Experiments with the Tenet Real-Time Protocol Suite on the Sequoia 2000 Wide Area Network

Anindo Banerjea, Edward W. Knightly, Fred L. Templin, and Hui Zhang
EECS Department, University of California at Berkeley
and
International Computer Science Institute

Abstract

Emerging distributed multimedia applications have stringent performance requirements in terms of bandwidth, delay, delay-jitter, and loss rate. The Tenet real-time protocol suite provides the services and mechanisms for delivering such performance guarantees, even during periods of high network load and congestion. The protocols achieve this by using resource management, connection admission control, and appropriate packet service disciplines inside the network. The Sequoia 2000 network employs the Tenet Protocol Suite at each of its hosts and routers making it one of the first wide area packet-switched networks to provide end-to-end per-connection performance guarantees. This paper presents experiments of the Tenet protocols on the Sequoia 2000 network including measurements of the performance of the protocols, the service received by real multimedia applications using the protocols, and comparisons with the service received by applications that use the Internet protocols (UDP/IP). We conclude that the Tenet protocols successfully protect the real-time channels from other traffic in the network, including other real-time channels, and continue to meet the performance guarantees, even when the network is highly loaded.

1 Introduction

High speed networks are enabling the coexistence of computer data traffic with new multimedia applications. The widespread use of video and audio tools over the Internet has demonstrated the potential advantages of integrated services packet switched networks [2]. Currently, most of these applications are based on the Internet Protocol Suite [13, 12, 11], which adopts a connection-less architecture in the network layer. In the current Internet, IP routers do not discriminate between packets from different conversations, thus IP can only provide a best-effort service. The network is vulnerable to congestion and the performance of each individual application is sensitive to the network load.

In order to build robust, user-friendly multimedia applications, the network needs to offer a service that provides end-to-end performance guarantees on a per-connection basis. The Tenet Group at the University of California at Berkeley has designed a protocol suite that supports guaranteed-performance communication in a heterogeneous internetworking environment. Tenet protocol suite differs from the DARPA Internet suite in that it is based on a connection-oriented and reservation-based architecture. The protocols have been implemented in a number of hardware and software platforms and are running on a variety of networks including the Sequoia 2000 wide area network. Several multimedia application programs have been built on top of the protocol suite.

In this paper, a measurement study evaluating the performance of the Tenet protocol suite is presented. A previous study [19] evaluated the performance of the data-delivery protocols in a local area network using synthetic workloads. This study overcomes many limitations of the prior work by extending the scenario to a wide area network, by using real multimedia applications to generate the workload, by measuring end-to-end throughput and packet inter-arrival characteristics, by measuring the *qualitative* effect of the improvement in service provided to the end users, and by measuring the overhead associated with channel establishment and admission control. Our conclusion is that the Tenet protocols protect the real-time connections, both from best effort traffic and other real-time connections, and continue to provide the performance required by the applications, even under high network utilization conditions.

In Section 2 we provide some background to this study by describing the Tenet schemes and the Sequoia 2000 network. In Section 3, we describe the motivation and methodology of the experiments. Next, in Section 4, we present measurements of the performance of the data delivery protocols (RMTP/RTIP) and the resource reservation and signaling protocol (RCAP). We also measure the performance of a video-conferencing application (vic) on our protocols, and compare it to the performance of the same application on the Internet (UDP/IP) protocols. In addition, we measure the effect of the performance improvement on the real users by taking an opinion poll. Finally, in the remainder of the paper, we discuss the implications of the measurements and conclude.

2 Background

This section provides the background information necessary to set the context for our experiments. We first describe the Tenet protocols and show how they allow the network to provide real-time services to the application. Next, we describe the Sequoia 2000 wide area network testbed along with its research environment and applications.

2.1 Tenet Protocol Suite

The Tenet protocol suite provides *real-time* or guaranteed performance communication services in an internetworking environment [7]. The protocol suite adopts a connetion-oriented and reservation-based architecture. The basic abstraction is the *real-time channel*, which defines communication services with guaranteed traffic and performance parameters in a packet-switched network [6]. A channel's traffic is characterized by peak rate, average rate, averaging interval and packet size information. The possible performance parameters include end-to-end packet delay, delay-jitter, buffer overflow probability and delay violation probability.

The primary components to providing these guarantees are admission control, priority scheduling in switches and routers, and a connection-oriented paradigm. With admission control, access to the network by real-time traffic is limited so that each connection receives its required service from the network (see [6, 18, 20]). The priority scheduling ensures that the promised service is actually delivered. While rate controllers provide the policing necessary to ensure that different real-time connections do not affect each other's service, priority queues provide the mechanisms necessary to provide different services to different connections (see [21]). Finally, the connection-oriented paradigm provides an efficient means to providing performance guarantees to each packet of a connection by explicitly reserving resources for individual connections.

In the Tenet scheme, a channel is *established* before data transfer. To establish a real-time channel, a real-time client first specifies its traffic characteristics and performance requirements to the network. Next, based on this information, the network determines the most suitable route for this channel. Finally, the network translates the end-to-end parameters requested by the client into local parameters at each node, and attempts to reserves resources at these nodes accordingly. Channel establishment thus involves a round trip from the source to the destination in which sufficient resources are reserved on the forward part of the trip. At the destination, the client's end-to-end requirements are compared with the aggregate performance bounds offered by the nodes along the route and the connection is appropriately accepted or denied. If accepted, on the reverse part of the trip each node's resources are relaxed to the minimum level so that the local performance bounds may be guaranteed.

To determine the availability of resources, each node uses the Tenet algorithms or admissions tests, which are based on the service discipline. Research by Ferrari [4] has shown that a broad spectrum of service disciplines may be used in the Tenet framework. In the routers of the Sequoia network, we have deployed Earliest Due Date (EDD) scheduling [9] and their corresponding admission tests [6].

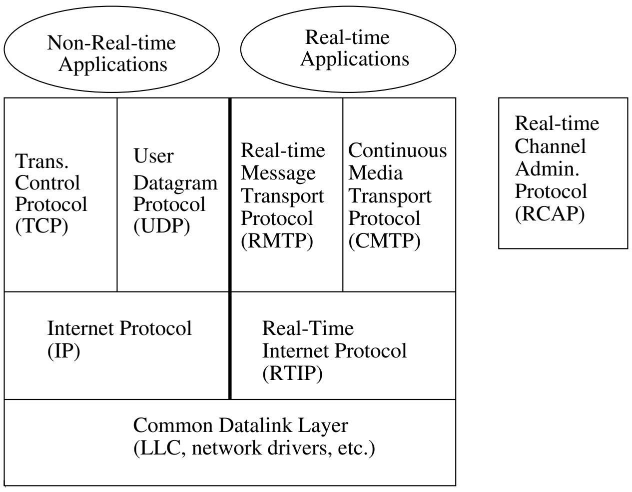


Figure 1: Tenet Real-Time Protocol Suite

The Tenet schemes are implemented in the Tenet protocol suite shown in Figure 1. The protocols are divided into data delivery and control protocols. The data delivery protocols consist of the Real-Time Internet Protocol (RTIP) [15, 22, 17] at the network layer, and the Real-time Message Transport Protocol (RMTP) [15, 17] and the Continuous Media Transport Protocol (CMTP) [16] at the transport layer. The control protocol is called the Real-time Channel Administration Protocol (RCAP) [1], and performs the tasks of channel establishment, tear-down, and status reporting. As shown in the figure, the Internet protocols, which are used to provide a best-effort service for non-real-time traffic, coexist with the Tenet protocols.

2.2 The Sequoia 2000 Network

The Sequoia 2000 network [10] provides the communications infrastructure for global change researchers and computer scientists involved in the Sequoia 2000 Project [14]. Sequoia scientists require networks which support real-time scientific visualization and video conferencing applications as well as high-speed data delivery services for the massive data sets characterizing global change applications [5]. To satisfy these requirements, Sequoia researchers are investigating methods for providing both real-time and best-effort services for voice, video, and data delivery on the Sequoia network. Currently, the Internet protocols [13, 11, 12] are used for best-effort data delivery, while the Tenet protocols are under investigation as a possible solution for real-time requirements.

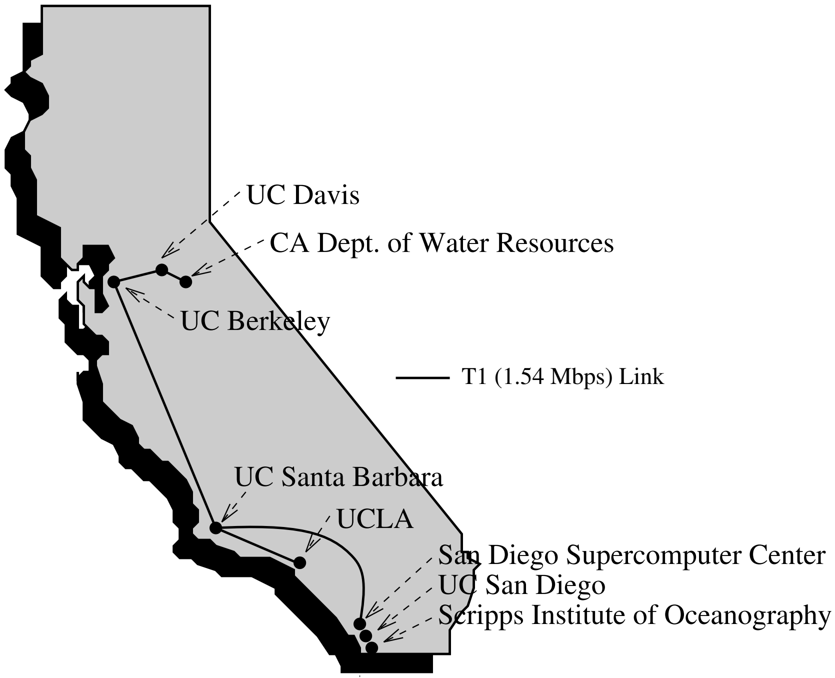


Figure 2: Sequoia 2000 Network Topology

The topology of the Sequoia network is shown in Figure 2. Its infrastructure consists of FDDI rings for local distribution with private T1 leased lines for wide-area services. DECstation 5000/240 general-purpose workstations interconnect the FDDI and T1 links and serve as network routers. These routers, along with many of the scientific workstations located at the individual Sequoia research sites, run a customized version of the ULTRIX^{TM} operating system which includes an implementation of the Tenet protocols. As shown in the figure, the Sequoia network provides service to the California Department of Water Resources, UC Davis, UC Berkeley, UC Santa Barbara, UCLA, the San Diego Supercomputer Center, UC San Diego, and Scripps Institute of Oceanography. T1 links interconnect all sites except those in San Diego, which are connected via a number of inter-campus FDDI rings.

2.3 The Sequoia Research Environment

Sequoia researchers at the sites shown in Figure 2 use scientific workstations in their studies of global change. Typically, global change researchers use their workstations to load, browse, and query objects such as satellite weather maps and global climate modeling data in the *Bigfoot* database [14] located at UC Berkeley using the Sequoia network as the transport medium. Additionally, many Sequoia workstations now support network transmission of digitally-encoded audio and video streams. In particular, several DECstation and Alpha workstations connected to the Sequoia network employ DEC's J-Video and J300 hardware compression/decompression cards which also provide live video and audio capture features. Cameras, speakers, and microphones attached to these cards provide audio and video data for multimedia applications. The Sequoia network is thus a good example of an integrated services network which supports these different types of traffic with vastly different characteristics and requirements.

3 Motivation and Methodology

The Sequoia network has become congested due to competitive workloads offered by scientific data loading and multimedia applications, leading to degradation in the audio and video quality observed by the users of the multimedia applications. Although the T1 wide-area links will soon be replaced with T3 (45 Mbps) links, we anticipate that the faster links will also shortly become saturated due to the rapid growth of the offered load on the Sequoia network. In the face of network congestion, Sequoia scientists require guaranteed quality of service for multimedia applications as well as best-effort services for scientific data loading applications. In [19], the Tenet protocols have been shown to address these needs in a local testbed. Thus, the motivation of this research is to study the behavior of the Tenet protocols on the Sequoia network as a possible alternative for the needs of Sequoia scientists. In our study, experiments take place on the actual Sequoia network using typical Sequoia multimedia applications and tools which emulate typical Sequoia scientific data loading operations. Our experiments are run on actual Sequoia workstations and span several distinct Sequoia research sites.

The performance measurements presented in this paper include packet queueing and forwarding performance in the kernels of the routers, throughput and packet inter-arrival measurements at the destination hosts, qualitative assessment of the video performance as perceived by end users with and without the Tenet protocols, and measurement of the latencies involved in channel establishment and admission control. The kernel measurements provide us with an understanding of the performance of our protocols at a single router in the network (chosen to be the congestion point). The measurements at the destination provide a quantitative evaluation of the end-to-end performance provided to a RTIP connection by the network, and allows us to compare this performance with that achieved by a UDP/IP session. The qualitative assessment allows us to judge the impact of the improved performance on the perceptual quality of a multimedia application. The measurements of the call establishment protocol allows us to allay fears about the impact of the admission control mechanism on the set up time required to start a multimedia session.

For the router measurement, we instrumented the kernel on the router. Timestamps were taken at various points inside the kernel for each arrival real-time packet, and stored in a circular buffer inside the kernel. A user level program periodically read the measurement data from the kernel and stored them on disk. In the measurement, we incorporated a technique developed by David Mills and enhanced the granularity of the DECStation 5000/240 system clock to 1 microsecond by taking advantage of an undocumented 25MHz hardware register.

To measure the end-to-end throughput and packet inter-arrival times we used `tcpdump` [8] to obtain packet traces at the destination host, and used post processing to separate the packet streams corresponding to the different streams. The qualitative assessment of the comparative video performance was based on an opinion poll.

To measure the performance of the control protocol (RCAP) we used the enhanced system clock to place microsecond accurate timestamps on the control packets themselves. We also instrumented the control protocol code to measure the delays on various code paths.

4 Measurement Experiments

As described in Section 2.1, the Tenet real-time protocol suite provides end-to-end per-connection performance guarantees in packet switched networks by adopting a connection-oriented architecture and using connection admission control, packet service disciplines hosts and routers. This section presents several different measurement experiments that demonstrate various aspects of the capabilities of the protocols.

Figure 3 shows the general scenario for the various experiments. As shown, the scenario involves the transport of live real-time video over the Sequoia Network using the program vic, a video conferencing tool developed by Steve McCanne of the University of California at Berkeley. vic uses DEC's J-Video hardware to do motion JPEG compression of the digitized live video input. At the receiver, vic sends the incoming video data to its local J-Video board for decompression and then displays the stream in a window of the receiver's workstation. vic allows the sender to choose a sending bandwidth by appropriately adjusting the quality factor, the frame size, and the frame rate of the transmitted video sequence. In the scenarios below, these values are appropriately limited (because of the link speed) to approximately 600 kbps and 10 frames per second. During the experiments, cross-traffic is introduced between a

UCLA workstation *blizzard.ucla* and a machine at UCB *grayling.berkeley*. This cross-traffic is representative of the load induced by researchers on the Sequoia testbed: the transfer of large images from remote mass storage devices to a local workstation for viewing and analysis. To stress the performance of the real-time connections, the cross-traffic was sent at 1.2 Mbps from the source. This traffic, in conjunction with the existing multimedia traffic, was enough to overload the T1 links (1.5 Mbps) of the network, and cause congestion. Some fraction of the cross-traffic successfully made it across to the destination, depending on the load on the network.

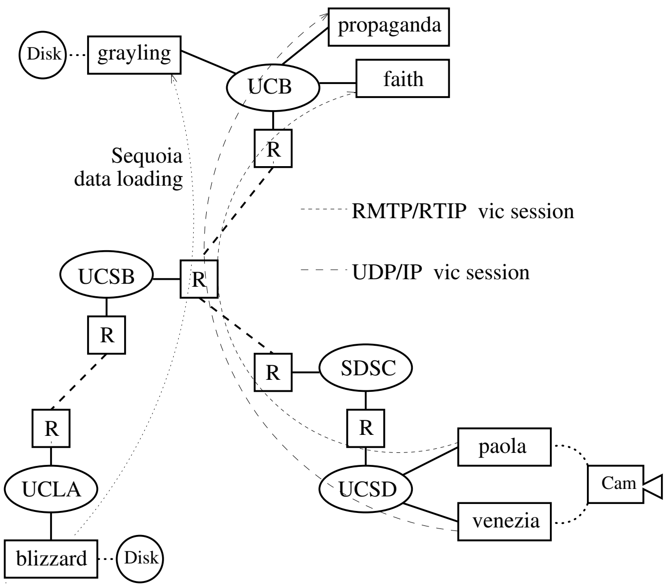


Figure 3: Measurement Scenario

The following experiments investigate the effect of the congestion in the multimedia traffic with the above scenario in progress. Various aspects of real-time data delivered by RMTP/RTIP or UDP/IP are explored. Clearly, when congestion occurs in the network, packets will be dropped due to buffer overflows, or delayed due to excessive queue sizes, unless special action is taken to protect the real-time connections from the effects of the congestion. Note that retransmissions are not feasible in real-time applications such as interactive video conferencing because of the excessive delay that this requires; in a real-time application, an excessively delayed packet must be considered lost. As well, real-time applications cannot in general be expected to throttle their rate because of conditions in the network. Specifically, many multimedia applications have some threshold below which they are unusable since the quality is so poor. Section 4.3 investigates the subjective aspects of the service provided by the network by analyzing the perceptual quality of video provided to end users. First we present quantitative measurements of the performance, in terms of the queue lengths and queueing delays at the bottleneck router, and the throughput and packet interarrival times at the destination host.

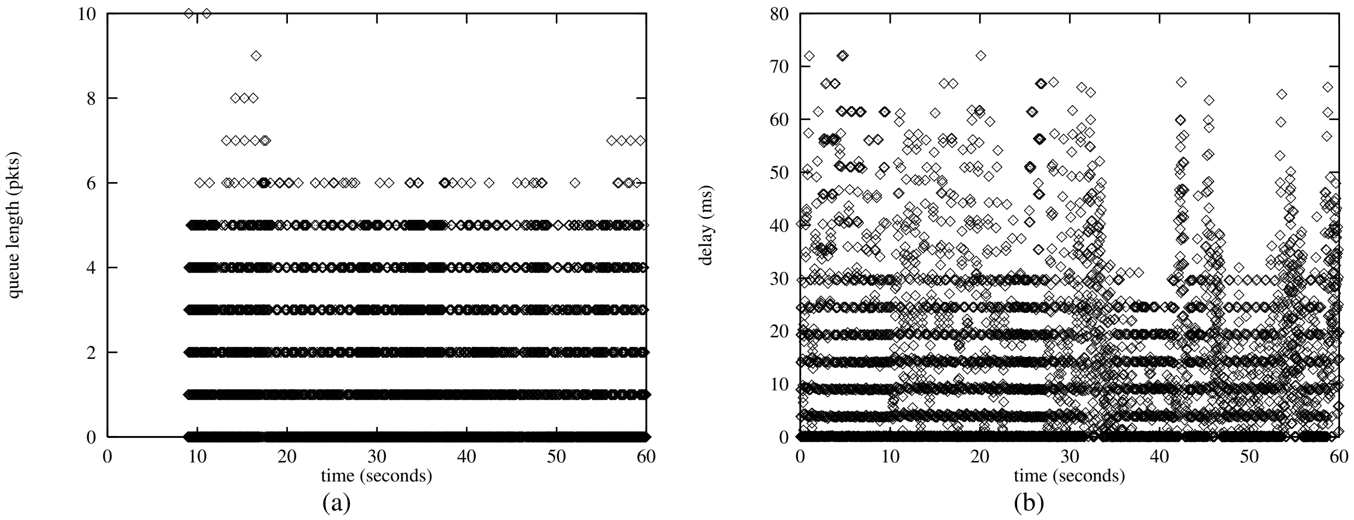


Figure 4: Queueing Time and Queue Length for an RTIP Channel

4.1 Experiment I - RMTP/RTIP vs. UDP/IP

This section investigates the performance of a video conference session using vic, transported over the RMTP/RTIP protocols. The connection was established from *faith.berkeley* to *paola.ucsd*. Measurements of the queueing behavior at the bottleneck router *sock.berkeley* are presented. Then the throughput and packet inter-arrival times as observed at the destination host (*paola.ucsd*) is compared to the performance of a vic session transmitted over UDP/IP from *propaganda.berkeley* to *paola.ucsd*.

4.1.1 Router Measurements

For the RMTP/RTIP vic session, Figures 4(a) and 4(b) illustrate the queueing behavior of the output link at the congested router. Figure 4(a) shows the queue *length* vs. packet arrival time for the measured RMTP/RTIP vic channel. Figure 4(b) shows the queueing *delay* vs. packet arrival time for the same channel. Since there is only one real-time channel traversing the router, and the service discipline is locally First-Come-First-Served for packets on the same real-time channel, the real-time packet queue length seen by an arriving real-time packet should correspond directly to the queueing time of the packet. (If there are multiple channels traversing the router, even if the real-time packet queue is short when a real-time packet arrives, it is possible that the packet experiences a longer delay due to the arrival of packets from other connections which have shorter delay bounds.) However, it can be seen from the figures that there are more packets experiencing longer queueing delays than packets arriving with longer real-time packet queues. The reason for this is that although real-time packets have higher priority than best-effort packets, this priority is non-preemptive. Thus if a real-time packet arrives when there is a best-effort packet being transmitted, the real-time packet must wait for the end of the transmission of the best-effort packet. In fact, most current network

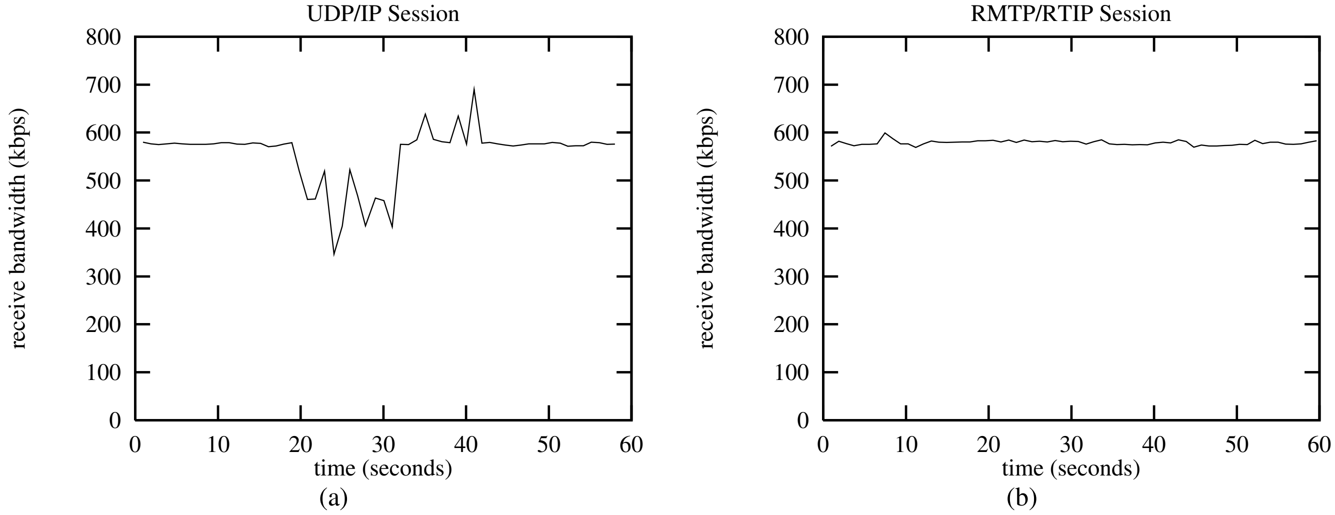


Figure 5: Receive Bandwidths for UDP/IP and RMTP/RTIP

interface cards, including the ones used in this experiment, are designed to hold multiple packets in the card's buffer, and newly arrival packets can only be transmitted when the interface empties its buffer. Thus, the worst case delay of a real-time packet depends not only on the real-time packet queue length, but also on the buffer size on the interface card. The admission control algorithms can take this into account by using a vacation model as described in [3, 17]. The scheduling policy successfully prevents the queueing delays from rising beyond the deadline, and also prevents buffer overflows, by ensuring that the real-time packets are served at the reserved rates. This can be seen from the graphs, where even at the most congested router, packet lengths for the real-time channel do not grow beyond 10, and the delay is always less than 80 ms.

4.1.2 Application Measurements

Figure 5 shows the effect of cross-traffic and network congestion on the received bandwidth of the two vic sessions. In the experiment, the data loading application is run between time $t \approx 20sec$ and $t \approx 30sec$. The vertical axes show the received bandwidth (averaged over 1 second intervals) in $kbps$ and the horizontal axes show $time$ in seconds. In both Figure 5(a) and 5(b) the received bandwidth is fairly smooth before $t = 20$. Although in general motion JPEG has higher peak-to-average rate ratios of two to three to one, this video is smoother for two reasons. Firstly, the camera is digitizing a person in an office. Without much action, the information in the scene does not change much and the frame sizes are smoother. This is not to say that there is *inter*-frame coding, merely that the information content in most frames of the office scene is fairly similar. Secondly, the plot shows the received bandwidth averaged over 1 second intervals which further smoothes the data.

At $t = 20$, Figures 5(a) and 5(b) diverge. At this time, the data-loading application from UCLA is contending for resources at the T1 link between UCSB and UCB. Since the aggregate data being transmitted by the applications

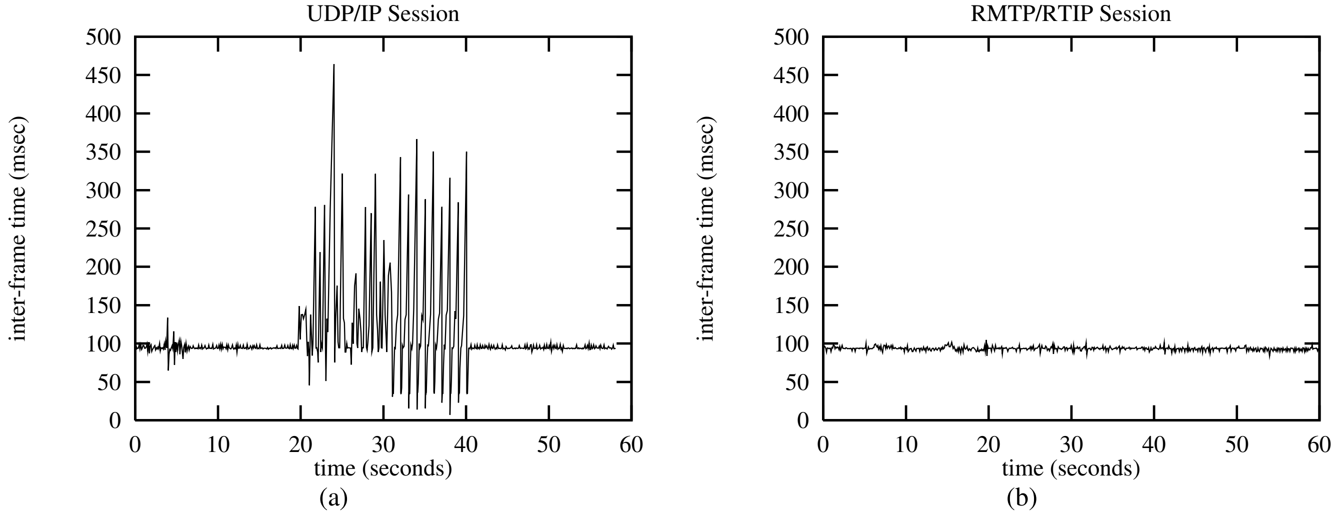


Figure 6: Receive Inter-frame Times for UDP/IP and RMTP/RTIP

is greater than the link speed, packets will be dropped due to buffer overflows. Figure 5(a) shows that the UDP/IP vic session suffers degradation due to the network congestion. Specifically, between $t = 20$ and $t = 30$ the UDP/IP vic session sees its average receive bandwidth drop 25% to 437 kbps from 580 kbps with a minimum of 310 kbps. Alternatively, the receive bandwidth of the RMTP/RTIP vic session in Figure 5(b) demonstrates that the RMTP/RTIP session was not at all affected by the congestion and buffer overflows at the UCSB router. The reason for this is that the Tenet protocols explicitly reserved network resources for the RMTP/RTIP connection. That is, even when the network routers are dropping packets because of excessive congestion and load, the RMTP/RTIP connection is unaffected since a certain portion of the bandwidth and buffers has been set aside for it.

Figures 6(a) and 6(b) further illustrate the performance of the protocols. In these figures, the horizontal axes represent time (in seconds) while the vertical axes represent the receive inter-frame time (in milliseconds), or the time difference between the arrival of two successive JPEG frames at the receiver. Since the senders are transmitting at 10 frames per second, variation from a receive inter-frame time of 100 msec is due to either delay introduced by the network due to queues at the routers or dropped frames due to buffer overflows at the routers. Figure 6 shows that when the network is unloaded ($t < 20$), both the UDP/IP and RMTP/RTIP vic sessions are received at a fairly constant rate of 10 frames per second. However, at $t = 20$ when the network becomes congested, the interframe time for the UDP/IP session becomes highly variable and reaches a third to almost a half a second. Further, for the UDP/IP session, even when the cross-traffic is eliminated at $t = 30$ seconds, the receive inter-frame time is still erratic for several more seconds. The reason for this is that the congestion is not alleviated instantly but rather queues will transiently be emptied. Finally, Figure 6(b) demonstrates that the RMTP/RTIP session was unaffected by the network load introduced between $t = 20$ and $t = 30$ since network resources have been explicitly set aside for this connection.

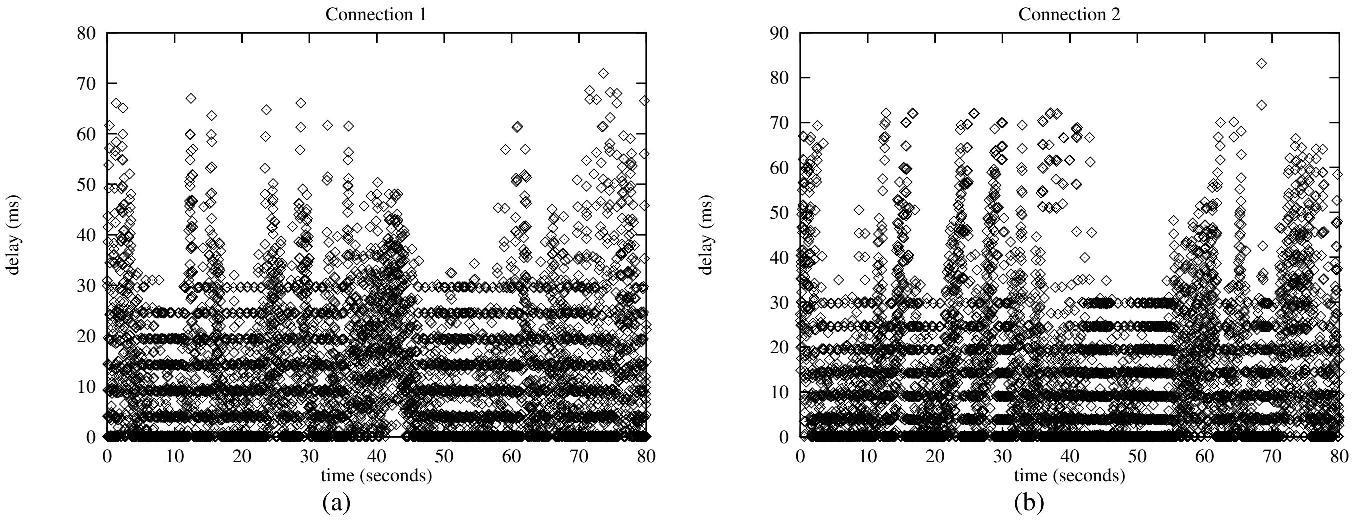


Figure 7: Queueing Time for Co-existing RTIP Channels

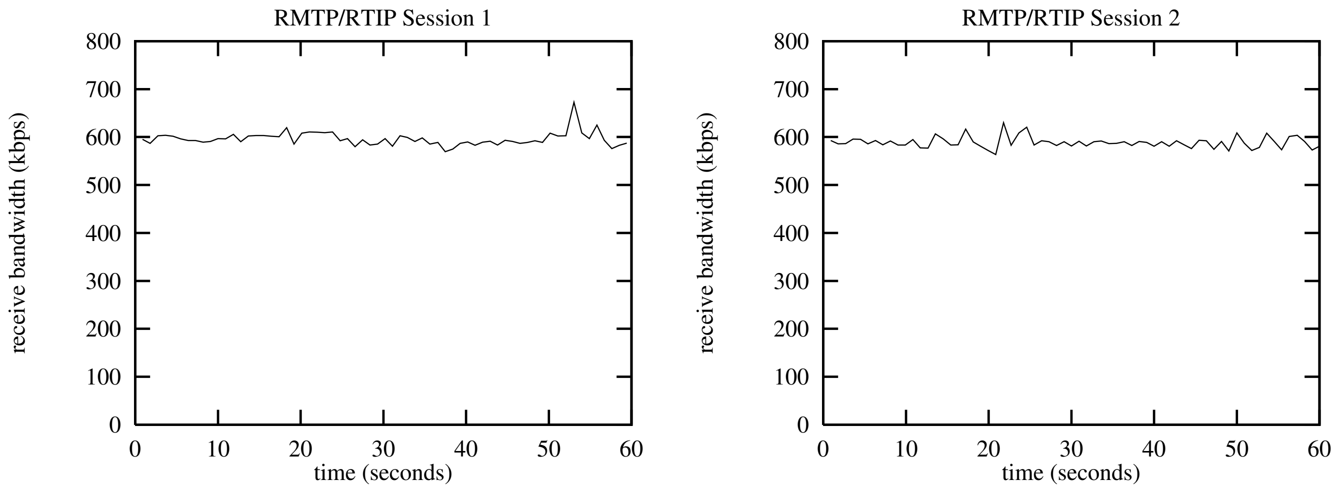
4.2 Experiment II - Two RMTP/RTIP Sessions

The previous experiment shows that the performance of RMTP/RTIP traffic was not affected by the presence of competing IP traffic. As well, it demonstrates that the UDP/IP vic session was severely degraded during periods of network congestion. This experiment investigates the simultaneous existence of multiple real-time channels with two established vic sessions, both using the Tenet Real-Time Protocol Suite. The two connections are established from *faith.berkeley* to *paola.ucsd* and from propaganda.berkeley to *paola.ucsd*. Competing IP traffic is again introduced to load the T1 link. Figures 7(a) and 7(b) show the queueing time vs packet arrival time graph for both channels. Compared to Figure 4(b), the co-existence of two real-time channels did not affect the performance of either of them thus illustrating the *protective* property of the Tenet protocols: RMTP/RTIP channels are protected not only from IP traffic, but other real-time traffic as well. Once a channel is accepted, its performance guarantees are met regardless behavior of other connections in the network.

Figure 8 shows the same effect from the perspective of the end application. The receive bandwidth for both RMTP/RTIP vic sessions are unaffected by each other as well as the IP data loading that occurs during the sessions.

4.3 Experiment III - Qualitative Investigations

The purpose of the Tenet protocol suite is to provide end-to-end performance guarantees to network clients that require such a service. The final evaluation of the effectiveness of the protocols to a multimedia application such as vic lies in actually viewing the video that is received from the network. Thus, a user's opinion of the video is critical. Although the figures of the previous section clearly demonstrate that the protocols are providing their promised service, it is an essential point that a user appreciate the difference between the quality of video transported by UDP/IP and video



(a) (b)

Figure 8: Receive Bandwidths for Co-existing RTIP Channels

transported by RMTP/RTIP.

There are two simultaneous vic sessions involved in the experiment: one is a UDP/IP session from the machine *venezia.ucsd* at UCSD to the machine *propaganda.berkeley* at UCB; the other is an RMTP/RTIP session from the machine *paola.ucsd* at UCSD to the machine *faith.berkeley* at UCB. Both sessions have an average rate of 580 kbps and 10 frames per second. This rate was purposely chosen so that both sessions can simultaneously exist on the T1 network (with a link speed of 1.54 Mbps) without adversely affecting each other. That is, when there is no congestion in the network and the two vic sessions have the capacity of the entire network to themselves, they are received by their respective destinations at their full sending frame rate (10 frames per second).

To address this problem of measuring the subjective perceptual quality of the video, we use the Mean Opinion Score (MOS). At the University of California at Berkeley's 1994 Industrial Liaison Program (ILP), over 50 visitors from industry and academia viewed a demonstration of the Tenet Protocol Suite over the Sequoia 2000 Network. The attendees were shown several scenarios of live video transmitted from UCSD to UCB and were asked to rate the video based on its perceptual quality. The different scenarios represented two different protocol suites, RMTP/RTIP and UDP/IP, transporting the video data under various levels of network load (as depicted in Figure 3). The attendees were not told which video was being transported by which protocol suite or what the different network loads were for the various scenarios.

The results of these opinion scores are as follows. With a 99% confidence interval, the opinion of the RMTP/RTIP video session did not change with network load. That is, when the network became congested due to the data loading application, the ILP attendee's opinion of the perceptual quality of the RMTP/RTIP video did not change. However, with the same confidence interval, the attendees' opinion of the UDP/IP video session *decreased* by a factor of 54% when load was introduced into the network. Thus, the results of the opinion score survey indicate that from a perceptual point of view, an RMTP/RTIP video session is impervious to network congestion while a UDP/IP video session severely degrades. Furthermore, these results are a qualitative confirmation of the data in the previous section indicating that use of the Tenet Protocol Suite is indeed significantly beneficial to the end user.

4.4 Experiment IV - Connection Setup Time

The control protocol used for channel establishment and resource reservation is called the Real-time Channel Administration Protocol (RCAP). RCAP is responsible for performing the admission control tests and maintaining the network state necessary for resource management. As described in Section 2.1, this is achieved by running admission control tests during the round-trip message exchange necessary for channel establishment. A major concern in such an approach is the speed of performing these admission control calculations and the effect this might have on the connection establishment time. To address this issue we measured the time to perform RCAP establishments both at the round-trip and the node-by-node level.

The time to perform an end-to-end round trip for channel establishment from UC Berkeley to UC San Diego, involving 6 machines and 5 links, is merely 80 to 90 ms, *independent* of the number of RTIP connections existing. To get a better idea of the break-up of the establishment latencies between link and node delays, we performed another experiment over a shorter path and put multiple time-stamps onto the RCAP packet. Space was allocated on the packet for four time-stamps per machine, and each machine time-stamped the packet once immediately after receiving it and once immediately before sending it, on both the forward and reverse path. Using these values, the delays incurred on the links and the nodes can be calculated. The results of a typical run are tabulated below. The node latency values are given separately for the forward and reverse directions, while the link delay presented is an average of the forward and reverse latencies.

Node	Forward (ms)	Reverse (ms)	Link(ms)
0	3.906	3.907	3.674
1	2.056	1.832	11.900
2	1.637	1.825	9.779
3	1.853	19.579	

Table 1: Node and Link Delays During Channel Establishment

Node 0 is a slower machine than the others so the higher processing times are understandable. On the faster machines (DECstation 5000/240) the processing times are between 1.5 to 2 ms. However the last value, on node 3 on the reverse pass, the delay is extremely high. This is a fairly consistent phenomenon, so we decided to investigate by printing out timestamps at several places in the code. We found that the major component of the delay was in a kernel call to set RTIP parameters. This delay (15-20 ms) was much larger than can be accounted for by the kernel boundary overhead and the processing required by the call. We postulate that the kernel uses this opportunity to preempt the RCAP process and schedule some other process. The time to actually perform the admission control tests is very small comparatively. The measured value was approximately 0.5 ms. The actual value is probably smaller since the time to execute `gettimeofday` and print the returned value to `stdout` is of the order of 0.4 ms, and this limits the accuracy of our measurements.

In conclusion, the time to perform establishment tests is dominated by the link delays and user-kernel overheads. The time to perform the actual admission control and resource reservation calculation is relatively small. The overheads could be reduced by moving the implementation of the signaling and control software into the kernel, but the current performance is more than satisfactory for our current environment.

5 Discussion

The above measurement scenarios show that with the appropriate queueing mechanisms in the routers, the RMTP/RTIP sessions are impervious to network congestion from the Internet protocols. The question arises, if there was a third RMTP/RTIP vic session, wouldn't the RMTP/RTIP connections suffer from dropped packets since their aggregate rate is greater than the link speed? The answer to this is that the third connection would not be allowed. That is, when the third vic tries to establish an RMTP/RTIP connection, RCAP, the signaling protocol, will block the call due to insufficient resources. Thus, admission control is a prerequisite to providing performance guarantees, the packet scheduling mechanism at routers alone is insufficient.

As well, one may ask, what would happen if the data-loading application uses RMTP/RTIP? First, the admission control algorithms would prevent the data-loading application from using any more than the remaining share of the network bandwidth. As well, if the data-loading application is accepted at (e.g.) 200 kbps (i.e. a real-time connection is established), the rate control mechanisms of RTIP will prevent the application from loading the network beyond the promised transmit rate of 200 kbps. Thus, rate control and policing are also necessary in providing performance guarantees in order to protect real-time connections from other, possibly malicious, real-time connections.

6 Conclusion

The Sequoia 2000 wide area network supports the coexistence of computer data and multimedia traffic. The requirements of the multimedia traffic make it necessary to provide network layer support for real-time communication. The Tenet protocol suite provides such support, enabling the multimedia applications to receive their required performance by protecting the connections from best-effort traffic as well as from other real-time connections. This allows the network to operate at high average utilization while still meeting the real-time requirements of the multimedia applications.

We presented measurements of the performance of the Tenet real-time protocols on the Sequoia network. The workload used in the experiments is generated by the Vic video-conferencing tool using motion JPEG hardware compression, and by tools which emulate typical Sequoia scientific data loading operations. The measurements presented are the queueing and packet forwarding delays at the bottleneck router, the throughput and packet inter-arrival traces at the destination, the qualitative assessment of the perceived video quality by end users, and latencies involved in admission control and channel establishment.

The measurements presented show that the queueing delays at the bottleneck router for the protected connections remains bounded even under extremely high data traffic load. The throughput and inter-arrival times remain smooth, as opposed to the performance of multimedia traffic carried by the UDP/IP protocol under the same network loading conditions. The Mean Opinion Score (MOS) of the participants in an opinion poll to evaluate the perceived quality of two video streams transported by the two protocol suites (UDP/IP vs. RMTP/RTIP) confirms that the quality of the UDP/IP stream degrades with network load, while the quality of the RMTP/RTIP stream is largely unaffected. Finally, we demonstrated that the latency involved in performing the admission control tests is negligible compared to the link delays and signaling overheads. The total latency of channel establishment is small (under 100 ms for a six node path).

In conclusion, the Tenet protocol suite has been demonstrated to provide guaranteed performance service to multimedia traffic, protecting the real-time channels from the effects of other traffic in the network, including other real-time channels. This protection allows the Tenet protocols to provide the multimedia application the required level of service, and guarantee this against all congestion conditions in the network. This goal has been met without introducing excessive overheads for call admission, or restricting the network to run at low utilization levels.

References

[1] A. Banerjea and B. Mah. The real-time channel administration protocol. In *Proceedings of the Second International Workshop on Network and Operating System Support for Digital Audio and Video*, pages 160–170, Heidelberg, Germany, November 1991. Springer-Verlag.

[2] Steve Casner and Steve Deering. First IETF internet audiocast. *ACM Computer Communication Review*, pages 92–97, 1992.

[3] Rene L. Cruz. A calculus for network delay, part I : Network elements in isolation. *IEEE Transaction of Information Theory*, 37(1):114–121, 1991.

[4] D. Ferrari. Client requirements for real-time communication services. *IEEE Communications Magazine*, 28(11):65–72, November 1990.

[5] D. Ferrari, J. Pasquale, and G. Polyzos. Network issues for Sequoia 2000. In *Proceedings of COMPCOM 92*, pages 401–406, San Francisco, CA, February 1992.

[6] D. Ferrari and D. Verma. A scheme for real-time channel establishment in wide-area networks. *IEEE Journal on Selected Areas in Communications*, 8(3):368–379, April 1990.

[7] Domenico Ferrari, Anindo Banerjea, and Hui Zhang. Network support for multimedia: a discussion of the Tenet approach. Technical Report TR-92-072, International Computer Science Institute, Berkeley, California, October 1992. Also to appear in *Computer Networks and ISDN Systems*.

[8] Van Jacobson, Craig Leres, and Steven McCanne. Tcpdump Man Page, January 1991.

[9] C.L. Liu and J.W. Layland. Scheduling algorithms for multiprogramming in a hard real-time environment. *Journal of ACM*, 20(1):46–61, January 1973.

[10] J. Pasquale, K. Fall, and J. Forrest. Sequoia 2000 network handbook. Technical Report 7, University of California at Berkeley, Berkeley, CA, June 1992.

[11] J. Postel. Internet protocol, September 1981. RFC 791.

[12] Jon Postel. User datagram protocol, August 1980. RFC 768.

[13] Jon Postel. Transmission control protocol, September 1981. RFC 793.

[14] M. Stonebraker. An overview of the Sequoia 2000 project. In *Proceedings of COMPCOM 92*, San Francisco, CA, February 1992.

[15] D. Verma and H. Zhang. Design documents for RMTP/RTIP, May 1991. unpublished internal technical report.

[16] Bernd Wolfinger and Mark Moran. A continuous media data transport service and protocol for real-time communication in high speed networks. In *Proceedings of the Second International Workshop on Network and Operating System Support for Digital Audio and Video*, pages 171–182, Heidelberg, Germany, November 1991. Springer-Verlag.

[17] H. Zhang. Service disciplines for integrated services packet-switching networks. PhD Dissertation. UCB/CSD-94-788, University of California at Berkeley, November 1993.

[18] H. Zhang and D. Ferrari. Improving utilization for deterministic service in multimedia communication. In *1994 International Conference on Multimedia Computing and Systems*, Boston, MA, May 1994.

[19] H. Zhang and T. Fisher. Preliminary measurement of RMTP/RTIP. In *Proceedings of the Third International Workshop on Network and Operating System Support for Digital Audio and Video*, San Diego, CA, November 1992.

[20] H. Zhang and E. Knightly. Providing end-to-end statistical performance guarantees with interval dependent stochastic models. In *ACM Sigmetrics'94*, Nashville, TN, May 1994.

[21] H. Zhang and K. Srinivasan. Comparison of rate-based service disciplines. In *Proceedings of ACM SIGCOMM'91*, pages 113–122, Zurich, Switzerland, September 1991.

[22] H. Zhang, D. Verma, and D. Ferrari. Design and implementation of the real-time internet protocol. In *Proceedings of the IEEE Workshop on the Architecture and Implementation of High Performance Communication Subsystems*, Tucson, Arizona, February 1992.