BROADCAST QUALITY VIDEO OVER IP NETWORKS: CHALLENGES AND SOLUTIONS

Santhana Krishnamachari, John Hartung, Junfeng Bai
EGT, Inc

Abstract

Deploying video services over IP networks is a significant hurdle because of the stringent requirements on bandwidth, loss rate, and delay jitter imposed by video traffic. Today, most of the video traffic over the Internet and private networks is in the form of data downloads in non real-time, or real-time transmission of low resolution video. Scaling this to a full-resolution high quality video and to a large number of flows continues to be a significant challenge. In this technical case study we present the results of experiments that we have performed to study the feasibility of transmitting broadcast quality video over the Internet. Several point to point communication links were set up with nodes at various geographical locations. Statistics pertaining to video transport, namely available bandwidth, delay jitter, and loss rate were collected over long periods of time. These statistics are presented in this paper and they show that although the network paths are reasonably well behaved for most of the time, intermittent variation in characteristics will require additional techniques such as passive measurement, path diversity, rate adaptation, and error recovery to achieve robust reception.

1. INTRODUCTION

The main motivators for distributing video content over IP networks for a multi-channel operator are : (i) Substantial cost reduction in dollars per bits transmitted compared to other modes of transmission, e.g., over a dedicated satellite link. (ii) Flexibility to converge multiple services such as high speed data, voice and video over a single network. This convergence offers better statistical utilization of the channel resource and allows the amortization of the cost of multiplexing, transport gear over different services. (iii) Deployment of switched digital video broadcast over the last mile is expected to increase consumer interest in niche, on-demand content. Bringing such niche content to the headend can be achieved in a cost effective manner over an IP network or the Internet.

The Internet in its current form and most private networks cannot guarantee the quality of service required for broadcast quality video. Over the last two years, we have been witnessing an increasing trend of using IP networks for transport of audio and voice for commercial purposes. Rhapsody and a multitude of radio stations on the Internet stream audio services. VoIP providers like Vonage and cable MSOs use the public or private network for voice traffic. Audio streaming can tolerate longer network delays that results in a start-up latency. VoIP has stringent limits on one way delay and loss rate so that the interactivity of conversation is not impaired. The expected evolution of using IP networks for video transport is significantly more difficult because of the higher bandwidth requirements, in addition to the restrictions on delay jitter and loss rate.

Prior experiments performed to measure internet characteristics show a wide variation in loss rate. It was reported in [1] that the average loss rate was a very low 0.42 %, but during the worst one-hour period the average loss rate was over 13%. Published reports in [2] indicate that the convergence of inter-domain routes can take upto tens of minutes in case of a link failure. Although the performance of private IP networks can be drastically improved by traffic engineering, as the loading on these network increases the performance characteristics will be similar to the Internet.

In addition to the loss rate characteristics discussed above, variations in delay jitter and available bandwidth in the network significantly affect the robustness of video transmission. In this paper we present the results of the experiment we have performed to measure the loss rate, delay jitter and available bandwidth of network links. Similar to the reported results on loss rate, other network characteristics also exhibit large variations over short time periods.

The paper is organized as follows: Section 2 presents the results of our measurements. Section 3 presents details on existing technologies that can be used to improve the quality of service and their limitations. Section 4 presents the additional supporting technologies that are required to achieve robust video transmission and concludes the paper.

2. NETWORK MEASUREMENTS

In this case study, we deployed 6 nodes globally to measure the statistics of the Internet links between different locations. The geographical locations of these nodes are illustrated in Figure-1. We collected the available bandwidth, delay jitter and loss rate statistics for all the links over a period of about one month. In this paper, we will show some of measurement results for the link from Gsnet.ch to EGT at Atlanta. Other links exhibit very similar characteristics.

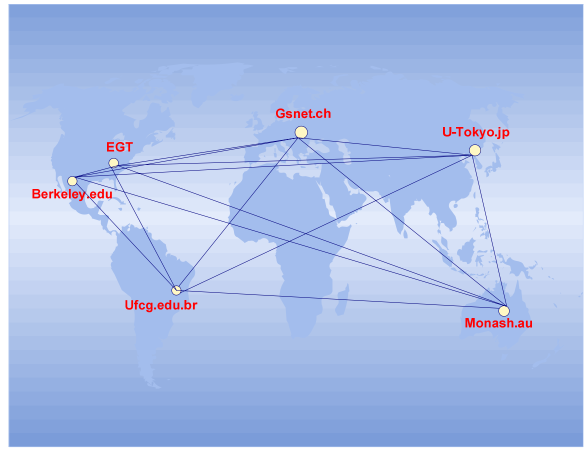

Figure 1. Geographical locations of nodes used for measurement

Packet Loss Rate Measurement

Figure-2 shows the packet loss rate experienced by a 1Mbps video stream from Gsnet.ch to EGT. The time duration of this experiment is from a Friday afternoon to the following Monday afternoon. The packet loss rate is collected every 1 second. Figure-3 shows the corresponding histogram for the loss rate.

The histogram shows that about 90% of the time the packet loss rate is less than or equal to 1%. Such small packet losses can be tolerated by video decoders by using advanced error concealment techniques. Forward Error Correction (such as Pro-MPEG [3]) techniques with about 5~10% overhead can also be used to counter small packet losses.

Figure-2 shows that the packet loss is relatively small during the weekend. For the period corresponding to Monday morning (between 65-75 hours in the plot) there are large bursty packet losses that continue for several hours. In such bursty loss cases, FEC or concealment techniques cannot recover all the lost packets.

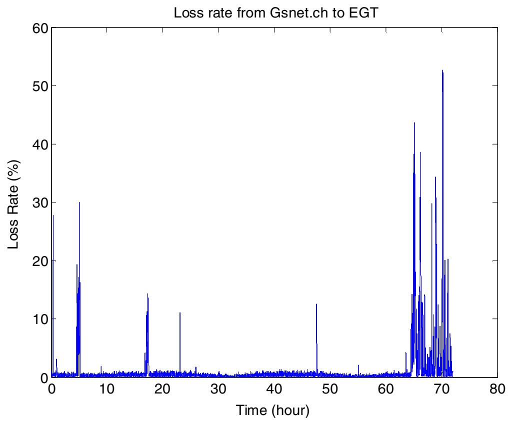

Figure 2. Packet loss rate statistics

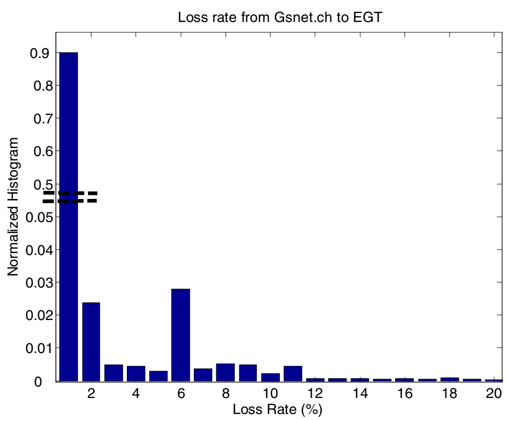

Figure 3. Histogram of packet loss rate

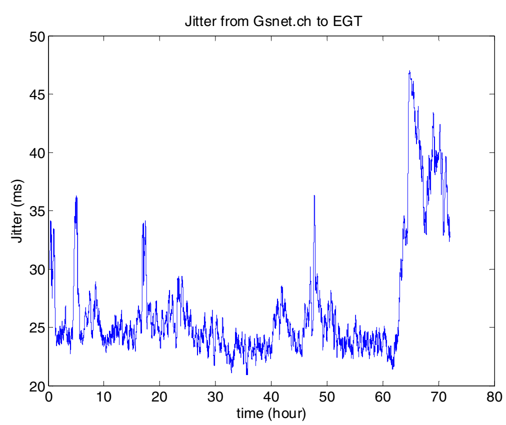

Figure 4. Delay jitter statistics

Delay Jitter Measurement

Figure-4 shows the jitter over same time duration and Figure-5 shows the corresponding histogram for the jitter. We measure the jitter at the receiving side with the following equation

$$Jitter = [Tr(i)-Tr(i-1)] - [Ts(i) - Ts(i-1)]$$

Here, Ts(i) is the sending time of the *i*th packet at the sending side and it is inserted into the packet header as a timestamp. Tr(i) is the time when receiving side receives the *i*th packet.

Each packet may experience different transit delay because of different queueing delay introduced by routers along the path from source to destination. We collect the maximum of absolute value of jitter for every 100 received packets and show the result in Figure-4. This maximum jitter is critical for determining the size of dejittering buffer at the receiving side to correctly recover time interval between packets while preventing the buffer from overflowing or underflowing.

It can be observed from Figure-2 and Figure-4 that as the network loss rate increases, the delay jitter also gets worse as expected.

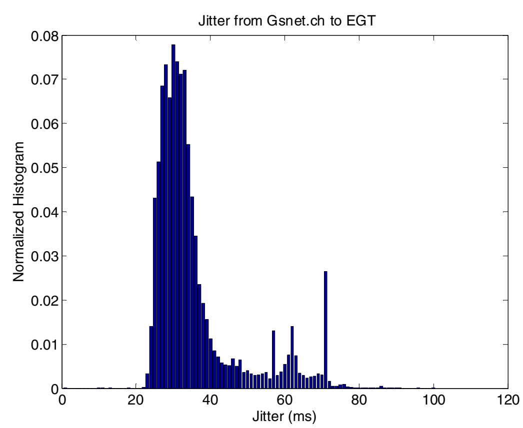

Figure 5. Histogram of delay jitter

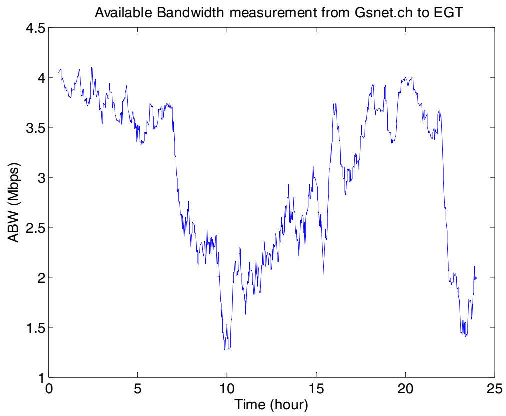

Figure 6. Available bandwidth statistics

Available Bandwidth Measurement

The available bandwidth estimation tool Pathload [4] was used to collect the bandwidth statistics. The result of available bandwidth measurement for 24 hours between Gsnet.ch to EGT is shown in Figure-6. The available bandwidth changes dramatically in the range of 1.5 – 4 Mbps. If the video rate is higher than the available bandwidth, this will result in increased congestion in the network links resulting in large packet losses and delay jitter. This plot clearly shows the need for adapting the video bitrate to match the available bandwidth.

3. EXISTING TECHNOLOGIES

Packet switched networks were originally developed to provide best effort delivery of data and achieve high efficiency through statistical multiplexing. However, the stochastic nature of traffic through these networks leads to unavoidable congestion at switches that is inconsistent with the requirements of high quality video delivery [5]. Two approaches have been taken to reduce or mitigate the effects of this congestion. The first accepts the performance of the network and attempts to provide acceptable QoS for video by compensating for network characteristics at the ingress and egress of the network. In the second approach new protocols have been standardized to allow prioritization of video flows through the switching elements of the network, thereby reducing the probability of congestion for that data. The following two sections give an overview of these two approaches.

Endpoint QoS

The available bandwidth in a default internet route varies widely, and the effect of exceeding this bandwidth is increased packet loss and a rapid degradation of video quality. One solution to this problem is source rate control where the sender adapts its rate to match the available bandwidth. This assumes that there is an accurate measurement of the bandwidth. This technique works for point to point transmission, however, for multicast streams the available bandwidth will usually vary for each endpoint. In this case receiver rate control is used to adapt the rate as a function of the available bandwidth to each receiver. This is implemented by encoding and packetizing the media in multiple layers, with a base layer providing the minimum acceptable quality. The receiver adjusts its rate by connecting to one or more layers whose sum satisfies the bandwidth constraint [6].

The transmission latency of packets between two endpoints varies when congestion occurs due to changing queueing delays in routers along the path. This delay jitter can lead to jerkiness in the playback and packet losses when a packet is delayed beyond its presentation deadline. The introduction of a playout buffer is used to relax the timing constraint, however, this leads to a delay in playout that can be unacceptable when a new stream is started.

Commercial streaming players like the Microsoft Windows player and the Real Network Real player typically introduce 5-15 seconds of delay. An alternative approach that minimizes latency and startup is adaptive media playout (AMP) [7]. In this scheme the rate at which the decoder buffer is emptied, and the media is presented, is varied in order to avoid losses due to missed presentation deadlines. This can combined with retransmission to avoid packet losses as described in the next section.

Packet losses can be dealt with using channel coding to recover from losses and error concealment and resilience to minimize its effect. There are two basic channel coding techniques, retransmission and forward error correction (FEC). Retransmission consists of detecting the lost packet at the receiver and signaling the sender of the loss. A minimum delay of one round trip time is incurred in addition to the time needed at the receiver to detect the loss. This technique has the advantage of using additional bandwidth only when losses occur, however, it requires a back channel that may not be available in applications such as multicast. Alternatively lost packets can be recovered without a back channel using forward error correction. This is accomplished by interleaving a group of packets and adding an FEC code to each. The FEC from the group can be used to recover a number of lost packets within the group. FEC has a disadvantage in that it incurs a rate increase due to the addition of the code words and a delay due to the interleaving of multiple packets even when there are no packet losses.

Error concealment makes use of spatial and temporal correlation to recover lost video information caused by packet loss. These techniques are not standardized and many techniques have been developed making use of motion and spatial information to improve the estimate. Error resilience attempts to encode and packetize the encoded video bitstream in order to minimize the effect of synchronization loss and error propagation. In general, a single bit error can prevent decoding of a video stream, due to variable length codewords, until the next synchronization word. Techniques such as application level framing (ALF) [8] allow the bitstream to be packetized so that each packet is independently decodable. This prevents individual packet losses from propagating errors to the following packet.

The above techniques can be combined to take advantage of the fact that not all bits in the video stream are of equal importance. For example, ALF can be used to form two packet types; one containing high importance data such as I-frames, and another containing lower importance B and P-frames. Unequal loss probabilities, and transmission cost, can be obtained for the two types by applying unequal error protection. This combined source and channel coding achieves lower distortion at an equivalent transmission rate as compared to a system using equal protection for all video bits.

Network QoS

Several types of network protocols have been standardized to enable end to end QoS capabilities in large scale networks (internet). An early protocol to support the requirements of individual flows was called intserv. Intserv works with a resource reservation protocol (RSVP) to set up a path through the network meeting the flow requirements. However, it scales poorly and two types of service aggregation protocols have been subsequently developed. The first type, called diffserv, provides a mechanism to label packets according to their required class of service (COS). The diffserv protocol uses an IP header field to label packets according to their transport requirements so that routers can apply different forwarding algorithms to meet those requirements. A second protocol that serves a similar purpose is multiprotocol label switching (MPLS). In addition to specifying the COS for the packet, this protocol also specifies the forwarding path.The forwarding path is set up in advance using a label distribution protocol (LDP). The primary difference between diffserv and MPLS is that diffserv uses the default routing (e.g. open shortest path first (OSPF)), while MPLS enables engineered routes to be specified.

These protocols have several shortcomings for both private networks and the public internet. The first is that traffic engineering is needed to allocate sufficient bandwidth for the aggregate traffic in each link. Because of the stochastic nature of the flows, however, large over-provisioning is needed. The second problem is that QoS guarantees are only possible if all routers in the network implement the protocols. It is possible to build private networks with these capabilities, however, due to the heterogeneity of the internet, this is unlikely to be supported for many years. In addition bridging protocols are needed to maintain the COS labeling across domains such as different internet service providers (ISPs) and types of facilities (e.g. DWDM and ethernet).

4. NEEDED SUPPORTING TECHNOLOGIES

As presented in the previous section, end-point and network QoS techniques offer improved performance, but still cannot deliver guaranteed resiliency against varying network characteristics. The following techniques or combination of technologies are required to achieve robust video transmission over IP networks.

Estimation of network characteristics: Robust measurement of network parameters is essential to proactively use the techniques discussed in Section 3. It should be clear from the plots of Section 2 that detecting and responding to variation in network characteristics can lead to frequent and in some cases long interruption in service. It is necessary to ensure that the measurement techniques have relatively short time constants to respond to short term variations and in addition be able predict the variations in network statistics as well.

Passive measurements: The measurement of network statistics needs to be achieved in a passive mode where no additional probing data is required for measurement. As the loading on the network links increases, the additional load introduced by the active measurement techniques should be minimized.

Path diversity: To avoid interruptions caused by catastrophic link failure it is necessary to introduce route diversity for video transmission that can be implemented in a fashion scalable to large number of streams. One way of achieving route diversity is to combine source coding technique such as multiple description coding (MDC) and use MPLS COS labeling and forwarding to ensure the transport of the two different descriptions is over paths without common links. MDC is an encoding technique where the video is coded into two or more streams each of which is independently decodable. Jointly decoding multiple descriptions increases the quality of the received video.

In summary, we have presented the results of network measurements between several nodes distributed over diverse geographical locations. These measurements clearly indicate that the characteristics essential for video transport vary widely over time. Two different existing approaches, namely the end-point QoS and network QoS, to improve the quality of video transport over IP networks were presented. Finally, a set of new techniques pertaining to measurement and path diversity that are required to provide broadcast quality distribution were presented.

REFERENCES

[1] D.G. Andersen, A.C. Snoren and H. Balakrishnan "Best Path vs multi-path Overlay Routing", in *Proc of IMC* 2003, Oct 2003.

[2] C. Labovitz, A. Ahuja, and F. Jahanian, "Delayed Internet Convergence", in *Proc. Of SIGCOMM* 2000, Aug 2000.

[3] Pro-MPEG Forum code of Practice #3 - release 2: Transmission of Professional MPEG-2 transport streams over IP networks. http://www.pro-mpeg.org

[4] M. Jain and C. Dovrolis, "Pathload: A Measurement Tool for End-to-End Available Bandwidth", *Proc of PAM* 2002, 2002

[5] Ion Stoica, "Stateless Core: A Scalable Approach for Quality of Service in the Internet", *Ph.D. thesis, Carnegie Mellon Univ., Pittsburgh*, PA, Dec. 15 2000.

[6] S. McCanne, V. Jacobsen, and M. Vetterli, "Receiver-driven layered multicast", *ACM SIGCOMM,* Aug. 1996.

[7] M. Kalman, E. Steinbach, and B. Girod, "Adaptive Media Playout for Low Delay Video Streaming over Error-Prone Channels", *IEEE Transactions on Circuits and Systems for Video Technology,* June 2004.

[8] D. Clark and D. Tennenhouse, "Architectural Considerations for a New Generation of Protocols", *ACM SIGCOMM,* September 1990.

Contact Information:
Santhana Krishnamachari
EGT
75 5th Street, NW
Atlanta GA 30308
santhana@egtinc.com