An Empirical Study of RealVideo Performance Across the Internet

Yubing Wang
EMC Corporation
Hopkinton, MA 01748, USA
wang_yubing@emc.com

Mark Claypool
Computer Science Department
Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609, USA
claypool@cs.wpi.edu

Zheng Zuo
EMC Corporation
Hopkinton, MA 01748, USA
zuo_zheng@emc.com

***Abstract*—The tremendous increase in computer power and bandwidth connectivity has fueled the growth of streaming video over the Internet to the desktop. While there have been large scale empirical studies of Internet, Web and multimedia traffic, the performance of popular Internet streaming video technologies from the user perspective and the impact of streaming video on the Internet is still largely unkown. This paper presents analysis from a wide-scale empirical study of RealVideo traffic from several Internet servers to many geographically diverse users. We find typical video quality to be high, achieving an average of frame rate of 10 fps and very smooth playout, but very few videos achieve full-motion video playout rates. Overall video performance is most influenced by the bandwidth of the end-user connection to the Internet, but high-bandwidth Internet connections are pushing the video performance bottleneck closer to the server. RealVideo traffic appears responsive to network congestion since much of RealVideo traffic uses TCP, and RealVideo traffic that uses UDP appears to have data rates similar to that of TCP over the duration of a video clip.**

I. INTRODUCTION

The growing number of users with high bandwidth connections to the Internet and the increasing power of desktop computers have fueled the use of the Internet to carry potentially high-quality video. Increasingly, Web sites are offering streaming videos of news broadcasts, music television and live sporting events. Users can watch these streaming video clips through a Web browser by simply clicking on a link and having the Web browser start up an associated video player.

Over the years, there have been a number of studies measuring the performance of Internet backbones and end-hosts [TMW97, Pax99], as well as detailed studies on the performance of Web clients [KW00, Mah97]. However, to the best of our knowledge there have not been wide-scale empirical measurement of video performance across the Internet. While the existing studies have been valuable in helping understand Internet performance, they are not sufficient for characterizing streaming video performance since video has application requirements different than the majority of Internet traffic.

Unlike typical Internet traffic, streaming video is sensitive to delay and jitter, but can tolerate some data loss. In addition, streaming video typically prefers a steady data rate rather than the bursty data rate often associated with window-based network protocols. For these reasons, streaming video applications often use UDP as a transport protocol rather than TCP, suggesting that video flows may not be "TCP-friendly" or, even worse, that video flows may be unresponsive to network congestion. Furthermore, while the performance of the Web is often determined by the response time in downloading an entire page, video traffic can be sent at a bit-rate adjusted to the end-host connection. Video traffic is often long-lived with even small clips lasting several minutes, so that while playing, the bandwidth required can be adjusted to prevailing network congestion conditions while still playing out the video in real-time. Thus, the arrival time of the last byte and even the total bandwidth are by themselves ineffective measures of video performance. Similarly, loss, a common measure of Internet performance, is not sufficient to characterize the performance of video traffic. While frame loss can have a severe impact on the perceptual quality of video, repair techniques to recover multimedia packet loss or ameliorate its effects [LC00, PHH98] are often applied to video streams.

RealNetworks[1] produces among the most popular streaming media clients and servers in the world [Jup01]. While RealNetworks includes guidelines for producing video clips that stream on the Internet with acceptable quality, the impact and effectiveness of RealVideo on the Internet is not well-researched. In previous work we measured system level statistics of streaming video [GK98], but lacking were performance measurements from industry quality video and user opinions on perceived video quality.

This study measures RealVideo performance across the Internet by playing video clips selected from a variety of geographically distributed Web servers to clients with many configuration parameters around

[1] http://www.real.com/

the world. We developed a customized video player called *RealTracer* that plays streaming RealVideo video clips and measures their performance. Through personal contacts and online forums, we solicited users to help in the study by downloading RealTracer and running it on their computers. During a two-week period in June 2001, over 60 users from 12 countries provided performance information on about 2800 streaming video clips downloaded from 11 servers in 8 different countries.

In analyzing our data, we make several contributions to better understanding the performance and impact of streaming video on the Internet. We find that overall, RealVideo videos on the Internet have very high quality on average. The correlation between geographic region and video performance is slight at the server-side, but quite noticable at the user-side. The increase in end-host computing power and network bandwidth is pushing the video performance bottleneck closer to the server. RealVideo appears to respond to congestion and appears to receive bandwidth comparable to that of TCP over the duration of the clip, even when using application-layer congestion control.

The rest of this paper is organized as follows: Section 2 presents background needed to help understand our results; Section 3 describes our approach to obtain a wide-scale set of Internet measurements; Sections 4 and 5 present and analyze the measurement data obtained; Section 6 introduces some related work; Section 7 summarizes our conclusions and Section 8 presents possible future work.

II. BACKGROUND

RealNetworks provides the most popular streaming media player, called *RealPlayer*, on the U.S. Internet. In January 2001, 25.9 million U.S. Internet users at home used a RealNetworks player, up 47.6 percent from January 2000; 21.5 million used Microsoft Windows Media Player, up 31.2 percent; and 7.3 million used Apple QuickTime, down 8.4 percent [Jup01]. RealNetworks also has the largest share of users at work: in January 2001, 10.5 million U.S. Internet users at work used a RealNetworks player, up 52.1 percent from 2000; 9.0 million used Media Player, up 39.9 percent from 2000; and 1.9 million used QuickTime, up 8.5 percent from 2000.

Content providers using RealVideo create streaming videos using a number of possible video codecs, convert it to RealNetworks' proprietary format and place it on a host running RealServer. During creation, content providers select target bandwidths appropriate for their target audience, and specify other encoding parameters such as frame size and frame rate appropriate for their content. A RealServer will stream the video to a user's RealPlayer client upon request.

A. Connections and Protocols

RealServer primarily uses Real Time Streaming Protocol (RTSP) [SRL98] to communicate with RealPlayer clients. However, earlier versions of Real server used Progressive Networks Audio (PNA) protocol and, for backward compatibility, newer real servers and players still support this protocol. Nearly all the video clips we selected for our study used RTSP. Occasionally, RealServer will use HTTP for metafiles or HTML pages, and it may also be used to deliver clips to RealPlayer clients that are located behind firewalls. However, we did not observe any HTTP connections for any of the users in our study.

RealServer uses two network connections to communicate with RealPlayer clients: one for communicating control information with the client, and one for communicating the actual data. RealServer uses the control connection to request client configuration parameters and to send information such as clip titles, and clients use the control connection to send instructions such as fast-forward, pause, and stop. The video clips themselves, on the other hand, are actually streamed over the data connection.

At the transport layer, RealServer uses both TCP and UDP for sending data. The initial connection is often in UDP, with control information then being sent along a two-way TCP connection. The video data itself is sent using either TCP or UDP. The actual choice of transport protocols used is determined by the RealPlayer and RealServer. This auto-configuration of protocols can be overridden by the user, but is the default and recommended setting for RealPlayers [Rea00b].

B. Buffering

For each video clip, RealPlayer keeps a buffer to smooth out the video stream because of changes in bandwidth, lost packets or jitter. Data enters the buffer as it streams to RealPlayer, and leaves the buffer as RealPlayer plays the video clip. If network congestion reduces bandwidth for a few seconds, for example, RealPlayer can keep the clip playing with the buffered data. If the buffer empties completely, RealPlayer halts the clip playback for up to 20 seconds while the buffer is filled again.

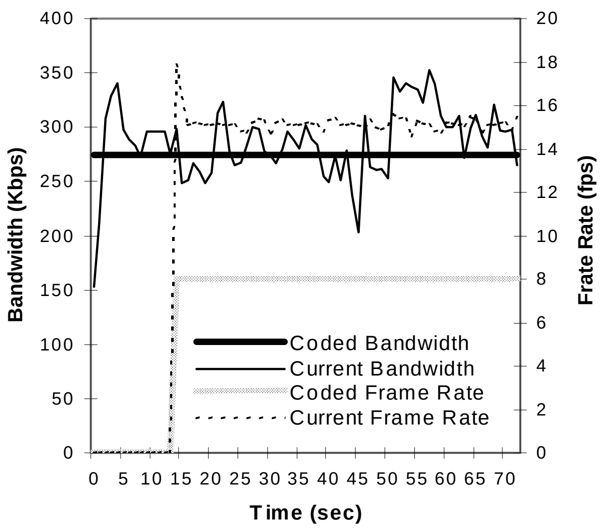

Figure 1. Buffering and Playout of a RealVideo Clip

Figure 1 depicts the buffering and start of the playout of a RealVideo clip. The horizontal axis represents time from when the clip is first downloaded. The vertical axis represents bandwidth on the left and frames per second on the right. The four lines depict the encoded bandwidth and framerate, specified by the server when the video was created, and the actual bandwidth and framerate recorded as the clip is playing out.

During the initial 13 seconds, the video clip is being downloaded and buffered but not played out. Once the playout of the video clip begins, the frame rate varies somewhat but is steadier than the actual bandwidth because of the initial buffering. The actual bandwidth is greatly influenced by the prevailing network conditions but the frame playout can rely on the buffer to smooth the playout of the video frames.

C. RealVideo Bandwidth Characteristics

RealSystem uses a technology called *SureStream* in which a RealVideo clip is encoded for multiple bandwidths [Rea00a]. A RealPlayer connects to a single video URL and the RealServer determines which stream to use based on the RealPlayer's specified minimum and average bandwidths. The actual video stream served can be varied in mid-playout, with the server switching to a lower bandwidth stream during network congestion and then back to a higher bandwidth stream when congestion clears.

A portion of a RealVideo clip's bandwidth first goes toward the audio, leaving the remainder of the track for the video. For example, a 20 Kbps RealVideo clip (typical for a 28.8 modem) with a 5 Kbps RealAudio voice codec will leave 15 Kbps for the video, while an 11 Kbps music codec, will leave only 9 Kbps for the video.

Most RealVideo streams are created with a *Scalable Video Technology* option that allows RealServer to automatically adjust the video stream according to the clients connection and computer processing speed [Rea00a]. If the clip is unable to play at the encoded frame rate on a client machine, it will gradually reduce the frame rate in a controlled fashion to maintain smooth video. The initial size of the video stream is based on the maximum client bit rate (a RealPlayer configuration parameter) and other video settings. If packets are lost during video delivery, special packets that correct errors are sent to reconstruct the lost data.

III. APPROACH

In order to empirically evaluate the performance of RealVideo across the Internet, we employed the following methodology:

- Build a customized player, called *RealTracer*, that plays RealVideo clips and records performance statistics, including user ratings (see Section 3.A).
- Select RealVideo servers from geographically diverse Web sites and choose diverse video clips from thoses sites (see Section 3.B).
- Solicit users to run our customized player and gather data (see Section 3.C).
- Analyze the results (see Sections 4 and 5).

A. RealTracer

We required a RealVideo player with a customized front-end interface to gather user end-host information and a customized back-end to record performance statistics. We designed and implemented such a player, called *RealTracer*, using the

RealPlayer core video playout engine and the RealNetworks Software Development Kit[2] (SDK). The SDK exposes the interfaces used in RealPlayer, enabling development of new tools and applications, and comes with documentation, header files and samples. The RealPlayer core is not included in the RealNetworks SDK, but comes with the latest basic version of RealPlayer (version 8.0 at the time of this study). We included instructions for users to download and install RealPlayer before using RealTracer in case their PC did not already have RealPlayer installed.

Upon startup, RealTracer requests country, state, and network configuration information from the user as depicted in Figure 2a. In addition, RealTracer automatically detects Operating System type, CPU type, RAM and IP address.

Upon clicking "OK" the main window pops up. The main window, depicted in Figure 2b, provides a playlist for video clip selection and allows users to start and stop playing the clip. Clicking "Play" begins playout of the first video clip in the playlist.

While the video is playing, RealTracer gathers system statistics: encoded bandwidth, measured bandwidth, transport protocol, encoded frame rate, measured frame rate, playout jitter, frames dropped and CPU utilization.

When each clip finishes playout, the user is solicited to assess the video quality by providing a numeric rating from 0-10 as depicted in Figure 2c.

The user data and specific clip statistics are then sent via both email and FTP to a server at Worcester Polytechnic Institute. The default behavior is to then proceed sequentially through the playlist to the end.

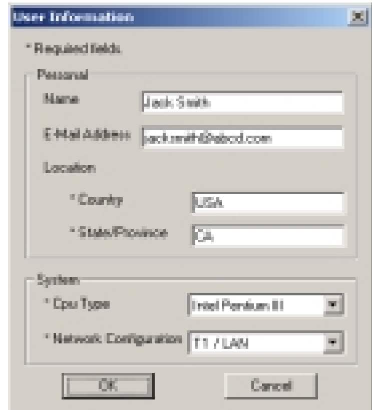

Figure 2a. RealTracer User Information Window

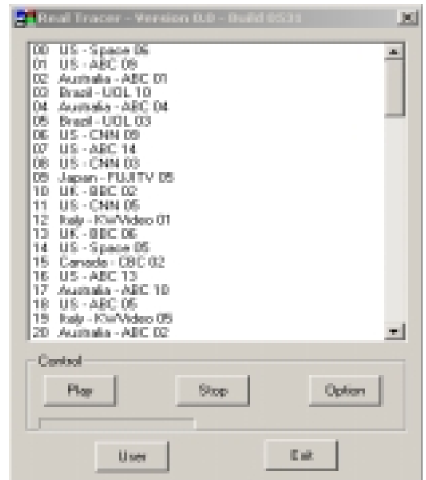

Figure 2b. RealTracer MainWindow

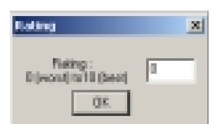

Figure 2c. RealTracer Clip Rating Window

If so desired, the user can control the length of the clip playout and the requests for quality ratings using the "Options" button. The defaults are to play the clip for 1 minute and request a rating for each clip, proceeding to the next clip after 10 seconds if no rating is given.

B. Video Selection

We chose RealServers accessible through Web pages from 6 geographic regions: Asia, Austrailia, Europe, Japan, North America, and South America. Within each region, popular RealNetworks sites were chosen using Yahoo[3]. The countries that were chosen include Australia, Brazil, Canada, China, Italy, Japan, the United Kingdom, and the United States. Figure 3

[2] http://www.realnetworks.com/devzone/sdk/index.html

[3] Yahoo: "News and Media" → "By Region" → "Countries"

depicts a geographic representation of the RealServer sites chosen. From each site, we tried to select a variety of video content among all the videos that were offered.

C. Solicit Users

Once the servers and videos were selected, we did beta testing with a few colleagues for about two weeks to try and catch and fix bugs in RealTracer and the data gathering process.

We solicited[4] friends, family and colleagues from various parts of the world to help in the study. Since it was fairly easy for us to obtain data points from inside Massachussetts, we asked friends and colleagues on campus and at work to solicit help from people they knew outside of Massachussets. We also posted messages asking for help to the `rec.video` newsgroup and `end2end-interest` mailing list.

We then gathered data from users running RealTracer for an 11 day period from June 4, 2001 to June 15, 2001. Figure 4 depicts a geographic representation of the locations of users that ran RealTracer.

Figure 3. Geographic Representation of RealServers

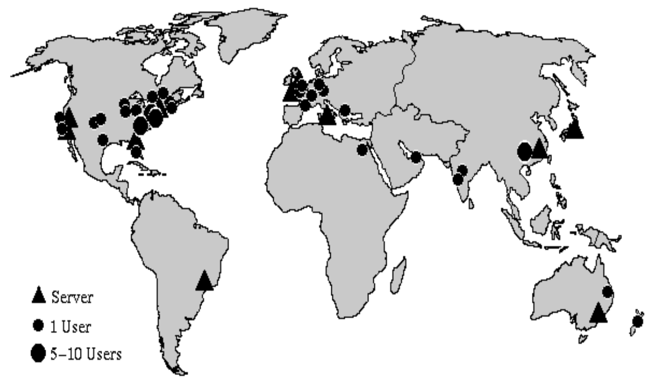

Figure 4. Geographic Representation of RealServers and Users

IV. RESULTS

A total of 63 users participated in the study, playing a total of 2855 clips and watching and rating a total of 388 clips. Figure 5 depicts a Cumulative Density Function (CDF) of the clips played per user. The playlist contained 98 video clips, the maximum that a user could have provided from one RealTracer run. However, as the playout of all videos in the playlist took about two hours and some users experienced troubles running RealTracer for some of the video clips, many users played out only a portion of the 98 clips. Still, from Figure 5 it can be seen that half the users played out 40 clips or more. Users were asked to rate 3-10 video clips or more, as their time and interest permitted. Figure 6 depicts a Cumulative Density Function (CDF) of the clips rated per user. Half the users provided ratings for 3 clips. Some users provided ratings for significantly more clips than requested while other users chose not to rate any clips.

---

[4] See `http://perform.wpi.edu/real-tracer/` for the specific instructions given to users

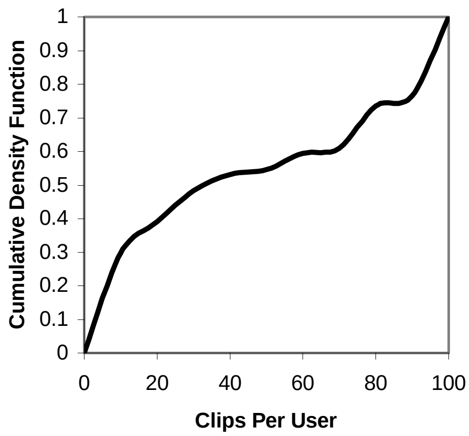

Figure 5. Cumulative Density Function (CDF) of Video Clips Played per User

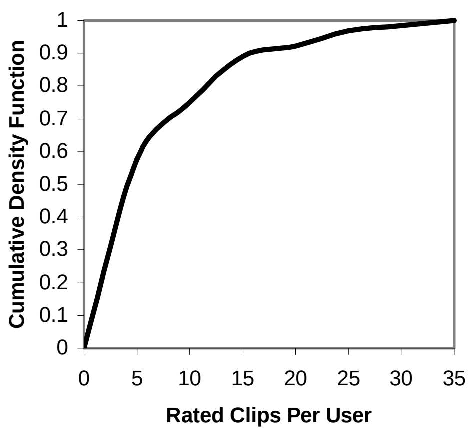

Figure 6. Cumulative Density Function (CDF) Video Clips Rated per User

The users provided data points from 12 different countries. The small circles in Figure 4 depict the geographic location of users that participated in the study, while the large circles depict clusters of 5-10 users. Figure 7 depicts the breakdown of the total clips that were played by users from each country while Figure 8 depicts the breakdown of the total clips that were served by RealServers from each country.

Figure 9 depicts a breakdown of the U.S. users per state. It is apparent we have considerably more data from users in Massachusetts than from other parts of the U.S. We did not include any data from our own RealTracer runs, but still naturally many people solicited for help with our study reside in Massachusetts, close to the authors of this paper. To see if this large body of users unduely biased the results, for the data analysis in Section 3, we briefly analyzed the overall frame rate in the U.S. by removing all data from Massachusetts users. The results indicate that the CDF for framerate without the Massachusetts users is nearly the same as the CDF for framerate with the Massachusetts users. Due to space constraints, we do not show this graph here. In subsequent results, we analyze all data gathered for completeness, unless otherwise noted.

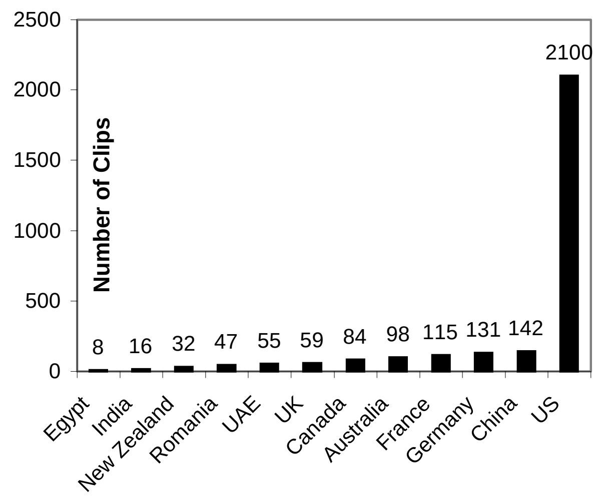

Figure 7. Video Clips Played by Users from Each Country

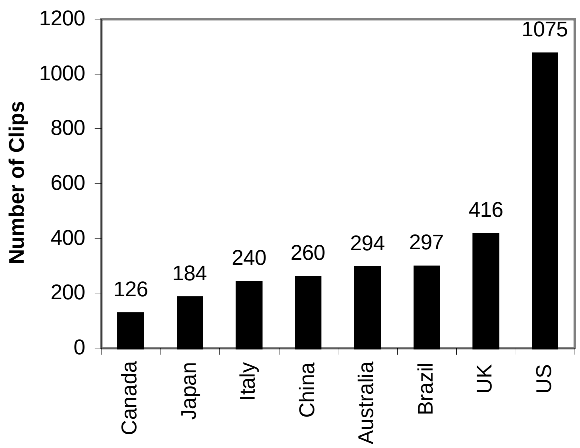

Figure 8. Video Clips Served by RealServers from Each Country

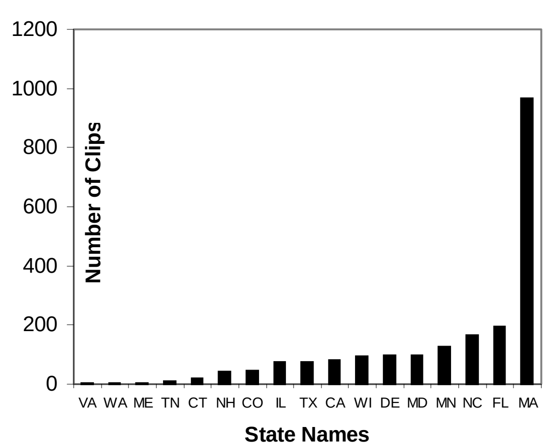

Figure 9. Video Clips Played by U.S. Users from Each State

There are notable absence of users from Japan, Korea, Southeast Asia, South America, Central America, Africa and perhaps some other European countries. Also notable is the sparsity of users from the Silicon Valley area. Future work would suggest trying to solicit help from users from those areas.

A surprising number of video clips in our playlist could not be accessed for short periods of time. Figure 10 depicts the fraction of clips from each server that were unavailable at the time a user tried to access them. Often, other clips on the same server could be accessed, so it is not necessarily a measure of server availability, but rather general RealVideo clip availability. We are are not certain why some clips were not available, but overall, on average about 10% of the time a video clip was unavailable. Note that there were several users that tried to participate in the study that were behind firewalls that did not allow RTSP packets through. Their data is not included in Figure 10 and we have removed their data from all analysis in this paper.

V. Analysis

A basic unit of video performance is the rate at which frames are played. Very low frame rates are perceived more like a slideshow of still images than of streaming video. The higher the frame rate, the smoother the motion. The key frame rates we observe are [Rea00a]:

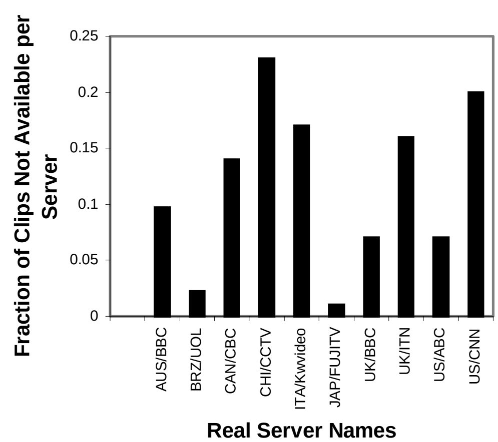

Figure 10. Fraction of Unavailable Clips

- The standard frame rate for full-motion video is 24 to 30 frames per second (fps). At this speed, the human eye perceives movement as continuous, without seeing individual frames.

- A common frame rate for computer video that approximates full-motion video is 15 fps. To most people, a 15 fps video flows smoothly, although for some videos, it will not appear quite as fluid as it would at a higher frame rate.
- Below 15 fps, a video looks choppy.
- Below 7 fps, a video looks very choppy.
- Below 3 fps, a video essentially becomes a series of still pictures.

In our analysis, we concentrate on frame rates of 3, 15 and 25 frames per second.

However, even a high frame rate can appear choppy if the frames are not played out at even intervals. In previous work [CT99], we found that variance, or *jitter*, in frame playout intervals can degrade perceived quality nearly as much as does frame loss. In this work, we measure jitter as the standard deviation of the inter-frame playout time over an entire video clip (1 minute long by default in RealTracer). Since human perception of delay for interactive applications is around 100 ms, we focus on the percentage of videos that have a jitter of 50 ms or less.[5] In addition, jitter events that are larger than the average inter-frame playout are most noticeable by users, so we also focus on the percentage of videos that have a standard deviations of about 300 ms (about the average inter-frame playout time for the minimum acceptable 3 fps rate) or greater, as this may be a reasonable upper bound on an acceptable amount jitter.

Even measures of frame rate and jitter alone are not always sufficient to determine the quality of the video as perceived by the user. During encoding, RealVideo adjusts the frame rate by keeping the frame rate up in high-action scenes, and reducing it in low-action scenes. Thus, an encoded video clip will intentionally not have just one frame rate, but a mix of frame rates that vary with the video scene content. In addition, our previous work [CT99, TC01] shows that the temporal aspect of a streaming video clip has an impact on the effects of reduced frame rate and jitter on perceptual quality. In this work, we record and analyze the ratings (from 0-10, see Section 2.1) for videos watched and rated by users to provide additional analysis of performance beyond measures of jitter and frame rate.

A. *Frame Rate*

We first analyze the performance of RealVideo clips across the Internet in general. Figure 11 shows a CDF of the frame rate for all the video clips played. The mean frame rate is 10 fps, above the range of really choppy video but well short of very fluid video. Approximately 25% of all videos played are under the minimum acceptable 3 fps, while the same number (25%) of videos are played at the approximate full-motion video rate, 15 fps. Only a very small fraction, less than 1%, of all videos achieve true full-motion video frame rates.

We next examine the frame rates achieved for different end-user network configurations. With the increase in high-speed Internet connections for home users, we may see more bottlenecks to performance in the server and not in the end-host network. Typical 56k modems can stream at rates up to 50 Kbps, but DSL and Cable modems can stream at rates up to 500 Kbps. Figure 12 depicts a CDF of frame rate for different end-user network configurations. The frame rates afforded by modem connections are clearly worse than the frame rates with higher speed connections. Over half of all videos streamed over modems play out at less than 3 fps, and less than 10% of videos streamed over modems achieve a smooth 15 fps. Contrast this to the higher speed connections in which only 20% of videos have frame rates less than 3 fps, while nearly 30% of videos play out at 15 fps. Also, high-speed home-Internet connections afforded by DSL and Cable modems provide nearly the same performance for streaming video as do higher-speed T1/LAN connections. This suggests that video performance bottlenecks are increasingly less likely to be the end-user connection.

[5] Based on the empirical assumption that approximately 95% of the playout times are within two standard deviations of the mean [DP93].

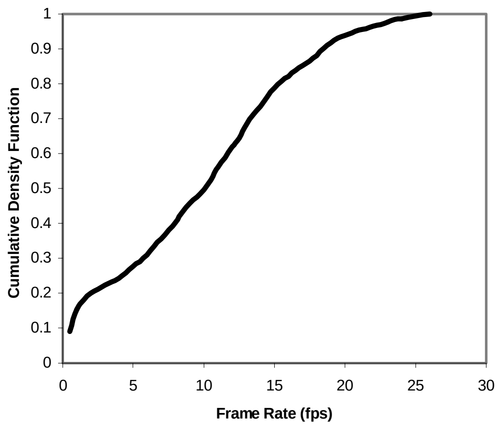

Figure 11. CDF of Frame Rate for all Video Clips

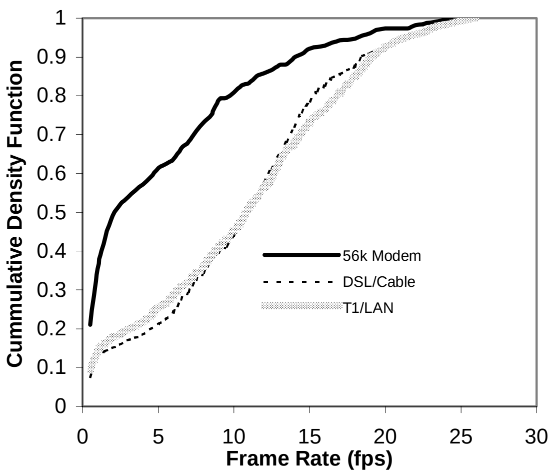

Figure 12. CDF of Frame Rate for Different End-Host Network Configurations

We further test this hypothesis by examining the bandwidth achieved by each class of end-host network configuration, depicted in Figure 13. Notice that DSL/Cable modems that can typically achieve throughputs from 256–512Kbps, operate near full capacity less than 10% of the time. This further suggests that the bottleneck to video bandwidth is beyond the end-network connection. By comparing Figure 12 with Figure 13, it can be seen that modem connections get a proportionally higher frame rate for their network bandwidth than do higher-speed connections.

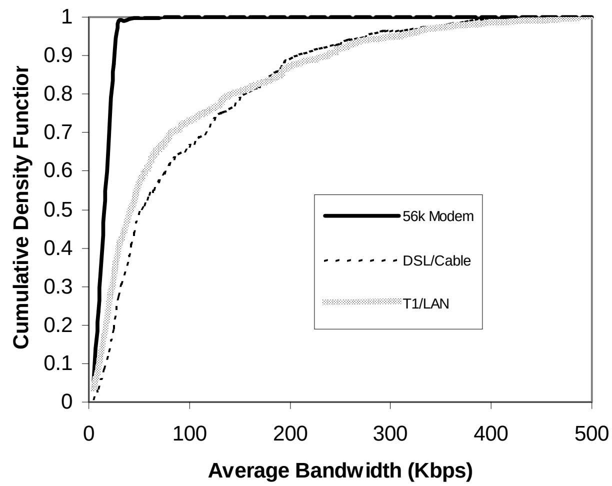

Figure 13. CDF of Bandwidth for Different End-Host Network Configurations

It may be expected that servers in “wired” geographic areas, say North America, will provide better streaming video performance than others, say Brazil. Figure 14 depicts a CDF of the frame rate for the servers used in our study, separated into 5 different geographic regions. The 5 regions all provide very similar frame rate distributions, with the mean of the best frame rate distribution about 13 fps and the mean of the worst frame rate distribution about 8 fps. Australia and Europe have the best frame rate distributions, with Europe providing a larger percentage of frame rates above 20%. Asia provides the worst frame rates, but the differences at very low frame rates is small, and Asia servers actually have a larger percentage of frame rates above 15 fps than do North America servers.

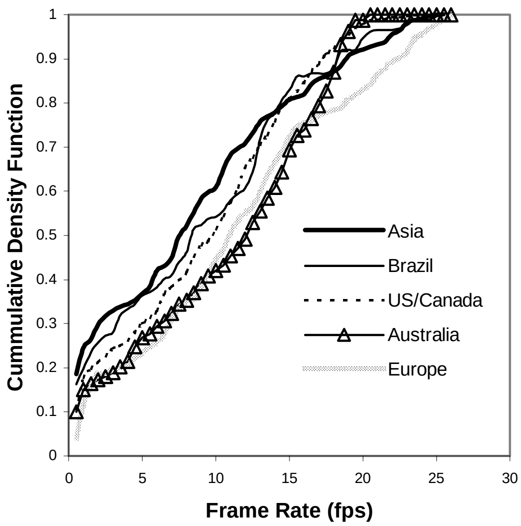

Figure 14. CDF of Frame Rate for RealServers in Different Geographic Regions

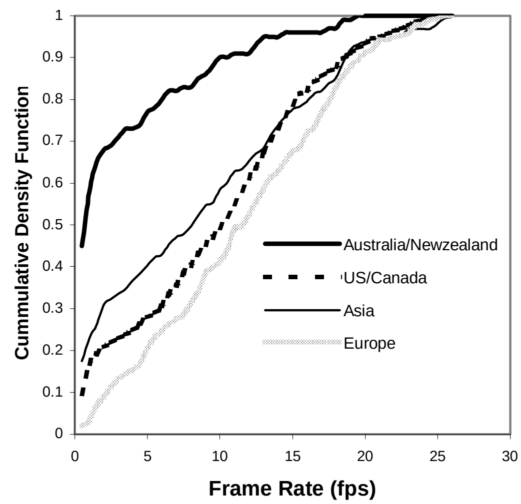

Figure 15. CDF of Frame Rate for Users in Different Geographic Regions

Similarly, it may also be expected that users in well "wired" geographic areas will observe better frame rates than users in more technologically remote areas. Figure 15 depicts a CDF of frame rate for the users in our study, separated into 4 geographic regions. In this case, geographic region appears to more clearly differentiate streaming video performance than it did in the case of servers. Australia/New Zealand provides the worst frame rates for all ranges, with 75% of videos having fewer than 3 fps and less than 10% of videos having more than 15 fps. Clips played in Europe have the best frame rates up to 15 fps, with only 15% of videos having less than 3 fps and 25% of videos getting more than 15 fps. North America is slightly better than Asia up to 15 fps. Both Europe, North America and Asia all provide about the same percentage of videos with frame rates above 20 fps.

There have been concerns raised about streaming multimedia applications using non-TCP friendly congestion control or, worse, being unresponsive to network congestion [FF98, FHPW00]. Figure 16 shows a breakdown of the network transport protocols observed among all recorded video clips in our study. Over 1/2 of RealVideo flows use UDP, indeed suggesting non-TCP congestion control. Still, a surprising fraction of RealVideo flows, 44%, use TCP, and should be well-behaved in the presence of network congestion.

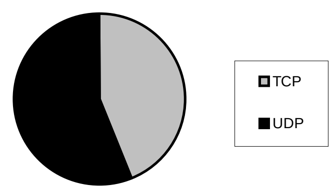

Figure 16. Fraction of Transport Protocols Observed

Figure 17 depicts a CDF of the frame rates observed for all the TCP and UDP flows. There is a slightly higher percentage of TCP flows with frame rates under 3 fps, about 28%, compared to UDP flows, about 22%. However, for the most part the frame rate distributions are nearly identical, suggesting that UDP, which provides applications with greater flexibility in controlling transmission rates and retransmissions than does TCP, does not necessarily provide better application frame rates than TCP.

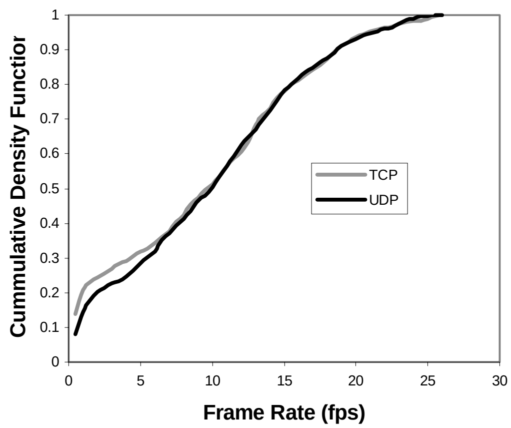

Figure 17. CDF of Frame Rate for Transport Protocols

Figure 18 depicts a CDF of the bandwidth observed for TCP and UDP flows. The bandwidth used by TCP and UDP data flows is very comparable, suggesting RealVideo uses application layer congestion control that is responsive to network congestion. However, for the most part, the UDP flows have a slightly higher bandwidth than do TCP flows, save for the very low bandwidth flows in which TCP flows have a slightly higher bandwidth than do UDP flows. This suggests that the application layer congestion control for the UDP flows may not be, in the strictest sense, TCP-friendly.

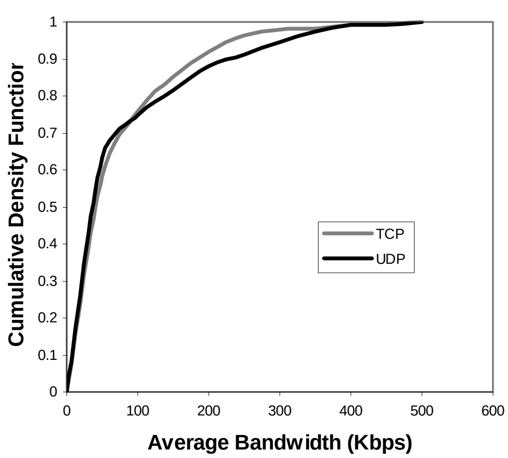

Figure 18. CDF of Bandwidth for Transport Protocols

Rather than the bottleneck being the network, the observed bottlenecks to streaming video performance could be in the end-user's PC itself. We combined the available memory with the CPU chip-type in an attempt to categorize the user PCs into different "power" classes. Figure 19 depicts a CDF of the frame rate observed for the different classes of PCs of the users. Clearly the slowest machines, older Pentium chips with limited memory, have the worst frame rate distributions. These slow machines provide frame rates above 3 fps only about 10%-20% of the time. For the other classes of machines the results are much less clear. Sometimes the seemingly more powerful machines provide lower frame rates and sometimes they provide higher frame rates. This suggests that except for very old generations of PCs, the PC itself is not the bottleneck to streaming video performance.

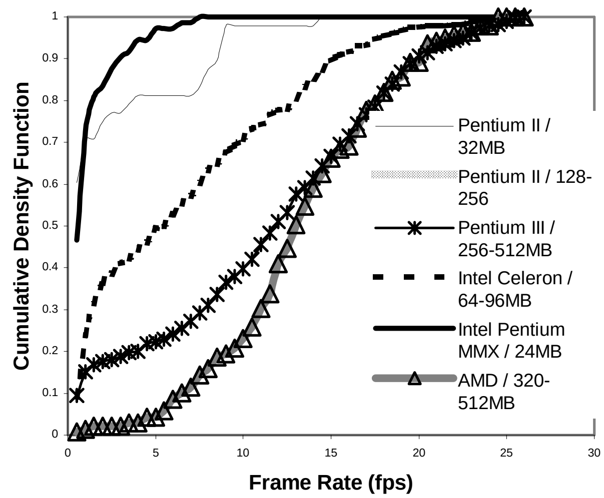

Figure 19. CDF of Frame Rate for Classes of User PCs

B. Jitter

Figure 20 depicts a CDF of jitter (standard deviation of inter-frame playout times) for all the video clips played. Just over 50% of all videos play with very little perceptible jitter. This high percentage of smooth videos is most likely due to the large initial buffer set by the RealPlayer core when the video connection is first made. Only about 15% of all videos play out with a potentially unacceptable 300 ms or more of jitter.

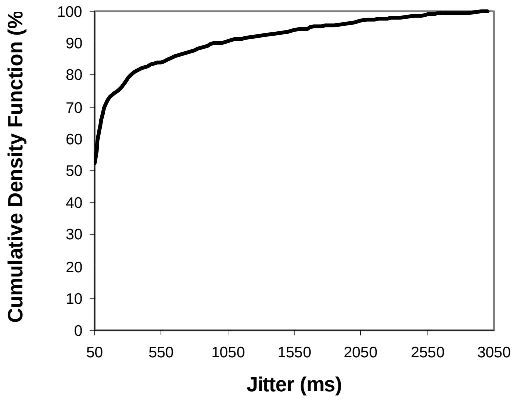

Figure 20. CDF of Overall Jitter

Our expectation is that the frame rate results for different end-host network configurations should hold for jitter, as well. We expect high-speed Internet connections to have less jitter than slower Internet connections. Figure 21 depicts a CDF of jitter for different end-user network configurations. From the graph, jitter in video played out over a modem is typically much greater than jitter over a higher-speed connections. Video clips played over a modem have no perceptible jitter only about 10% of the time and have potentially unacceptable jitter nearly 45% of the time. DSL/Cable modems and T1/LAN connections have a nearly identical percentage of perceptually jitter-free streams, while DSL/Cable modems also have a smaller percentage of potentially unacceptable amounts of jitter (15% vs. 20%, respectively). Overall, DSL/Cable modems have better jitter distributions, possibly because users contend with fewer other users for bandwidth, causing less variance in bandwidth than occurs on corporate LANs.

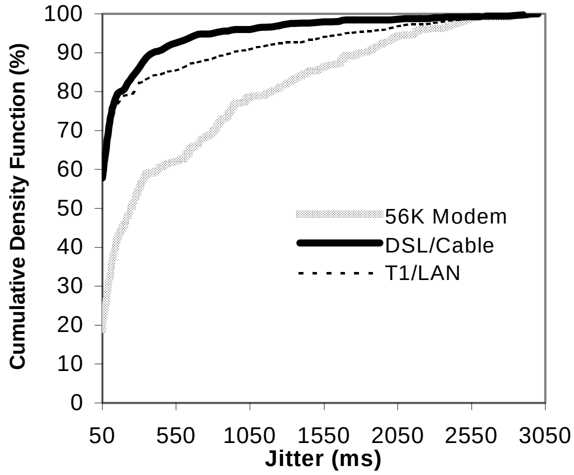

Figure 21. CDF of Jitter for Different Network Configurations

We next examine whether the results for video frame rates played from different servers hold for jitter, as well. Figure 22 depicts a CDF of jitter for the servers used in our study, separated into the same 5 different geographic regions used earlier. Asia provides clips with the most jitter, with only 45% of the clips having imperceptible jitter compared with about 55% for North American, Brazilian and Australian servers. This ranking is consistent with the frame rate raking in that Asian servers provided the worst frame rate distribution, too. Europe, on the other hand, had one of the the best frame rate distributions but has the second worst distribution of jitter overall. However, for the cutoff of imperceptible jitter at 50 ms and the potential upper bound of jitter at 300 ms, all servers, except Asia, are comparable.

Similarly, we examined observed jitter from our previous 4 different geographic user regions. Figure 23 depicts a CDF of jitter for the users in our study. As in the frame rate analysis, geographic region appears to clearly differentiate streaming video performance. Australia/New Zealand again provides the worst performance over both the imperceptible and tolerable limits of jitter. Asia provides the next worst performance, and video clips played out in Europe and North America have comparable jitter.

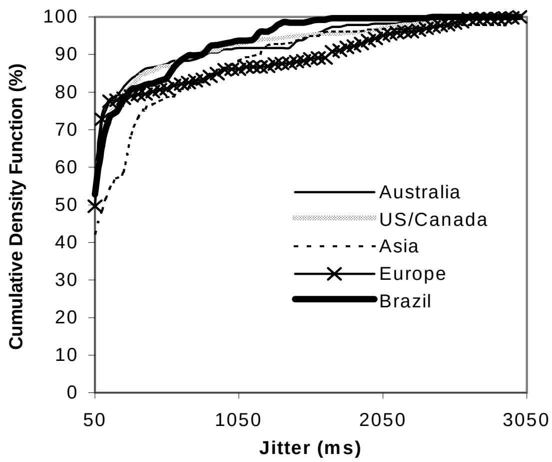

Figure 22. CDF of Jitter for RealServers in Different Geographic Regions

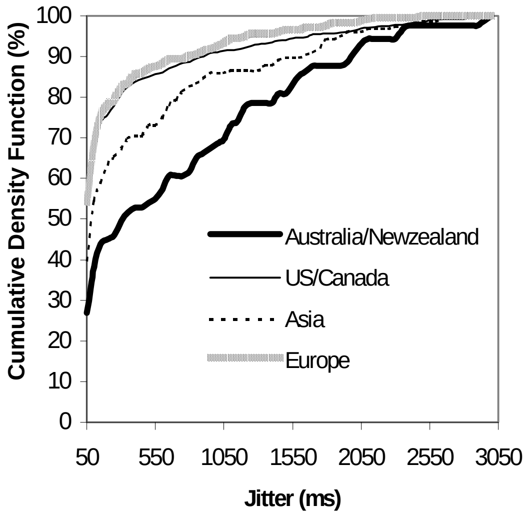

Figure 23. CDF of Jitter for Users in Different Geographic Regions

Figure 24 depicts a CDF of jitter observed for all the TCP and UDP flows. Similar to the frame rate CDF for the different protocols, both UDP and TCP provide nearly identical smoothness of video playout.

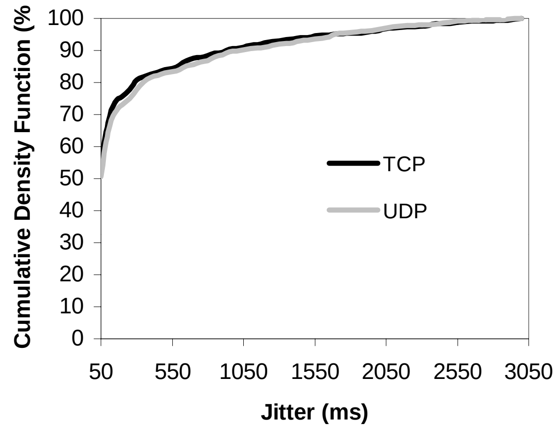

Figure 24. CDF of Jitter for Transport Protocols

Figure 25 depicts a CDF of jitter for different network bandwidths recorded. There is strong correlation between the bandwidth in the connection and the jitter in the video playout. Low bandwidth connections play out jitter free videos only 10% of the time compared with the 80% jitter free playout of high bandwidth connections. Only 20% of low bandwidth connections have an acceptable level of jitter, compared with nearly 95% of high bandwidth connections. This result of frame-level jitter is consistent with our previous measurements of packet-level jitter for different end-host network bandwidths [CR99].

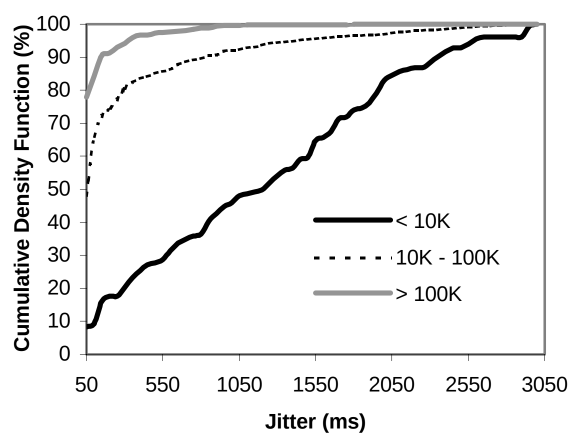

Figure 25. CDF of Jitter for Observed Bandwidth

C. Perceptual Quality

As discussed earlier, even measures of frame rate and jitter are not always sufficient to determine how a video will be perceived by the user, thus we analyze the perceptual quality ratings given by users for the video clips they watched and rated. Our purpose in this analysis is two-fold. First, as in previous sections, we wish to measure video performance across the Internet, only in this section we concentrate on a more user-centric measure of performance. Second, we wish to determine if there is a clear relationship between the system measurements of frame rate and jitter with perceptual quality. If we can discover such a relationship, we can perhaps develop an accurate mapping of system level measurements to user perceptual quality. We have only begun to analyze the perceptual quality scores captured in our study. For this reason, and due to space constraints, we only present our preliminary analysis.

We first analyze the perceptual quality of performance of all RealVideo clips across the Internet, depicted as a CDF in Figure 26. The mean perceptual quality rating is about 5 and the distribution line is very uniform. This suggests there may be a “normalization” of ratings that causes users to provide an average rating of 5 for the video clips they watch, regardless of the system conditions. This would also suggest that any mapping of system level measurements to perceptual quality may have to be developed on a per user basis.

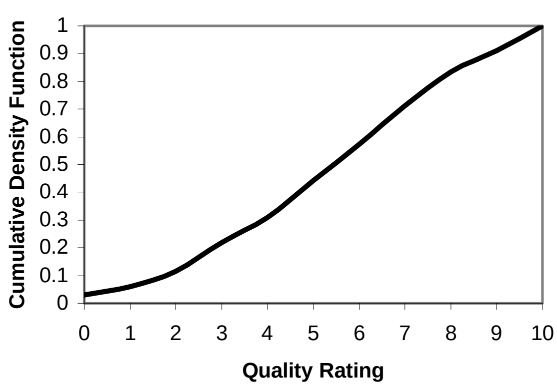

Figure 26. CDF of Overall Quality

The end-host network configuration has one of the biggest impacts on video frame rate and jitter. We expect the impact of end-host network configuration to hold for perceptual quality ratings, too. Figure 27 depicts a CDF of quality rating for different end-user network configurations. The end-host network has a large impact on perceptual quality. The average video watched over a modem is only about half as good in perceived quality as the average video watched on a DSL/Cable modem. DSL/Cable modems have better perceptual quality distributions than do LAN/T1 connections. This difference was not evident in the frame rate CDF for network configuration (Figure 12) but was evident in the jitter CDF for network configuration (Figure 21), suggesting that jitter is differentiating the video quality between the two configurations.

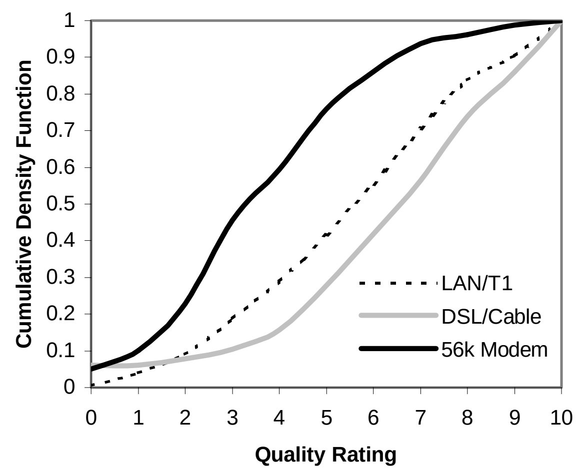

Figure 27. CDF of Quality for Different End-Host Network Configurations

We next briefly examine if there is a clear relationship between perceptual quality and systems measures of video perfomance. Overall, we have found there to be very little visual correlation between system measurements and configuration and perceptual quality when taken over all users. As an illustration, Figure 28 depicts a scatter plot of quality rating versus network bandwidth recorded for each clip. There is no strong visual correlation between quality and bandwidth, except most notably the lack of low quality ratings for videos that played out at a high network bandwidth. There also does appear to be a slight upward trend in quality ratings as network bandwidth increases.

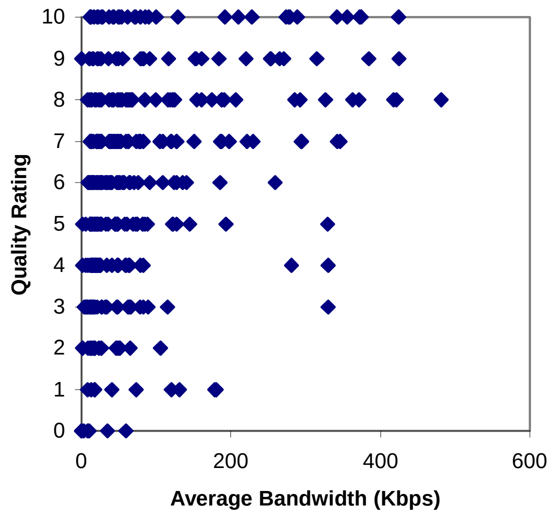

Figure 28. Quality Rating vs. Network Bandwidth

The surprising lack of correlation between quality and system measurements or configurations may be explained by the method we used to gather quality ratings. After doing the study, several users voiced concerns about the video rating criteria. Users who asked were uncertain if they were supposed to rate the video quality only or the audio and video quality together. During encoding, RealServer preserves the encoding rate of the audio and provides the remaining bandwidth to the video, suggesting those users that rated the video quality alone would have different results than those that used audio and video, especially for low bandwidth clips. This may explain some of the clustering in the upper left corner of the graph in Figure 28. Users were also uncertain of what rating to supply if a video clip started poorly but got better. Using a slider to gather instantaneous ratings, as recommended in [WS98], may help users rate the quality of video as it streams. Similarly, some users expressed confusion about how the subject matter should influence the rating, such as whether the video clip was artistic or supplied interesting information. Despite all such confusions, even users who expressed confusion stated they came up with a set of quality rating criteria of their own which they applied to the video, which implicitly happens to users watching video anyway. This suggests that there may be strong per user relationships between perceptual quality and system measurements that we can still discover. For now, we leave this as future work.

VI. Related Work

[MH00] presents the results of a brief study examining the traffic emanating from one popular Internet audio service using RealAudio. They found UDP to be the dominant download transport protocol, suggesting non-TCP congestion control. They observed consistent audio traffic packet sizes and rates that perhaps can be used for identifying flows or doing RealAudio simulations. We seek to build upon such work in measuring RealNetwork traffic by measuring RealVideo performance. In addition, instead of measuring only network flow characteristics, we focus on more user-centric methods of performance evaluation.

[KW00] details an extensive study carried out from many client sites geographically distributed around the world to a collection of about 700 servers to which a majority of Web traffic is directed. We conduct experiments similar in that we use geographically distributed clients and servers, but instead of Web traffic we use RealVideo traffic which has very different bandwidth requirements and quality of service constraints.

[CWVL01] presents and analyzes a week long trace of RTSP packets from the University of Washington. They analyze session length, session size and time of day correlations and the potential benefits from caching using their trace data and simulation. Instead of having clients at one location, we provide analysis of traces of clients from geographically diverse locations and concentrate on system impact and user-centric performance of RealVideo rather than general RTSP-based multimedia traffic.

[MCCS00] describes the *mmdump* tool for parsing typical multimedia control protocols. Although the emphasis of their work is on presenting the tool itself, in demonstrating mmdump's utility they present results from monitoring live RTSP and H.323 traffic on At&T's WorldNet IP network. Instead of clients from one ASP, we provide analysis from users across multiple ASPs, and focus on video performance for those users.

VII. Conclusion

The growing Internet and World Wide Web are fueling the growth of streaming high-quality video around the world. Previous empirical studies of wide-scale Internet performance have concentrated on general Internet traffic or have focused on Web-specific traffic. Previous empirical studies of wide-scale multimedia performance on the Internet have primarily looked at general multimedia traffic or have focused on audio. Empirical measurements of video performance on the Internet can provide insight into the impact of streaming video on the network, providing valuable information for research into the next generation of the Internet. In addition, empirical measurements of video performance on the Internet can provide insight into the bottlenecks to video performance, providing valuable inforamtion for research into next generation streaming video technology.

In this work, we present an empirical study of RealVideo, one of the most popular commercial video technologies, across the Internet. To gather data for our study, we built a customized video player called RealTracer that plays RealVideo from a series of pre-selected, geographically diverse servers. For each video played, RealTracer records user-centric video performance information, inlcuding frame rate, jitter and user ratings, and transmits the information to WPI for analysis. During a two-week period in June 2001, about 60 users ran RealTracer, playing about 2800 video clips from 11 servers world-wide and watching and rating the quality of about 400 of those same video clips.

From analysis of the data, we find that the average RealVideo clip streamed over the Internet has good quality, playing out at 10 fps (somewhat less than a very-good 15 fps) and, aided by a large, initial delay buffer, with nearly imperceptible amounts of inter-frame jitter. Users connecting to the Internet with modems and/or slow computers still have their PC or their network connection as the video performance bottleneck, while typical new computers connecting to the Internet via DSL or Cable modem achieve even slightly better performance than corporate network connections to the Internet. This suggests that increasing broadband connections for home users are pusing the bottlenecks for video performance closer to the server.

The RealTracer users came from 12 different countries and accessed servers in 8 different countries, providing data to compare video performance across geographically diverse parts of the Internet. We found there is very little difference in streaming video that is served from different countries, but there are distinct performance differences from video that is received in different countries.

Recently there have been concerns about possible network congestion collapse from streaming video that uses unresponsive or non-TCP-Friendly network protocols [FF98], but RealVideo itself does not appear to warrant these concerns. Nearly half of RealVideo flows use TCP to stream video data and the other half that uses UDP appears to respond to network congestion, but perhaps not quite in a TCP-friendly manner.

VIII. FUTURE WORK

We intentionally selected pre-recorded video clips to help ensure consistency in the videos played out by each user. Live content, captured and served directly from a video camera or television, has been shown to have different characteristics than does pre-recorded content [LH01]. Future work could be to measure the performance of live RealVideo content on the Internet and compare it to that of the pre-recorded RealVideo content in our study.

Our study had a notable absence of users from some countries in Asia and Europe and the West Coast of the United States. In addition, video clips were served from a diverse, yet limited set of countries. Future work could be to continue our current work and seek to broaden the data set of both users and servers. In doing so, however, we believe the data gathering techniques would need to move beyond requests for voluntary help, since we believe we nearly reached the limit of the number people willing to help and the number of clips they were willing to play and watch.

The major commercial competitor to RealNetwork's RealPlayer is Microsoft's MediaPlayer[6]. Developing similar tools to RealTracer for MediaPlayer, perhaps a *MediaTracer*, would enable an empirical study of more general video performance on the Internet.

The work in this paper did not explore the relationship in detail between frame rate and jitter with perceptual quality of video. Understanding these relationships may enable user-centric measures of performance that are easier to obtain and do not require users to rate videos. Currently, we are

[6] http://www.microsoft.com/windows/windowsmedia/en/default.asp

conducting ongoing work to carefully measure the impact of jitter on the perceptual quality of RealVideo as the first step in closely identifying this relationship.

NOTES

The RealTracer Web site[7] contains the latest version of RealTracer, the complete playlist used in this study and a list of the IP addresses of users that participated. It is our intent to release a customizable version of RealTracer, an accompanying analysis tool called *RealData*, and all the data we used in this study once we have completed our analysis.

We would like to thank all those who helped in debugging RealTracer duing our beta testing and those that ran RealTracer and rated videos. Without them this study would not have been possible.

REFERENCES

[CR99] M. Claypool and J. Riedl. The Effects of High-Speed Networks on Multimedia Jitter, In *Proceedings of SCS Euromedia Conference (COMTEC)*, Munich, Germany, April 25-28, 1999.

[CT99] M. Claypool and J. Tanner. The Effects of Jitter on the Perceptual Quality of Video, In *Proceedings of the ACM Multimedia Conference*, Vol. 2, Orlando, Florida, USA, November 1999.

[CWVL01] M. Chesire, A. Wolman, G. Voelker, and H. Levy. Measurement and Analysis of a Streaming Media Workload, In *Proceedings of the USENIX Symposium on Internet Technologies and Systems (USITS)*, San Francisco, CA, USA, March 2001.

[DP93] J. Devore and R. Peck, *Statistics - The Exploration and Analysis of Data*, 2nd edition, Wadsworth, Inc., 1993.

[FF98] S. Floyd and K. Fall. Promoting the Use of End-to-End Congestion Control in the Internet, *IEEE/ACM Transactions on Networking*, February 1998.

[FHPW00] S. Floyd, M. Handley, J. Padhye and J. Widmer. Equation-Based Congestion Control for Unicast Applications, In *Proceedings of ACM SIGCOMM Conference*, pp. 45-58, 2000.

[GB98] J. Gerek and W. Buchanan. MMLib - A Library for End-to-End Simulation of Multimedia over a WAN, *Major Qualifying Project MQP-MLC-MJ98*, Spring 1998. (Advisor Mark Claypool).

[KW00] B. Krishnamurthy and C. Wills. Analyzing Factors that Influence End-to-End Web Performance. In *Proceedings of the Ninth International World Wide Web Conference*, Amsterdam, Netherlands, May 2000.

[LC00] Y. Liu and M. Claypool. Using Redundancy to Repair Video Damaged by Network Data Loss, In *Proceedings of ACM Multimedia Computing and Networking (MMCN)*, San Jose, California, USA, January 2000.

[LH01] K. Lan and J. Heidemann. Understanding the structure of RealAudio traffic. (*Submitted for review*) http://www.isi.edu/~kclan/paper/180kl0202011550.ps

[Mah97] B. Mah. An Empirical Model of HTTP Network Traffic, In *Proceedings of ACM SIGCOMM Conference*, pp. 301-313, 1997.

[MCCS00] J. van der Merwe, R. Cáceres, Y-H. Chu, C. Sreenan. Mmdump - A Tool for Monitoring Internet Multimedia Traffic. *ACM Computer Communication Review*, 30(4), October 2000.

[MH00] A. Mena and J. Heidemann. An Empirical Study of Real Audio Traffic. In *Proceedings of the IEEE Infocom*, pp. 101-110. Tel-Aviv, Israel. March, 2000.

[Pax99] V. Paxson. End-to-end Internet Packet Dynamics, *IEEE/ACM Transactions on Networking*, Vol. 7, No. 3, pp 277-292, 1999.

[PHH98] C. Perkins, O. Hodson and V. Hardman. A Survey of Packet-Loss Recovery Techniques for Streaming Audio, *IEEE Network Magazine*, Sep/Oct, 1998.

[Rea00a] Real Networks Incorporated, RealProducer User's Guide, copyright 2000. http://www.service.real.com/help/library/guides/producerplus85/producer.htm

[Rea00b] Real Networks Incorporated, RealPlayer 8 User Manual, copyright 2000. http://service.real.com/help/player/plus_manual.8/rppmanual.htm

[TC01] A. Tripathi and M. Claypool. Improving Multimedia Streaming with Content-Aware Video Scaling, *Technical Report WPI-CS-TR-01-02*, CS Department, Worcester Polytechnic Institute, February 2001.

[TMW97] K. Thompson, G. Miller and R. Wilder. Wide-Area Interent Traffic Patterns and Characteristics, *IEEE Network*, November/December 1997.

[WS98] A. Watson and M. A. Sasse. Measuring Perceived Quality of Speech and Video in Multimedia Conferencing Applications, In *Proceedings of the ACM Multimedia Conference*, pp. 55-60, Bristol, UK, September 1998.

[7] http://perform.wpi.edu/real-tracer/