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 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 RealVideos to have high quality, achieving an average frame rate of 10 frames per second and very smooth playout, but very few videos achieve full-motion frame 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.**

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 access 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 [Pax99, TMW97], 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 empirical 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 has the opportunity to 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, packet 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 [GB98], 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

---

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

variety of geographically distributed Web servers to clients with many configuration parameters around the world. We developed a customized video player called *RealTracer* that plays streaming RealVideo 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 over 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 RealVideo videos on the Internet have good 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.

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 discusses our findings; Section 7 briefly describes some related work; Section 8 summarizes our conclusions; and Section 9 presents possible future work.

II. REALVIDEO 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 Internet 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 an Internet 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 then streams 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 the Progressive Networks Audio (PNA) protocol and, for backward compatibility, newer RealServers 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 protocol used is automatically 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 RealPlayer [Rea00b].

B. Buffering

For each video clip, RealPlayer keeps a buffer to smooth out the video stream in case 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.
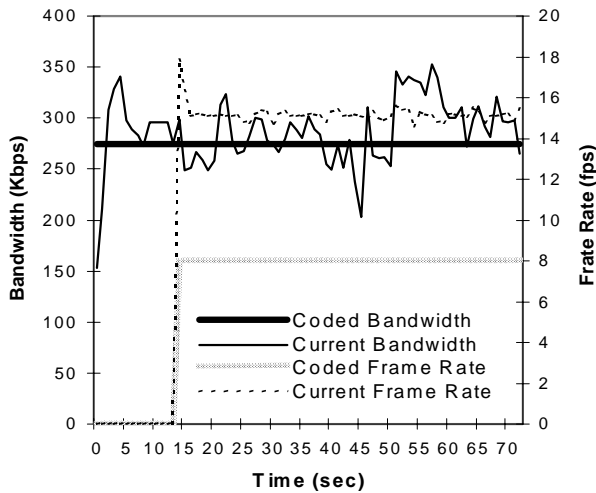

Figure 1. Buffering and Playout of a RealVideo Clip

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 axes represent 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. Bandwidth Characteristics

RealSystem uses a technology called *SureStream* in which one 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 clips are created with a *Scalable Video Technology* option that allows RealServer to automatically adjust the video stream according to the client's connection and computer processing speed [Rea00a]. If the client machine is unable to play the clip at the encoded frame rate, the frame rate will gradually reduce 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 (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 the 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).

---

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

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 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, CPU type and network configuration information from the user, as depicted in Figure 2 (top). The network configuration information is solicitied because it is difficult to automatcially detect, but the actual network settings used by the RealPlayer core are specified by the user during RealPlayer installation. RealTracer automatically detects Operating System type, available RAM and IP address.

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

While the video is playing, RealTracer uses the SDK to gather video statistics: encoded bandwidth, measured bandwidth, transport protocol, encoded frame rate, measured frame rate, playout jitter and frames dropped. RealTracer uses system calls to also gather the CPU utilization.

When each clip finishes playout, the user is solicited to assess the video quality by providing a numeric rating from 0-10 with labels only for the best and worst values, as depicted in Figure 2 (bottom). This unlabeled (save for the endpoints) scale avoids some of the pitfalls [WS96] associated with using the labeled scales specified in the International Tele-communications Union standard, a standard designed primarily for measuring television quality video. While slider tools such as QUASS [BS98] have been designed to support continuous user evalation for streaming media, the short duration of the video clips in our study limits the need for a changing evaluation,
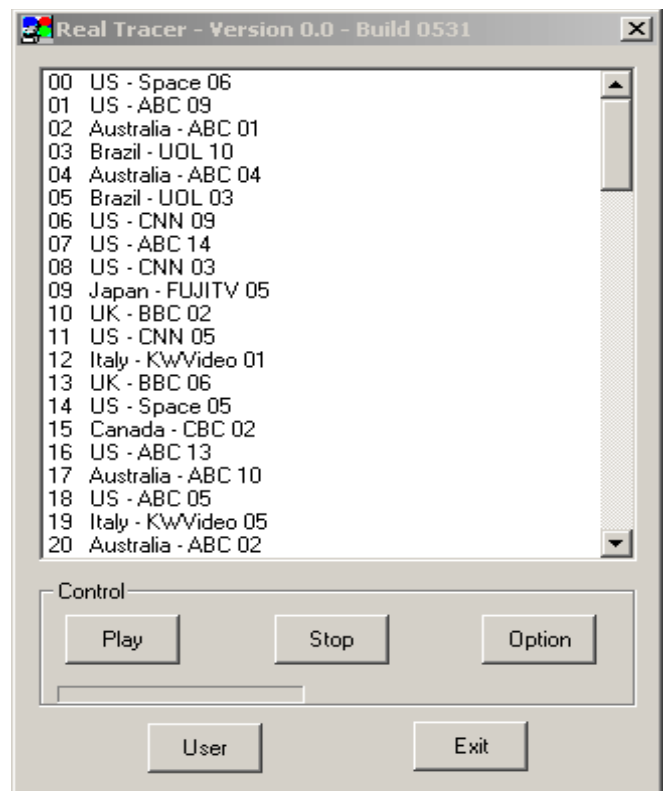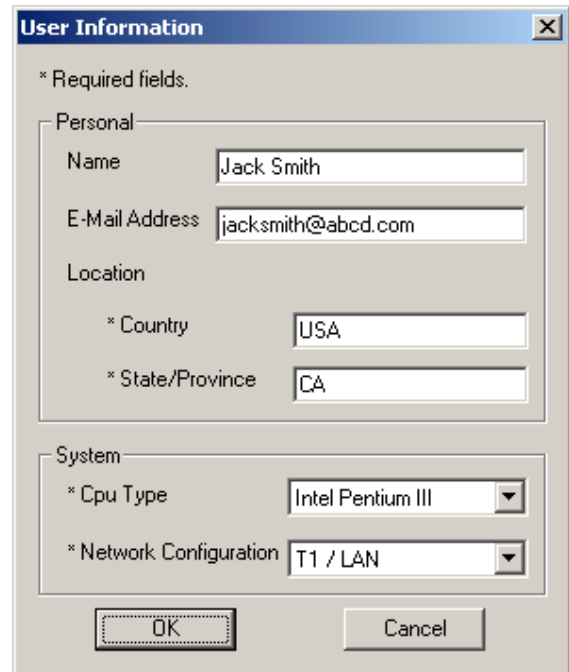

Figure 2. RealTracer Windows

as well as avoids the recency effect in which the last part of the video influences the perceived quality for the entire clip [WS98]. Furthermore, in our past work using 5-point scales for evaluation, users often wanted an "in-between" value, hence we chose a larger scale for this study.

For each clip played, the user data and specific clip statistics are sent via both email and FTP to a server at Worcester Polytechnic Institute. The default behavior for RealTracer is to then proceed sequentially through the playlist to the end. The videos in the playlist are the same for each user, but the playout order is shuffled across servers for consecutive videos.

If so desired, the user can control the length of the clip playout and the requests for quality ratings using the "Options" button. Performance statistics are collected over the length of the clip playout. 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, Australia, 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 4 depicts a geographic representation of the RealServer sites chosen. From each site, we selected 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 Massachusetts, we asked friends and colleagues on campus and at work to solicit help from people they

---

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

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

knew outside of Massachusetts. 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 3 depicts a geographic representation of the locations of users that ran RealTracer.
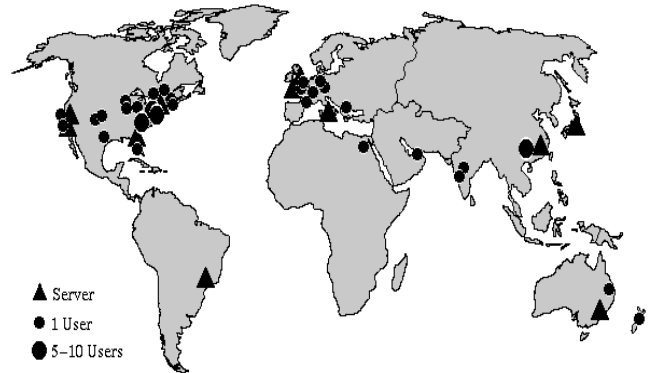

Figure 3. Geographic Depiction of RealServers and Users

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. Several users that tried to participate in the study were behind firewalls that did not allow RTSP packets through, so their data is not included in the analysis in this paper. Figure 4 depicts a Cumulative Density Function (CDF) of the number of 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 some of the 98 clips. Still, from Figure 4 it can be seen that the median user played out about 40 clips. Users were asked to rate 3-10 video clips or more, as their time and interest permitted.

Figure 5 depicts a Cumulative Density Function (CDF) of the clips rated per user. The median user provided ratings for 4 clips. About 10% of the users rated 15 or more clips, so analysis of the ratings (Section 5.3) may be largely dependent on the ratings given by this minority.
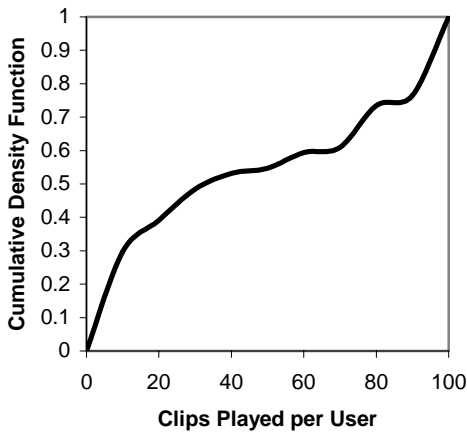
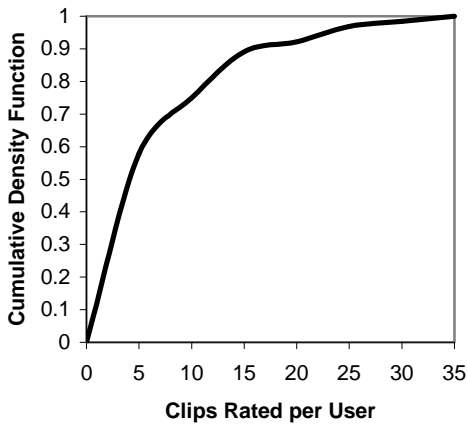**Figure 4. Cumulative Density Function (CDF) of Video Clips Played per User**

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

The users provided data points from 12 different countries. The small circles in Figure 3 depict the geographic location of users that participated in the study, while the large circles depict clusters of 5-10 users. Figure 6 depicts the breakdown of the total clips that were played by users from each country while Figure 7 depicts the breakdown of the total clips that were served by RealServers from each country. The servers may have employed load balancing, using, for example, DNS or Akamai. Determining when such balancing is done is difficult and we did not attempt to do so from RealTracer. However, from the URLs chosen, it appeared that only the video clips chosen from the CNN site were served by Akamai servers.

Figure 8 depicts a breakdown of the U.S. users per state. It is apparent that 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 that graph here. In subsequent results, for completeness we analyze all data gathered, unless otherwise noted.
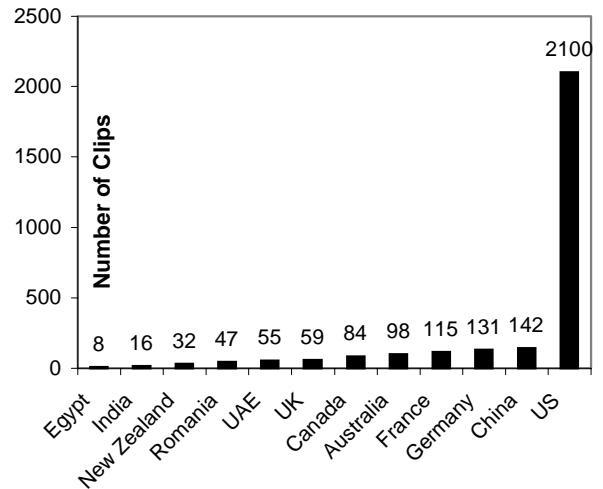

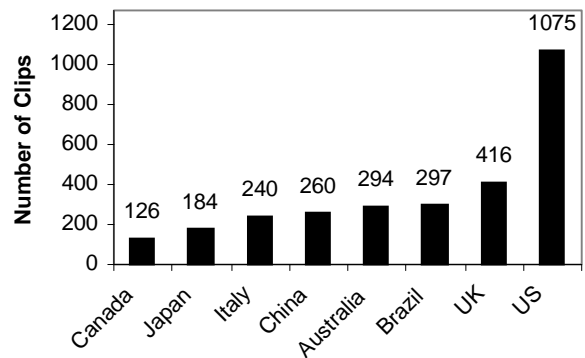**Figure 6. Video Clips Played by Users from Each Country**

Figure 7. Video Clips Served by RealServers from Each Country
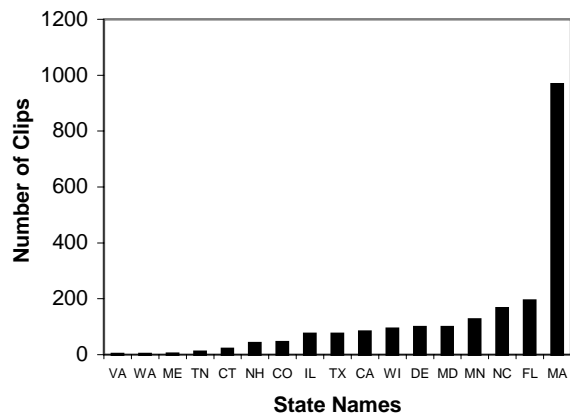
Figure 8. Video Clips Played by U.S. Users from Each State

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 in California, U.S.A. Future work would suggest trying to obtain 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 9 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 of general RealVideo clip availability. On average about 10% of the time a video clip was unavailable. Since the data we gathered does not allow us to determine why some clips were not available at some times for some users, we leave that as future work.
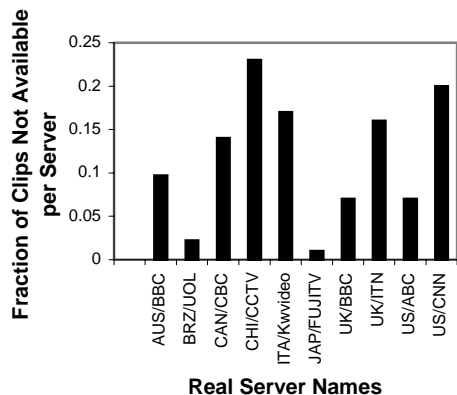

Figure 9. Fraction of Unavailable Clips

V. ANALYSIS

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

- 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). While there has been substantial literature on using delay buffering to ameliorate the effects of jitter [RKTS94, SJ95] and some literature on the effects of jitter on perceived audio-video synchronization [Ste96], to the best of our knowledge, there have been no quantitative measures of the impact of standard deviation of frame rate on perceptual quality. Thus, we build upon our previous experience [CT99] and place jitter measurements into three categories:

- Since human detection for synchronization between audio and video is around 100 ms [Ste96], we classify videos having jitter of 50 ms or less[5] as smooth.

---

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

- We classify videos having jitter between 50 ms and 300 ms as average.

- In our experience, jitter events that are larger than the average inter-frame playout are most noticeable by users, so we classify 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 more as rough.

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 high in high-action scenes, and lowering the frame rate 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 ratings (from 0-10, see Section 2.1) of 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 frame rate performance of RealVideo clips across the Internet in general. Figure 10 shows a CDF of the frame rate at the receiver for all the video clips played. The median 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, 30 fps.

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, while DSL and Cable modems can stream at rates up to 500 Kbps. Figure 11 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 from the end-user connection.
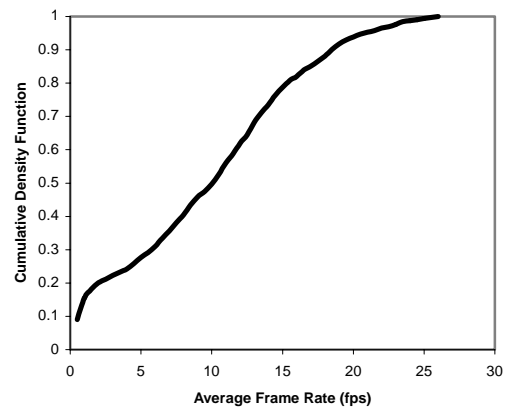

Figure 10. CDF of Frame Rate for all Video Clips

We further test this hypothesis by examining the bandwidth, as reported by the SDK, achieved by each class of end-host network configuration, depicted in Figure 12. 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 11 with Figure 12, it can be seen that modem connections get a proportionally higher frame rate for their network bandwidth than do higher-speed connections.
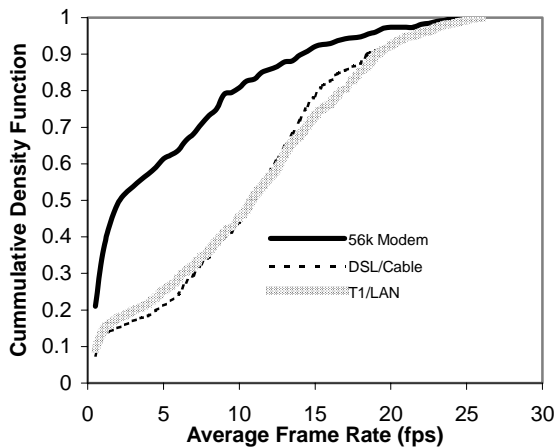
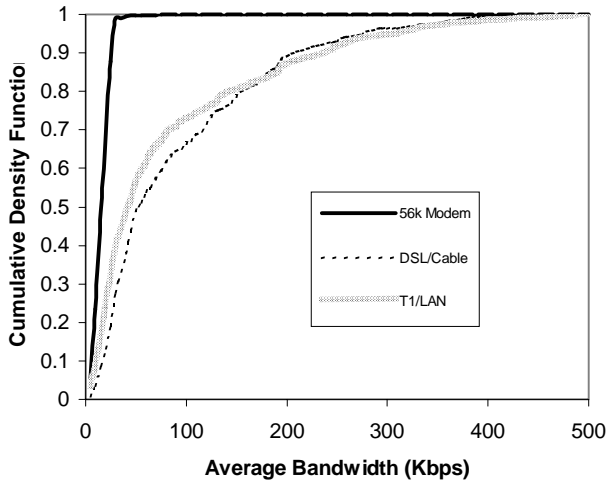**Figure 11. CDF of Frame Rate for Different End-Host Network Configurations**

Figure 12. 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 other locations, say Brazil. Figure 13 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 distribution shapes, although the median of the best frame rate distribution is about 13 fps and the median of the worst frame rate distribution is about 8 fps. Australia and Europe have the best frame rate distributions, with Europe having the best frame rates for the top 20% of the distribution. Asia provides the worst frame rates, but the differences at very low frame rates is fairly small with about 30% of Asia servers

having frame rates below 3 frames per second compared with about 20% for the best servers, and Asia servers actually have a larger percentage of frame rates above 15 fps than do North America servers.

While the peering richness of a client's ISP may largely determine video performance, this is difficult to measure. Thus, we consider geographic region, which is easy to measure, in place of client ISP information. Similarly to servers, 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 14 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 the 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, 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 the 15 fps rate. Both Europe, North America and Asia all provide about the same percentage of videos with frame rates above 20 fps.
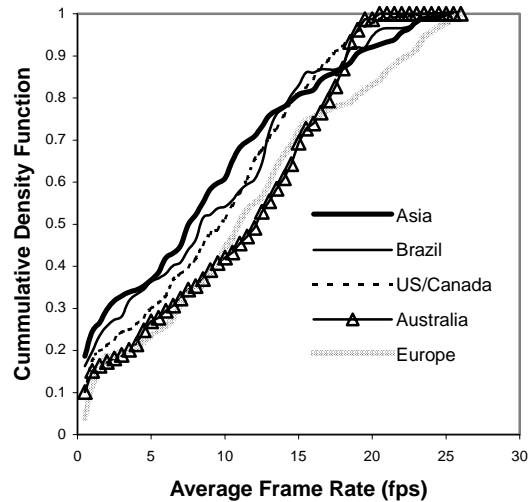

Figure 13. CDF of Frame Rate for RealServers in Different Geographic Regions

There have been concerns raised about streaming multimedia applications using non-TCP friendly congestion control or, worse, being unresponsive to network congestion [FF98]. Figure 15 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 unresponsive traffic or at least 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.
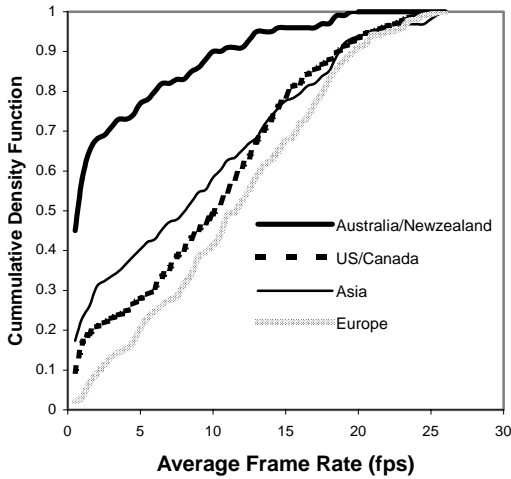

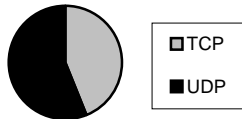**Figure 14. CDF of Frame Rate for Users in Different Geographic Regions**

Figure 15. Fraction of Transport Protocols Observed

Rather than the bottleneck to streaming video performance being the network, the bottleneck 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 16 depicts a CDF of the frame rate observed for the different classes of PCs of the users. Clearly the slowest machines, with older Pentium chips and 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.
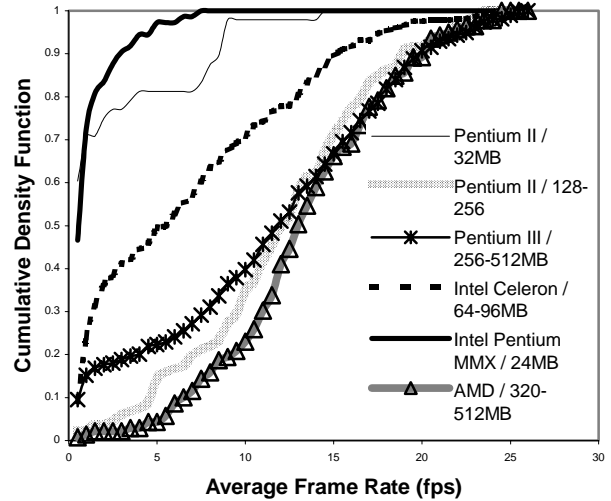

Figure 16. CDF of Frame Rate for Classes of User PCs

B. Jitter

Figure 17 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.
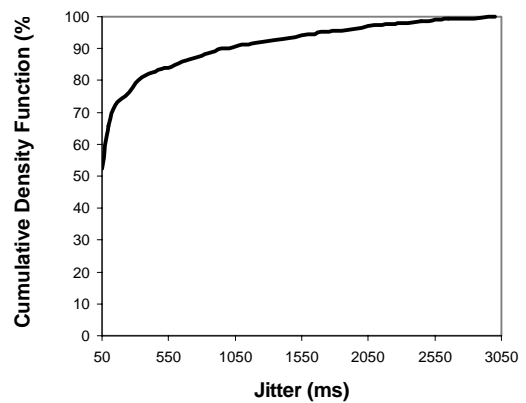

Figure 17. CDF of Overall Jitter

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 18 depicts a CDF of jitter for different end-user network configurations. From the figure, jitter in video played out over a modem is

typically much greater than jitter over a higher-speed connection. 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.
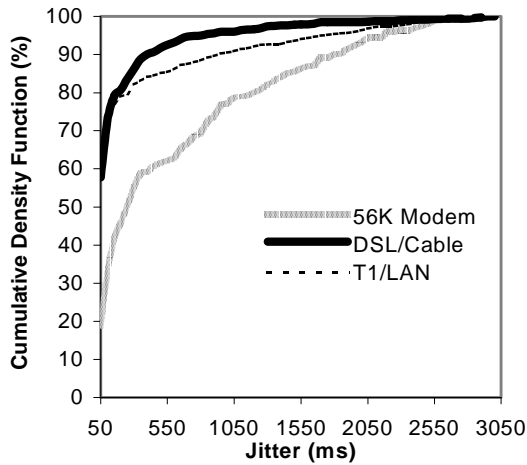

Figure 18. CDF of Jitter for Different Network Configurations

We next examine whether the results for video frame rates played from different servers hold for jitter, as well. Figure 19 depicts a CDF of jitter for the servers used in our study, separated into the same 5 different geographic regions used earlier. Asia servers provide clips with the most jitter, with only 45% of the clips having imperceptible jitter compared with about 55% for North America, Brazil and Australia servers. This ranking is consistent with the frame rate ranking in that Asia 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 jitter distribution 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 20 depicts a CDF of jitter for the users in our study. As in the frame rate analysis, user 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.
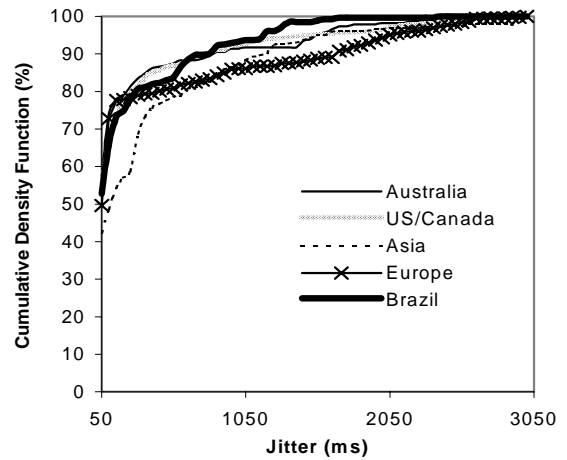

Figure 19. CDF of Jitter for RealServers in Different Geographic Regions

Figure 21 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].
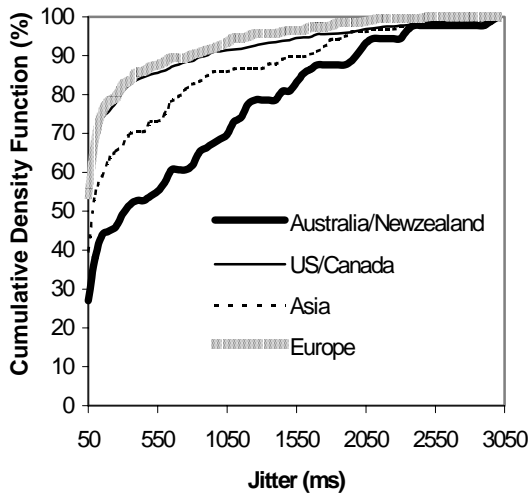
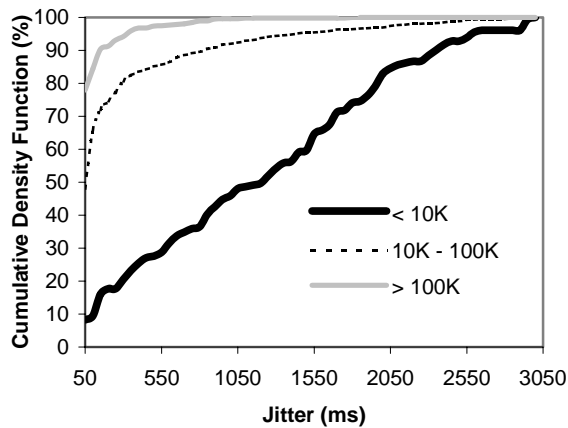**Figure 20. CDF of Jitter for Users in Different Geographic Regions**

Figure 21. CDF of Jitter for Observed Bandwidth

C. Perceptual Quality

As discussed earlier, 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 note that we have only a small number (388) of samples compared with the number of independent variables and considerably more data would need to be collected before results are conclusive or even statistically valid. In addition, user perception of video is based on many factors: image size, frame rate image quality, audio quality, language, content, etc. While the ratings we gather are designed to let users indicate video quality using their personal criteria, we note that during testing, some users expressed some confusion on what the rating is to be based. Thus, the brief results we present in this section should be viewed as preliminary only.

We first analyze the perceptual quality of performance of all RealVideo clips across the Internet, depicted as a CDF in Figure 22. The median perceptual quality rating is about 5 and the distribution is very uniform across all quality rating values. 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 mappings of system level measurements to perceptual quality may have to be developed on a per user basis.

In Sections 5.2 and 5.3, we saw that 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 23 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 11) but was evident in the jitter CDF for network configuration (Figure 18), suggesting that jitter is differentiating the video quality between the two configurations.
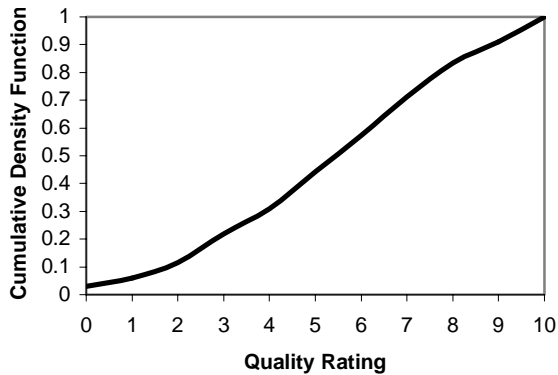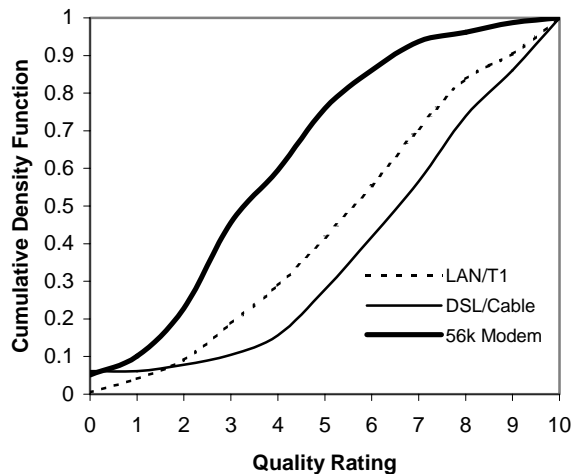
Figure 22. CDF of Overall Quality

Figure 23. CDF of Quality for Different End-Host Network Configurations

VI. DISCUSSION OF RESULTS

In this work, the video clip playout data is sliced along a number of independent variables: client geographic location, server geographic location, network configuration and PC type. In addition, there are at least several independent variables not analyzed: choice of user playback length, time of day, video clip encoding bandwidth, etc. This large number of variables suggests that additional measurements may help with confidence in the results. It is intended that the methodology used in this work be described in sufficient detail to allow the experiments to be reproduced. However, in repeating the work, the volatile nature of Internet and specific distribution of users that voluntarily participated in the study may provide somewhat different results. Thus, results in this work may be viewed as a first

step in a larger body of needed measurements on Internet video performance.

Despite this, the results presented here should be useful to network practitioners that want insight into the impact from the growing volume of video traffic, and to video practitioners that want insight into expected performance of video they provide or intend to provide. Network researchers should be able to use the results to produce more realistic video traffic for popular simulators such as NS[6], and streaming video developers should be able to use the results to target specific bottlenecks to performance.

VII. 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 analyzed 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.

---

[6] http://www.isi.edu/nsnam/ns/

[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 ISP, we provide analysis from users across multiple ISPs, 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 examined 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, including 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 rate) 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 newer 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 pushing 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.

VIII. FUTURE WORK

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 as used in this work, 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 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 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.

Past personal experience has suggested some perceived time of day correlations with both frame rate and jitter for RealVideo streams. However, doing time of day correlations with the data in this paper is difficult because both the servers and the users cover many different time zones and our preliminary

analysis shows only slight correlations between GMT time for servers and users and frame rate and jitter. Future work might include a more detailed study of time of day effects and video performance.

The major commercial competitor to RealNetwork's RealPlayer is Microsoft's MediaPlayer[7]. Developing similar tools to RealTracer for Media Player, perhaps a *MediaTracer*, would enable an empirical study of more general video performance on the Internet.

The video clips in this study are all pre-recorded. Moreover, RealVideo is not designed for interactive video, such as for a videoconference. Future work would include devising experiments to emprically measure the performance of live and interactive video across the Internet.

NOTES AND ACKNOWLEDGEMENTS

The RealTracer Web site[8] 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 during our beta testing and those that ran RealTracer and rated videos during our study. Without them this work would not have been possible. We would also like to thank the IWM'01 program committee and reviewers for their comments on our work and especially Mark Allman for detailed discussions on a previous version of this paper.

REFERENCES

[BS98] A. Bouch and M. A. Sasse, QUASS – A Tool for Measuring the Subjective Quality of Real-Time Multimedia Audio and Video, In *the HCI'98 Conference Companion*, pp. 94-95, 1998.

[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 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.

[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*, May 1998. (Advisor Mark Claypool).

[Jup01] Jupiter Media Metrix. "Users of Media Player Applications Increased 33 Percent Since Last Year", *Press Release*, April 2001. http://www.jup.com/company/pressrelease-.jsp?doc=pr010403

[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, CA, USA, January 2000.

[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*, September/October 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

[RKTS94] R. Ramjee, J. Kurose, D. Towsley and H. Schulzrinne, Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks, In *Proceedings of IEEE Computer and Communications Societies on Networking for Global Communication*, Toronto, Ontario, Canada, 1994.

[SJ95] D. Stone and K. Jeffay, An Empirical Study of Delay Jitter Management Policies, *ACM Multimedia Systems*, Vol. 2, No. 6, pp. 267-279, January 1995.

[SRL98] H. Schulzrinne, A. Rao and R. Lanphier. Real Time Streaming Protocol (RTSP), *Internet Request for Comments RFC 2326*, April 1998.

[Ste96] R. Steinmetz, Human Perception of Jitter and Media Synchronization, *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 1, January 1996.

[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.

[WS96] A. Watson and M. A. Sasse, Evaluating Audio and Video Quality in Low-Cost Multimedia Conferencing Systems, *Interacting with Computers*, Vol. 8, No. 3, pp. 255-275.

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

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