Trace Traffic Integration into Model-Driven Simulations

Sponsor: Sprint

Kert Mezger
David W. Petr

Technical Report TISL-10230-10

Telecommunications and Information Sciences Laboratory
Department of Electrical Engineering and Computer Sciences
University of Kansas

May 1995

Contents

1 Introduction

1.1 Trace Traffic Background

In Phase 2 of the Sprint Evaluation of Broadband Networking Technologies Project [2], traffic models were used which were statistical in nature. The models attempted to capture the statistical properties of the source based on source characteristics (e.g. burstiness, interarrival distribution, etc.). While this helps simplify the process of simulating the network, real trace traffic captures the effects of real traffic characteristics. Therefore, when presented with the opportunity of using real trace traffic by obtaining trace Ethernet data files and trace video files from Bellcore, this proved an ideal situation in which to compare and contrast them with the existing statistical models. The Ethernet files [1] consisted of four files entitled “pAug”, “pOct”, “OctExt”, and “OctExt4”. The first two files are internal Ethernet traffic traces. They are two different trace routines which were taken from the same point on the network but at different times. They measure an average rate of 1.1 Mbps and 2.9 Mbps over about 3143 seconds and 1760 seconds respectively. The last two files are external Ethernet traffic traces. They again are two different trace routines which were taken from the same point on the network but at different times. They measure an average rate of 9 Kbps and 27 Kbps over about 122798 seconds and 75943 seconds respectively. The format of these files is such that each line consists of a time stamp and the number of bytes (packet size) that had passed through since the last time stamp.

As for the video trace files [1], 60 files were found in which each file consisted of 2 minutes worth of discrete-cosine-transform coded video trace information. The whole of the 60 files are the first two hours of the movie “Star Wars”. The format comes in two styles. One style is the frame style. This style gives the number of bytes per frame time that are produced. There is one frame per line. With 24 frames per second, the frame time is roughly 42 milliseconds. The second style is the slice style. A slice is defined as 1/30th of a frame. Therefore, each line has the number of bytes per slice time that are throughput. The slice time roughly corresponds to 1.4 milliseconds.

In order to incorporate these files into BONeS, the file structures were changed somewhat, but the integrity of the traces was preserved.

1.2 Trace Ethernet Data Analysis

In analyzing the trace Ethernet data files, the purpose was to compare the characteristics of the trace streams to the statistical model used in Phase II. The characteristics used in the statistical model were packet size distribution and interarrival time distribution. A bimodal distribution was used by the statistical model for packet size, and an exponential distribution was used for silence times. In looking at the packet size distribution of the “pOct” file, it was found that the distribution was bimodal, but the distribution was not continuous, but discrete at the packet size values. These values mainly concentrated at the sizes of 1518 bytes, 1090 bytes, 162 bytes, and 64 bytes. To see a continuous distribution, any packets which were separated by two milliseconds or less were grouped together as one packet and the interarrival times were added together. This gave more of an exponential distribution to the packet sizes. So, without knowing more about the nature of the discrete packet sizes, it is assumed that the discrete bimodal distribution justifies the statistical continuous bimodal distribution. As for interarrival times, Figure 1 and 2 show an exponential-looking distribution without packet grouping which justifies the statistical modeling of an exponential interarrival time. As a note, the file had problems with "over-lapping" LAN packets. The silence time between some consecutive packets were calculated to be negative. Since these consecutive overlapping packets were very rare occurrences, the packets were concatenated to make one packet. This preserves the amount of data represented by the file.

The model used in Phase 2 assumed that packet arrivals were independent. However, further analysis of the trace data traffic showed that the burstiness of the traffic is much greater than that of the model-driven data traffic. Specifically, packet arrivals tended to be highly correlated, indicating that large source bursts were being divided into many consecutive packets. Two tests were performed which concatenated packets with less than 2 milliseconds and 5 milliseconds of silence time between them to reconstruct the original source bursts. The maximum burst size (MBS) of the resulting traffic stream was recorded. The model-driven source

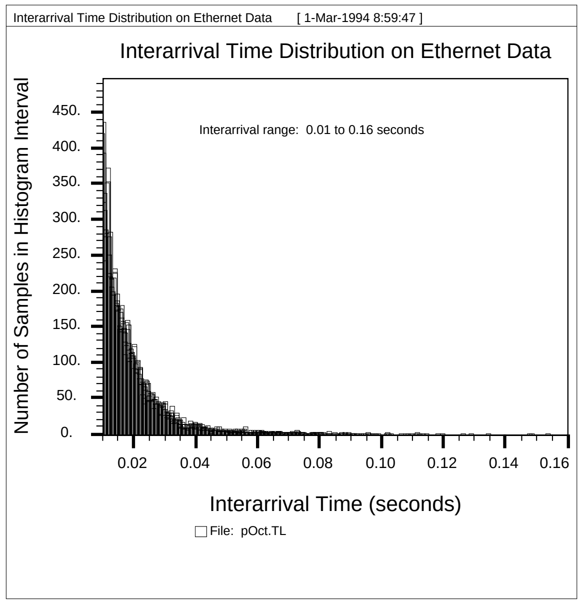

Figure 1: Interarrival Time Distribution on Ethernet Data (large scale)

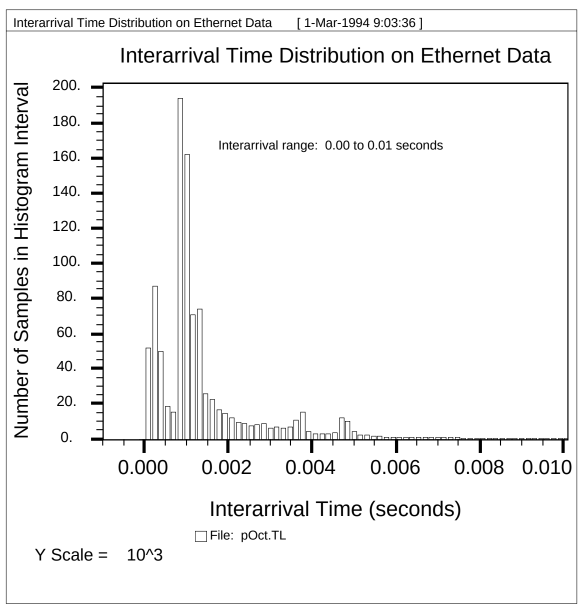

Figure 2: Interarrival Time Distribution on Ethernet Data (small scale)

used a MBS value of approximately 100 cells to obtain a 5% marking by a leaky-bucket policer [2]. The concatenation tests resulted in a MBS of 3375 cells for the 2 millisecond test and 10431 cells for the 5 millisecond test. These results are at least an order of magnitude larger than the policing parameter in the model-driven case. Therefore, it was decided that direct simulation comparison between the modeled data source and the trace data source would be meaningless.

1.3 Trace Video Analysis

The trace video underwent a similar analysis. Since the interarrival time of the trace was fixed at a constant value depending on whether the traffic was by the slice or by the frame, the only analysis could be based on packet size (frame or slice) distribution. Distributions for the first two video files are shown in Figures 3 and 4. The packet sizes are bits per video slice. It should be noted that distributions are not consistent even within the same movie. This is not surprising given the wide variety of scenes in such a movie. The problem with comparing these distributions with those of the statistical sources is that the statistical source holds the packet size constant and varies the interarrival time based on an exponential distribution. Therefore, simulations are necessary to compare the performance of the two sources.

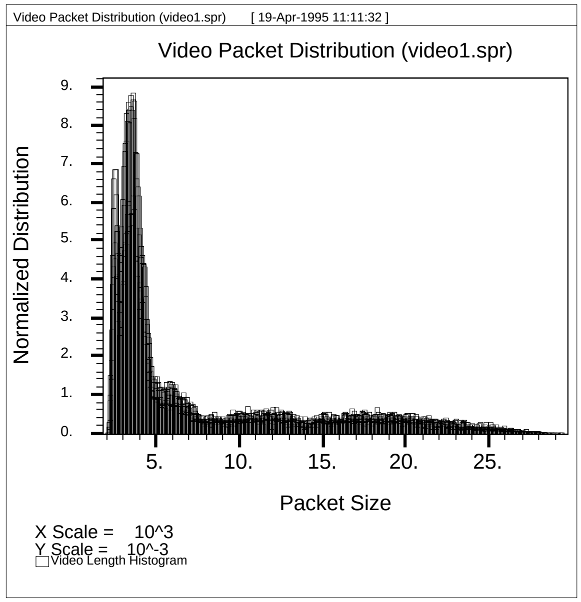

Figure 3: Video Packet Distribution (video1.spr)

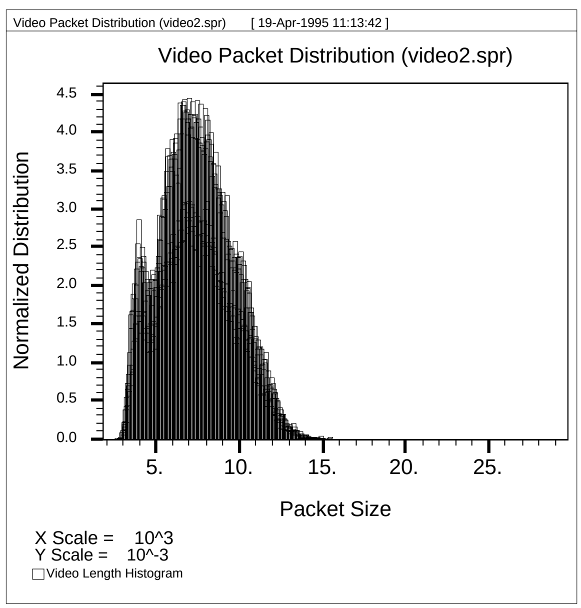

Figure 4: Video Packet Distribution (video2.spr)

2 Simulation Comparison Study

2.1 Trace Video Simulation Background

The previous Phase 2 study involved ATM networks with heterogeneous traffic in which the sources (voice, video, data, and image) were based on statistical models of the respective traffic types [2]. The results of this work form the basis of the present work. Using the model shown in Figure 5, new source models were created for the video traffic which exactly mimicked the traffic flow as described by the trace information. These new sources were then substituted for the old statistical models and the top level simulation was rerun.

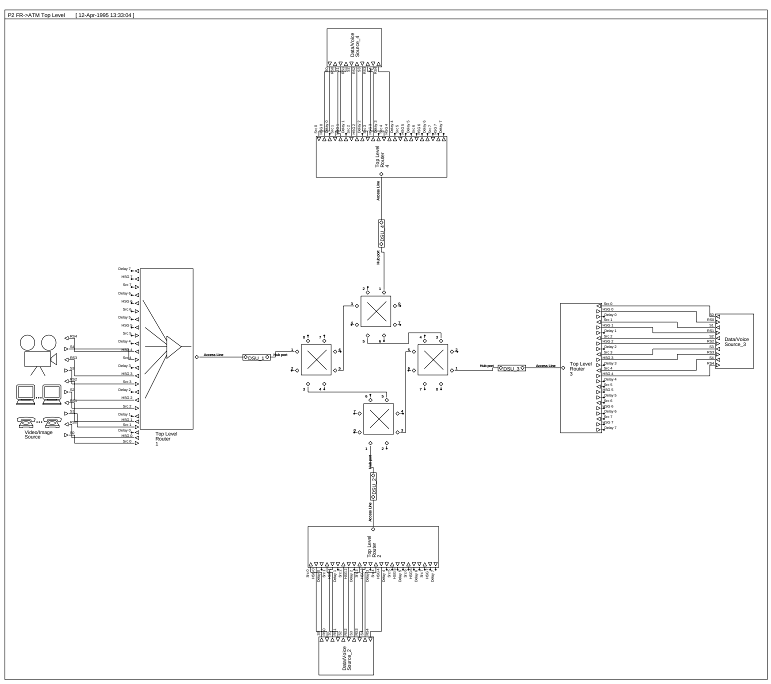

Figure 5: Top Level Diagram of Phase II

Three sets of simulations were performed. The only difference between the sets was the video traffic source. All other parameters, including the parameters of the voice, data, and image sources, were held constant between the sets. The first set of four simulations (with different queue buffer sizes) involved the statistically modeled source with a leaky bucket policer which limited the burstiness of the video traffic to a maximum burst size (MBS) of 5.1 cells. Traffic exceeding this burst size was marked with a Cell Loss Priority (CLP) of 1; conforming traffic was assigned CLP = 0. The choice of MBS = 5.1 cells yielded approximately 5% of the cells marked CLP = 1. These simulations essentially repeated those done in [2].

The other simulations involved the trace driven source with a leaky bucket policer which also marked as CLP = 1 video traffic exceeding a maximum burst size. Four of these simulations had a maximum burst size of 5.1 cells as in the original study, but this produced a CLP = 1 marking rate of 35%. The others had a maximum burst size of 27.0 cells which was chosen to return the marking rate to approximately 5%. Each of the three simulation sets consisted of four simulations corresponding to different buffer sizes. The parameters of the simulations were the same as those of the 98% load case in [2] except for the video source rate. In order to match the statistically modeled source with the trace driven source, the average video modeled source rate was changed from 5 Mbps to 5.28 Mbps.

It was discovered in the course of this study that significant traffic shaping occurred at the router. The video traffic rate into the router segmenter was limited to

5 Mbps. Since the average video traffic rate is very close to this, much smoothing is done. To see how burstiness affects parameter set-up, a test was run on the modeled and trace video traffic with a 20 Mbps segmenter input rate.

2.2 Results

For the initial comparison (including the video traffic smoothing), a cell loss ratio was calculated for each of the sources as the quality of service measure. This measure was used to determine if the burstiness of the trace video information was similar to that of the statistical video traffic. Figures 6 through 13 show the results of the simulations.

Figure 6 shows the cell loss of the image traffic to be slightly higher due to interference from the modeled video traffic as compared to the trace traffic. Similarly, the trace traffic causes slightly less cell loss to the “standard” data traffic than does the modeled source, as seen in Figure 7. Reserved data with CLP=1 (as marked by its policer) shows similar trends in Figure 8. At higher buffer sizes, though, the cell loss ratios are almost the same. The reserved data with CLP=0 (not marked by its policer) has an overall lower cell loss ratio than with CLP=1 (as seen in Figure 9) due to priority discarding, but the difference between the three lines is again small. In looking at Figures 10 and 11, the modeled source does tend to show a higher cell loss ratio for the voice traffic but the difference compared to the trace driven source is small. Again, CLP=1 shows a higher cell loss ratio than does CLP=0. This also occurs for the video traffic itself in Figures 12 and 13. Both figures show the modeled source with a slightly higher cell loss ratio than the trace source, but the trace source with a maximum burst size of 27.0 cells more closely matches the cell loss ratio of the modeled source. Further analysis of the simulations showed that the *actual* video throughput of the trace simulation tended to be somewhat smaller than the modeled video throughput. We believe that this is the primary reason that the trace simulation loss rate are smaller than the modeled source simulation loss rates.

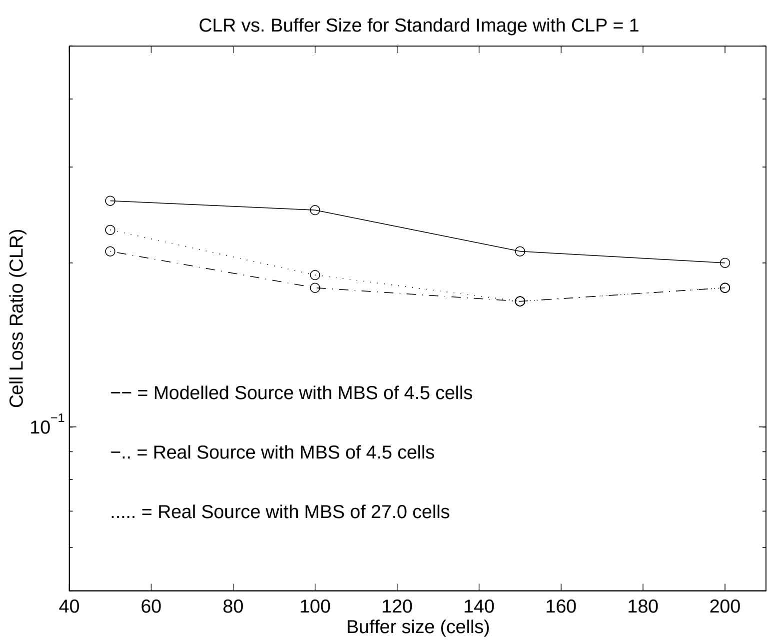

Figure 6: CLR vs. Buffer Size for Standard Image with CLP = 1

When the shaper rate was increased to 20 Mbps, the MBS of the video traffic (whether trace or modeled) which yielded a marking rate of approximately

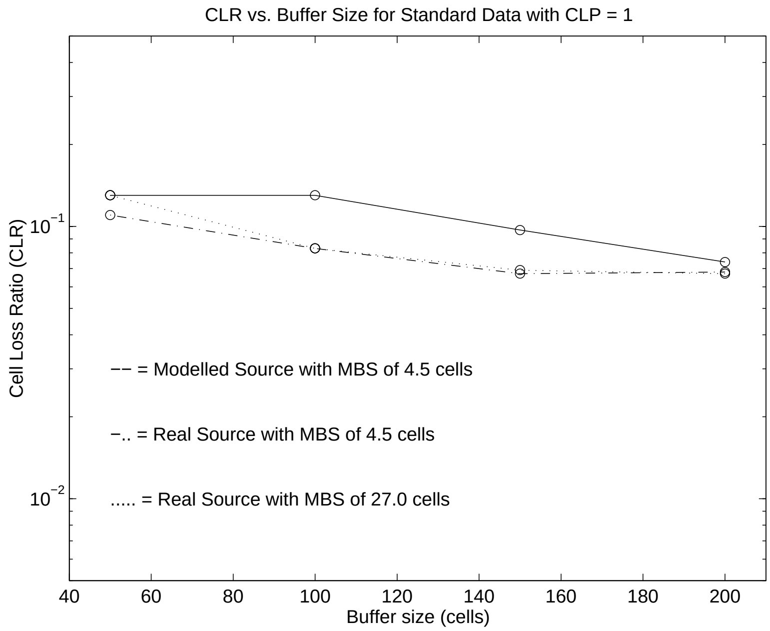

Figure 7: CLR vs. Buffer Size for Standard Data with CLP = 1

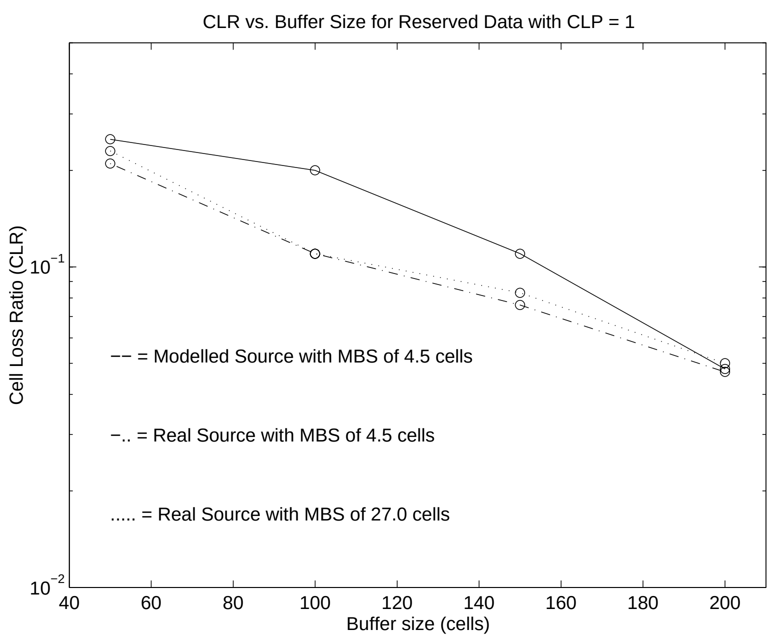

Figure 8: CLR vs. Buffer Size for Reserved Data with CLP = 1

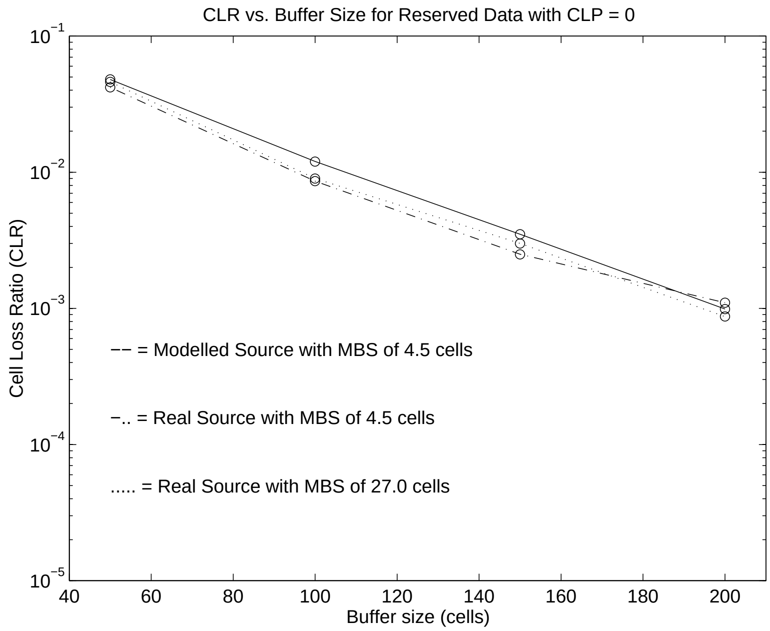

Figure 9: CLR vs. Buffer Size for Reserved Data with CLP = 0

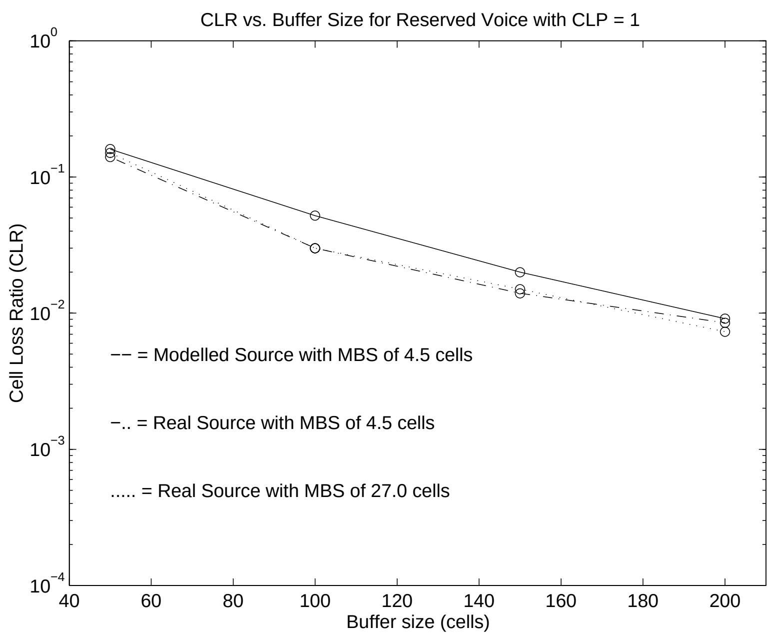

Figure 10: CLR vs. Buffer Size for Reserved Voice with CLP = 1

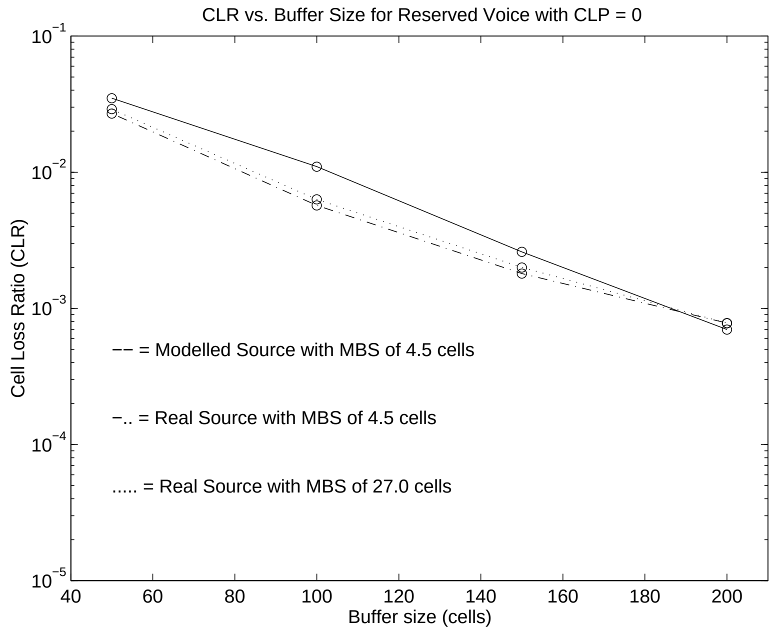

Figure 11: CLR vs. Buffer Size for Reserved Voice with CLP = 0

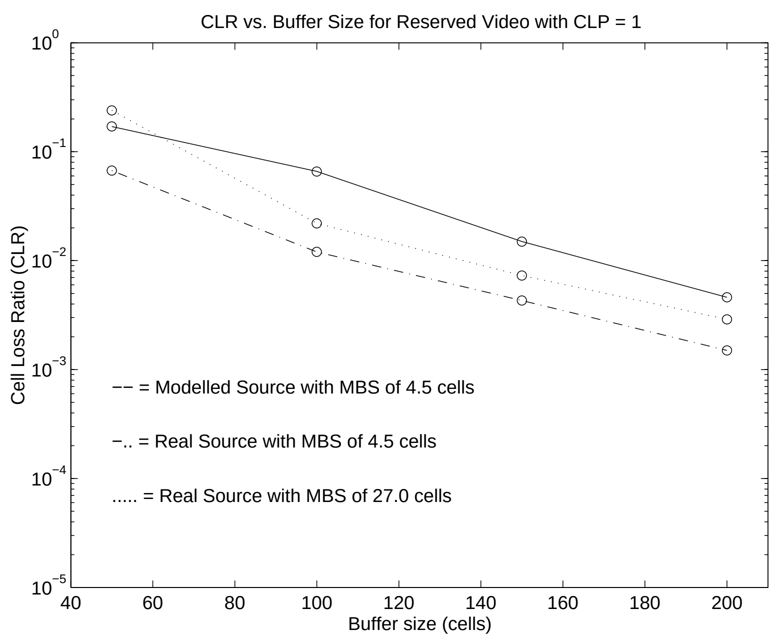

Figure 12: CLR vs. Buffer Size for Reserved Video with CLP = 1

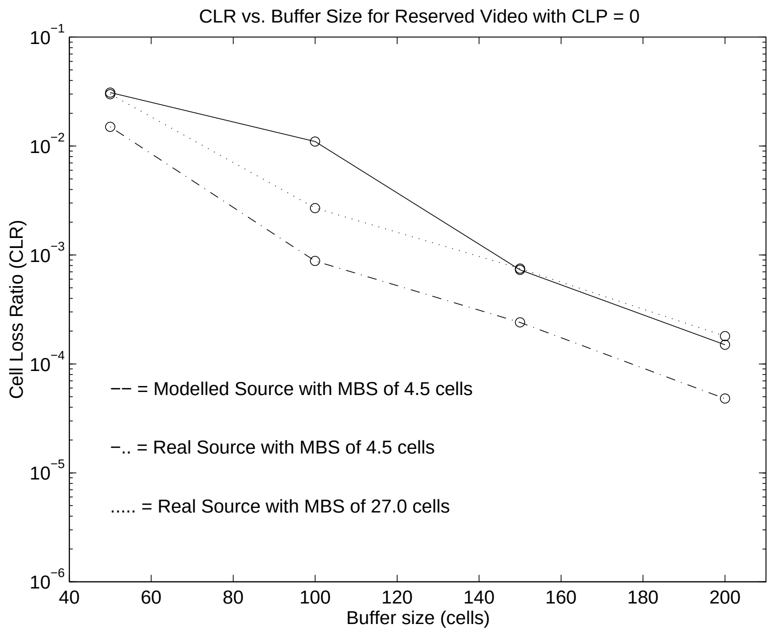

Figure 13: CLR vs. Buffer Size for Reserved Video with CLP = 0

5% increased in size. The modeled video traffic MBS increased from 5.1 cells to 15.0 cells. The trace video traffic MBS increased from 27.0 cells to well over 5000.0 cells. The reason for this dramatic increase experienced by the trace traffic is due to video traffic rate over time. The average video rate for 1 second intervals for the first video trace file is shown in Figure 14.

To improve the marking rate while lowering the MBS requirement, the maximum sustainable rate was increased to 10 Mbps plus overhead. The 10 Mbps is associated with the large video rates in the interval between 60 seconds and 80 seconds. This lowered the MBS to 85.0 cells to get a marking rate of approximately 5% during this peak rate period. Of course, the marking rate would be considerably

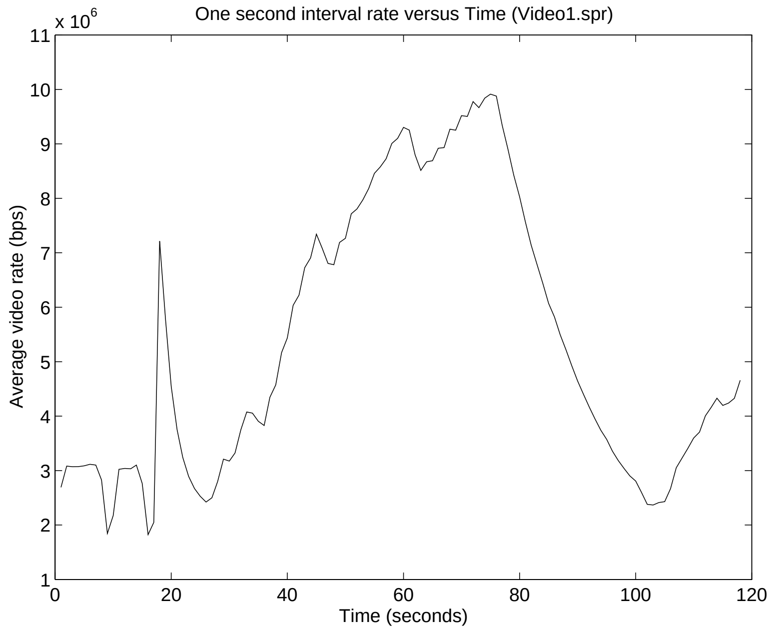

Figure 14: One second interval rate versus Time (Video1.spr)

less for the other portions of this file. This illustrates the problems associated with analysis of sources that have such wide and sustained variations in their traffic characteristics. This behavior is characteristic of source traffic that possesses long-range dependency [1].

2.3 Conclusions

With the Phase 2 parameters (i.e. the 5 Mbps shaping for the video traffic), it should be noted that the video traffic is severely shaped. Under this condition, the traffic model used does not matter. That is, the cell loss ratios will be similar regardless of the model.

When the shaping parameter is increased, the burstiness of the traffic is kept more intact. Therefore, shaping drastically changes the MBS associated with different marking rates. The long-range dependence exhibited in the trace traffic is not preserved in the modeled source. This shows that the modeled source used in Phase 2 is very poor in capturing the burstiness of real video traffic.

More work is required is required to understand how best to deal with sources like the video and LAN trace data, especially in a simulation environment.

References

[1] Yong-Qing Lu, David W. Petr, Victor S. Frost. *Characterization and Modeling of Long-Range Dependent Telecommunication Traffic*. TISL Technical Report TISL-10230-04

[2] David W. Petr, Victor Frost, Ann Demirtjis, Cameron Braun. *Evaluation of Broadband Networking Technologies: Phase II Report*. TISL Technical Report TISL-9750-4.