IDENTIFICATION AND DATA EXTRACTION

MBone is an acronym for IP Multicast Backbone with the purpose to define common means for the distribution of multimedia using the protocol IP Multicast. MBone has been allocated the IP multicast addresses in the range 224.2.xxx.xxx.[1]

There can be many MBone sessions simultaneously active around the world. Details about these sessions are multicasted among MBone servers, along with information about which server to connect to in order to receive the media streams. Further, information about each session can be obtained by means of software similar to the so-called Session Directory (SDR).[2] SDR lists all sessions announced by a specific server together with the time and date when they are active. In addition, the SDR Graphical User Interface (GUI) allows the user to easily join or leave a session.

MBone acts only as a transportation entity for a given data set. What data is transported and how the data is interpreted is not MBone's concern. The software running on top of MBone has the responsibility to fill enough information inside the session announcement in such a manner that a receiving part can identify the data streams. Based on this identification, SDR can automatically launch specific applications that interpret the data streams and present the information to the user.

2.1. MBone Protocols

There are two common ways to announce MBone sessions. One is the Session Announcement Protocol (SAP)[1,3] and the other is the Session Initiation Protocol (SIP).[4] The main difference between them is that while SAP lists all the announced sessions to everybody interested, SIP sends this information to a specific set of users only.

Each session announcement contains relevant information, such as session name, session owner, time and date when the session is active, and most important where to connect to in order to receive the media streams. All this information is encoded using the Session Description Protocol (SDP).[5]

The media streams use the RTP/RTCP protocol to deliver their content to the receiver.[6] The Real-time Transport Protocol (RTP) stream is the actual multimedia carrier, and the RTP Control Protocol (RTCP) stream carries receiver reports to the sender. The receiver reports contain information about the quality of reception.

2.2. Measurement Infrastructure

The NIKSUN NetVCR^{TM} monitoring and analysis system has been used for traffic measurements. NetVCR[7] is a non-intrusive system for network monitoring and analysis able to collect and storage data from link layer up to application layer. The system can be connected to all the interfaces in a specific local/wide area network (LAN/WAN) and collect all data traffic in that specific LAN or on the probed WAN interfaces. The observed data traffic can be recorded for later reference on a storage device such as hard disk or tape. Individual packets are automatically classified into link to application layer sets and automatically analyzed (e.g., Frame Relay, Ethernet, ATM, IP, TCP, UDP, WWW, MBone). NetVCR's application programming interface (API) allows retrieving of stored data traffic based on some criteria (e.g., time interval, packet type, IP address, port number, etc).

We used the NIKSUN NetVCR system to collect data from a T1 WAN link connecting Stanford University, in San Francisco, USA with the Royal Institute of Technology (KTH), in Stockholm, Sweden. This dedicated network education link, known as the "SSVL" (Stockholm Silicon Valley Link) [†], carries network education traffic between Stanford University and Swedish Universities. Many multimedia sessions were monitored on this link using the NIKSUN NetVCR system and the results from one particular session are reported in this paper.

2.3. Session Decoding

NetVCR was extended by us so that it is capable of analyzing MBone sessions.[8] Using the API, an extension to NetVCR has been built that collects information about all MBone sessions and the belonging data streams during a specific time interval. This software (named *sdrdump*) works similarly to SDR and has a functionality that is somehow similar to that of the well-known *tcpdump*.

Using *sdrdump*, we were able to extract all the information necessary to connect to a live session. In this case, the sessions were actually past and had been recorded by NetVCR into its internal store. What we still required was the capability to play back (with a high degree of accuracy) any session previously recorded by NetVCR. Thus, a special module for NetVCR called *play_mbone* was developed. This software is able to accurately play back (i.e., with a very high degree of inter-packet accuracy) a stored session to a client. When NetVCR records data traffic, it adds a header to every stored packet. The header contains, among others, a timestamp for the time when the packet was recorded. Using these timestamps, one can accurately reconstruct the same inter-arrival delays at playback as experienced when the packet was recorded. Thus the client does not require any special software other than one of the standard conferencing tools such as Video Conferenceing Tool (VIC) and Robust Audio Tool (RAT). The interface for the client to have the playback capability is through a standard web page that makes the selection simple for an end user.

3. TRAFFIC MEASUREMENTS

The network traffic data used for measurements is a video and audio session containing a forty minutes long seminar. This is a typical session for network distributed education. The session was recorded on a T1 link (data rate of 1.536 Mbps). The NetVCR system was connected to one end of the link, close to the MBone source.

QoS measurements performed at the application layer with the MBone playback system are reported. Our goal is to accurately present the actual QoS experienced by the end user.

[†] http://ssvl.stanford.edu/

The MBone extension to NetVCR additionally processes multimedia streams to generate statistics that are important inputs for analyzing the measurable quality of multimedia sessions. For each session, information about packet arrival times, packet sizes, packet loss, inter-arrival delay, delay jitter, etc., are reported on multi-timescales. That is, NetVCR can report statistics that range from packet-by-packet statistics to aggregate statistics over user specified intervals (e.g., per milli-second, second, minute, half-hour, etc).

Multimedia packets are both loss and delay sensitive. While video might still be able to present a frame with acceptable quality if no vital packets have been lost, audio will most likely break up or generate disturbing "pops" at the receiver if packets are lost. This is also true for delays except that even video may rapidly lose image quality. This is a result of the so-called *playback point* effect.[9] Multimedia applications are designed to adapt to certain variations in network conditions; this adaptation is generally carried out by designing of an appropriate dimensioned packet buffer. As an example, consider a packet carrying 100ms of speech. In a real-time session scenario, if the packet is delayed by 300ms then the application will most likely discard that packet and play the ones situated in time closest to the present. To avoid these situations the application keeps a buffer where it stores a certain amount of data before starting to play. This way, even if packets are delayed in the network, there may still be packets in the buffer that can be played. The buffer size, counted in units of time, is called the *playback point*.

An important question is related to the value of the playback point, which should be based on a trade-off between memory size and backup considerations (according to the QoS experienced). Delays in excess of the playback point would most likely result in degraded quality as the video compression algorithm can not cope with this situation. This algorithm optimizes the network traffic by sending only traffic that corresponds to the change between previous and current frame. It cannot do much about delays and the decoder will most likely wait passively for the next packet carrying frame information and keep the display in the same state until it gets it. This may result in jerky movements.

For a network designer wishing to setup a network to ensure that all constraints are satisfied for a good end-to-end QoS of multimedia traffic, this can be a nightmare. Due to the bursty nature of multimedia traffic, the network resources can be best utilized by using the silence periods between bursts to send data. Unfortunately, the bursty behavior of multimedia traffic makes it very hard to actually discover these silence periods. Actually, there are traffic models based on fractal mathematics that better characterize multimedia traffic flows than the conventional methods but there still remains a lot of work to be done before a workable solution can be obtained.

The focus of our experiments is on understanding the relationship between packet loss, delays and other measurables on the subjective quality of the multimedia session. Another issue is to find out the importance of the bandwidth usage and if there is any correlation between any of these factors and the perceived performance. To do this a small scale QoS test was conducted. One person listened to a MBone session of 40 minutes length and marked the time intervals when the sound quality was bad. The test was conducted first using 60 seconds time intervals and then stepped down to 10 seconds intervals. In spite of the fact that the test is subject to the person ranking the audio quality, it can still be used as a crude QoS measurement at the application level before better QoS metrics are found.

3.1. Bandwidth Usage

For the 40 minute seminar that has been analyzed in detail and presented in this paper, the audio stream was encoded with PCM μ-law encoding. Figure 1 shows the graph of the bandwidth used by the audio stream. The audio encoding should generate Constant Bit Rate (CBR) at 64 Kbps and one can notice a variation around 8000 bytes/sec. The graph shows that the CBR traffic was affected to a large degree by the effect of other traffic, transiting through routers, and application scheduling. The specific effects of each are not studied in this paper.

Figure 2 shows the bandwidth used by the video stream that is encoded using the H.261 standard.[10] Note the highly bursty behavior as illustrated by the high peaks and low averages. Such a behavior is typical of Variable Bit Rate (VBR) encoded streams. For this stream, the minimum bandwidth utilization was approximately 10 KBytes/sec (approx. 5% link utilization) with a maximum usage of about 140 KBytes/sec (approx. 68% utilization). The average utilization over the entire interval was 54 KBytes/sec corresponding to about 28% link utilization. These figures indicate a high peak to mean ratio of approximately 2.6.

This example illustrates the difficulty in dimensioning a link for VBR encoded multimedia traffic. If the link is dimensioned at even 2 times the average required, the resulting peaks will be clipped. The clipping of the peaks

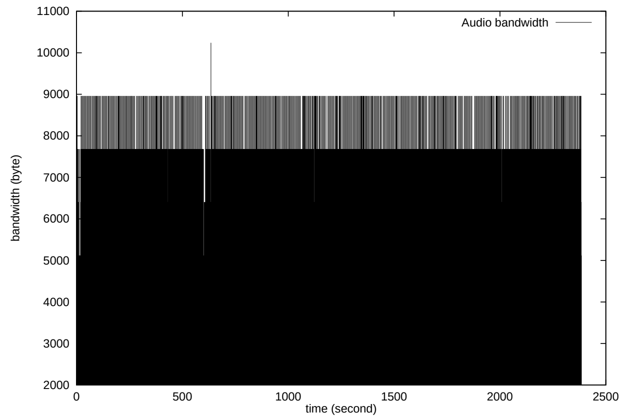

Figure 1. The bandwidth used by a 64 Kbps μ-law encoded PCM audio stream. Note that the CBR stream suffered variations around the target rate of 8 KBytes/sec due to effects of transiting through routers and application scheduling.

can be quite damaging to video quality as most of the information is contained in these burst of traffic (generally scene changes or fast motion in the video). Statistical multiplexing may be one solution, but the high variability suggests that effective multiplexing gains may be only obtained when a large number of streams are multiplexed (thus reducing the peak to mean ratios), with the consequence that very high bandwidths will be required.

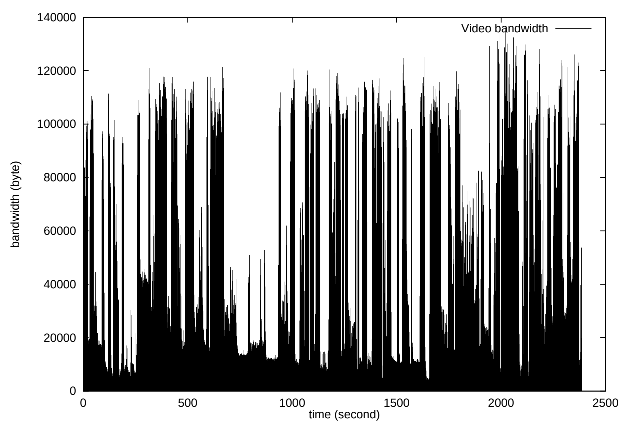

Figure 2. Bandwidth utilization of the VBR encoded video stream. Note the high variability of VBR traffic making it difficult to utilize the link effectively without large increases in bandwidth to effectively obtain high quality and statistical multiplexing gains.

3.2. Long-Range Dependence Properties

One of the most important parameters in the case of bursty traffic is related to the degree of Long-Range Dependence (LRD) existent in the data traffic, i.e., if the traffic is Short-Range Dependent (SRD) or LRD.[11] A first experiment was done in this case for the video stream by zooming into the peak at seconds 610-670 and looking for signs of self-similarity in the video stream (Figure 3). It is noticed however that Figure 3 is not very similar to Figure 2, at least not visually. It is observed that the bandwidth plot in Figure 3 lacks bursts of the type seen in Figure 2 with the consequence that the degree of self-similarity is low.

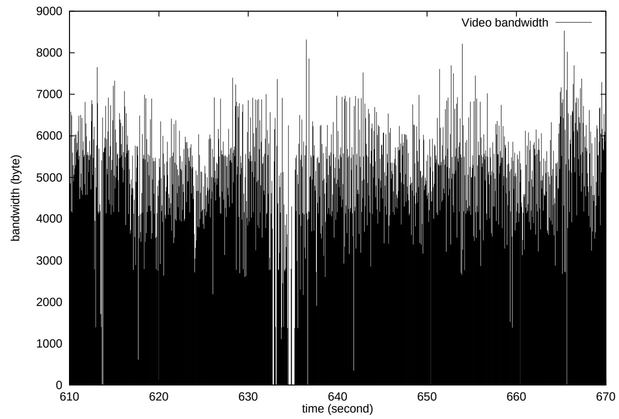

Figure 3. Zoomed video bandwidth

It is clear however that there is a high degree of temporal correlation in the video stream, as it was shown earlier by Willinger et. al[12] as well as that VBR video exhibits long-range dependence.

Further, a quantile-quantile plot was used to test whether the packet size has a Gaussian distribution or not (Figure 4). The horizontal axis shows the quantiles of the standard normal (Gaussian) distribution. Quantiles of the video packet size are plotted on the vertical axis. The S-shaped curve proves that the data has a non-Gaussian distribution. Also, the central limit theorem does not hold for the data set in question, mainly because the number of samples is too low ($\approx 127,000$).[8] Thus, techniques based upon Robinsons method, such as the Whittle estimator can not be used to measure the degree of self-similarity.

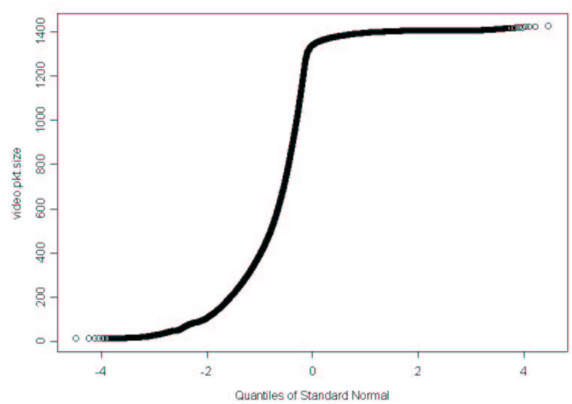

Figure 4. Quantile-quantile plot for the video stream. For a distribution matching the standard normal distribution the plot should show an (almost) straight line crossing the figure from the lower left corner to the upper right corner.

The degree of long-range dependence was estimated by using a variance-time plot. This method is not as accurate as the Whittle estimator but, on the other hand, it does not make any assumptions on the data distribution. Figure 5 shows the variance-time plot for the video packet size. The long line crossing the plot from the top left corner down to the right bottom corner marks the short dependency domain where the Hurst parameter is equal to 0.5. The slope estimate of the tail for the variance data determines the Hurst parameter to approximately 0.69. This is an indication that the video stream exhibits long-range dependency, but not self-similarity.

The almost flat slope at small lags in Figure 5 shows the presence of strong short range correlation in the VBR stream. This characteristic is typical of VBR sequences that have been studied by many other authors.

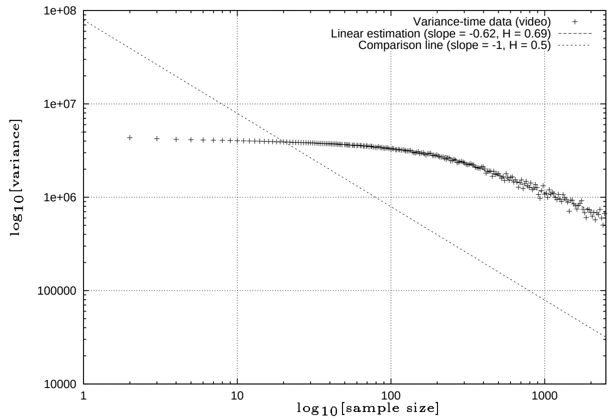

Figure 5. Variance-time plot for the video stream (H = 0.69) and for a SRD process (H = 0.5).

4. QOS METRICS FOR MULTIMEDIA COMMUNICATION

Due to the variable bandwidth requirements of various multimedia coding schemes and the statistical sharing of resources along a path, a set of parameters characterizing the requirements for a multimedia session must be negotiated at connection initiation on an end-to-end basis. The negotiated parameters must appropriately map the subjective QoS requirements with actual measurements from the network. Some examples of important measurements are essential bandwidth, bounds on (end-to-end) delay, delay jitter, packet loss rate, etc. The relationship between these measurements and the subjective quality is therefore of utmost importance in dimensioning resources to meet the required QoS.

The most important parameters that affect the QoS of a multimedia session are delay and loss. We therefore report the results of our experiments that attempt to relate subjective multimedia quality to inter-arrival delay and packet loss for a 40 minute MBone audio and video session.

4.1. Inter-Arrival Delay

The inter-arrival delay is defined as the delay between two consecutive packets arriving at one specific point in the network. This delay is a combination of two parts, namely a delay inserted by the source and another delay component that is due to the network. For instance, large inter-arrival delays can be accounted for by lost packets, network congestion, or silence suppression mechanisms at the sender.

The MBone session reported here was recorded close to the transmitter (the lecturer in the distributed network education environment). During the analysis phase of the session, we artificially divided the time interval for the audio session into a number of non-overlapping segments of equal length, i.e., ten seconds and one minute. As the audio session was played back, test subjects listening to the audio marked the time intervals that they considered were subjectively bad.[8] For each audio segment, the average inter-arrival delay was plotted on a graph and the users subjective criteria (good or bad quality) was superimposed on the same graph.

Fig. 6 shows the average inter-arrival delays for the audio stream, with one minute time average. Fig. 7 zooms inside the region containing minutes 9-12 (seconds 480-720) of the audio stream where the delays are averaged over 10 seconds. Each sample in the zoomed graph covers the time segment starting ten seconds back in time and ending where the sample is. That is the sample at second 600 covers the time segment 591–600.

The average audio inter-arrival delay is a little over 160 ms. The maximum inter-arrival delay is found at minute 11, where the plot above shows a delay larger than 162 ms. The second plot shows three spikes over 162 ms with

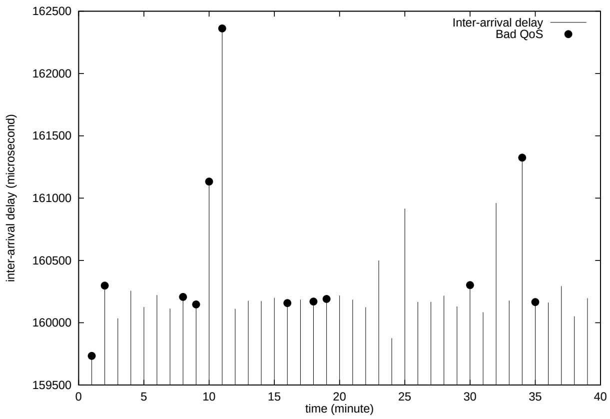

Figure 6. Average inter-arrival delay for audio packets with time average one minute

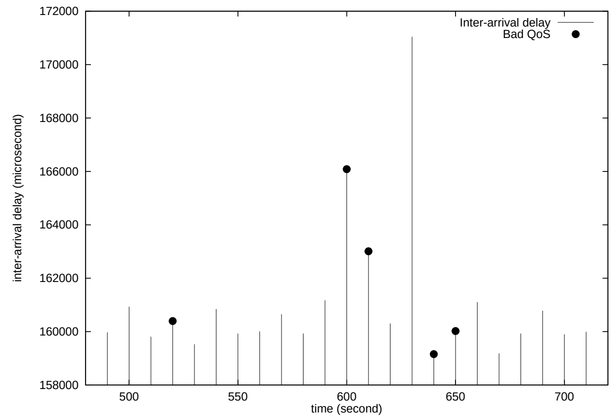

Figure 7. Average inter-arrival delay for audio packets with time average ten seconds, zoomed region minutes 9-12

one of them going as far as 170 ms. Note that the audio stream was coded as CBR and hence the variability in the inter-arrival delays are within a relatively narrow range.

A first investigation indicates that only 5 samples out of 24 are marked as bad in the zoomed region. For instance, the plot reveals that a bad segment of 10 seconds, the one situated at 520 seconds on the horizontal axis, accounted for a whole bad minute, minute 9 on Fig. 6. The same thing happens at minute 10 (second 541-600), which is all good with the exception of the last sample. Minute 11 on the other hand (seconds 601-660) is "Bad QoS" on both plots. Furthermore, a fractal like behavior can be observed by looking at the both time scales.

Overall it appears that the average delay is not a good indicator of the subjective QoS. For CBR audio the inter-arrival delays should be constant when no silence suppression mechanisms are present. End-to-end delays of 200 ms have been shown to be commercially acceptable and the goal is to limit these delays to the range 100-200 ms.[13] These measurements show that the delays experienced by the session are in range, the problem however is that these are delays at the originating point only. At far away destination, where many hops are involved, the inter-arrival delay and delays will generally increase due to diverse reasons.[14]

Fig. 8 shows the maximum inter-arrival delays for every minute of the audio stream. Fig. 9 zooms inside the region containing minutes 9-12 of the audio stream, and using 10 seconds long time segments. It is observed that the largest spike, almost 800ms, is at minute 10, and this effectively results in the user marking this segment with "Bad QoS".

Such long delays are either the result of silence suppression mechanisms or lost packets. If silence suppression mechanisms were working during minute 10 it is unlikely that the listener would have noticed any audio quality degradation. The consequence therefore is that this large delay is accounted for by lost packets. It also appears that the maximum inter-arrival delay has a better relation with the subjective QoS.

Similar experiments and analysis are done with the video stream. Fig. 10 shows the video sequence with averages over one minute and the Fig. 11 zooms in over minutes 9-12 at a finer resolution (10ms). It is important to mention that the marks for "Bad QoS" still belong to the audio stream. The reason is that it is much harder to judge the QoS for the video part of the session. Some viewers would say blurry frames mean bad video quality while others may consider the frame rate more important. The goal is to see if there is any connection between bad quality for the audio stream and the behavior of the video stream.

It is observed the specific bursty behavior of the VBR source, with inter-arrival delays highly variable. The medium delay is around 22 ms, almost eight times smaller than the medium delay for audio. The zoomed sequence has averages going over 30 ms.

Larger inter-arrival delays of the video stream means that the video stream requires less bandwidth, and hence positively effects the audio QoS. During the periods of high video activity (e.g., scene change, large motion), the video stream takes up more bandwidth and hence adversely effects the audio stream causing a high jitter between

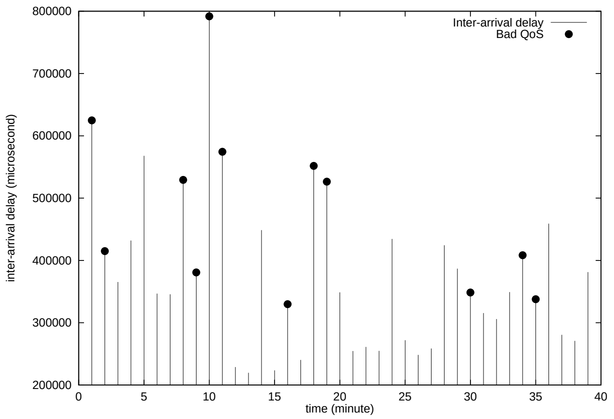

Figure 8. Maximum inter-arrival delay for audio packets with time interval one minute

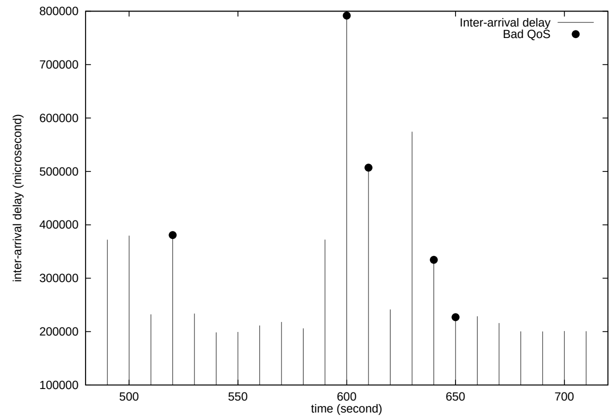

Figure 9. Maximum inter-arrival delay for audio packets with time interval ten seconds, zoomed region minutes 9-12

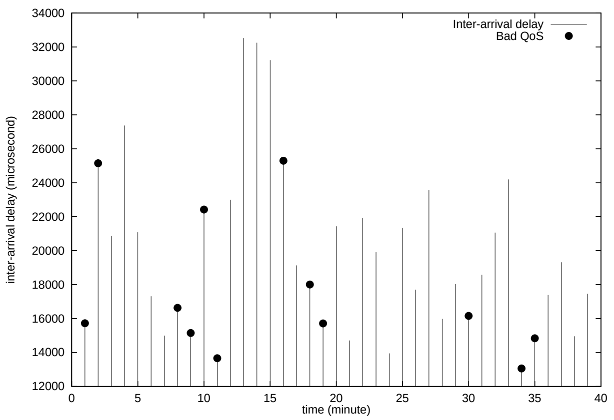

Figure 10. Average inter-arrival delay for video packets with time average one minute

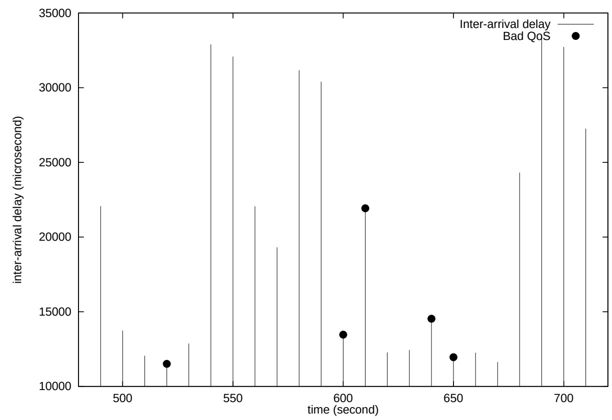

Figure 11. Average inter-arrival delay for video packets with time interval ten milliseconds, zoomed region minutes 9-12

the audio packets (and may also result in some lost audio packets). We are going to issue a follow up paper where we statistically analyze these observed (inverse) correlations.

4.2. Packet Loss

The impact of packet losses on the audio QoS of the MBone session is investigated. Similar type of QoS assessment is used, i.e., same "Bad QoS" marks are superimposed on the plots in Fig. 12 and Fig. 13.

An interesting observation is that all minutes marked "Bad QoS" show a packet loss greater than 0.5%. The audio codec sends on average 375 packets per minute. A true statement is that, *for this specific MBone session* any audio packet loss greater than 0.5% per minute (i.e., more than 2 packets per minute) will result in bad QoS. Furthermore, this also confirms the above-mentioned observation that the packet loss is the main reason behind excessive inter-arrival delays. Second 541-600 (minute 10) shows a packet loss of over 45% concentrated on a ten seconds segment (i.e. second 591-600).

There is also a "Bad QoS" mark at the second 650 where there is no packet loss. This mark could be explained by the "Bad QoS" mark in the previous time segment. On average, 63 audio packets are sent during 10 seconds. It is possible that the 4% loss rate at the second 640 includes some packet sent from the end of the time segment. The effect of the lost packet was then observed in the next time segment.

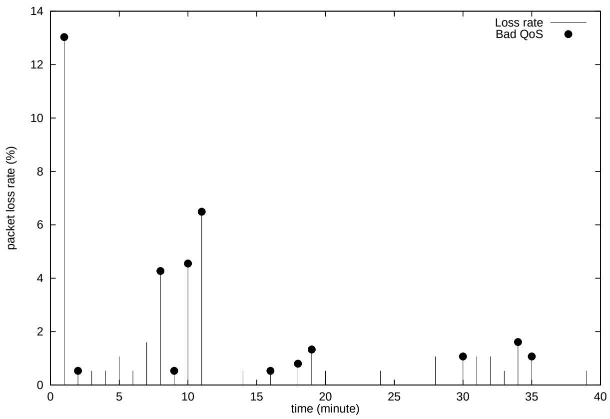

Figure 12. Average loss rate for audio packets with time average one minute

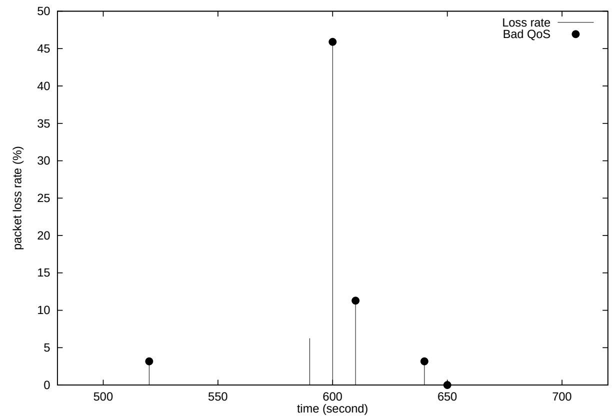

Figure 13. Average loss rate for audio packets, zoomed region minutes 9-12

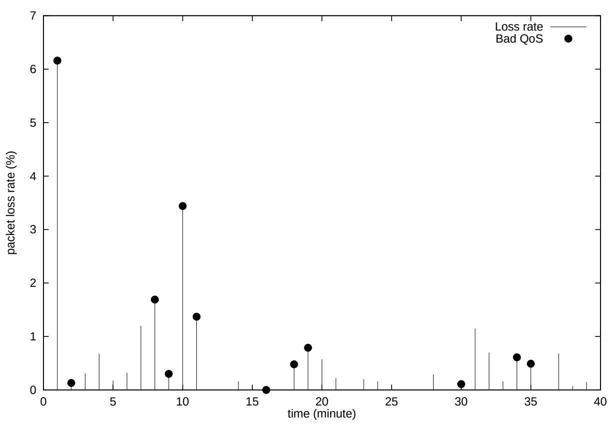

Figure 14. Loss rate for video packets

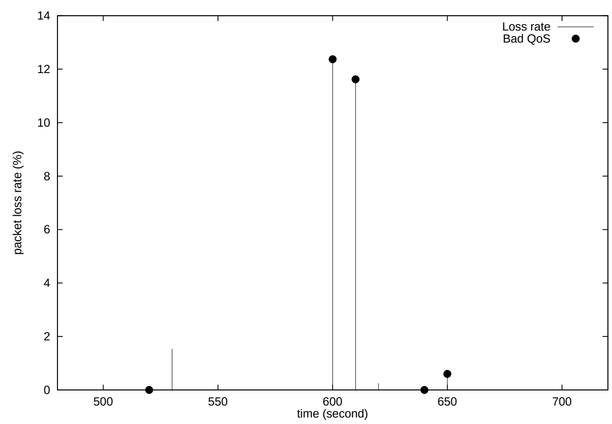

Figure 15. Loss rate for video packets, zoomed region minutes 9-12

Fig. 14 and Fig. 15 show similar results for the video session. The average loss rate for the video packets is approximately 0.2% per minute. The reason for the bid difference in the audio packet loss rate and the video packet loss rates must be related to the average packet size. The average (and constant) audio packet was larger by about 25% than the average (and variable) video packet.

The first minute shows the greatest packet loss for the whole session. One contributing factor is again because of the codec. It simply needs a certain amount of time before it adjusts to the network impulse response. At the same time the audio codec is also doing the same thing and it could be possible that both of them are using a large amount of bandwidth, probably more than it is available. This process results in increased packet loss. The problem of bandwidth utilization could also cause the increased packet losses inside time segment 9-12.

5. CONCLUSIONS

MBone multimedia provides a good insight into the traffic that is likely to be observed in future networks. Today's networks are not properly engineered to deliver high Quality of Service for multimedia traffic. Due to the subjective nature of the audio and video quality it still remains difficult to relate the QoS to measurements from the network.

Our experiments show that both delay jitter and loss effect the subjective quality of the audio session. By far, the more significant contributor to the QoS is packet loss. This suggests that designers should perhaps design their (de)coders to cope with even low rates of network losses (e.g., 0.5% as observed in this paper). This will result in higher overhead in both packet sizes as well as packet processing time at both the coder and decoder, and it is hoped that the benefits of robust coding schemes can more than offset the drawbacks due to the resulting overheads.

In our further work we will statistically relate the loss rates and packet delay jitter. We will also try to build engineering guidelines for end-to-end dimensioning of network resources in order to deliver high QoS.

REFERENCES

1. V. Kumar, "What is the MBone (or IP multicast)," 1997. http://www.mbone.com.
2. "Mbone related software." http://www-mice.cs.ucl.ac.uk/multimedia/software/.
3. P. Kirstein, G. Montaser-Kohsari, and E. Whelan, "SAP Security Using Public Key Algorithms," Internet Draft, IETF, September 1998. Work in Progress.
4. M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: Session Initiation Protocol," Internet Draft, IETF, September 1998. Work in Progress.
5. M. Handley and V. Jacobson, "SDP: Session Description Protocol," Standards Track RFC 2327, IETF, April 1998.
6. H. Schulzrine, S. Casner, R. Frederik, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," Standards Track RFC 1889, Network Working Group, January 1996.
7. "NIKSUN NetVCR." http://www.netvcr.com.
8. D. Ilie, "Multimedia protocol decoding," Tech. Rep. MEE 98-10, University of Karlskrona/Ronneby, Sweden, January 1999.
9. L. L. Peterson and B. S. Davie, *Computer Networks: A Systems Approach*, Morgan Kaufmann, USA, 1996.
10. ITU, "Video codec for audiovisual services at x 64 kbit/s," Tech. Rep. Recommendation H.261, ITU-T, Geneva, Switzerland, June 1998.
11. J. Beran, *Statistics for Long-Memory Processes*, Monographs on Statistics and Applied Probability, Chapman & Hall, One Penn Plaza, New York, NY 10119, USA, 1994.
12. J. Beran, R. Sherman, M. S. Taqqu, and W. Willinger, "Long-Range Dependence in Variable-Bit Rate Video Traffic," *IEEE Transactions on Communications* **43**, pp. 1566–1579, February/March/April 1995.
13. D. Minoli and E. Minoli, *Delivering Voice over IP Networks*, John Wiley & Sons, USA, 1998.
14. D. Devadhar, P. Pruthi, and M. Gerla, "Delay in public frame relay networks: Measurements, modeling and simulation," in *11th ITC Specialist Seminar: Multimedia and Nomadic Communications*, (Yokohama, Japan), October 1998.