Fixed Versus Variable Packet Sizes in Fast Packet-Switched Networks

Mahmoud Naghshineh and Roch Guérin

IBM Thomas J. Watson Research Center

ABSTRACT

In this paper we investigate various performance measures of interest, when comparing fast packet-switched networks that operate with either fixed or variable packet sizes. These performance measures include queue length distribution, packet loss probability, as well as user frame loss probability. The focus of the paper is on identifying key parameters that influence the outcome of this comparison, and on quantifying the potential benefits of each approach.

I. Introduction

In fast packet-switching networks, the subject of fixed vs. variable size packets has become an important design and architecture issue (see [1-3] and [6-10] for examples of networks that support fixed and variable size packets, respectively.) One aspect of the comparison between the two approaches, is the impact of packet format, i.e., fixed size or variable size, on the buffer sizes required to achieve a given frame or packet loss probability. Throughout this paper, the term frame is used to denote an end-user data unit, while the term packet refers to the network data unit.

The objective of this work is to provide some insight into the factors that influence the outcome of this comparison. In particular, we want to show when and why the choice of a particular packet format may be advantageous. This study is carried out by means of simulations, which allow us to define and model a general network environment without making overly restrictive assumptions on the traffic.

Traffic sources accessing the network are represented by means of ON-OFF sources which alternate between active and idle periods. When active, a source may generate one or more consecutive user frames. Frames go through an adaptation layer when entering and leaving the network. This adaptation layer performs the necessary mapping to an from the network packet format and the user frame format. In particular, this layer is responsible for the segmentation and reassembly functions. Note that even a network which allows transmission of variable size packets may require such functions, in cases where the user frame size exceeds the maximum network packet size.

As mentioned earlier, the objective of this work is to identify the key factors that influence the performance of networks using different packet formats. Similar studies have already been conducted in the past [4, 5, 13], which investigated some of the trade-offs involved in selecting a particular packet format. Some of the aspects considered included the impact of the header overhead for different frame size distributions, and the potential better buffer utilization of small packets[1]. We extend the scope of these studies and show where and why the differences in packet formats become significant. This is done essentially by means of simulations, but some simple analytical results are also provided, that help shed some light on some of the phenomena observed. The study focuses on three aspects: 1) The impact of header overhead and network link utilization, 2) The effect of source activity and pipelining on buffer efficiency, 3) The sensitivity of frame loss as a function of data loss and source characteristics.

The paper is organized as follows. We first describe the simulation models and the environment used to compare packet formats. This is followed by a simple analytical model, that illustrates the type of trade-off

[1] Note, that we assume here that cut-through is not permitted within network nodes, i.e., packets must be fully reassembled before they can be forwarded on the next link.

0743-166X/93 $03.00 © 1993 IEEE

that can be expected when comparing fixed versus variable size packet formats. The next three sections are devoted to the study of the three aspects mentioned earlier, and it is shown when and why different packet formats can yield better performances. Finally, a brief conclusion summarizes the findings of this paper.

II. Simulation Models

In this section we describe the simulation models, that were built to compare the impact of fixed and variable size packet formats on the buffer requirements in the network. An overview of these simulation models is provided in Figure 1 and Figure 2.

II.A Network Model

The purpose of this first model is to observe a single intermediate network link, which merges traffic generated from different sources after it has traveled through different paths in the network. The model has four major components. From left to right in Figure 1, we distinguish between traffic sources, adaptation layer, intermediate network nodes, and the tagged network link. As mentioned earlier, the traffic sources generate data according to an ON-OFF process, which is characterized by its peak data generation rate R_{peak}, and the distribution of its ON and OFF periods. For illustration purposes, these distributions were taken to be exponential, with mean duration T_{on} and T_{off}, respectively. Because of the flexibility of a simulation-based approach, other choices are clearly possible.

The second component of the simulation model is the adaptation layer (AL) that separates the traffic sources from the network. This layer is responsible for the conversion from the user frame format into the network packet format, and conversely. In particular, the AL performs segmentation and reassembly of user frames, as well as bit padding whenever necessary, i.e., in the last cell of a user frame in the case of a fixed size packet format. The time Δ_i out of the segmentation unit between two consecutive packets from the same ON period, is proportional to the ratio of the packet size to the source peak rate. For example, in the case of an ATM network, packets from the same ON period are generated $\Delta = P/R_{peak}$ sec. apart[2], where $P = 48$ bytes and R_{peak} is the source peak rate. In the reassembly phase, the AL must ensure that all packets generated from a user frame have been properly received. In this study, we assume that a user frame is "lost" if one or more of its segments, i.e., network packets, is in error or missing. In other words, we assume that error recovery is performed above the network layer.

In the case of an ATM-type network with fixed size packets or cells, we assumed a 48-byte data payload and a 5-byte header. Note, that this ignores the additional overhead introduced by the ATM adaptation layer itself, which ranges from 0 to 4 additional bytes [1-3]. A 10-byte header with a 2-Kbyte maximum packet size were chosen for networks that allow variable length packets (denoted by VLP model in the rest of this paper.) Note that although the raw data generated by sources and passed to the AL is the same for both ATM and VLP models, the differences in header overhead result in different outputs into the network.

The third component of the model is an "intermediate" network node, which attempts to account for the random delay that packets encounter when going through several stages of a network. A possible approach to capture this effect, would have been to actually simulate a large network and the corresponding traffic so as to effectively recreate the interactions that take place within a network. Such an approach is unfortunately computationally too costly. Instead, we use a single *"wait node"* to represent the effect of network interactions on a packet. The wait node adds a random delay to packets entering the network.

Specifically, the wait node is modelled as an M/M/1 (or M/D/1 for fixed size packets) queue with fixed load ρ, e.g., $\rho = 85\%$. The node receives two traffic streams: the traffic generated by the source itself, and a background traffic of intensity ρ', where ρ' is chosen so that the aggregate traffic has intensity ρ. The delay seen by a source packet arriving at the wait node depends on the position of the packet within the corresponding ON period of the source. The first packet generated during an ON period is assumed to see a random system upon its arrival and it, therefore, sees an unfinished work at the wait node equal to the average unfinished work in an M/M/1 (M/D/1) queue.

[2] With possibly the exception of the last packet

This yields [11] a delay (including the service time) for this first packet of the form $W_1 = \rho/\mu(1-\rho) + P_1/S \quad (W_1 = \rho/2\mu(1-\rho) + P_1/S)$, where $1/\mu$ is the average service time of a packet ($1/\mu = P/S$, with P the average network packet size and S the speed of a network link) and P_1 is the size of the first packet in the ON period.

Subsequent packets generated during the same ON period are then taken to see a delay, which reflects the dependency on the system seen by the first packet and the impact of additional background traffic that may have joined the queue since then. In other words, assuming that the first packet arrived at the wait node at $t = 0$, the i-th packet generated during an ON-period will leave the wait node at time $t_i = W_1 + \Sigma_{j=2}^{i} W_j$, where W_j is the expected amount of work that arrived at the wait node (including the new packet itself) between the arrivals of the j-th and $(j-1)$-th packets. The quantities W_j are easily found to be given by $W_j = \rho' \Delta_j + P/S$, where ρ' and Δ_j are as previously defined. This wait node model provides us with a simple means to randomize the network delay seen by packets, while keeping some of the correlation that exists between the network states seen by consecutive packets from the same source[3].

The last component of the simulation model is the tagged network node (and link), where the traffic from all the sources is merged. The inputs to this tagged node are the outputs of all the wait nodes, and the quantity being monitored is the queue length or buffer content distribution. This distribution is used to compare the buffer requirements of the ATM and VLP formats.

II.B Frame Loss Model

The second simulation model shown in Figure 2, was developed to study user frame loss probability, and in particular its sensitivity to the network packet format and the ratio between source and link speeds. Recall that a user frame is assumed lost if any of the corresponding network packets is lost. The model consists of three components: a traffic source characterized by its peak rate and the distributions of the ON and OFF periods[4], an adaptation layer identical to what has been previously described, and a network link carrying both the source traffic and some background traffic. The background traffic was assumed to be either Poisson or generated by ON-OFF sources. User frames are produced by the tagged source (and the background traffic) and are assumed to consist of all the packets generated during an ON period. Individual packets coming from the tagged and background sources are aggregated and offered to the network link. Reassembly of user frames from the tagged source is carried out at the output of the network link.

As mentioned above, the quantity of interest is the frame loss probability corresponding to the data loss probability induced by the finite size of the link buffer. Here, we define data loss probability as the ratio of the number of dropped or lost bytes to the number of bytes transmitted on the tagged link[5]. We are interested in the impact of network packet format and speed of the network link on the user frame loss. In particular, we want to capture the impact of the link speed on the correlation between queue sizes seen by packets from the same frame. This investigation is carried out by varying the speed of the network link, while keeping the characteristics of the tagged source constant, and adjusting the intensity of the background traffic to keep the data loss constant. This is done for both the ATM and VLP packet formats and additional details on this investigation are provided in section V.

III. Impact of Packet Overhead

In this section, we investigate the significance of the per packet overhead under different scenarios. This overhead corresponds to the header and possible padding bits (for fixed size packets), that must be added to each packet carried by the network. Its impact is clearly a function of the frame size distribution and the packet format. For example, a 1 kbyte packet only carries a 10-byte overhead in a VLP network with a maximum packet size greater than 1 kbyte, while the overhead is 113 bytes in an ATM network

[3] In particular for high-speed sources, where $\Delta_i \simeq 0$.

[4] Exponential distributions have again been selected in the examples, but other choices are possible.

[5] Note, that in the case of fixed size packets this is also the packet loss probability.

(21 packets each with a 5-byte header, and 8 bytes of padding in the last packet.) Figure 3 illustrates the difference in overhead for VLP and ATM networks as a function of the average message length or frame size. Exponentially distributed frame sizes were assumed and it should be noted that under this assumption, the overhead due to padding bits results in the ATM format being always less efficient (on average) than the VLP format. This may not hold with other frame size distributions, for example small constant (e.g., 48 bytes!) frame sizes, where the smaller per packet header of ATM network can yield better efficiency. (See [5] for similar results with other frame size distributions.)

The difference in per packet overhead between ATM and VLP networks results in different link utilizations for the same amount of carried user data. Typically, imposing fixed size packets results in higher link utilization. This may in turn influence the buffer content distribution, and therefore the packet loss within the network. The impact of this effect is, however, strongly dependent on the load at which network links are being operated. In particular, the higher utilization imposed by fixed size packets is typically not significant and packet loss is dominated by other factors when the link utilization is relatively low.

A simple "illustration" of this behavior can be obtained by plotting the average buffer content in both an M/M/1 and an M/D/1 system as a function of the original data load [12]. Despite a higher link utilization, the M/D/1 system outperforms the M/M/1 system except at high loads. Intuitively, this is due to the more regular service in the M/D/1 system, which results in better buffer utilization. Specifically, an idle server in the M/D/1 system starts serving an incoming packet earlier because of its fixed small size. At higher loads, incoming packets hardly ever find an idle server, and the the difference in link utilization becomes significant and the performance advantage of the M/D/1 system disappears.

This phenomenon is further illustrated for a more realistic system in Figure 4, which plots for different link utilizations the buffer content distribution at an intermediate node in ATM and VLP networks. The results are obtained by means of simulations, using the network model described in section II.A. The traffic on the tagged network link is generated from 10 identical ON-OFF sources, each with a peak rate of 16 Mbps, an average ON period of 0.5 msec (1 kbytes average frame size), and a utilization of 10%. Both ON and OFF periods are assumed to be exponentially distributed. The distribution of the buffer content is obtained for different link speeds, and therefore utilizations. Because of the impact of other factors such as the correlation between successive arrivals from the same source, the outcome is not as pronounced as in the simplified comparison of M/M/1 and M/D/1 systems. However, it can again be seen that at low link utilizations, the additional overhead imposed by the fixed packet size of ATM is more than compensated for by the better buffer efficiency. This is again reversed when link utilization increases beyond a value[6], above which the price of the higher overhead of ATM becomes the dominant factor. These observations are in line with results from previous studies [4].

The impact of packet overhead and more specifically the padding bits (for fixed size packets) also depends on the burst length or message size. As can be seen in Figure 3, the effective[7] message length of ATM cells increases sharply as the mean message size decreases (i.e., assuming an exponential frame size distribution.) This means that the buffer length distribution is also heavily influenced by the mean frame size. As a result, the outcome of the comparison between the buffer length distributions of ATM and VLP systems can change in favor of VLP as the average frame size decreases. This is because for very small average message sizes the effective message size, as a result of padding bits of ATM, brings the link utilization above the threshold where the VLP starts to perform better. This effect is illustrated in Figure 5, which shows simulation results for a set of 10 ON-OFF sources with utilization of 50% and a peak rate of 16 Mbps. The raw link utilization is about 50-55%. We observe, that for small message sizes (mean message size of 20 Bytes,) the effective link utilization of the ATM system becomes so high that the link becomes almost saturated. This is a by-product of padding bits. This deficiency of the ATM

[6] Of the order of 60-65% for this example.

[7] By this we mean the original data length plus the added overhead.

system vanishes as the mean message size increases, up to a point where pipelining efficiency becomes the dominant factor so that ATM systems outperforms VLP ones.

IV. Impact of Source Utilization

In this section, we study the impact of another parameter that influences the outcome of the comparison between ATM and VLP formats. Specifically, we investigate the effect of varying source utilization while keeping network link utilizations constant. Here, source utilization ρ denotes the fraction of time the source is active generating data at its peak rate (i.e., $\rho = T_{on}/(T_{on} + T_{off})$.) Obviously, for a given link utilization, higher source utilizations result in a link speed closer to the aggregate peak rate of the sources whose traffic is routed on that link. This should in turn favor networks operating with small packet sizes since on such links packets will essentially be served as they arrive, without being able to accumulate in the link buffer. Intuitively, smaller packets result in a more progressive arrival of data. This makes the system resemble more closely a fluid-flow model, which has minimal queueing when the service rate and the aggregate input rate are close. Conversely, low source utilizations allow for link rates much smaller than the aggregate peak rate of all sources with traffic routed on the link. The advantage of the better buffer efficiency of the ATM format should then be much less significant.

In order to observe these behaviors we simulated a system similar to the one used in the previous section. As before, it consists of 10 identical ON-OFF sources sharing a common network link, each with a 16 Mbps peak rate and a 0.5 msec average ON period (1 kbytes average frame size). The raw link utilization was kept constant at about 80%, while source utilization was varied from 10 to 70%. The buffer content distribution was again obtained for different cases by means of simulations, and the results are shown in Figure 6. As expected, the potential for better buffer utilization of the ATM format is realized at high source utilizations. Conversely, the VLP format shows better results when the source utilization is low and the effect of the smaller overhead becomes the dominant factor. Note, however, that the crossover point, i.e., the source utilization above which ATM outperforms VLP, clearly depends on other parameters such as link utilization, etc.

V. Frame and Data Loss Probabilities

In this section, we investigate the relation between the frame loss probability experienced by a source, and the data loss probability provided by the network. The first quantity is the one of real interest to end-users, while the second is typically the only one the network can control and monitor. We are interested in comparing the frame loss seen by an end-user sending its data over either an ATM or a VLP network, each ensuring the same data loss. In other words, given that two different networks -one with fixed and another with variable length packets- provide comparable packet or data loss, how do they compare in terms of frame loss, i.e., do the different packet formats have any effect.

For the purpose of this comparison, we use the simulation model described in section II.B, where we vary the ratio of the source to link speeds. As mentioned earlier, this ratio is expected to have a significant impact on the performance, i.e., frame loss probability, seen when using different packet formats. Intuitively, segmentation is expected to have limited impact when the ratio of source to link speed is large. In such cases, the high speed of the source ensures that all segments (packets) of frame arrive nearly back-to-back in the network, emulating, therefore, the appearance of a single packet carrying the full frame. The impact of the adaptation layer is then, except for the different overheads, similar for both packet formats. In the case of low speed sources, the spacing between packets of the same frame introduced by the segmentation layer of the ATM format, produces a packet stream significantly different from that of the VLP format. This is likely to result in different frame loss probabilities.

In particular, when the source speed is much lower than the network link speed, packets from the same frame will essentially see uncorrelated network queues (the time between two packets is large compared to the time constant of the network queue). Assuming a packet loss probability of p (in the case of ATM networks this is equal to the data loss probability), the loss probability of a frame consisting of n packets is approximately equal to np. As n is much larger in

ATM than in VLP networks, it is clear that for the same data loss probability p, low speed sources will see a worse frame loss probability when going over a network that enforces a small fixed size packet format. This is the "avalanche" effect mentioned in [5]. More problematic than the sheer degradation of frame loss probability which also depends on a number of additional factors (see below), is the fact that this behavior makes it very difficult to translate network performance into user performance in ATM type networks. In other words, for a given network performance (data loss probability), end-users can see drastically different quality of service (frame loss probability) depending on their own peak rate.

This aspect is illustrated in Figure 7, which reports the results of a set of simulations based on the model described in section II.B. The tagged source has a peak rate of 16 Mbps, an average ON period of 0.5 msec, and a utilization of 10%. The link speed was varied from a fraction (around 20%) to several times (6 to 7) the source peak rate, while its utilization was kept constant at about 80% by changing the intensity of the Poisson background traffic. The data loss at the link buffer was kept approximately constant (between 5×10^{-4} and 7×10^{-4}) for both network types, by appropriately adjusting the size of the link buffer. As expected, the frame loss probability seen by an end-user shows a sharp increase when its source peak rate decreases in ATM type networks, while it is quite insensitive to this parameter in VLP type networks. This implies, that the difficulty in ATM type network to translate network performance into user performance, is not present in VLP type networks. This difference is mainly due to the fact that the latter ones allow users to *transparently* send their data through the network[8], with minimum constraint on the packet format to be used. This feature is likely to significantly simplify the Grade-Of-Service (GOS) negotiations between users and the network.

The previous scenario compared frame loss probabilities for ATM and VLP networks carrying the same traffic and ensuring the same data loss probability. This was achieved by properly adjusting the buffer sizes needed in each case to achieve the desired data loss. This comparison, however, did not capture the fact that for the same user traffic and buffer sizes, ATM and VLP networks typically yield different data loss probabilities. This was illustrated in the previous sections, where it was shown that ATM and VLP networks have different buffer content distributions, that vary as a function of source characteristics and link utilizations. In particular, ATM type networks have the potential for lower data loss probabilities when link utilizations are not too high. In such cases, it is possible that the lower data loss probability offsets the penalty imposed on the frame loss probability by the fixed size packet format.

This was investigated in [12] through a simple analytical model. A more realistic investigation would combine the influence of finite link speeds as illustrated in Figure 7, and the impact of the difference in data losses due to packet format. In other words, the questions we should ask are: How do the frame losses of ATM and VLP compare for networks with equal buffer sizes, finite speed links, and various link utilizations? When does the fact that ATM network packets see decorrelated buffer lengths for very high link speeds dominates ATM's potential advantage of lower data loss, and result in a higher frame loss?

This investigation is again carried out using the simulation model of section II.B, where both networks now have the same finite buffer space. Traffic is generated from 10 ON-OFF sources all with a utilization of 10% and an ON-time of 0.5 msec. One of the sources is tagged and assigned a peak rate of 16 Mbps, while the link speed is varied. The link utilization is kept constant at about 30% by changing the peak rate of the 9 other sources accordingly. For each set of simulations the buffer size of the ATM simulation was adjusted to give a data loss of about 5×10^{-4} at the link. This buffer size was then kept for the VLP system. The frame loss of both systems was obtained and the results are shown in Figure 8.

From the figure, one can conclude that as the ratio of the link speed to the source peak rate increases, the frame loss of ATM increases so that at some point it crosses the frame loss line of the VLP and becomes worse even though the data loss of VLP always remains higher. Therefore, despite its lower packet loss ATM's user frame loss can often be worse than in VLP networks in particular for low speed users.

[8] See [10] for additional discussions on the benefits of network transparency.

VI. Conclusion

In this paper, we have studied the influence of a number of parameters on the performance of networks operating with either ATM or VLP packet formats. It was shown that the outcome heavily depends on the assumed operating conditions and source characteristics. In particular, by appropriately adjusting parameters such as link utilization, frame size distribution, buffer sizes, source utilization, etc., it is possible to claim that either approach outperforms the other. A conclusion of this work is, therefore, that network performance cannot be used as decisive factor when comparing networks using different packet formats. Rather, the emphasis should be on other aspects such as network complexity and flexibility in supporting various user requirements. From this perspective, VLP type networks may have some advantage because of the more direct relation between network data loss and user message loss. This can help simplify the mapping from network into user performance.

[1] "Special Issue on Asynchronous Transfer Mode," *Int. J. Digital and Analog Cabled Systems*, vol. 1, no. 4, 1988.

[2] "Draft - General B-ISDN Aspects," *CCITT Study Group XVIII, Report R 34*, June 1990.

[3] J.-Y. Le Boudec, "The Asynchronous Transfer Mode: A Tutorial," *IBM Research Report*, no. RZ 2133, May 1991.

[4] J.-Y. Le Boudec, "About Maximum Transfer Rates for Fast Packet Switching Networks," *IBM Research Report*, no. RZ 2111, April 1991.

[5] I. Cidon, J. Derby, I. Gopal, and B. Kadaba, "A Critique of ATM from a Data Communications Perspective," *Proc. ICCC'90*, November 1990.

[6] I. Cidon and I. Gopal, "PARIS: An Approach to Integrated High-Speed Private Networks," *Int. J. Digital and Analog Cabled Systems*, vol. 1, no. 2, pp. 77-86, April-June 1988.

[7] I. Cidon, I. Gopal, P. M. Gopal, R. Guérin, J. Janniello, and M. Kaplan, "The plaNET/ORBIT High Speed Network," *IBM Research Report*, no. RC 18270, August 1992. (Available from the authors).

[8] I. Cidon, I. Gopal, G. Grover, and M. Sidi, "Real-Time Packet Switching: A Performance Analysis," *IEEE J. Select. Areas Commun.*, vol. SAC-6, no. 9, pp. 1576-1586, December 1988.

[9] D. Clark, B. Davie, D. Farber, I. Gopal, B. Kadaba, D. Sincoskie, J. Smith, and D. Tennenhouse, "The AURORA Gigabit Testbed," *Computer Networks and ISDN*, 1991.

[10] I. Gopal and R. Guérin, "Network Transparency: The plaNET Approach," *Proc. INFOCOM'92*, pp. 590-601, Florence, Italy, May 1992.

[11] L. Kleinrock, *Queueing Systems, Vol. 1: Theory*, New York, John Wiley & Sons, 1975.

[12] M. Naghshineh and R. A. Guérin, "Fixed Versus Variable Packet Sizes in Fast Packet-Switched," *IBM Research Report*, 1992.

[13] S. Parekh and K. Sohraby, "Some Performance Trade-Offs Associated With ATM Fixed-Length vs Variable-Length Cell Formats," *Proc. ICC'88*, November 1988.

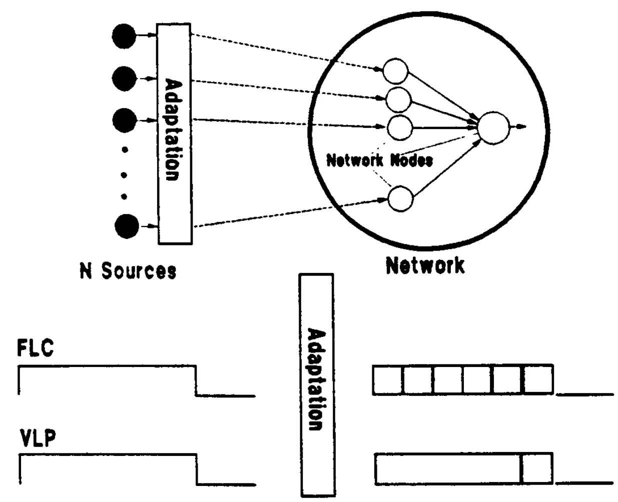

Figure 1. Network Simulation Model

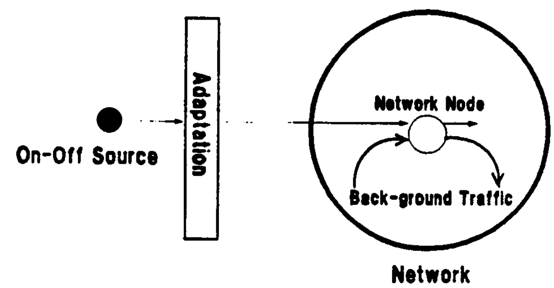

Figure 2. Frame Loss Simulation Model

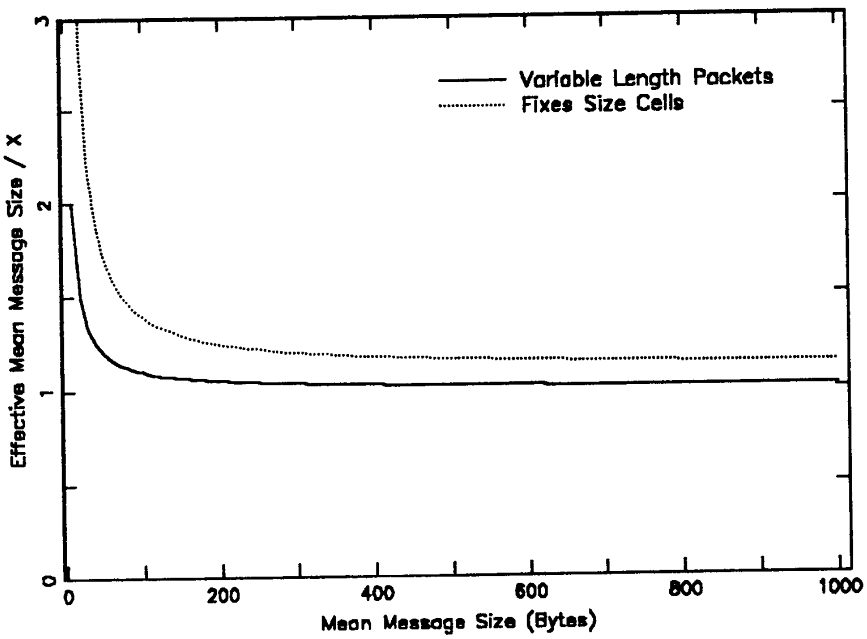

Figure 3. Network Overhead as A Function of Average Frame Size

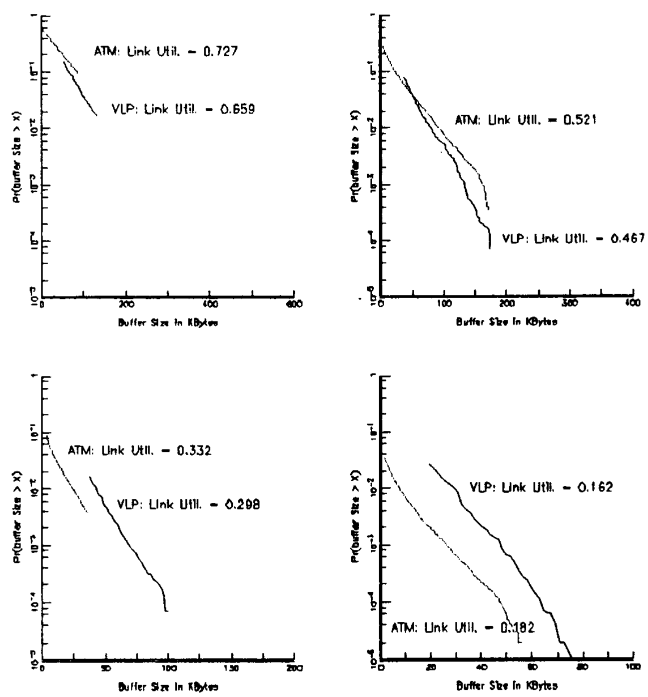

Figure 4. Impact of Link Utilization

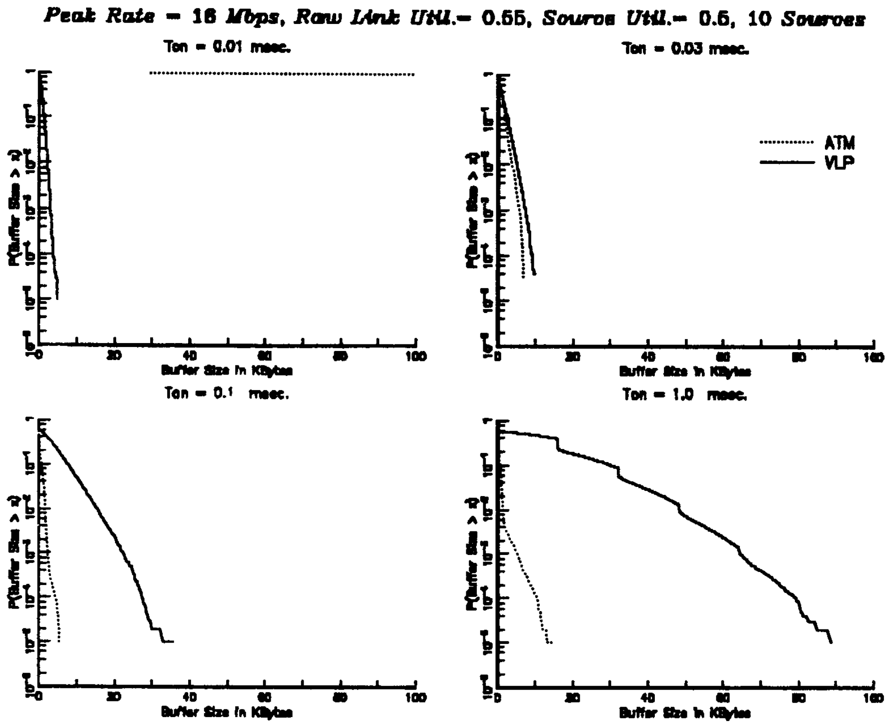

Figure 5. Impact of Frame Size, Header and Padding Bits Overhead

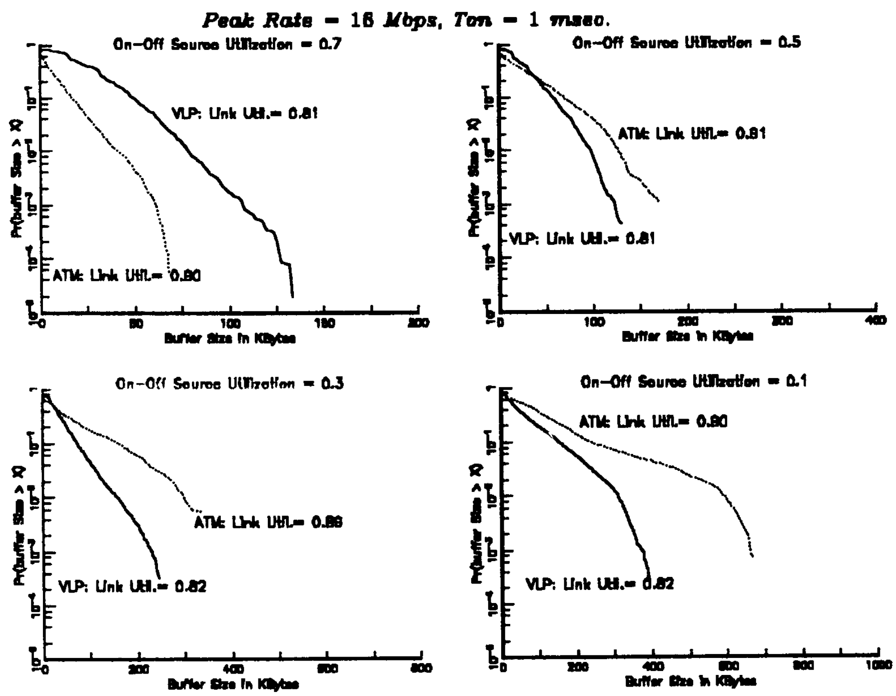

Figure 6. Impact of Source Utilization

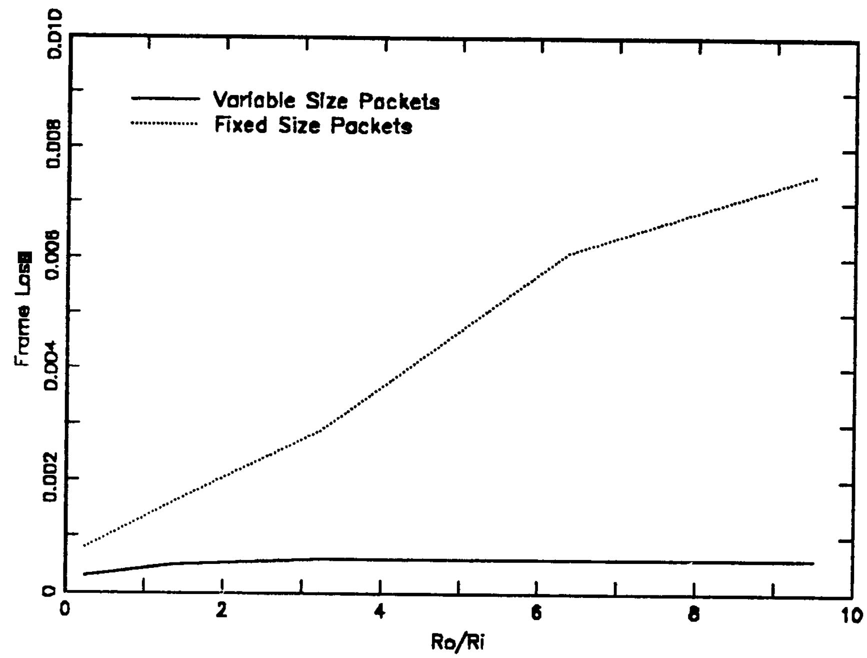

Figure 7. Sensitivity of Frame Loss to Ratio of Link Speed to Source Peak Rate for fixed GOS: Data loss is kept constant, R_o is link speed, and R_i is source peak rate.

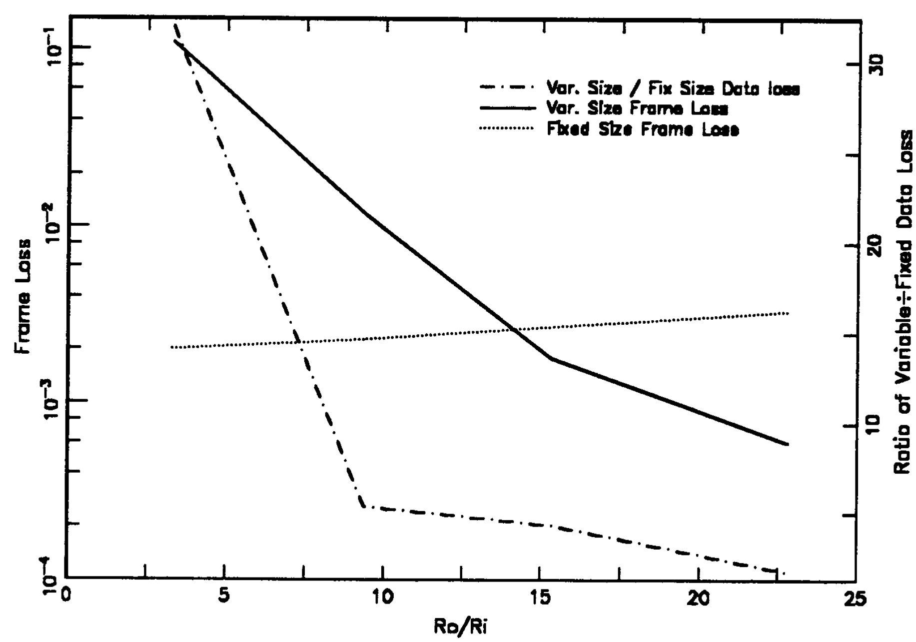

Figure 8. Sensitivity of Frame Loss to Ratio of Link Speed/Source Peak Rate for Fixed Buffer

2c.2.10