Statistical Multiplexing of Layered Video Streams over ATM Networks with Leaky-Bucket Traffic Descriptors

Arthur W. Berger

AT&T Bell Laboratories
Holmdel, New Jersey 07733-3030

Samuel P. Morgan

AT&T Bell Laboratories
Murray Hill, New Jersey 07974-0636

Amy R. Reibman

AT&T Bell Laboratories
Holmdel, New Jersey 07733-3030

ABSTRACT

We consider the transmission of variable-bit-rate (VBR) video over Asynchronous Transfer Mode (ATM) networks using statistical multiplexing. The encoder output is assumed to comply with a leaky-bucket traffic descriptor, and the network uses cell-level but not burst-level buffering at switches. Feasible multiplexing levels are determined based on given requirements of cell loss and the duration of periods of consecutive cell loss. Single-layer VBR encoding is compared with two-layer encoding, where the enhancement layer is coded VBR, and the base layer is coded using various bit-rate constraints.

Compared to a one-layer constant bit-rate (CBR) encoding, statistical multiplexing gains for VBR of between 1.5:1 and 2:1 appear to be possible for teleconferencing connections over a 155 Mb/s network. Two-layer encoding with a peak-bit-rate constraint on the base layer is sometimes superior to one-layer encoding, and it is superior to two-layer encoding with a constant-bit-rate base layer. However, for the encoding algorithms considered, we did not demonstrate a significant quantitative advantage of 2-layer encoding over 1-layer VBR encoding.

August 5, 1993

Statistical Multiplexing of Layered Video Streams over ATM Networks with Leaky-Bucket Traffic Descriptors

Arthur W. Berger

AT&T Bell Laboratories
Holmdel, New Jersey 07733-3030

Samuel P. Morgan

AT&T Bell Laboratories
Murray Hill, New Jersey 07974-0636

Amy R. Reibman

AT&T Bell Laboratories
Holmdel, New Jersey 07733-3030

1. INTRODUCTION

Traditionally, video has been transmitted across circuit-switched channels that have a constant bit rate. Since video encoders typically produce a variable bit rate, in part due to the use of a variable-length code to improve compression, a buffer is used to translate the variable-rate output of the encoder into the constant rate required by the channel. From time to time, the video quality may have to be reduced to ensure that the encoder buffer does not overflow. If video were transmitted across an inherently variable-rate channel such as might be provided by an Asynchronous Transfer Mode (ATM) network, one could in principle achieve either constant video quality, reduced buffering and delay requirements at the encoder and decoder, or statistical multiplexing gains.

It is straightforward, via an appropriate queueing discipline, to multiplex a single variable-bit-rate (VBR) video channel with a delay-insensitive service such as data traffic. However, if a substantial fraction of the traffic on an ATM network consists of VBR video, it may be necessary to statistically multiplex the video channels with each other to achieve reasonable utilization. Statistical multiplexing will result in a nonzero probability that data is lost or, what is the same thing, delayed until the decoder can no longer use it. Without some form of protection, such as a suitably designed coding scheme, one lost ATM cell could significantly affect the image quality of many subsequent frames. Therefore, before ATM networks can be successfully used to transmit VBR video, methods to minimize the effect of lost cells must be found.

One promising technique to minimize the effect of lost cells is to transmit using two priorities. The ATM network allows the end-terminal to mark each cell as either high or low priority. If congestion develops, the network drops low-priority cells before high-priority cells. Video is well suited to this, since many techniques are possible to separate the video data into more and less important parts [1-11]. The more important part is essential for the decoder to produce a minimally acceptable image, and is commonly called the base layer. The less important part is used to improve the quality of the base-layer image; it is called the enhancement layer. The base layer is transmitted at high priority and the enhancement layer is transmitted at low priority.

However, there is usually some penalty associated with encoding video using two layers, such that the base layer has acceptable quality all by itself. Coding and decoding are somewhat more complicated. Also, the combined average bit rate of the two layers is generally larger than the average bit rate of a similar quality one-layer encoding. If so, the per-channel bit rate allocation by the network may have to be larger for a two-layer encoding than for a one-layer encoding, particularly if the number of multiplexed channels is sufficiently high. Recent work using a base layer with constant bit rate [6, 7, 12] has suggested that at least for some test sequences, two-layer encoding is less efficient for the network than one-layer encoding.

Morgan and Reibman [13] have made a systematic comparison of the multiplexing behavior of one- and two-layer video codecs for teleconferencing. However, they used the conventional frame-by-frame multiplexing model, in which independent sequences of temporally synchronized frames are served from a common multiplex buffer. This model may accurately represent situations where several colocated video encoders are under central management [14], but it is less realistic when a network wants to multiplex traffic from independently managed encoders. In the latter case, studies of potential multiplexing gains [13, 15 et al.] have assumed that some kind of source policing would be done, but have not specified details.

In recent years, a consensus has emerged [16, 17] as to how traffic will be carried in

Broadband Integrated Services Digital Networks (B-ISDN) based on ATM. Namely, when a new connection is set up, the user agrees to comply with a negotiated traffic descriptor (TD) that characterizes the traffic that it intends to submit to the network. The network enforces the TD via the Usage Parameter Control (UPC) function, informally known as the policing function. ATM cells that do not comply with the TD may be dropped by the UPC, or alternatively they may be tagged as low-priority cells and discarded within the network if congestion arises.

For its part, the network undertakes to deliver compliant traffic subject to agreed-on bounds on loss ratio, delay, and/or delay jitter. The standards do not suggest how this performance is to be achieved. Similarly, the Connection Admission Control (CAC) function, which determines whether a new connection with a given TD can be safely accepted in the presence of existing connections, is still under active research. Connection admission control is outside the scope of this paper.

The standardization of TDs is a subject of current work in CCITT, T1S1 and other standards bodies. To date, CCITT has only defined a peak cell rate [16]. The definition is made specific via a measurement procedure that is based on the leaky-bucket (LB) algorithm. A likely additional TD is one for the sustainable cell rate, or, informally, the "negotiated average rate", which would bound the possible compliant long-term average rates. Although this has not yet been agreed to in a standards body, an industry forum, the ATM Forum, has specified the sustainable cell rate in terms of, again, the LB algorithm. The LB algorithm also specifies a "maximum burst size" for ATM cells. Reibman and Berger [17] have computed LB parameters for samples of free-running one-layer video encoder output, and Reibman and Haskell [18] have shown how to make a one-layer encoder comply with any preassigned LB parameters.

In the present paper we investigate traffic descriptors and potential statistical multiplexing gains for a one-layer and a variety of two-layer video encoders. The comparison is based on a one-layer constant-bit-rate (CBR-1) encoding that has equivalent quality. In the two-layer encoders, the enhancement layer is always assumed to have a variable bit rate, but as in [13] there are various possibilities for the base layer. Namely, the base layer can be encoded using a variable bit rate, leading to a VBR-VBR algorithm, or using a constant bit rate (CBR-VBR), or using a peak-bit-rate constraint (PBR-VBR). In contrast to [13], we consider traffic-shaped rather than frame-synchronized sources, and in contrast to [13] and [17], we impose conditions not only on the cell loss ratio, but also on the duration of intervals of consecutive cell loss.

The paper is organized as follows. Section 2 describes the one- and two-layer video codecs that we have simulated. Section 3 describes leaky-bucket traffic descriptors and usage parameter control for ATM networks. In Section 4 we discuss cell delay jitter and cell loss. We estimate the mean cell loss ratio of a link carrying video calls that the codec has shaped to be compliant with a leaky-bucket traffic descriptor, as well as the mean length of congestion periods, that is, periods of consecutive cell loss due to overload. In Section 5 we show frame-size sequences that are produced when the various encoding algorithms are applied to two 5-minute samples of a videoconferencing session. In Section 6 we compute sets of traffic-descriptor parameters with which the empirical traffic streams are compliant, and in Section 7 we estimate the maximum statistical multiplexing gains that might be achieved with these streams. Section 8 contains conclusions and discussion.

2. VIDEO ENCODING ALGORITHMS

A generic packet video system is shown in Figure 1. A video signal is applied to the video encoder, which produces an encoded video bit-stream. The number of encoded bits produced by frame i is E_i. The encoded bit-stream is stored in the encoder buffer before being transmitted via the channel interface to the network. The rate control device selects R_i, the number of bits transmitted on the channel during frame period i, such that no video buffer constraints will be violated and no traffic that exceeds the negotiated TD parameters is submitted to the network [18]. In addition, the rate control device also selects the quantizer step size used by the encoder. After being transmitted across the network, the video bit-stream is stored in the decoder buffer. It is then input to the video decoder, which outputs a video signal.

2.1. One-layer Video Encoder

The specific one-layer codec that we have considered has syntax compatible with H.261 [19], which uses conditional replenishment, motion compensation and the discrete cosine transform (DCT). The codec uses exhaustive motion estimation with ±15 search range, a constant quantizer step-size, and intra/inter/motion-compensation decisions for each macroblock as in Reference Model 8 (RM8) [20]. The first frame is coded intraframe, and all remaining frames are coded predictively. Within each frame, 3 macroblocks are transmitted intraframe to avoid accumulation of inverse DCT mismatch between encoder and decoder. To generate a VBR output, the codec uses a constant quantizer step-size of $Q_{\min}=8$, which produces output video that has essentially constant quality.

2.2. Two-layer Video Encoder

The two-layer algorithms are identical except for the rate control. The basic two-layer algorithm is a requantization-type algorithm, as shown in Figure 2 [3]. The base-layer encoder appears inside the dashed box; the enhancement-layer encoder is outside the box. In a requantization algorithm, the base layer coarsely quantizes the original data after motion compensation, while the enhancement layer more finely quantizes the difference between the original data and the base-coded data. No motion compensation is used to reduce the amount of enhancement information, so that errors due to enhancement-layer cell losses, which may be orders of magnitude more frequent than base-layer cell losses, will affect only individual frames. Here, we use H.261 syntax in both layers, although the requantization algorithm can also be implemented using the emerging MPEG-2 standard using spatial scalability with two identical-resolution layers and no temporal prediction in the enhancement layer.

Two-layer coding using the requantization algorithm is less efficient than one-layer coding with the same quality. Both the one-layer algorithm and the base layer of the two-layer algorithm use interframe coding with motion compensation, while the enhancement layer is generated with a less efficient intraframe algorithm. Therefore, to obtain the same quality, the two-layer algorithm generally requires a larger total bit rate than a one-layer algorithm.

For the two-layer algorithm, we examine three different rate constraints on the base layer: no constraint (VBR-VBR), a constant-rate constraint (CBR-VBR), and a peak-rate constraint (PBR-VBR). In the latter two cases, the buffer shown in Figure 2, together with a rate control algorithm, is used to smooth the base bit rate transmitted onto the network. Depending on the constraint, the quantizer step size is increased to reduce the base-layer bit rate when the buffer fills, and it is decreased to increase the base-layer bit rate when the buffer empties. In all cases, the enhancement layer has no rate constraint.

For VBR-VBR, there is no constraint on the bit rate of the base layer. If Q_{min} is the quantizer step size of the enhancement layer, then we choose the quantizer step size of the base layer to be $Q_{min}+2$, since this is the next allowable step size that is immediately larger than Q_{min}. Since both the base and enhancement layers have constant quantizer step size, they each have essentially constant quality. Furthermore, since the enhancement layer uses the same Q_{min} as the previous one-layer algorithm, this two-layer algorithm and the one-layer algorithm have comparable quality when there are no cell losses.

In the second case, CBR-VBR, the base layer is constrained to have a constant bit rate. As in H.261, we define the rate using the variable P, where a constraint of P means constant bit rate of $P\times 64$ kilobits per second (kbps). The rate control in the base layer is identical to that in RM8, where the quantizer step size is selected once every 11 macroblocks based solely on the fullness of the buffer.[1] Therefore, the quantizer step size may be smaller than Q_{min}. However, the enhancement layer still uses a constant quantizer step size of Q_{min}. Therefore, no DCT data remains to be coded in the enhancement layer if the base layer uses a quantizer step size less than Q_{min}.

Because the quantizer step size in the base layer may be less than Q_{min}, the quality of the CBR-VBR video may occasionally be better than the other cases we examine. However, we have selected $Q_{min}=8$ because this produces acceptable quality to most viewers. It is difficult to see

[1] More sophisticated rate control algorithms are possible; however, RM8 provides a straightforward and generally available basis for comparison.

any visual improvement when the quantizer step size decreases below 4.

The third rate constraint we examine for the base layer is a peak constraint; hence, the abbreviation PBR-VBR. The base layer is constrained to have bit rate no larger than a peak value. Again, we use P to define the constraint. The rate control in the base layer is identical to that in RM8, except that the quantizer step-size is not allowed to be smaller than Q_{min}. Therefore, the instantaneous base bit rate may be lower than its peak. The enhancement layer is coded with constant quantizer step size of Q_{min}.

3. TRAFFIC DESCRIPTORS AND USAGE PARAMETER CONTROL

The only traffic descriptor that we consider here is the leaky bucket, since for video sources the leaky-bucket algorithm appears to be superior to other algorithms such as sliding window [17, 21]. The leaky-bucket algorithm can be defined in several equivalent ways. Here we consider the leaky bucket to be an imaginary FIFO buffer of size B^{max} bits with constant drain rate $\bar{R}$ bits per frame time. The content of the imaginary buffer is incremented in proportion to the arriving traffic up to the buffer capacity, B^{max}. The portion of traffic that would cause the bucket capacity to be exceeded is noncompliant and is assumed to be discarded or marked for lower priority. Note that no traffic is actually queued in the imaginary buffer.

Let B_i be the content of the imaginary buffer (measured in bits) at the end of frame period i. We assume that the arriving bits are spread uniformly over the frame time; generalization to more bursty arrival processes would be straightforward. Then, allowing for the possibility of underflow or overflow,

$$B_i = \min(B^{max}, \max(B_{i-1} + R_i - \bar{R}, 0)) \qquad B_0 = 0 . \tag{1}$$

The bucket does not overflow, that is, the traffic is in compliance with the leaky-bucket TD defined by $\bar{R}$ and B^{max}, if

$$B^{max} \geq B_{i-1} + R_i - \bar{R} , \quad \forall\, i . \tag{2}$$

In practice, a TD will include not only a bucket size B^{max} and a negotiated average rate $\bar{R}$, but also a peak rate $R^{\max}$. For purposes of this paper we do not need to use the leaky-bucket definition of peak rate in [16], but rather can simply say that the peak bit rate is $R^{\max}$.

Before entering the network proper, the offered traffic must pass through a Usage Parameter Control owned by the network. The UPC determines whether the traffic is in compliance with the negotiated TD. A cell determined to be compliant with the TD passes through the UPC without modification and without delay. A cell determined to be noncompliant is either discarded by the UPC, or marked with a bit in its header that permits it to be discarded within the network if congestion exists.

In practice, it is expected that a video user will request a connection having TD parameters that are known, on the basis of earlier statistical studies, to be approximately what the encoder requires for typical traffic of the given class (e.g., videoconferencing). There may be some negotiation with the network before the call is accepted, but once having been assigned TD parameters, the encoder will emulate the operational definition of the TD and will ensure that the output of its encoder buffer never exceeds the TD limits [18].

For a two-layer encoder, as in Figure 2, we expect that the stream of base cells will be subject to a negotiated peak rate $R^{\max}_{\text{base}}$, a negotiated average rate $\bar{R}_{\text{base}}$ and a negotiated bucket size $B^{\max}_{\text{base}}$. Similarly we expect that the merged stream of base cells and enhancement cells will be subject to a negotiated peak rate $R^{\max}_{\text{sum}}$ where $R^{\max}_{\text{sum}} \ge R^{\max}_{\text{base}}$, a negotiated average rate $\bar{R}_{\text{sum}}$ where $\bar{R}_{\text{sum}} \ge \bar{R}_{\text{base}}$, and a negotiated bucket size $B^{\max}_{\text{sum}}$. It is essential to the video system that the UPC, which polices both the base stream and the merged stream, neither discard or demote base cells, inasmuch as base cells may carry information whose loss would disrupt the decoding sequence for a considerable time. It will therefore be necessary for the encoder to monitor both the base-cell stream and the merged stream, so as never to present any noncompliant cells to the UPC. However we shall not propose specific implementations.

4. DELAY JITTER AND CELL LOSS

Cell delay between the output of the encoder buffer and the input of the decoder buffer in Figure 1 includes a constant part due to the finite speed of light together with fixed processing times at network nodes, and a variable part due to queueing delays while the cells of one connection wait for the cells of other connections to be served. In this paper we shall ignore the constant part, which for a transcontinental connection amounts to a few tens of milliseconds. However, the variable part of the delay, that is, the delay jitter, cannot be ignored for ATM networks. The jitter has to be absorbed in the decoder buffer of Figure 1. The decoder buffer must allow for delay variations equal to the maximum tolerable delay jitter. Cells having larger delay variations cannot be used by the decoder, so there is a tradeoff between maximum tolerable delay jitter and effective cell loss ratio.

In this section we discuss cell delay jitter, distinguishing, as is customary for ATM networks, between cell-level and burst-level effects. We also estimate the mean cell loss ratio of a link carrying traffic compliant with a leaky-bucket traffic descriptor, assuming no burst-level buffering, in terms of the number of multiplexed connections. Finally, we estimate the mean length of congestion periods, that is, periods of consecutive cell loss due to overload, in terms of leaky-bucket parameters.

4.1. Cell-Level Jitter

Cell-level jitter occurs when the cells of the connections sharing a link arrive at constant or nearly constant rates, and the capacity of the link is sufficient to handle the merged arrival rate. A cell belonging to a given connection can still be delayed while the cells of other connections are served, but the maximum delay will be comparable to the average intercell interval on the given connection. The maximum cell-level jitter can be rigorously bounded for certain queueing disciplines [22, 23]. An example of such a discipline is Framed Round Robin.[2]

Framed Round Robin divides the capacity of a high-speed link into channels whose rates

[2] Framed Round Robin is a special case of Hierarchical Round Robin, as described in [22].

are rational fractions of the link rate, while at the same time allocating any unused capacity to best-effort traffic. In Framed Round Robin, a group of N video connections sharing an ATM link are given a total of N cell services during each successive link frame time T_F. T_F is a network quantity, not to be confused with the video frame time, usually 1/30 second, which we shall call T_V. Effectively, the multiplexed video connections share a link whose rate is N/T_F cells per second. The bit rate is not perfectly uniform, as it would be over a synchronous link, but the requisite cell services are always completed by the end of the link frame time. If, as is usually the case, the link can transport more than N cells per link frame time, the remaining cell service slots are given to other services that need not concern us.

Suppose that N_V is the average number of cells in a video frame and T_V is the video frame time. If, on the average, we expect each link frame to carry one cell per video connection, the link frame time must satisfy

$$T_F = T_V/N_V . \tag{3}$$

Since it can be shown [22] that the maximum delay jitter introduced by a Framed Round Robin switching node is $2T_F$, the maximum delay jitter $J^{\max}$ introduced by a chain of H nodes satisfies

$$\frac{J^{\max}}{T_V} \leq \frac{2HT_F}{T_V} = \frac{2H}{N_V} . \tag{4}$$

As will appear in Section 5, average video frame sizes were roughly 120 ATM cells (43200 bits) in one of the two sessions we recorded, and roughly 60 ATM cells (21600 bits) in the other. So if there are 15 switching nodes, which seems like a generous estimate, along a given connection, the maximum cell-level jitter for Framed Round Robin queueing would be from a quarter to a half of a video frame time. It should be feasible to absorb this much jitter in the decoder buffer.

The worst-case delay jitter in Framed Round Robin queueing is proportional to the number of nodes along the path. However, since the delay jitter across a Framed Round Robin connection with a realistic number of nodes appears to be substantially less than one video frame time, there may be no need for the additional complexity of disciplines that provide an edge-to-edge delay jitter independent of the number of switches, at the expense of a larger constant component of delay within the network [24].

4.2. Burst-Level Jitter

Burst-level congestion occurs when the merged arrival rate to a network link exceeds the capacity of the link during a period of several intercell intervals on a typical connection. If there is burst-level congestion, several cells must be buffered for a typical connection at the given node, and the buffered cells may experience queueing delays amounting to several intercell intervals. Since queueing delays are small when congestion is not present, the maximum burst-level queueing delay is also a measure of the burst-level jitter.

In situations where burst-scale effects dominate cell-scale effects by orders of magnitude, fluid models are often used to calculate burst-scale behavior. There are also models that, in principle, permit the calculation of cell-scale and burst-scale effects in a single framework [25, 26], but these models tend to be computationally expensive.

Burst-level queueing is tolerable for delay-insensitive (e.g., data) traffic and may in fact be the preferred way of statistically multiplexing such traffic, since data traffic can ordinarily stand delay fluctuations of the order of seconds. However, if burst-level queueing can occur on a video connection, the corresponding fluctuations in delay will have to be absorbed in the decoder buffer, leading to an increase in the time interval between encoding and decoding operations. For interactive video, it may be desirable to limit such an increase to a few tens of milliseconds. We shall not consider burst-level queueing in this paper, but shall study a model of video transmission in which burst-level effects result in cell losses rather than in delay fluctuations.

4.3. Cell Loss Ratio

Statistical multiplexing of independent sources always involves some nonzero probability of cell loss. According to the scenario of Section 3, each of the multiplexed video streams complies with three traffic-descriptor parameters, namely a peak rate $R^{\max}$, a negotiated average rate $\bar{R}$, and a leaky-bucket size $B^{\max}$. We want to calculate the mean cell loss ratio CLR, that is, the fraction of cells lost from a given stream when some number N of video connections are multiplexed over a link of capacity C bits/frame time and there is no burst-scale buffering. There must, of course, be cell-scale buffering as discussed in Section 4.1. However, we take the conservative viewpoint that the buffer space is small compared to the size of a burst, and assume that if the instantaneous combined rate of a fluid model of the sources is greater than the link rate, then cell losses occur.

To estimate the CLR, we do not use a detailed model of the video source. Rather, we take the approach of a network operator whose only information about the source is the traffic descriptor and who presumes that the traffic will be the "worst case" that is still compliant with the traffic descriptor.

Doshi [27] has recently proved that under reasonably general assumptions, the maximum loss ratio for unbuffered, randomly phased, leaky-bucket-controlled sources occurs when each source alternates between being ON at the peak rate long enough to fill the bucket and then OFF until the leaky-bucket content is zero. The ON and OFF periods and the fraction of time p that each source is ON are given by

$$T_{ON} = \frac{B^{\max}}{R^{\max} - \bar{R}}, \quad T_{OFF} = \frac{B^{\max}}{\bar{R}}, \quad p = \frac{T_{ON}}{T_{ON} + T_{OFF}} = \frac{\bar{R}}{R^{\max}} . \tag{5}$$

Suppose that N ON-OFF sources each go through the same deterministic cycle, except that the starting points of the cycles are randomly distributed. Then the probability that m sources are ON at a random instant is given by $\binom{N}{m} p^m (1-p)^{N-m}$, and the expected cell loss ratio over any interval of length $T_{ON} + T_{OFF}$ is

$$\mathrm{CLR} = \frac{1}{Np} \sum_{m=n_0+1}^{N} (m - C/R^{\max}) \binom{N}{m} p^m (1-p)^{N-m} , \tag{6}$$

where $n_0 = \lfloor C/R^{\max} \rfloor$ is the maximum number of sources that can be ON simultaneously without overloading the link. We shall use Eq. (6) in Section 6.

4.4. Length of Congestion Periods

Many different sets of TD parameters can characterize a given traffic stream. In particular, there are tradeoffs between the negotiated average rate $\overline{R}$ and the smallest compatible bucket size $B^{\max}$. $B^{\max}$ is a decreasing function of $\overline{R}$, being infinite (for an infinitely long traffic sample) when $\overline{R}$ is less than the actual average traffic rate R^{av}, and zero when $\overline{R}$ is greater than the peak rate $R^{\max}$. On the other hand, the average cell loss ratio CLR given by Eq. (6) is an increasing function of $\overline{R}/R^{\max}$, and does not depend on $B^{\max}$. If we were just interested in minimizing CLR, we should pick $\overline{R}$ as small as possible.

The bucket size defines the time scale of the ON-OFF process. Doubling the bucket size doubles the lengths of the ON periods and the OFF periods. It doubles the lengths of the congestion intervals during which cells are being dropped, and it doubles the lengths of the intervals between the congestion intervals, so that the long-term CLR remains constant. Cell losses during a congestion interval may be distributed among all of the multiplexed connections; nevertheless, it may be that frequent short periods of cell loss are more tolerable to users than occasional long periods of cell loss. It is therefore useful to estimate the average length of a congestion period in terms of the parameters of the system. This may be done as follows.

We represent N ON-OFF sources as N arcs each of length T_{ON}, randomly placed around a circle of circumference $T_{ON} + T_{OFF}$. A congestion period is a period during which at least $n_0 + 1$ sources are ON. We assume that $N > n_0$, and to avoid congestion periods of infinite length, we assume that $N\overline{R} < (n_0+1)R^{\max}$.

The mean length of a congestion period is

$$\overline{T}^{\mathrm{cong}} = \overline{L}/\overline{N}\,, \tag{7}$$

where $\overline{L}$ is the expected total length of congestion periods, and $\overline{N}$ is the expected number of congestion periods. $\overline{L}$ is just the circumference of the circle times the probability that there are more than n_0 active users, that is,

$$\overline{L} = (T_{ON} + T_{OFF}) \sum_{m=n_0+1}^{N} \binom{N}{m} p^m (1-p)^{N-m}\,. \tag{8}$$

$\overline{N}$ is N times the probability that a random source starts a congestion period by going ON when n_0 out of $N-1$ sources are already on, since there are N sources and each source has the same probability of starting a congestion period in the presence of $N-1$ other sources. Accordingly,

$$\overline{N} = N \binom{N-1}{n_0} p^{n_0} (1-p)^{N-1-n_0} . \tag{9}$$

Substituting into (7) and carrying out some algebra using (5) gives

$$\overline{T}^{\text{cong}} = \frac{B^{\max}}{R^{\max} - \overline{R}} \sum_{m=n_0+1}^{N} \frac{(N-n_0-1)!\,n_0!}{(N-m)!\,m!} \left[\frac{\overline{R}}{R^{\max} - \overline{R}} \right]^{m-n_0-1} . \tag{10}$$

It can be shown that Eq. (10) is valid for general distributions of ON-OFF periods so long as the probability of infinitely long congestion periods is zero. If $N = n_0 + 1$, the equation takes the intuitively reasonable form

$$\overline{T}^{\text{cong}} = \frac{B^{\max}}{(R^{\max} - \overline{R})(n_0 + 1)} = \frac{T_{ON}}{n_0 + 1} . \tag{11}$$

5. FRAME-SIZE SEQUENCES

The video used in these examples was recorded at an actual meeting. Two sequences are used here, each 5 minutes in length. Since the camera recorded 30 frames per second, this corresponds to 9000 frames. Each sequence contains one active participant. In sequence 1, the subject is moving constantly, while in sequence 2, the subject moves only occasionally. In both the sequences, intervals with high activity do not necessarily imply that the subject is speaking. Often, a high activity region corresponds to a period in which the subject is silent, and a low activity region corresponds to when he or she speaks.

Tables 1 and 2 show frame-size statistics resulting from a variety of encoding algorithms. The algorithms include single-layer (VBR-1) encoding, two-layer (VBR-VBR) encoding, and several cases each of CBR-VBR and PBR-VBR encoding with different values of P, where $P\times64$ kbps is the fixed bit rate of the CBR base layer, or the peak bit rate constraint of the PBR base layer.

The units of Tables 1 and 2 are bits per frame. A rate of R bits per frame is equal to a rate of

$30R$ bits per second, exclusive of ATM overhead. Assuming that the video ATM Adaptation Layer (AAL) uses three bytes, and given the five bytes of ATM header in the 53-byte cell, the ATM overhead is $8/45 \approx 0.178$. Since each ATM cell carries 45 bytes of video data, a frame of size R bits occupies $R/360$ ATM cells.

In the tables, the columns headed $R^{\text{av}}_{\text{base}}$ and $R^{\text{av}}_{\text{sum}}$ show the empirical mean rate for the base layer and for the sum of the base and enhancement layers for the two sequences. The mean rates in bits per frame are plotted in Figure 3. For each sequence, the lowest combined mean rate is that for VBR-1. For CBR-VBR, the base rate increases linearly with P, as would be expected, and carries the sum rate up with it. For PBR-VBR, the mean bit rate for the sum channel decreases steadily toward the mean bit rate for VBR-1 as P increases. This again is to be expected, since as the peak-bit-rate constraint is raised, it becomes effectively no constraint at all and the coding algorithm approaches VBR-1.

In Tables 1 and 2, two maximum frame sizes are shown for each 5-minute sequence and for the base and combined layers. The frame sizes labeled $R^{\max}_{\text{base,raw}}$ and $R^{\max}_{\text{sum,raw}}$ are those measured at the output of the encoder prior to the smoothing buffer, while those labeled $R^{\max}_{\text{base,smooth}}$ and $R^{\max}_{\text{sum,smooth}}$ are measurements after the smoothing buffer. The latter two are the smallest maximums possible given the delay constraints imposed by the encoder and decoder buffers, which are determined using the algorithm presented in [17]. The delay has a duration of three frames, comparable to that required in H.261 [20].

Here, to produce comparable quality between a one-layer constant-bit-rate stream (CBR-1) and both the VBR-1 and two-layer bit streams, we set the peak rate for CBR-1 equal to the maximum frame size $R^{\max}_{\text{base,smooth}}$ for VBR-1. Conceptually, then, the CBR-1 bit stream fills in the valleys between VBR peaks by reducing the quantizer step size to increase the instantaneous bit rate without markedly influencing the video quality.

In the following, for clarity of the figures, we present only the multiplexing potential of the sequences with smoothing. We also examined the sequences without smoothing in the enhancement layer and found little difference for the two-layer algorithms. However, [17] shows that smoothing does have considerable impact on the multiplexing potential of VBR-1.

6. COMPUTING PROCEDURES

The first step is to obtain the parameters of a range of leaky-bucket TDs with which the empirical traffic streams are compliant, and then to calculate the expected loss ratios and congestion-period lengths corresponding to these TDs at various levels of multiplexing.

For the TD of a single-layer encoder, or for the TD of each layer of a two-layer encoder, we take the peak rate R^{max} equal to the empirical peak rate of the given layer, and we vary the negotiated average rate $\bar{R}$ between the empirical mean rate R^{av} and the empirical peak rate R^{max} of the layer. Thus for a two-layer encoder, $R^{av}_{base} \le \bar{R}_{base} \le R^{max}_{base}$ and $R^{av}_{sum} \le \bar{R}_{sum} \le R^{max}_{sum}$. Varying $\bar{R}$ for each layer yields a range of bucket sizes B^{max} for that layer. For a two-layer encoder, it is automatically true that $R^{max}_{sum} \ge R^{max}_{base}$. For overall consistency, we must also choose $\bar{R}_{sum} \ge \bar{R}_{base}$.

Figure 4 shows bucket-size behavior for VBR-1 as $\bar{R}$ increases from R^{av} to R^{max}, for the smoothed frame-size sequences. Although there are some differences in the shapes of the curves for other cases in Tables 1 and 2, they all have the characteristic property that B^{max} is very large, of the order of tens of thousands of ATM cells or hundreds of video frames, if $\bar{R}$ is close to R^{av}. On the other hand, B^{max} decreases by orders of magnitude as $\bar{R}$ approaches R^{max}.

To determine the optimal multiplexing potential of a given video stream within the current idealized framework, we need to find the pair of TD parameters $\bar{R}$ and B^{max} that correspond to the maximum number N of streams through a channel of capacity C, subject to two conditions: namely, the cell loss ratio and the mean length of congestion periods should be less than preassigned values. In symbols,

$$\mathrm{CLR} < \varepsilon \quad \text{and} \quad T^{cong} < \tau\,, \tag{12}$$

where ε and τ are given.

The computational procedure is conceptually straightforward, although it involves several steps. The simplest case is a single traffic stream, for example a VBR-1 encoding, having empirical values of R^{av} and R^{max}. For each value of $\bar{R}$ in the range $R^{av} \le \bar{R} \le R^{max}$, there is a corresponding smallest value of B^{max} with which the traffic stream is compliant. If the link capacity is C and if $\bar{R}$ is fixed, both CLR and T^{cong} are increasing functions of the multiplexing level N, according to Eqs. (6) and (10). We find the largest value of N that satisfies both of the conditions (12). We repeat the process for different values of $\bar{R}$ to get the largest value of N for the given encoding. The largest value, which we call N^{mult}, is the desired multiplexing potential of the given traffic stream.

For a two-layer encoding, the base and enhancement layers each have their own parameters and their own limits on CLR and $\bar{T}^{cong}$. We first calculate the multiplexing potentials N^{mult}_{base} and N^{mult}_{sum} separately. Then the multiplexing potential of the two-layer encoding is, provisionally,

$$N^{mult} = \min(N^{mult}_{base}\ ,\ N^{mult}_{sum})\ . \tag{13}$$

If the values of $\bar{R}_{base}$ and $\bar{R}_{sum}$ resulting from this calculation satisfy $\bar{R}_{sum} \geq \bar{R}_{base}$, we are done. If the inequality is not satisfied, we have to do a constrained optimization in which $\bar{R}_{base}$ varies freely while $\bar{R}_{sum}$ is constrained to satisfy $\bar{R}_{sum} \geq \bar{R}_{base}$. N^{mult} is the largest number of connections that can be multiplexed onto a given link with both layers satisfying conditions of the form (12) and also $\bar{R}_{sum} \geq \bar{R}_{base}$.

We assume that a cell loss ratio ε_{base} is acceptable for one-layer encodings and for the base layer of two-layer encodings. Similarly we assume that a cell loss ratio ε_{enh} is acceptable for enhancement cells, where $\varepsilon_{enh} >> \varepsilon_{base}$. We further assume that the queueing discipline imposes almost all of the cell losses on the sum stream and, in particular, on enhancement cells, so that to a good approximation the overall cell loss ratio of the merged cell stream satisfies

$$\varepsilon_{sum} = \varepsilon_{enh} R^{av}_{enh}/R^{av}_{sum} = \varepsilon_{enh}(R^{av}_{sum} - R^{av}_{base})/R^{av}_{sum}. \tag{14}$$

Acceptable values of ε_{base}, ε_{enh}, and τ are not yet well known. Ultimately they will depend on the coding techniques that are used to avoid or conceal data loss due to dropped ATM cells, and on subjective quality measurements that are still to be made. It is usually assumed that mean cell loss ratios in the range from 10^{-6} to 10^{-9} will be acceptable for the base layer, and from 10^{-1} to 10^{-4} for the enhancement layer. If we also assume that relatively frequent short congestion periods are less objectionable than relatively infrequent long congestion periods, then τ should probably be at most a few frame times.

7. STATISTICAL MULTIPLEXING POTENTIAL

In the examples to follow, we have assumed a B-ISDN network. In B-ISDN/ATM a physical layer of 155.52 Mb/s (e.g., SONET STS-3c) provides a transmission capacity of 149.760 Mb/s to the ATM layer. Allowing for five bytes of ATM header and three bytes of video ATM Adaptation Layer (AAL) in the 53-byte cell, the bit rate available for the video information, C, is $149.760 \times 45/53 = 127.155$ Mb/s. Since we assume the CBR-1 rate is equal to the smoothed peak rate $R^{\max}_{\text{base,smooth}}$ for VBR-1 in Tables 1 and 2, the number of CBR-1 connections that can share a channel is

$$n_0 = \lfloor 127.155 \times 10^6 / 30 R^{\max}_{\text{base,smooth}} \rfloor = \lfloor 4.2385 \times 10^6 / R^{\max}_{\text{base,smooth}} \rfloor . \qquad (15)$$

For example, 81 CBR-1 connections can be supported for sequence 1, while 141 CBR-1 connections can be supported for sequence 2

Figures 5-8 show the maximum multiplexing potential for each of the encodings in Tables 1 and 2, for different assumed values of $\varepsilon_{\text{base}}$, ε_{enh}, and τ. We have calculated ε_{sum} from Eq. (14), using the average rates for the base and sum layers shown in Tables 1 and 2. The units of τ are frame times, provided that $B^{\max}$ in Eq. (10) is expressed in bits and $R^{\max}$ and $\bar{R}$ are expressed in bits per frame time.

Each plot is arranged as follows. The multiplexing potential for PBR-VBR as a function of P is shown using solid curves, while the multiplexing potential for CBR-VBR is shown using dashed curves. The multiplexing potential for both VBR-VBR and VBR-1 are plotted as points on the right boundary. Finally, the number of CBR-1 sources is indicated by a horizontal line segment at the left of the plot. These line segments are the same in all four figures.

In Figure 5, we assume $\varepsilon_{\text{base}} = 0$ (no base cell losses), $\varepsilon_{\text{enh}} = 10^{-3}$ (constant in all our examples), and $\tau = \infty$ (no restrictions on congestion intervals). As discussed in Section 5, an appropriate reference level for computing multiplexing gain is the number of CBR-1 connections that the link can support. In Figure 5, the highest multiplexing level for sequence 1 is 105 connections, achieved by PBR-VBR at $P = 17.5$. This corresponds to a multiplexing gain over CBR-1 of $105/81 = 1.30$. For sequence 2, the highest multiplexing level is 238 connections, achieved by PBR-VBR at $P = 7.5$, and the gain over CBR-1 is $238/141 = 1.69$. The multiplexing levels achieved by CBR-VBR are lower than those achieved by PBR-VBR and occur at lower values of P, that is, with smaller amounts of information being carried by the base layer. VBR-VBR is also not competitive with PBR-VBR. In this example VBR-1 cannot be statistically multiplexed because the condition $\varepsilon_{base} = 0$ implies that each VBR-1 connection must be allocated its peak (smoothed) rate.

Figure 6 differs from Figure 5 by adding a restriction on τ, namely $\tau = 5$ frame times. The effect is to lower the multiplexing potential of all the two-layer encodings, and to shift the maxima to somewhat larger values of P, corresponding to more information in the base layer. The greatest multiplexing gain for sequence 1 is $85/81 = 1.05$ for PBR-VBR with $P = 22.5$, and for sequence 2 it is $183/141 = 1.30$ at $P = 10$. CBR-VBR and VBR-VBR are again not competitive.

In Figure 7, we take $\varepsilon_{base} = 10^{-6}$, $\varepsilon_{enh} = 10^{-3}$, and $\tau = \infty$. This is like Figure 5 except that we now permit some cells to be lost from the base layer, which allows statistical multiplexing of VBR-1. Now, in this example, VBR-1 dominates the two-layer disciplines. Its multiplexing gain over CBR-1 is $112/81 = 1.38$ for sequence 1, and $280/141 = 1.99$ for sequence 2. PBR-VBR is almost but not quite as good at large values of P, corresponding to having almost all the information in the base layer, while CBR-VBR and VBR-VBR are out of the running.

Figure 8 is like Figure 7 except that we add a restriction on congestion periods, namely $\tau = 5$ frame times. This restriction reduces the multiplexing potentials, but VBR-1 still slightly surpasses PBR-VBR. The maximum VBR-1 gains are $90/81 = 1.11$ for sequence 1 and $210/141 = 1.49$ for sequence 2.

The general effects of varying ε_{base} and τ are apparent from these figures. There is some range of parameters, near $\varepsilon_{base} = 0$ and $\tau = \infty$, where PBR-VBR encoding surpasses both VBR-1 and CBR-1, so far as statistical multiplexing gain is concerned. Increasing the allowable cell loss ratio ε_{base} increases the potential multiplexing gain, and for the considered video examples and encoding schemes, VBR-1 yields the highest multiplexing gain. Decreasing the mean congestion period τ decreases the potential multiplexing gain, and for the considered video examples and encoding schemes, VBR-1 again yields the highest multiplexing gain.

8. CONCLUSIONS AND DISCUSSION

In this paper, we have compared the multiplexing performance of one- and two-layer video compression algorithms, using two 5-minute samples of teleconferencing video. Three different rate constraints were assumed for the base layer of the two-layer codec: no constraint (VBR), a constant-rate constraint (CBR), and a peak-rate constraint (PBR). We have computed leaky-bucket traffic descriptors for the sample sequences, and have calculated the potential statistical multiplexing gains based on given cell loss ratios and congestion period durations. For a 155 Mb/s B-ISDN network, we have estimated statistical multiplexing gains between 1.5:1 and 2:1, in general agreement with earlier work.

The most effective two-layer encoders, from the statistical multiplexing viewpoint, use a peak bit rate control on the base layer, with the base-layer peak bit rate chosen so that most of the information is carried by the base layer and the two-layer overhead is small. Lower gains are offered by CBR-VBR encoders, and their optimal operating point has a smaller fraction of the total information in the base layer. For the encoding algorithm and sample video sequences considered, single-layer VBR encoders offer the best multiplexing potential, provided their output is permitted the same amount of smoothing that is used in conventional CBR systems.

Intuitively, PBR-VBR outperforms CBR-VBR because statistical multiplexing allows the network to use the gaps, produced when the base layer of PBR-VBR is less than the peak rate, to carry enhancement data from other sources. Therefore, while CBR-VBR may be advantageous for interworking with existing equipment, PBR-VBR may provide a cheaper alternative for the user who does not demand interworking.

We have also looked at the question of end-to-end delay jitter in an ATM network. We estimate that cell-level jitter across a geographically large ATM network will be less, but not orders of magnitude less, than one video frame time. Cell-level delay jitter is unavoidable in packet networks, and has to be absorbed in the decoder buffer. In this paper we have considered the case of no burst-level jitter. However, it is unknown how much additional multiplexing gain could be achieved if burst-level jitter were allowed.

This work raises a number of issues relating both to video encoding and to network design. We envision that traffic descriptor parameters appropriate to a specific application, such as videoconferencing, will be negotiated between the encoder and the network. The network will agree to the parameters and the encoder will then undertake to comply with them, so that the network will have no occasion to downgrade or reject cells that the encoder regards as essential.

Much more work needs to be done to determine realistic TD parameters for given applications. For example, in the present study two traffic streams with substantially different average bit rates were obtained from two 5-minute segments of the same videoconferencing session. Presumably the user will not wish to renegotiate TD parameters with the network during the course of a call (at least, not often). An alternative is to have representative TD parameters associated with a particular encoder and application, so that under normal conditions the encoder can comply with the TD without drastically "cramping its style". To obtain representative parameter values may require orders of magnitude larger statistical samples than have yet been obtained for state-of-the-art algorithms. Software simulations of encoding algorithms are notoriously time-consuming; good statistics may depend on the availability of high-performance 2-layer hardware encoders. It is conceivable that, if substantial differences persist among samples of VBR-encoded traffic derived from superficially similar situations, the need for conservative engineering may nullify any potential multiplexing gains projected on the basis of a particular sample.

In addition, it would be useful to study the statistical multiplexing capabilities of other two-layer encoding algorithms. In the present study, we chose an enhancement-layer compression algorithm that is very robust to cell losses but is inefficient. An alternative algorithm can trade off robustness for efficiency, through the use of temporal prediction in the enhancement layer.

On the network side, we have used Framed Round Robin queueing as a means of estimating cell-level delay jitter. More detailed study of potential queueing disciplines is needed, in order to determine how easily queueing disciplines that require substantial processing at switching nodes can be scaled to higher speeds.

Another network issue has to do with estimating end-to-end delay jitter and cell loss. Most models of delay and loss (for example, the cell loss ratios calculated in this paper) apply to a single node. Parekh [23] has computed simple and rigorous end-to-end delay bounds for a particular model, but his bounds may not always be tight. Most proposed end-to-end analyses [28, 29] are computationally expensive. On the other hand, enhancement cells are designed to be disposable. Perhaps if the system is designed to guarantee not to lose base cells, accurate end-to-end loss calculations will not be essential for two-layer encoders.

In summary, two-layer encoding is still worth studying for videoconferencing services. A peak-bit-rate constraint on the base layer is better than a constant or variable-bit-rate constraint, from the viewpoint of statistical multiplexing gain. Substantial statistical work needs to be done to determine representative parameters for traffic descriptors, in order that variable-rate video encoders will be able to work effectively with the types of standards that are expected for ATM-based networks.

ACKNOWLEDGMENTS

We are indebted to B. T. Doshi for the result cited in Section 4.3, and for a thoughtful reading of an earlier draft of this paper. E. N. Gilbert provided a general proof of Eq. (10).

Att.
References
Tables 1-2
Figures 1-8

REFERENCES

[1] F. Kishino, K. Manabe, Y. Hayashi, and H. Yasuda, "Variable Bit-Rate Coding of Video Signals for ATM Networks," IEEE J. Select. Areas Commun., Vol. 7, No. 5 (June 1989), pp. 801-806.

[2] D. Beaumont and G. Morrison, "Results from a Prototype 2-Layer VBR Video Codec," 4th International Workshop on Packet Video, Kyoto, Japan (August 1991).

[3] M. Ghanbari and J. Azari, "Comparison Between the CBR and VBR at the Base Layer of a Two-Layer Video Codec," 4th International Workshop on Packet Video, Kyoto, Japan (August 1991).

[4] K. Joseph, J. Zdepski, R. Siracusa, and D. Raychaudhuri, "Design and Performance Evaluation of a Variable Bit-Rate (VBR) Video Compression Algorithm for ATM Networks," GLOBECOM '91, Phoenix, AZ (December 1991), pp. 9-15.

[5] T. Kinoshita, T. Nakahashi, and M. Takizawa, "Variable Bit-Rate HDTV Coding Algorithm for ATM Environments in B-ISDN," SPIE Visual Communication and Image Processing '91: Visual Communication, Vol. 1605 (1991), pp. 604-613.

[6] J. R. Louvion, "2-Layer Versus 1-Layer Video Codecs: A Network Performance Approach," 4th International Workshop on Packet Video, Kyoto, Japan (August 1991).

[7] G. Morrison and D. Beaumont, "Two-Level Video Coding for ATM Networks," Signal Processing: Image Communication, Vol. 3, Nos. 2-3 (June 1991), pp. 179-195.

[8] M. Nomura, T. Fujii, and N. Ohta, "Layered Coding for ATM Based Video Distribution Systems," Signal Processing: Image Communication, Vol. 3, No. 4 (September 1991), pp. 301-311.

[9] A. R. Reibman, "DCT-Based Embedding Coding for Packet Video," Signal Processing: Image Communication, Vol. 3, Nos. 2-3 (June 1991), pp. 231-237.

[10] S. Tubaro, "A Two Layers Video Coding Scheme for ATM Networks," Signal Processing: Image Communication, Vol. 3, Nos. 2-3 (June 1991), pp. 129-141.

[11] M. Ghanbari, "An Adapted H.261 Two-Layer Video Codec for ATM Networks," IEEE Trans. Comm., Vol. 40, No. 9 (September 1992), pp. 1481-1490.

[12] J. W. Roberts, editor, "Performance Evaluation and Design of Multiservice Networks," COST 224 Final Report, Commission of the European Communities, Brussels (1992).

[13] S. P. Morgan and A. R. Reibman, "Statistical Multiplexing Comparison of One- and Two-Layer Video Codecs for Teleconferencing," 5th International Workshop on Packet Video, Berlin, Germany (March 1993).

[14] B. G. Haskell and A. R. Reibman, "Multiplexing of Variable Rate Encoded Streams," IEEE Transactions on Circuits and Systems for Video Technology, to appear.

[15] B. G. Haskell, "Buffer and Channel Sharing by Several Interframe *Picturephone®* Coders," BSTJ, Vol. 51, No. 1 (January 1972), pp. 261-298.

[16] "Traffic Control and Congestion Control in B-ISDN", CCITT SGXVIII Draft Recommendation I.371 (June 1992).

[17] A. R. Reibman and A. W. Berger, "Traffic Descriptors for VBR Video Teleconferencing over ATM Networks," GLOBECOM '92 (December 1992).

[18] A. R. Reibman and B. G. Haskell, "Constraints on Variable Bit-Rate Video for ATM Networks," IEEE Transactions on Circuits and Systems for Video Technology, Vol 2, No. 4 (December 1992), pp. 361-372.

[19] M. Liou, "Overview of the px64 kbit/s Video Coding Standard," Comm. ACM, Vol. 34, No. 4 (April 1991), pp. 59-63.

[20] "Description of Reference Model 8 (RM8)", Document No. 525, CCITT SGXV Working Party XV/4 (1989).

[21] T. Tanaka, K. Manabe, and H. Haseyama, "A Study on Comparison Between VBR and CBR Video Service in ATM Environment," SUPERCOMM/ICC '92, Chicago, IL (June 1992), pp. 551-555.

[22] C. R. Kalmanek, H. Kanakia, and S. Keshav, "Rate Controlled Servers for Very High-Speed

Networks," GLOBECOM '90, San Diego, CA (December 1990), pp. 12-20.

[23] A. K. J. Parekh, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks," MIT PhD thesis, LIDS-TH-2089 (February 1992).

[24] S. J. Golestani, "Congestion-Free Communication in High-Speed Packet Networks," IEEE Trans. Comm., Vol. 39, No. 12 (December 1991), pp. 1802-1812.

[25] I. Norros, J. W. Roberts, A. Simonian, and J. T. Virtamo, "The Superposition of Variable Bit Rate Sources in an ATM Multiplexer," IEEE J. Select. Areas Commun., Vol. 9, No. 3 (April 1991), pp. 378-387.

[26] C. Rasmussen, J. H. Sørensen, K. S. Kvols, and S. B. Jacobsen, "Source-Independent Call Acceptance Procedures in ATM Networks," IEEE J. Select. Areas Commun., Vol. 9, No. 3 (April 1991), pp. 351-358.

[27] B. T. Doshi, "Deterministic Rule Based Traffic Descriptors For Broadband ISDN: Worst Case Behavior and Equivalent Bandwidth," GLOBECOM '93, Houston, TX (November 1993).

[28] H. Kröner, M. Eberspächer, T. H. Theimer, P. J. Kühn, and U. Briem, "Approximate Analysis of the End-to-End Delay in ATM Networks," INFOCOM '92, Florence, Italy (May 1992), pp. 978-986.

[29] J. Kurose, "On Computing Per-Session Performance Bounds in High-Speed Multi-Hop Computer Networks," Performance Evaluation Review, Vol. 20, No. 1 (June 1992), pp. 128-139.

Table 1

Sequence 1 Statistics

Algorithm	P	R^{av}_{base}	$R^{max}_{base,raw}$	$R^{max}_{base,smooth}$	R^{av}_{sum}	$R^{max}_{sum,raw}$	$R^{max}_{sum,smooth}$
VBR-1		28330	112549	52036			
VBR-VBR		21971	93580	40805	44944	120414	70414
CBR-VBR	5	10667	10667	10667	54827	106921	96193
	10	21333	21333	21333	43264	92605	84071
	12.5	26666	26667	26667	41892	87126	79540
	15	32000	32000	32000	42410	83338	76492
	17.5	37333	37333	37333	44422	81475	74419
	20	42667	42667	42667	47596	79989	72123
	22.5	48000	48000	48000	51267	76261	70045
	25	53333	53333	53333	55407	75986	71192
	27.5	58667	58667	58667	59945	75345	71219
	30	63999	64000	64000	64743	76225	72860
PBR-VBR	5	10583	10667	10667	54311	106163	95877
	10	19132	21333	21333	39214	92931	84027
	12.5	22173	26667	26667	35254	87224	79482
	15	24555	32000	32000	32552	83503	76277
	17.5	26387	37333	37333	30637	80815	73924
	20	27587	42667	42667	29453	78507	71189
	22.5	28201	48000	48000	28835	74035	67859
	25	28327	53333	53333	28695	68243	63501
	27.5	28336	58667	58667	28678	63751	60024
	30	28336	64000	64000	28678	67014	64839

Table 2

Sequence 2 Statistics

Algorithm	P	R^{av}_{base}	$R^{max}_{base,raw}$	$R^{max}_{base,smooth}$	R^{av}_{sum}	$R^{max}_{sum,raw}$	$R^{max}_{sum,smooth}$
VBR-1		11704	78004	29908			
VBR-VBR		8667	65826	22464	23954	83617	42787
CBR-VBR	2.5	5333	5333	5333	26261	60646	52172
	5	10667	10667	10667	20948	57945	46538
	6.25	13333	13333	13333	20820	56506	44994
	7.5	16000	16000	16000	21447	55203	44353
	8.75	18667	18667	18667	22492	52478	43690
	10	21333	21333	21333	23973	51639	43262
	11.25	24000	24000	24000	25860	51122	41381
	12.5	26667	26667	26667	28033	48670	39493
	13.75	29333	29333	29333	30354	44570	39817
	15	32000	32000	32000	32741	43901	40640
	16.25	34666	34667	34667	35194	45244	41697
PBR-VBR	2.5	5172	5333	5333	25591	60646	52257
	5	8318	10667	10667	17656	57945	46538
	6.25	9322	13333	13333	15878	56506	44994
	7.5	10095	16000	16000	14614	55203	43972
	8.75	10686	18667	18667	13641	52478	43244
	10	11117	21333	21333	12939	51639	42429
	11.25	11431	24000	24000	12448	51122	40691
	12.5	11610	26667	26667	12160	48670	38723
	13.75	11679	29333	29333	12050	44570	37310
	15	11707	32000	32000	12016	40982	35701
	16.25	11709	34667	34667	12013	42305	36976

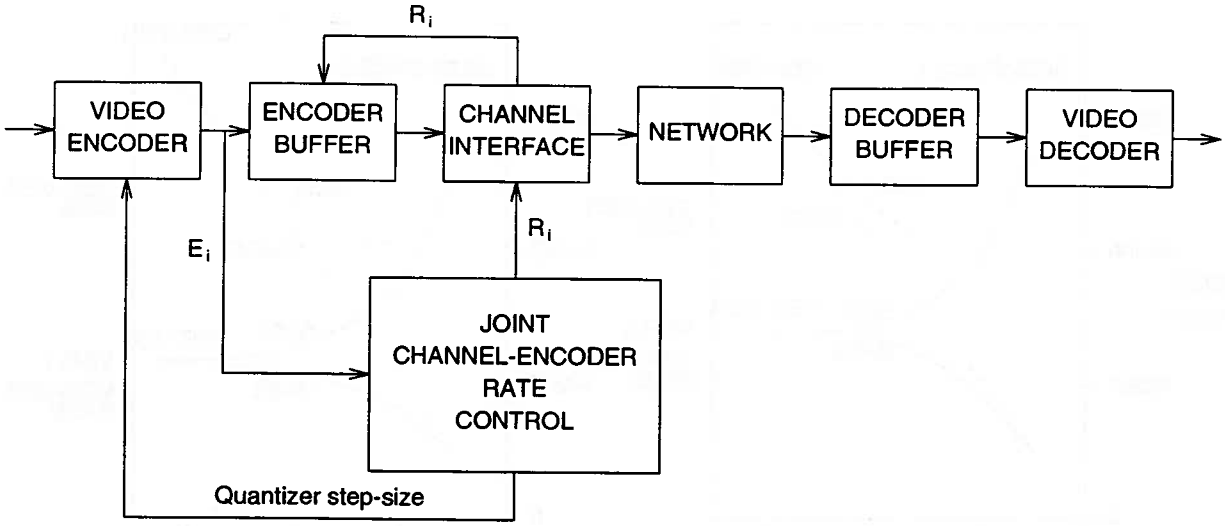

Figure 1. Generic packet video system.

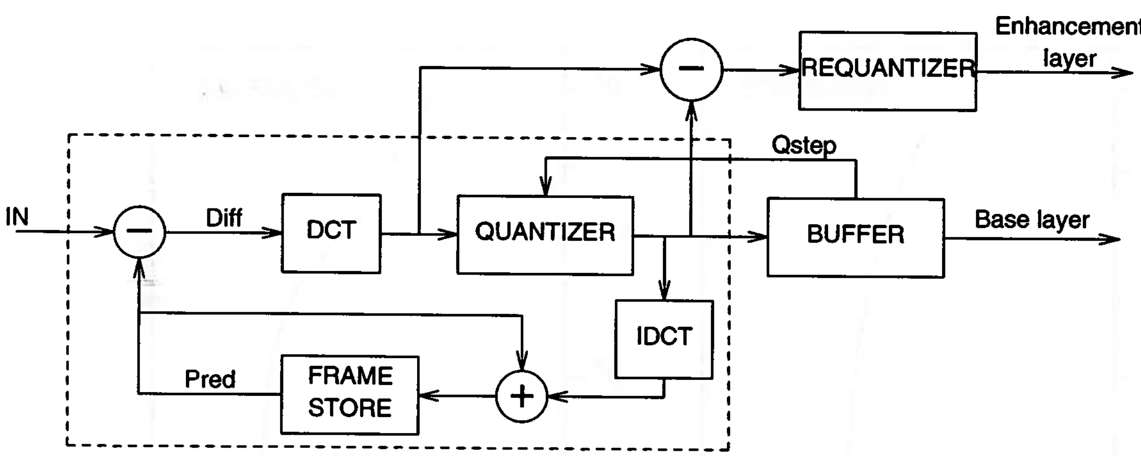

Figure 2. Block diagram of two-layer coder.

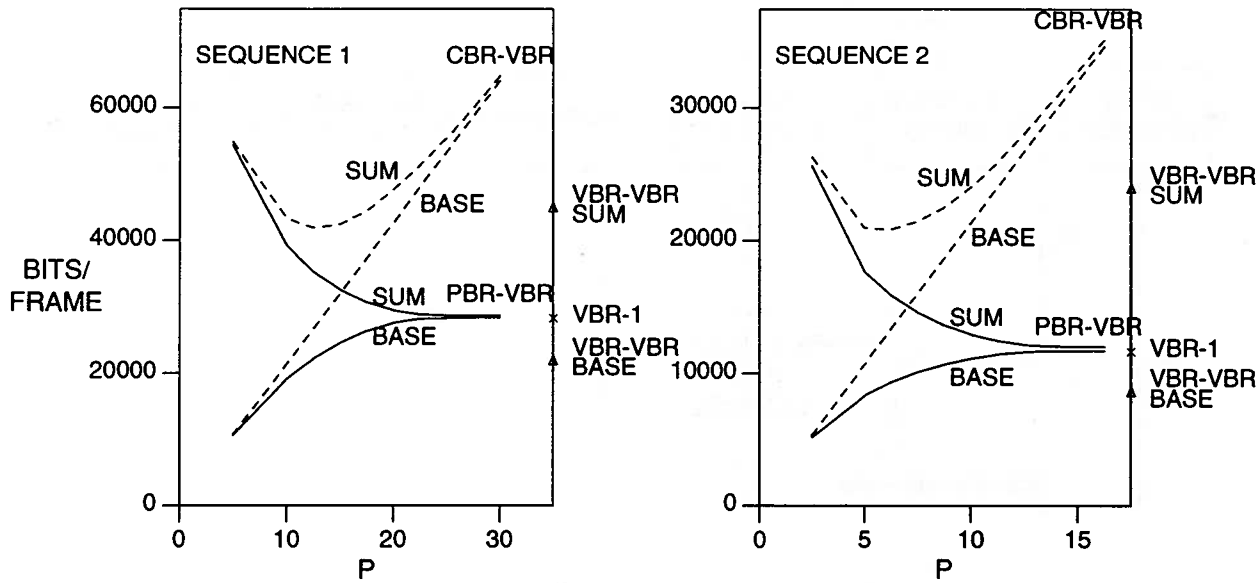

Figure 3. Empirical mean frame sizes R^{av}_{base} and R^{av}_{sum}.

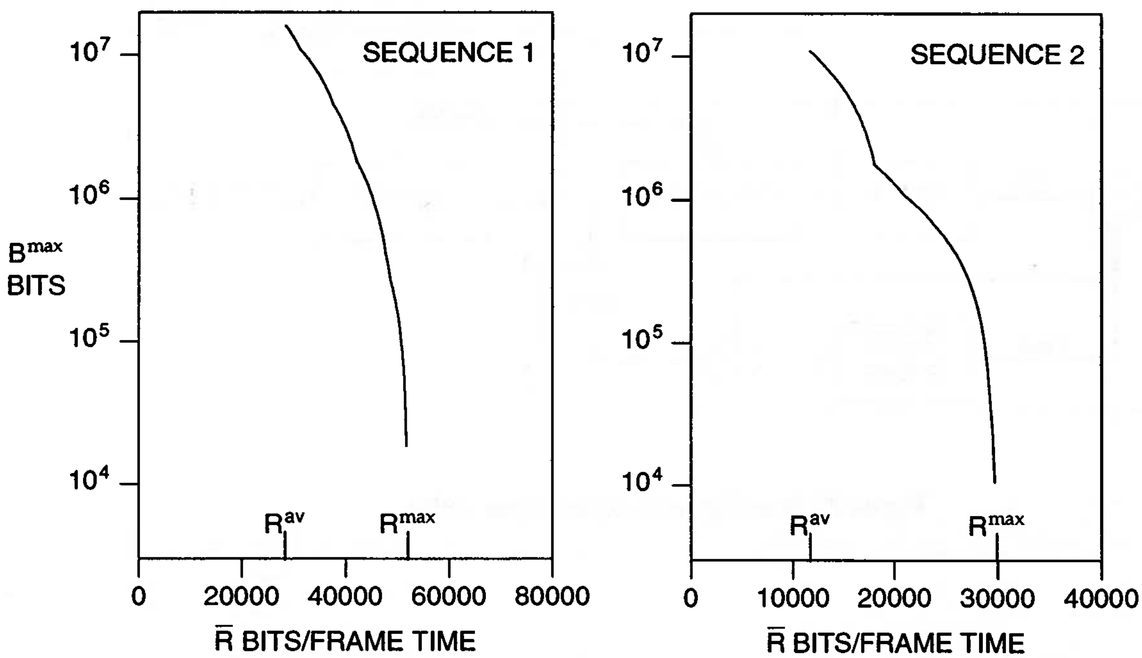

Figure 4. Empirical bucket size B^{max} as a function of $\bar{R}$ for smoothed VBR-1.

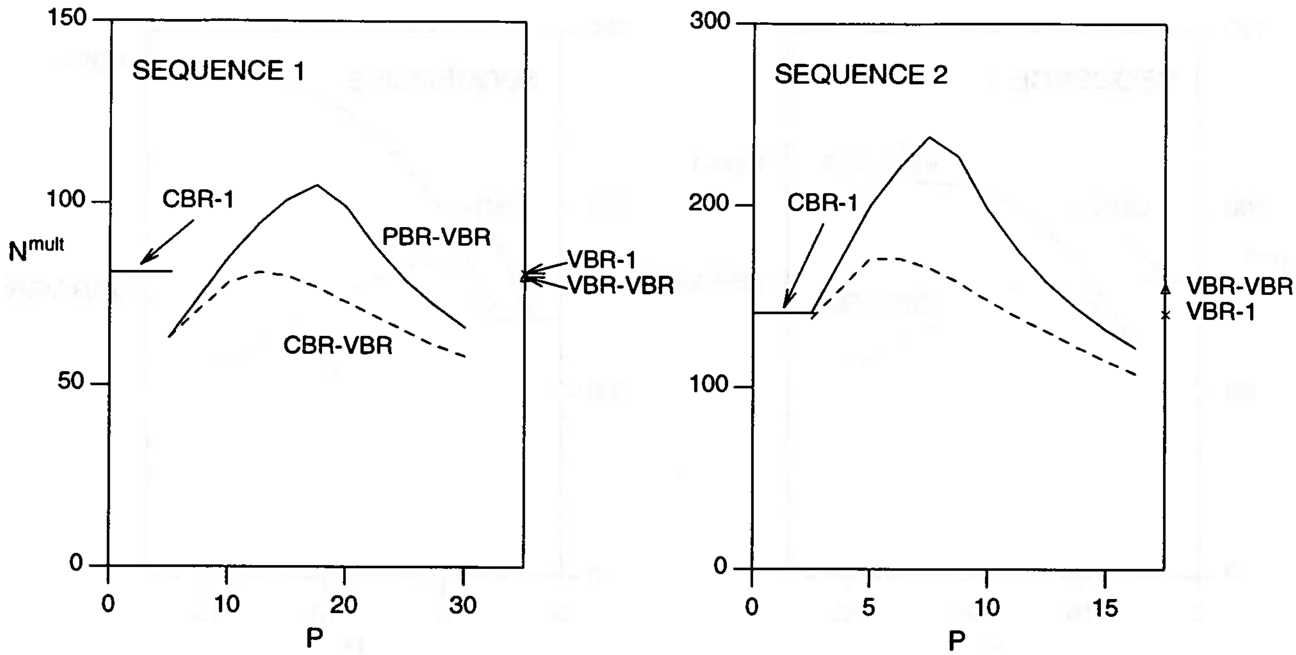

Figure 5. Maximum multiplexing potential for $\varepsilon_{base} = 0$, $\varepsilon_{enh} = 10^{-3}$, $\tau = \infty$.

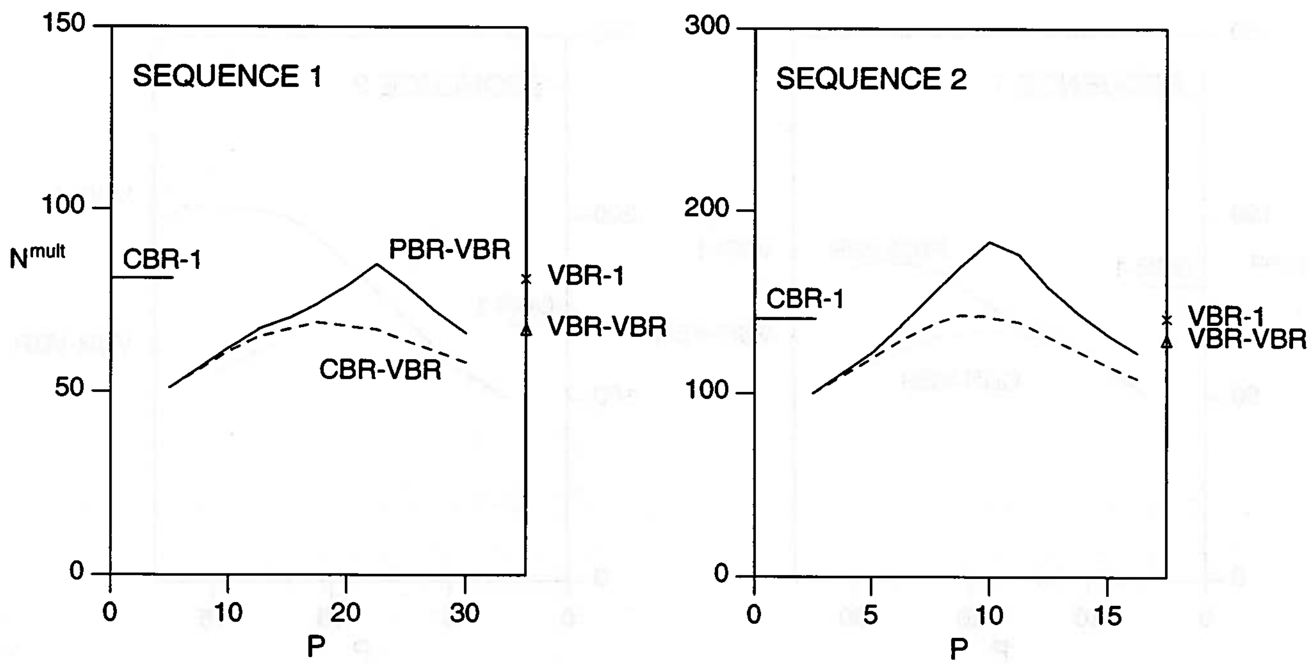

Figure 6. Maximum multiplexing potential for $\varepsilon_{base} = 0$, $\varepsilon_{enh} = 10^{-3}$, $\tau = 5$.

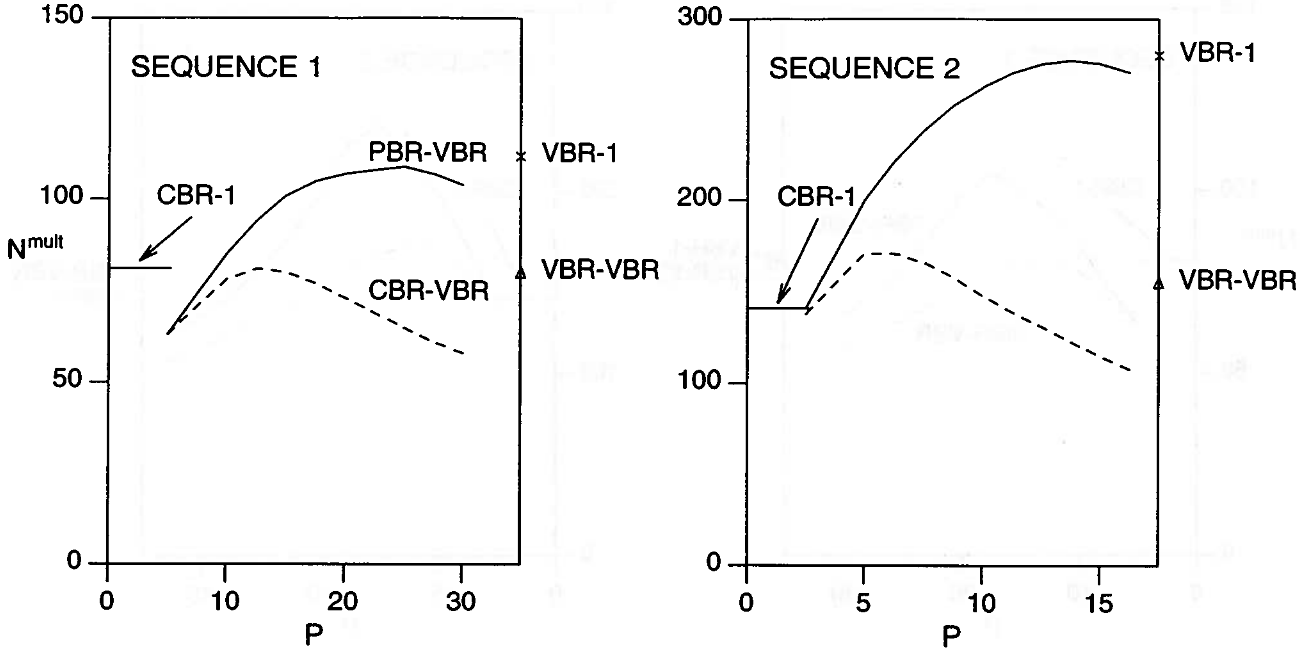

Figure 7. Maximum multiplexing potential for $\varepsilon_{base} = 10^{-6}$, $\varepsilon_{enh} = 10^{-3}$, $\tau = \infty$.

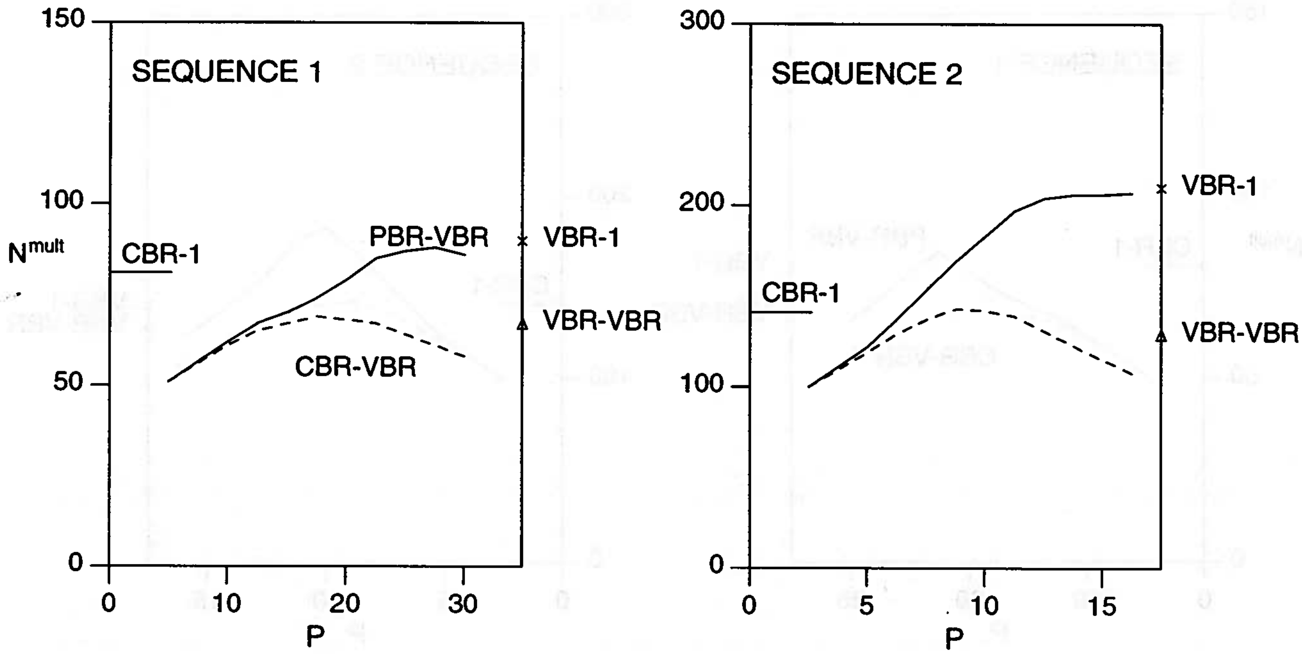

Figure 8. Maximum multiplexing potential for $\varepsilon_{base} = 10^{-6}$, $\varepsilon_{enh} = 10^{-3}$, $\tau = 5$.