Evaluation of 10Base-T and 100Base-T Ethernets Carrying Video, Audio and Data Traffic *

İsmail Dalgıç William Chien† Fouad A. Tobagi

Department of Electrical Engineering
Stanford University
Stanford, CA 94305

Abstract

Multimedia applications integrate a variety of media, namely, audio, video, images, graphics, text, and data. These media have different bandwidth, delay, and loss requirements. Data traffic is predominantly bursty and it requires reliable service from the network, but it can tolerate delay. In contrast, audio and video traffic is stream-oriented, it can tolerate some loss of data, but it requires delivery within a bounded delay. To support multimedia communications, it is desirable to use existing network infrastructures, which are designed specifically for data applications. In such existing network infrastructures, Ethernet is the most popular LAN scheme used. In this work we have simulated 10Base-T and 100Base-T Ethernet segments carrying constant bit rate audio/video streams and data traffic (bursty and non-bursty). Given the stream rate, stream delay and loss constraints, as well as the data load and burst size, we have determined the number of streams supportable.

1 Introduction

Multimedia applications integrate a variety of media, namely, audio, video, images, graphics, text, and data. These media have different bandwidth, delay and loss requirements. Data traffic is predominantly bursty and it requires reliable service from the network, but it can tolerate delay. In contrast, audio and video traffic is stream-oriented, whereby a stream is a continuous flow of information that has to be delivered in a timely fashion; in audio-video traffic, some loss of data can be tolerated.

*This work was supported in part by NASA under grant NAGW-419, by NSF under grant NCR-9016032, and by Pacific Bell. İsmail Dalgıç was also supported by a graduate scholarship from TÜBİTAK/NATO between 1988 and 1990.

†Currently with C-Cube

The delay tolerance of an audio/video application depends on its interactiveness requirements. For example, one-way broadcasting can tolerate end-to-end delays upwards of one second. On the other hand, interactive applications such as video conferencing require end-to-end delays to be less than 100-200 ms. If a video stream travels multiple hops, the delay constraint per hop will be a fraction of the end-to-end delay requirement (e.g., 20 ms). Interactive multimedia applications can nevertheless communicate over networks that cannot guarantee bounded delay, owing to the ability of users to tolerate some loss. Packets that experience delays in excess of the delay bound are considered lost. Thus, the challenge to the network is to provide bounded delay and loss to stream traffic.

There exist a wide variety of data applications which generate traffic with different burstiness characteristics. For example, text editing, electronic mail, and remote logins are likely to generate sporadic, single packet data traffic. In contrast, file transfers and page swapping generate bursts of packets. As a result of the advances in today's personal computer and workstation speeds, optimization of protocol implementations, and more efficient network interfaces, stations are able to generate high peak loads and send packets back-to-back over the network. This results in a highly bursty data traffic. Accordingly, the burstiness of data traffic has to be taken into account in a model that can reflect today's environments.

To support multimedia communications, it is desirable to use existing network infrastructures, which are designed specifically for data applications. In existing network infrastructures, Ethernet [1] is the most popular LAN scheme used, owing to its simplicity and efficiency. However, it does not provide bounded delays, and it does not distinguish between different traffic types such as video and data. Nevertheless, for millions of Ethernet users, the alternative of changing to a new technology, and consequently replacing the wiring, hubs, and network interfaces would be very costly and would almost certainly delay the use of multimedia communications. In this work, the interest is in evaluating the ability of Ethernet to support multimedia traffic despite its shortcomings. The popular 10Base-T Ethernet standard specifies that stations are connected to a hub (i.e., multiport repeater) using unshielded twisted pair wires with a maximum cable length of 100 meters. Thus, in this study, we focus on such cable lengths.

In this work we also consider 100 Mb/s Ethernets, given the current efforts to incorporate a 100 Mb/s Ethernet standard (to be named 100Base-T) into the IEEE 802.3 [2][3]. In 100Base-T, the transmission speed will be increased to 100 Mb/s, while the 802.3 MAC protocol is retained. All the MAC protocol attributes will remain unchanged, with the exception of the interpacket gap, which is reduced from 9.2 μs to 0.92 μs (all the other timing specifications are given in number of bits, and thereby automatically scale with the transmission speed). In 100Base-T, stations are connected to a central hub using a maximum cable length of 100 meters, similar to 10Base-T.

The objective of this research is to evaluate the performance of 10 Mb/s and 100 Mb/s Ethernets for supporting multimedia traffic, in particular audio and video streams. Our primary performance measure is the number of streams that can be supported given the stream rate, stream delay and loss constraints, data traffic load and data burst size. We consider streams alone, as well as streams and data together, considering both bursty and non-bursty data traffic. We have chosen to achieve our objective via computer simulation, since the complexity of the problem makes it impossible to use accurate analytical techniques.

We are not the first to study the performance of CSMA networks for supporting audio or video traffic. Notable prior studies involving audio are [4] and [5]. We are aware of only two prior papers that study video on Ethernets [6] [7]. An important difference between our work and all the prior studies is our consideration of mixing streams with *bursty* data traffic. Also, we consider a wide range of stream rates and delay constraints representative of typical audio/video applications. Furthermore, we focus on the contemporary 10Base-T and 100Base-T networks. Finally, we choose not to consider any modifications to the 802.3 protocol, since any such modification is bound not to be applied given the wide deployment of Ethernets today.

In a future paper, we will also evaluate the performance of multiple Ethernet segments connected by a switching hub for supporting multimedia traffic.

Section 2 describes our models for the network and the traffic. Section 3 presents the results of this study and Section 4 discusses implications of this work.

2 Network and Traffic Models

2.1 Network Configuration and Protocols

Our model of an Ethernet segment consists of a number of stations connected to a central hub using twisted pair wires, as defined by 10Base-T and 100Base-T. We assume that a station generates either stream or data traffic, but not both. We use N_v to denote the number of stream sources, and N_d to denote the number of data sources. W denotes the network bandwidth (either 10 or 100 Mb/s).

In the simulator, the stream and data traffic sources generate messages that are passed to a simulated protocol stack. We do not model the application, presentation, and session layers. Furthermore, since the study is limited to single Ethernet segments, the network layer has no effect, and it is not modelled.

The functions of our transport layer are to perform packetization and de-packetization, as well as to provide buffering. The transport layer maintains a queue of packets waiting to be transmitted; this queue has infinite capacity. We do not account for transport protocol processing overhead, or header overheads. No retransmission or flow control schemes are modelled.

The MAC and LLC layer parameters are modelled according to the 10Base-T standard [1] and the 100Base-T proposals [2][3]. We ignore any MAC level processing overhead, as it is negligible.

At the physical layer, all stations are assumed to be 200 meters from each other. This corresponds to the maximum station-to-hub cable length of 100 meters in 10Base-T and 100Base-T Ethernets. Propagation delays are computed using 0.6 times the speed of light, as appropriate for twisted-pair wires.

2.2 Traffic Assumptions

2.2.1 Stream Traffic

In Figure 1, we depict the block diagram of the system. A frame is taken in the video camera, and sent as an analog signal into the frame grabber, where it is digitized. Then it is compressed by the encoder. The encoder is followed by a rate buffer whose purpose is to smooth out the variations in the encoder's output rate, producing a constant bit-rate stream. The buffer occupancy is used by the encoder as feedback to control the encoder output rate (and hence quality) so that the rate buffer does not overflow or underflow. The constant bit rate stream is passed from the rate buffer onto the main memory through the system bus. In order to be sent over the network, streams are packetized by the sending station. The packets are sent over the network to the destination station. The destination station buffers the received packets to compensate for the delay variation due to the network. The contents of the buffer are passed to the decoder, which decompresses the stream and delivers it to the display.

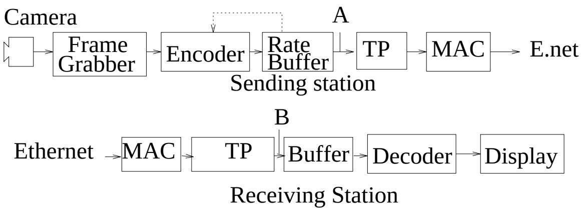


Figure 1: Block diagram of the system

We consider that the streams are encoded according to a coding standard suitable for interactive communications, such as H.261 [8]. In H.261, a frame is composed of a number of Group Of Blocks, or GOBs (depending on the number of pixels in a frame, it is divided into either 3 or 12 GOBs). Each GOB is in turn divided into 33 macroblocks. A macroblock contains information for an area of 16×16 pixels. In H.261 it is the smallest unit that can be encoded/decoded without any future information. Encoding schemes other than H.261 use similar hierarchical structures as well.

There are several hypotheses one can make regarding the operation of the CODEC. In the following, we describe these hypotheses, and state our assumptions:

i) Video CODECs can generate either variable bit rate, or constant bit rate streams. In this study, we only consider constant bit rate CODECs. We have studied streams with the following rates: V=64 kb/s for PCM voice and low quality video conferencing, V=384 kb/s for medium quality video conferencing, and V=1536 kb/s for high quality video conferencing.

ii) Encoders may operate on a frame-by-frame, GOB-by-GOB, or macroblock-by-macroblock basis. We assume that they operate on a macroblock-by-macroblock basis so as to minimize the encoding latency. We also assume that the macroblocks are delivered by the encoder at regular intervals (i.e., the encoding time is constant for each macroblock).

iii) Decoders may wait until they receive an entire GOB or a frame before they start the decoding process, or they may start decoding a macroblock as soon as they receive it. We assume the latter, so as to minimize the decoding latency.

We assume that streams are packetized by the transport layer using a constant packet size (denoted by P_v). Accordingly, the stream packets are generated at regularly spaced intervals. Note that the Ethernet packet sizes are large enough to accommodate a number of macroblocks for the stream rates that we consider (e.g., macroblock sizes for V=1536 kb/s range from 10 bits to about 150 bytes, with an average of about 16 bytes).

As we mentioned before, streams require delivery within bounded delays. As a simplification, we only consider the delay between points A and B as shown in Figure 1. For a given packet i in the simulation, we denote this delay by D_i. D_i is the sum of two components, the packet formation time, T_f, and the network delay, D_i^n. The packet formation time is the delay experienced by the first macroblock in a packet while the rest of the packet is being formed (i.e., $T_f = P_v/V$). The network delay is the time elapsed from when the packet has been formed, until it completes transmission. We denote the maximum delay that stream traffic can tolerate by D_{max} (again, measured between points A and B). We have chosen D_{max} values of 20 ms and 100 ms.

Since streams can tolerate some loss, the stream delay is bounded by discarding those packets that exceed the delay bound. Note that packet loss is due to two reasons: (i) exceeding the delay bound D_{max}, (ii) exceeding the maximum number of collisions as determined by the MAC protocol. (A packet is discarded after it collides 16 times, as per 802.3.) We assume that the loss tolerance of a stream can roughly be represented by a maximum tolerable packet loss rate, L_{max}. The relationship between packet loss rate and quality is not straightforward, as it depends on the stream syntax and rate. In this study, we consider maximum tolerable packet loss rates of 0.01 and 0.001, chosen somewhat arbitrarily. We are currently in the process of investigating the effects of packet loss on video quality; some of our preliminary results are given in section 3.3.

Even though we do not consider the delays before point A or after point B in our study, in Figure 2, we provide the timing diagram for the end-to-end delay, and in the following, we discuss the ranges of values that the delays can take in each component of the system:

Camera: Between the camera and the frame grabber, there is no delay, since the frame grabber digitizes the analog signal as it comes.

Frame Grabber: Let the delay in the frame grabber be D_g. Even if the frame grabber delivers the digitized bytes with zero delay, the encoder has to wait at least until the first macroblock is fully received before it can start encoding. Therefore, the minimum value of D_g is 16 lines' time (1.8 ms for CIF, 3.7 ms for QCIF). On the other hand, the frame grabber may wait until it digitizes a whole frame, and then pass it to the encoder, in which case, D_g=33 ms.
Encoder: Let the encoding delay be D_e. If the encoder operates on a frame-by-frame basis, then $D_e \leq F$, where F is the time between successive frames. If it operates on a macroblock-by-macroblock basis, then $D_e \leq F/M$, where M is the number of macroblocks in a frame (e.g., $M = 396$ for CIF).
Rate Buffer: Let the rate buffer size be B, and the delay of a macroblock i in the rate buffer be D_i^r. Then, clearly, $0 \leq D_i^r \leq B/V$. As an example, H.261 specifies that B/V can be as large as 250 ms.
Receiver Buffer: The receiving station buffers each received packet so that the delay from the input of the rate buffer to the output of the receiver buffer is equal to a delay bound $D_{max}^{'}$ ($D_{max}^{'} > B/V$).
Decoder: Let the decoding delay be D_{dec}. If the decoder operates on a frame-by-frame basis, then $D_{dec} \leq F$, and if it operates on a macroblock-by-macroblock basis, then $D_{dec} \leq F/M$.
Display: Let the delay due to displaying a decoded frame be D_{disp}. The decoded macroblocks may immediately be displayed, in which case, D_{disp} is between 0 and 33 ms, depending on the timing relationship between the display scanning and the placement of the macroblocks in the frame buffer. On the other extreme, a frame may be displayed only when it is decoded in its entirety, resulting in $F \leq D_{dec} + D_{disp} \leq F + 33\ ms$.

Thus, the end-to-end delay is $D_g + D_e + D_{max}^{'} + D_{dec} + D_{disp}$. It can be as small as 1.8 ms + $D_{max}^{'}$, when the frame grabber delay is 16 lines time, the encoder and decoder delays are one macroblock time, and the decoded macroblocks are displayed immediately. At the other extreme, for D_g=33 ms, $D_e = F$, and $D_{dec} + D_{disp}$ = F + 33 ms, the end-to-end delay can be as large as 132 ms+$D_{max}^{'}$ for F=33 ms (30 frames/s), and 330 ms+$D_{max}^{'}$ for F=132 ms (7.5 frames/s).

2.2.2 Data Traffic

We assume that data sources generate fixed size messages with uniformly distributed interarrival times[1]. A

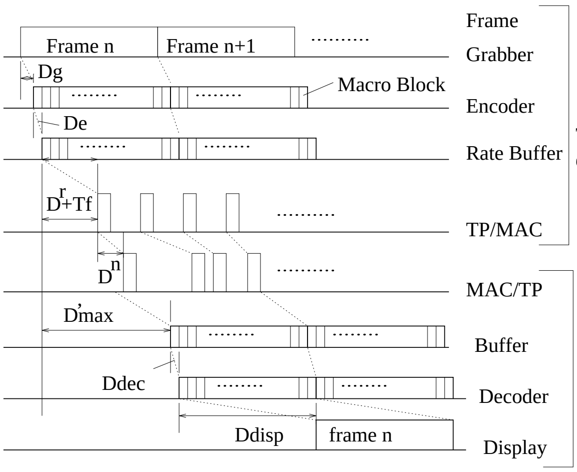


Figure 2: Timing diagram for video streams.

data source is characterized by its message size (M_d) and its average message interarrival time (u).

If a message is larger than a full size packet (1500 bytes), it is divided into as many full-size packets as needed. In that case, the packets belonging to the message are placed into the transmit queue as a bulk arrival.

In this study, we have limited ourselves to 3 message sizes. To generate non-bursty data traffic, we let M_d=1 kbytes, in which case each message fits in a single packet. To generate bursty data traffic, we choose M_d=10 kbytes and M_d=100 kbytes. After packetization at the transport layer, these message sizes correspond to 7 and 67 back-to-back packets respectively.

The average interarrival time, u, is varied to adjust the total data load on the network (denoted by G_d). The number of data sources in the network was fixed at two in the 10Base-T Ethernet simulations and twenty in the 100Base-T Ethernet simulations [2].

3 Results

In this section, we present our numerical results first for 10Base-T and then for 100Base-T Ethernets. We first present our findings when there are only streams in the network, where we show the existence of an optimum packet size, and discuss the effects of

[1] The simulator was also run with exponentially distributed interarrival times, and exponentially distributed message sizes. The results did not differ significantly from those obtained with the uniform distribution with fixed size messages. The uniform distribution was chosen over the exponential distribution for its lower complexity and faster convergence properties.

[2] We have also simulated some test cases with 20 data sources for 10Base-T, and 2 data sources for 100Base-T. For a given data load and burst size, the number of streams supportable did not change as we changed the number of data sources from 2 to 20.

stream rate and delay/loss constraints on the optimum packet size and the maximum number of streams supportable. We then present our results when both stream and data traffic are present in the network, discussing the effects of data load and burst size on the optimum packet size and the maximum number of streams supportable. Finally, we present some results of our ongoing work for relating packet loss rate with video quality.

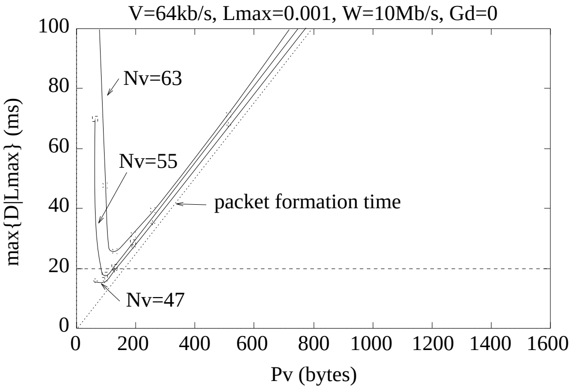
Throughout this section, we assume that for a given run of the simulation, all streams have the same rate, packet size, delay and loss requirements.

3.1 Results for 10Base-T

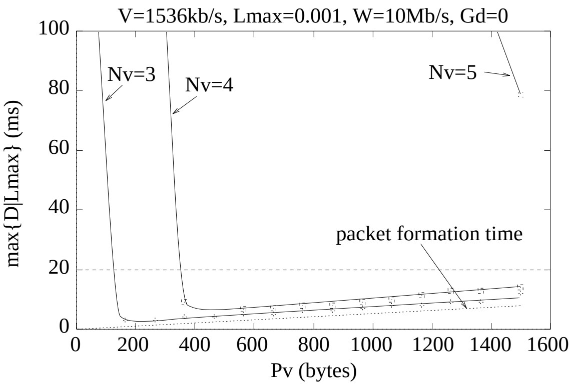
3.1.1 Streams Only

We first discuss the effect of packet size on the delay performance. Recall our assumption that CODECs operate on a macroblock-by-macroblock basis. Since a macroblock is typically much smaller than a packet, this allows us to treat the packet size as a continuous variable. Also recall that we take into account only the network delay. Therefore, we are able to follow the same treatment done in [4] to show the existence of an optimum packet size. Let k denote the number of packets sampled in the simulation. Recall that D_i $(i = 1, 2, \ldots, k)$ is the delay incurred by the i'th packet. Let $D_1^{'} < D_2^{'} < \ldots < D_k^{'}$ be the ordered sequence of delay samples. We let $\max\{D|L_{max}\} \triangleq D^{'}_{k(1-L_{max})}$, that is, the value of delay which is exceeded by exactly a fraction L_{max} of all samples. We can similarly define $\max\{D^n|L_{max}\}$. Clearly, $\max\{D|L_{max}\}=T_f+\max\{D^n|L_{max}\}$. When P_v is large, T_f is proportionally large. On the other hand, as P_v gets smaller, the number of packets contending increases, causing $\max\{D^n|L_{max}\}$ to increase. The combination of these two effects lead to an optimum packet size for which $\max\{D|L_{max}\}$ is minimum. This effect is illustrated in Figure 3. The solid lines show $\max\{D|L_{max}\}$ versus P_v, for $V = \{64, 1536\}$kb/s, $L_{max}=\{0.001, 0.01\}$, and various values of N_v. The packet formation time, T_f, versus P_v is shown as a dotted line. The figure shows that given V, D_{max} and L_{max}, there exists a unique optimum value of P_v (denoted by P_{opt}) and a maximum value of N_v (denoted by N_{max}) which satisfy the constraint $\max\{D|L_{max}\} \leq D_{max}$. N_{max} is our principal performance measure.

In Figure 4, we plot the maximum number of streams supportable versus stream packet size, for W=10 Mb/s, V=384 kb/s, L_{max}=0.001, and D_{max}={20,100}ms. The figure illustrates that the choice of packet size has an important effect on the maximum number of streams supportable by the network. Also, as we can see from the figure, the maximum number of streams is achieved over a range of packet sizes. Moreover, the range of feasible packet sizes can be limited by the maximum allowable Ether-



(a) V=64 kb/s



(b) V=1536 kb/s

Figure 3: $\max\{D|L_{max}\}$ vs. P_v for W=10 Mb/s, L_{max}=0.001, G_d=0

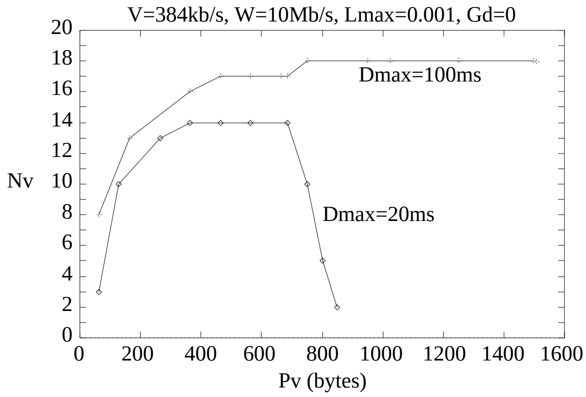


Figure 4: N_v vs. P_v for W=10 Mb/s, V=384 kb/s, L_{max}=0.001

D_{max}(ms)	L_{max}	64 kb/s	384 kb/s	1536 kb/s
20	0.001	100	450	800
20	0.01	120	650	1500
100	0.001	380	1250	1500
100	0.01	480	1500	1500

Table 1: P_{opt} (bytes) for W=10 Mb/s, G_d=0.

net packet size, particularly for larger values of D_{max}.

In Table 1, we show P_{opt} for V = {64, 384, 1536}kb/s, D_{max}={20,100}ms, and L_{max}={0.01, 0.001}. Observe from the table that P_{opt} increases with increasing stream rate. The reason is clear; with higher rate streams, a given packet formation time is attained by using a larger packet size. This allows the higher rate streams to use larger packets, which are handled more efficiently by the CSMA/CD protocol. Also note that for a small D_{max}, small packets need to be used due to the limit imposed on the packet formation time. As D_{max} increases, larger packets can be used, which are more efficiently handled by the CSMA/CD protocol. Thus, for a given stream rate, P_{opt} increases with increasing D_{max}.

In Table 2 we show N_{max} (and the corresponding maximum achievable network utilization in parenthesis) for D_{max}={20,100}ms, L_{max}={0.1%, 1%}, and V={64,384,1536} kb/s. As shown in the table, Ethernets can adequately support audio and video streams when the delay and loss constraints are within the ranges given above. It is also clear that more streams can be supported when the delay and/or loss constraints are more relaxed.

Recall that higher rate streams have a larger optimum packet size. Thus, they achieve the same network utilization with fewer packets. Also, with higher rate streams, there are fewer stations contending over the network for the same load, reducing the likelihood of contention and the average time to resolve contentions. As a result of these two effects, the CSMA/CD protocol operates more efficiently at high stream rates. Thus, at higher stream rates, even though fewer users are supported, more of the network bandwidth is utilized for useful traffic. As shown in Table 2, this is especially apparent for small D_{max}, because the low rate streams are forced to use very small packets.

3.1.2 Stream and Data Traffic

Our simulation results indicate that with non-bursty data traffic, the optimum packet size is independent of the data load. This result implies that if data traffic on the network is not bursty, the packet sizes shown in Table 1 can be used regardless of the actual data load on the network.

D_{max} (ms)	L_{max}	64 kb/s	384 kb/s	1536 kb/s
		N_{max}	N_{max}	N_{max}
20	0.001	55 (35%)	14 (54%)	4 (61%)
20	0.01	64 (41%)	17 (65%)	5 (77%)
100	0.001	89 (57%)	18 (69%)	5 (77%)
100	0.01	104 (67%)	20 (77%)	5 (77%)
Bandwidth limit		156	26	6

Table 2: N_{max} and the corresponding network utilization for W=10 Mb/s, G_d=0

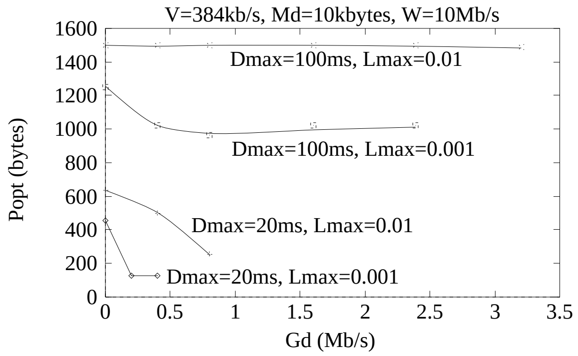


Figure 5: P_{opt} vs. G_d for W=10 Mb/s, V=384 kb/s, M_d=10 kbytes

On the other hand, with bursty data traffic, we have observed that the optimum packet size for streams decreases as the data load or the burst size increases. This result is illustrated in Figure 5 where P_{opt} versus G_d is plotted for V=384 kb/s, M_d=10 kbytes, and four combinations of D_{max} and L_{max}[3]. We explain the cause of this effect as follows. Introducing bursty data load causes streams to experience higher delays due to contention. To combat this delay, smaller stream packets are used, which allow more time for contending over the network because of shorter packet formation times.

In Figure 6, we plot N_{max} versus the data load for V=384 kb/s and for various values of M_d (at each data point, the corresponding P_{opt} is used). Figure 6(a) is with stringent delay and loss constraints (i.e., D_{max}=20ms, L_{max}=0.001), and Figure 6(b) is with relaxed delay and loss constraints (i.e., D_{max}=100ms, L_{max}=0.01). As we can see from the figure, introducing some non-bursty data load in the network causes

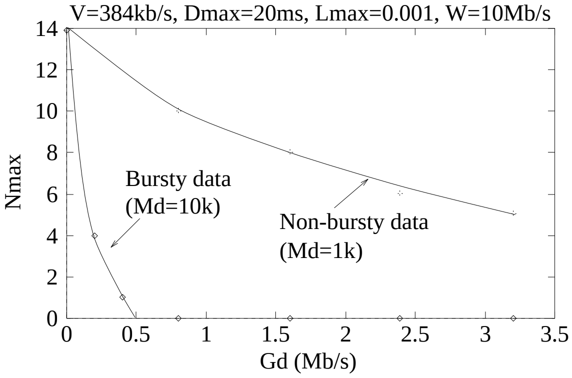
[3]Note that for V=384 kb/s, D_{max}=100 ms and L_{max}=0.01, the optimum packet size remains constant at 1500 bytes, the maximum packet size for Ethernets. In this case, the actual optimum packet size would have occurred at a point greater than 1500 bytes if there were not such a limit, and it would still have decreased as the data load is increased.

the number of streams to decrease by an equivalent load. That is, if we denote the change in data load by ΔG_d, and the corresponding change in the number of streams by ΔN_{max}, then $\Delta N_{max} V \approx -\Delta G_d$. This result, as well as the result that P_{opt} is independent of the data load for non-bursty data traffic, can be explained by the similarity of non-bursty data traffic and stream traffic. Even though non-bursty data traffic is less regular than stream traffic because of its uniformly distributed interarrival times, this irregularity seems unimportant.

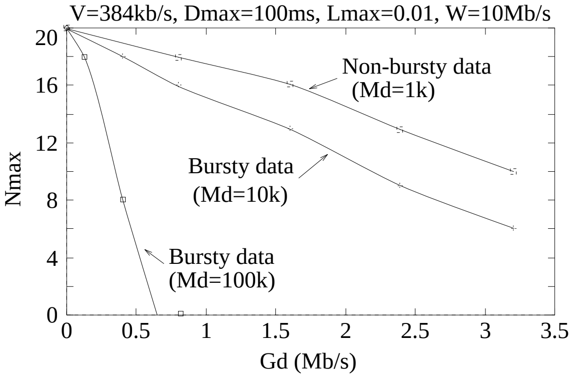
On the other hand, Figure 6 also indicates that with bursty data traffic, for the same data load, the number of supportable streams decreases as the burst size increases. For example, as shown in part (a) of the figure, when the data message size is 1 kbyte, 11 streams can be supported for a data load of 0.5 Mb/s, whereas when the data message size is 10 kbytes, no streams can be supported when the data load is 0.5 Mb/s or more.

A comparison between Figure 6(a) and Figure 6(b) indicates that the degradation due to increasing burst size is less apparent with less stringent delay and loss constraints. For example, as shown in Figure 6(b), when D_{max}=100ms, L_{max}=0.01, and G_d=0.5Mb/s, 18 streams can be supported with a data message size of 1 kbyte. For the same data load, and for a data message size of 10 kbytes, 17 streams can be supported.

In Figure 7, we plot N_{max} versus D_{max} for V=384 kb/s, G_d={0, 0.8, 1.6, 3.2}Mb/s and M_d=10 kbytes. As shown in the figure, N_{max} becomes insensitive to the delay constraint as the delay constraint gets larger than about 400 ms. The reason is that when D_{max} is larger than around 400 ms, most of the packet loss occurs due to excessive collisions (as opposed to exceeding the delay constraint). That makes it ineffective to relax the delay constraint any further; the same amount of packets will be lost. Also observe that N_{max} is very sensitive to D_{max} in our main range of interest: [20 ms–100 ms].



(a) D_{max}=20 ms, L_{max}=0.001



(b) D_{max}=100 ms, L_{max}=0.01

Figure 6: N_{max} vs. G_d for W=10Mb/s

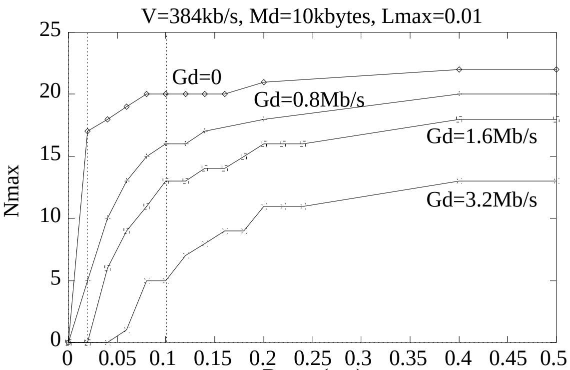


Figure 7: N_{max} vs. D_{max} for W=10 Mb/s

3.2 Results for 100Base-T

Our simulations of a 100Base-T segment yielded similar results to those found for 10 Mb/s Ethernets. In the following, we summarize the numeric results for 100Base-T.

In Table 3, we show P_{opt} for V={384,1536}kb/s, D_{max}={20,100}ms, and L_{max}={0.01, 0.001} when there are only streams on the network. Comparing Table 3 with Table 1, we observe that the optimum packet sizes for 100Base-T are about the same as those for 10Base-T for given V, D_{max}, and L_{max}.

In Table 4, we show N_{max} (and the corresponding maximum achievable network utilization in parenthesis) for D_{max}={20,100}ms, L_{max}={0.1%, 1%}, and V={384,1536} kb/s. By comparing this table with Table 2, we observe that 100Base-T can support about 10 times more streams compared to 10Base-T. This result is not immediately obvious from the fact that 100Base-T has a 10 times higher transmission speed.

D_{max}(ms)	L_{max}	384 kb/s	1536 kb/s
20	0.001	500	1000
20	0.01	650	1500
100	0.001	1500	1500
100	0.01	1500	1500

Table 3: P_{opt} (bytes) for W=100 Mb/s, G_d=0.

D_{max}	L_{max}	384 kb/s	1536 kb/s
(ms)		N_{max}	N_{max}
20	0.001	138 (53%)	43 (66%)
20	0.01	160 (61%)	49 (75%)
100	0.001	180 (69%)	49 (75%)
100	0.01	205 (79%)	52 (80%)
Bandwidth limit		260	65

Table 4: N_{max} and the corresponding network utilization for W=10Mb/s, G_d=0

There are several competing factors, some working for the advantage of 100Base-T, and some to the disadvantage of it. First of all, 100Base-T has a 10 times smaller slot size compared to 10Base-T. This works to the advantage of 100Base-T, since for a given D_{max}, 10 times more slots are available in 100Base-T for a packet to resolve contention and be transmitted successfully. On the other hand, in 100Base-T, since packets are transmitted 10 times faster, the parameter a (i.e., propagation delay divided by the packet transmission time) is 10 times larger compared to that in 10Base-T for a given network span, and a given packet size. In CSMA/CD networks, performance degrades as a increases. Moreover, to achieve the same network utilization, 10 times more streams must exist in 100Base-T. For a given traffic load, the performance of a CSMA/CD network decreases as the number of stations which generate that load increases.

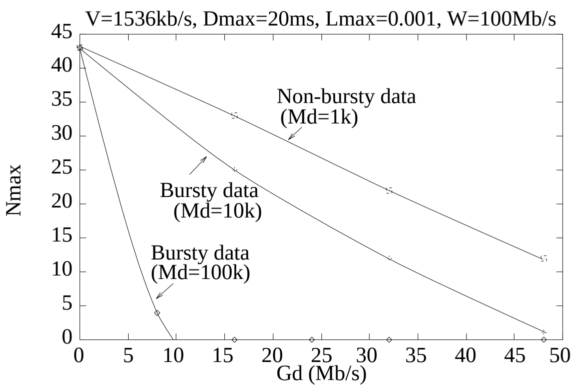


Figure 8: N_{max} vs. G_d for W=100 Mb/s

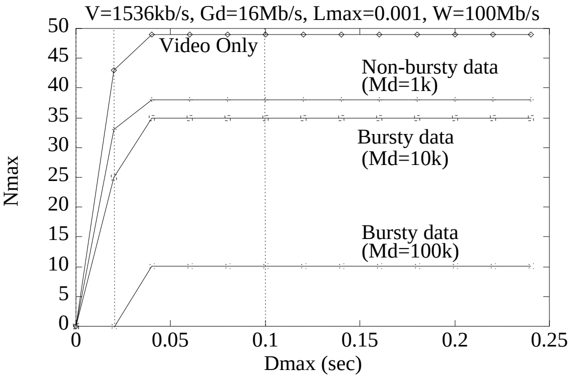


Figure 9: N_{max} vs. D_{max} for W=100 Mb/s

In Figure 8, we plot N_{max} versus G_d for V=1536 kb/s, D_{max}=20 ms, L_{max}=0.001, and M_d={1,10,100}kbytes. Comparing Figure 8 with Figure 6(a), we can see that in 100Base-T, about 10 times larger burst sizes than those in 10Base-T cause an equivalent degradation. This is due to the 10 times larger transmission speed and 10 times smaller slot size, causing the same size burst to be transmitted about 10 times faster, and thereby affecting stream packets for a time about 10 times shorter in 100Base-T. Note that the parameter a does not have an important effect while a burst is being transmitted; regardless of the value of a, the i'th packet of the burst will always collide with the packets that were deferring to the $i-1$'th packet of the burst owing to the 1-persistence of the Ethernet MAC protocol.

In Figure 9 we plot N_{max} versus D_{max} for V=1536 kb/s, L_{max}=0.001, G_d={0,16}Mb/s, and M_d={1,10,100}kbytes. As indicated by the figure, in 100Base-T, the number of streams supportable becomes insensitive to the delay constraint as it gets larger than around 40-50 ms. The reason is that due to the 10 times smaller slot size in 100Base-T, when D_{max} is larger than about 40-50 ms, most of the packet loss occurs due to excessive collisions. Consequently, relaxing the delay constraint any further will be of no help. This result suggests that if the application can tolerate a maximum delay of 100 ms or larger, retransmitting lost packets at the transport layer is conceivable.

3.3 Relating the Packet Loss Rate with Video Quality

As we mentioned before, packet loss rate is only a rough measure of quality of video streams, since the effect of packet loss on quality depends on the stream syntax and rate. We are in the process of investigating the effects of packet loss on video quality, taking into account the stream syntax, as well as the delays due to the encoder and the decoder. As a preliminary step of this investigation, we consider the following. Using actual video sequences, we generate H.261 streams of rate V using a rate buffer size B. Instead of measuring the delay between points A and B as shown in Figure 1, here we measure it between the input of the rate buffer and point B. We denote the delay of a macroblock i in the rate buffer by D_i^r $(0 \leq D_i^r \leq B/V)$. We simulate N_{max} streams with the packet size of P_{opt}, subject to $\max\{D^n + T_f \mid L_{max}\} \leq D_{max}$, same as before. However, here we assume that a packet j is lost if $D_{j_1}^r + D_j^n + T_f > D_{max} + B/V$, where j_1 is the first macroblock in the packet j. We measure the resulting packet, macroblock and GOB loss rates, as well as glitch rates, where we define a glitch as a number of consecutive frames such that each frame contains at least one lost macroblock. We consider buffer sizes such that B/V={50,100,250}ms.

For V = 384 kb/s, B/V=50 ms and L_{max}=0.001, the resulting packet loss rates are between 10^{-5} and 10^{-4}, and the glitch rates are between 0.04/min to 0.1/min (depending on the video sequence and on D_{max}). When L_{max}=0.01, the resulting packet loss rates are between 10^{-4} and 3×10^{-3}, and the glitch rates are between 0.75/min to 2.3/min. In all cases, the GOB loss rate is about 3-5 times the packet loss rate, and the macroblock loss rate is about 80% of the GOB loss rate. As the rate buffer size increases, the loss rates decrease. For V=1536 kb/s, the loss rates are at least an order of magnitude smaller than their counterparts for V=384 kb/s given D_{max}, L_{max}, and B. Those results apply regardless of the data load and burst size.

4 Conclusions

We have found that Ethernets are adequate for supporting audio and video streams with the delay and loss constraints that we considered. However, to ensure that streams meet their delay and loss constraints, admission control must be exercised so that no more than N_{max} streams are transmitted at the same time on a segment.

Also, since the number of streams supportable decreases significantly as the data burst size is increased, the data traffic must be throttled to limit the data load and burst size.

If the data traffic is not bursty, the optimum packet size depends only on the stream rate and delay/loss bounds. This information could be stored at each station. However, if the data traffic is bursty, the optimum packet size depends on the data load and burstiness as well. A conservative means of operating such a system is to choose the packet size corresponding to the highest expected data load and burst size.

We have seen in several places that Ethernet's maximum packet size, 1500 bytes, places a limit on the optimum packet size. However, given that at such operating points, the network utilization is typically as high as 70-80%, we do not expect a significant increase in the performance even if there were not a limit on the maximum packet size.

References

[1] "Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications," ISO/IEC 8802-3, (4th edition, 1993-07-08).

[2] "100Base-X Physical Layer Specification for Fast Ethernet," Sponsored by the Fast Ethernet Alliance. (Version 1.0, October 15, 1993).

[3] "4T+ Physical Layer Specification for UTP Category 3/4/5 Wiring (Specification Overview)," Sponsored by the Fast Ethernet Alliance. (December, 1993).

[4] F. A. Tobagi and N. Gonzalez-Cawley, "On CSMA-CD local networks and voice communication," in *Proceedings of the Conference on Computer Communications (IEEE Infocom)*, (Las Vegas, Nevada), pp. 122–127, IEEE, March/April 1982.

[5] T. A. Gonsalves and F. A. Tobagi, "Comparative Performance of Voice/Data Local Area Networks," *IEEE Journal on Selected Areas in Communications*, pp. 657–669, June 1989.

[6] J. Zdepski, K. Joseph, and D. Raychaudhuri, "Packet transport of VBR interframe DCT compressed digital video on a CSMA/CD LAN," in *Proceedings of the Conference on Global Communications (GLOBECOM)*, (CA), pp. 886–892 (25.3), IEEE, Dec. 1989.

[7] K. M. Nichols, "Network Performance of Packet Video on a Local Area Network," in *IEEE 11th Annual International Phoenix Conference on Computers and Communications*, (Scottsdale, Arizona), pp. 659–666, IEEE, April 1–3 1992.

[8] "H.261: Video CODEC for Audiovisual Services at p×64 kbit/s," ITU/CCITT, (Geneva, 1990).