Partial versus Early Packet Video Discard

Ahmed Mehaoua
University of Cambridge
10 Downing Street
Cambridge, United Kingdom
am296@ccsr.cam.ac.uk

Raouf Boutaba
University of Toronto
10 King's College Road
Toronto, Canada
rboutaba@nal.utoronto.ca

Youssef Iraqi
University of Montreal
C.P. 6128, succ. Centre-Ville
Montreal, Canada
iraqi@iro.umontreal.ca

ABSTRACT

In this paper, we propose and compare two video slice-based discard schemes, namely adaptive-PSD and Adaptive-ESD, for the transmission of MPEG video streams overATM best effort services. The schemes perform adaptive and selective cell burst discard at the level of MPEG video slices and intelligently adjusts drop policies to switch buffer occupancy and video cell payload types. In comparison to previous approaches, the performance evaluation have shown a significant reduction of the bad throughput crossing the network and a better protection of critical Intra- and Predictive-coded pictures at both cell and video slice levels.

Keywords: ATM, Best Effort, Packet Video, MPEG, Cell Discard.

I. INTRODUCTION

With increasing interest in the transmission of MPEG-compressed video streams over unreliable ATM best effort services (ABR, UBR+), efficient video-oriented packet dropping mechanisms have to be designed which attempt to gracefully control picture quality degradation during network congestion.

These video applications will extensively make use of MPEG video compression standards to save network resources. MPEG defines a video stream as a hierarchy of data structures ordered by increasing spatial size : pixel, 8x8 pixel Block, 16x16 pixel MacroBlock, Slice, Frame, Group of Pictures and Sequence [1]. Two of them have significant impacts on the decoding/displaying process and thereby on the picture quality perceived by the end users.

Video slice is the main coding processing unit in MPEG. Coding and decoding of blocks and macroblocks are feasible only when all the pixels of a slice are available. Besides, encoding of a slice is done independently from its adjacent slices, making it the smallest autonomous decoding unit. Consequently, it serves as resynchronization point in case of problems.

Frame or picture is the basic unit of display. Three picture types may be present in a MPEG video stream. They differ from the encoding method used: Intra-coded (I) picture, Predictive-coded (P) picture and Bidirectionally predictive-coded (B) picture. I- and P-encoded pictures are essential and have to be preserved from corruption during transmission. Due to error propagation at the decoding layer, a corrupted or non-available reference picture (e.g. I- or P-frame) leads on perceptible picture degradation. I-frame impairments will affect all the subsequent frames on the same Group Of Picture (GOP). Similarly, the impairment of P-frames will affect the following P- and B-frames until the next I-frame. Only B-frame impairments have no adverse effects on other frames.

According to the above statements, three obvious remarks stand out. First, the smallest transmission data unit is rather video slice than ATM cell, AAL PDU or MPEG multiplex packet (e.g. Transport Stream or Program Stream). Secondly, in situation of congestion, dropping video cells indiscriminately can cause serious degradation in picture quality. Intra- and Predictive frames have to be better protected from errors during transmission. Finally, without intelligent FEC and Error concealment mechanisms at destination, forwarding partially corrupted video slices is wasteful and may even worsen the congestion in the upstream nodes. Thus, the question arises which is the best cell dropping policy that ensure the highest network bandwidth utilization while minimizing the video slice loss probability at the application layer.

To address this problem, we present and compare the performance of two packet video drop policies for use with ATM best effort services (i.e. ABR, UBR+, GFR). The schemes are referred to Adaptive and Partial video Slice Discard (A-PSD) and Adaptive and Early video Slice Discard (A-ESD).

The paper is organized as follows. In section 2, we briefly review some previous works in this area.. Section 3 is devoted to the description of the two proposed video cell discarding schemes. Section 4 introduces the network simulation model and discuss the performance results. Finally, we conclude in Section 5.

II. RELATED WORKS

Various data protection and recovery techniques have been proposed to cope with the problem of transmitting compressed video streams over lossy networks. These schemes attempt to minimize picture quality degradation induced by data loss during network congestion.

Some are running at the adaptation layer or above such as layered video encoding with data prioritization which is one of the most popular approach [2][3]. Forward Error detection and Correction (FEC) techniques associated with byte interleaving and error concealment mechanisms at the destination have also been proposed to address this issue [4][5].

At the network level, smoothing algorithms attempt to reduce the burstiness and the peak bandwidth of video streams prior to transmission by applying complex shaping and buffering techniques at the source [6][7]. This should minimize switch buffer oscillation, ease cell scheduling and thereby reduce cell loss probability.

In addition, to cope with the problem of packet fragmentation and poor transmission performance of traditional packet services (IP, Frame-relay, …) over ATM, some mechanisms have been designed to preserve packet integrity and achieve higher good throughputs.

Packet Tail Discard or *Partial Packet discard* (PPD) has been proposed first to address this problem [8]. If a cell is dropped by a switch, the subsequent cells of the packet are also dropped. Romanov and al. have shown that PPD improves network performance to a certain degree, but it is still not optimal. Therefore, they proposed a new mechanism called *Early Packet Discard* (EPD) that achieves better throughput performance but does not guarantee fairness among the connections [9]. When the switch buffer queue reaches a threshold, entire newly arriving packets (e.g. AAL5 PDU) are preventively discarded. To improve its fairness, selective packet drops based on per-Virtual Circuit accounting have been introduced by Heinanen and Kilkki and referred as *Fair Buffer Allocation* (FBA).

Since MPEG video is also a packet-oriented service, it may be interesting to apply such approach to the video streams. Nevertheless, none of the previously mentioned smart data packet discard schemes are taking into account the specific properties of MPEG-encoded video. Therefore, we are proposing in the following section some enhancement to PPD and EPD in order to gracefully degrade picture quality during network congestion and optimize network resource utilization.

III. PACKET VIDEO DISCARD SCHEMES

Adaptive Partial Video Slice Discard

Adaptive-PSD runs per-Virtual Circuit and defines four operating modes in respect to the shared buffer's queue length (Q_L). With the two-level ATM CLP-based priority mechanism, A-PSD makes use of two thresholds (i.e. a low threshold (L_T) and a high threshold (H_T). As depicted in Figure 1, these thresholds define four operation modes :

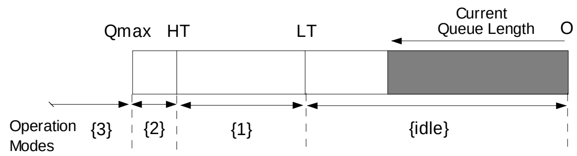

Figure 1 - Adaptive-PSD and Adaptive-ESD Operation Modes

- *Mode {Idle}*: if $Q_L < L_T$ then no incoming cells are discarded.
- *Mode {1}:* if $L_T \le Q_L < H_T$, then for every VC currently emitting a low priority video slice (i.e. belonging to B-frames with CLP bit set to 1), A-PSD starts to drop their arriving cells until the reception of an End Of Slice (EOS) cell. This cell is always preserved from elimination since it permits to detect slice delineation at ATM layer. Other high priority incoming cells (i.e. from I- and P- frames with CLP bit set to 0) are accepted in the buffer.
- *Mode {2}*: if $H_T \le Q_L < Q_{\max}$, then incoming video cells are dropped regardless to their priority (except EOS cells). This mode stops when the current queue length drops below H_T . Nevertheless, the scheme keeps discarding the tail of every slices that have been partially corrupted. Two variables per VC are used for that purpose.
- *Mode {3}:* if $Q_L \ge Q_{\max}$ then all incoming cells are discarded and slice integrity are checked for partial discard.

Adaptive Early Video Slice Discard

Network performance can be significantly improved by prevently dropping entire video slices prior to congestion rather than only a portion of them as it may happend with A-PSD. In addition, if the early discarded slices belong to low priority frames (B-frames), then the switch can better accommodate and forward the most important video slices from referenced pictures (e.g. I- and P-frames). This stratege has been implemented in [11] and referred to Adaptive and Early video Slice Discard. A-ESD can be implemented with any MPEG data unit (block, macroblock, frame). According to MPEG definition, we believe that the video slice layer is the most appropriate.

Similarly to the previous scheme, A-ESD adjusts dynamically drop severity to the network load and gives priority to the I- and P-cells (CLP bit set to 0) over B-cells (CLP bit set to 1). A-ESD differs to A-PSD during the operation modes {1} and {2} as fellows:

- *Mode {1}*: If $L_T \le Q_L < H_T$, then only newly arriving low priority slices (e.g. reception of the first cell of the slice) are entirely discarded. Other slices are accepted in the buffer. That way, A-ESD preserves buffer space to slices that have already entered the buffer or slices that are assigned a higher priority. Thus, it optimizes the probability of transmitting successfully complete slices and favors I- and P- frames over B-frames. This mode stops when Q_L falls down to L_T.
- *Mode {2}:* if $H_T \le Q_L < Q_{\max}$, then newly entering video slices are entirely dropped regardless to their priority (except EOS cells). This mode stops when the current queue length drops below H_T.

The others operation modes are similar to A-PSD.

IV. PERFORMANCE EVALUATION

Simulation Model

We consider a network simulation model composed with an ATM switch (*SW*), an OC-1 bottleneck link *(L)* and three VBR MPEG-1 video connections. The distance between the sources and the switch are constant and set to 0.2 km (e.g. 0.125 miles). '*L*' is initialized to 2.5 Km (e.g. 1.56 miles) to emulate a LAN. We assume a link propagation speed of 2.5×10^8 m/s, and a propagation delay between the switch and the Broadband Terminal Equipment (*BTE*) of 10×10^{-6} ms. The Round Trip Time (*RTT*) between the sources and destinations is then set to 23.2×10^{-6} ms.

The video streams are generated using three different MPEG-1 frame traces : '*Star-Wars*', '*Tennis*' and '*Soccer*'. The main statistics of the MPEG sequences are summarized in Table 1. and 2.

	Star Wars	Tennis	Soccer
Compression Rate (X:1)	130	121	106
Quantizer scale	I : 10	P : 14	B : 18
GOP pattern (N=12 M=3)	IBBPBBPBBPBB		

Table 1 - Video Coding parameters

We are running the UBR+ service and we carried out the simulations with seven switch buffer configurations. For each of them, the same method is applied to determine the threshold values. H_T and L_T are respectively set to $1.0\,Q_{\max}$ and $0.8\,Q_{\max}$ and $Q_{\max}$ is ranging from 40 Kbits to 165 Kbits while 10% of the buffer resource are always reserved to accommodate signalling and data control cells (i.e.EOS cells).

	Star Wars	Tennis	Soccer
Mean Cell Rate (Mbps)	0.36	0.55	0.63
Peak Cell Rate (Mbps)	4.24	1.58	2.29
Peak/Mean ratio	11.7	2.87	3.63

Table 2 - Traffic descriptors

Every connection starts transmitting at different times in the range of [0, 41.6 and 83.3 ms.] to avoid I-frames overlapping. We also assume 15 uniformly distributed slices per video frame and a simulation time of 1.43 min (i.e. 37080 slices). Cells are smoothed and transmitted in piece-wise CBR during each frame slot (1/24 sec.).

Let us define the Cell Bad throughput *(CB)* as the ratio of the number of dead cells vs. the total transmitted cells. It is a performance parameter evaluated at the ATM layer. We consider a dead cell, as a correctly received cell that belongs to a corrupted slice.

Let us define the video Slice Loss Ratio (*SLR*) as the number of corrupted slices received vs. the number of transmitted slices. The *SLR* is measured at the application layer and is applied to the aggregate stream and for each individual frame substream (I-, P, and B-).

Let us also define the end-to-end Cell Transfer Delay (*CTD*) as the time between the departure of cell K from the source node (t_{iK}) and its arrival at the destination node (t_{0_K}) : $D_K = t_{0_K} - t_{i_K}$.

To estimate the benefit of packet-based drop strategies over cell-based ones, the proposed A-PSD and A-ESD are also compared with two following discard mechanisms.

1. *Random Cell Discard* (RD) where all video cells are assigned a high priority. Incoming cells are dropped randomly when H_T is exceeded and elimination stops when Q_L drops below H_T.
2. *Enhanced Partial Buffer Sharing* (Enh-PBS) [3], where cells from I- and P- frames are assigned a high priority, while B-frame cells are assigned a low priority. Enh-PBS is a video-dedicated Adaptive and Selective Cell Discard (A-SCD) scheme, which proactively drops low priority B-cells during light congestion, and gives the buffer space to the higher priority video cells (I- and P-cells). If the congestion worsens then incoming high priority cells are gradually submitted to elimination. More implementation details can be found in [3].

Results Analysis

Figures 2, 3 and 4 describe the variation of the slice loss rate (*SLR*) per frame type for the four cell drop schemes.

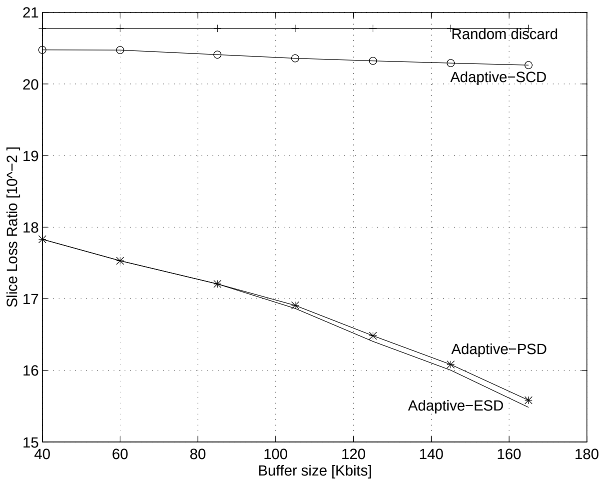

Figure 2 - (I)ntra-frame Slice Loss Ratio

In comparison to Random Discard (RD) and enhanced Partial Buffer Sharing (enh-PBS), both A-PSD and A-ESD show better results in protecting referenced Intra- and Predictive encoded video frames.

In addition, a significant reducing of the Slice Loss Ratio of Bidirectionnally Predictive encoded frames (*i.e. SLR_b*) is also experienced by the two proposed schemes. This is explained by the minimization of the adverse effect of error propagation at the encoder side since referenced frames are less concerned by errors.

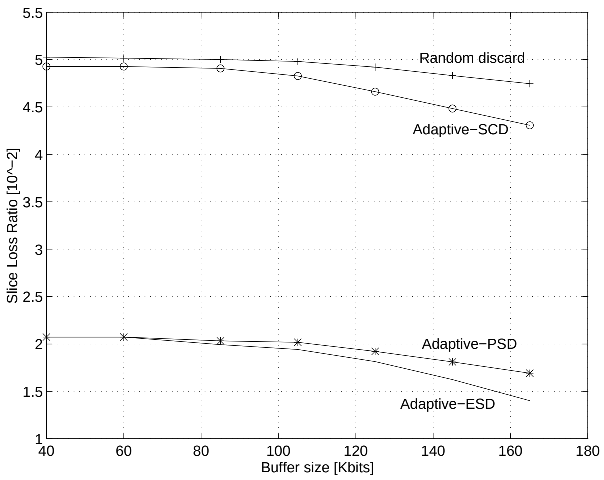

Figure 3 - (P)redictive-frame Slice Loss Ratio

Performance difference between A-PSD and A-ESD is non-significant when the buffer size ($Q_{\max}$) is small. When $Q_{\max}$ increases and reaches 165 Kbits, the *SLR_i*, is estimated to $1.5x10^{-1}$ with a slight advantage to A-ESD. This means that the schemes are more accurate and efficient when the buffer is well dimensioned to accommodate simultaneously entire slices and is then able to discriminate between them. In comparison *SLR_i* well exceeds $2x10^{-1}$ for both RD and enh-PBS. With these schemes, S*LR_i* seems to be constant and independent to the buffer's queue size.

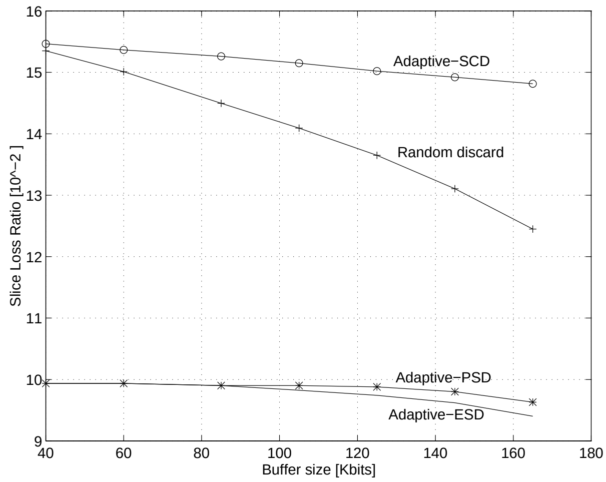

Figure 4 - (B)idirectional Predictive-frame Slice Loss Ratio

Similar remarks can be done to the *SLR_p* and *SLR_b*. From Figure 4., Enh-PBS experienced the highest *SLR_b* with about $1.5x10^{-1}$ (i.e. from $1.54x10^{-1}$ to $1.48x10^{-1}$). The source priority assignation scheme and the preventive drop policy that concentrate losses in B-cells explain these results.

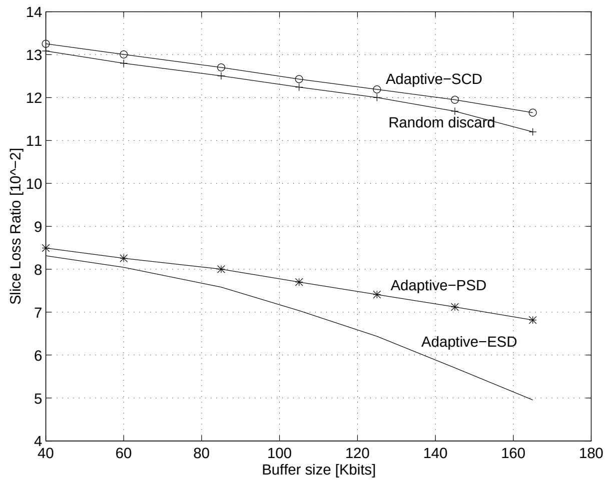

Figure 5 - Slice Loss Ratio for the aggregate video stream

In comparison with RD and enh-PBS, the slice-based drop mechanisms significantly improve the overall probability of receiving non-corrupted slices at destination. From Figure 5., in the best case scenario (i.e. the largest buffer size), the slice loss ratio for the aggregate video stream *(SLR_agg),* which measures the number of partially corrupted or totally lost slices, is estimated to $6.8x10^{-2}$ for A-PSD and $4.9x10^{-2}$ for A-ESD. In the other hand, RD and enh-PBS performed equally at the Application layer with an estimated *SLR_agg* of $11.5x10^{-2}$.

These simulation results demonstrate that burst discard strategy provides better performance at the application layer since more complete video slices are correctly received at the destination.

To compare the performance at the network level, we have evaluated the gain of applying group cell discard into the effective throughput. As explained in the introduction, video slices can be decoded and displayed by the recipients only if all the data of the slice are correctly received. Without any error recovery mechanisms at the destination, it is useless to keep transmitting a partially discarded slice.

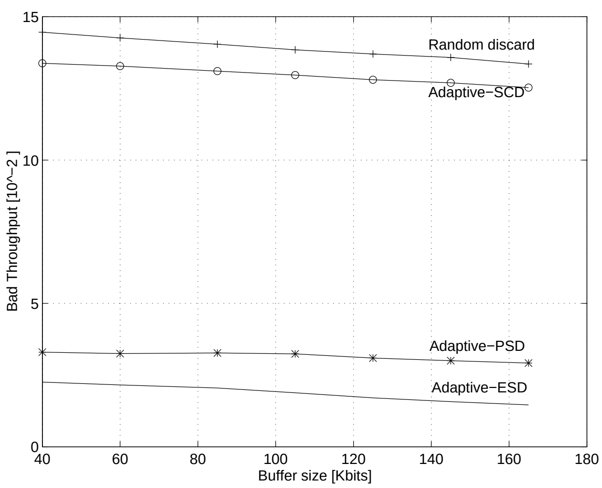

Figure 6 - Cell Bad Throughput

From [11], A-ESD and A-PSD demonstrated the highest cell loss ratio regardless to the picture type. Because they stop elimination only at the reception of a mark of end of video slices (i.e. EOS cell) while the other schemes stop as early as the congestion ends, they show the worst performance at the cell level.

Nevertheless, as illustrated in Figure 6, the cell bad throughput *(CBT)* crossing the network and measured at the application layer is dramatically reduced by a factor of 10 when running A-ESD and of five with A-PSD.

Except with A-ESD, *CBT* seems to be independent to the buffer configuration and remains constant during the simulation. When buffer's queue size reaches 165 Kbits, *CBT* values are $1.33x10^{-1}$, $1.25x10^{-1}$, $3x10^{-2}$ and $1.4x10^{-2}$ for respectively RD, enh-PBS, A-PSD and A-ESD.

One may expect that A-ESD should encouter no cell bad throughput. We noticed that when Intra-coded frames are received the buffer is quickly filled up and the scheme switchs to the mode {4} and performs like A-PSD by dropping already served video slices. When increasing $Q_{\max}$, the bad throughput is progressively reduced to the minimum.

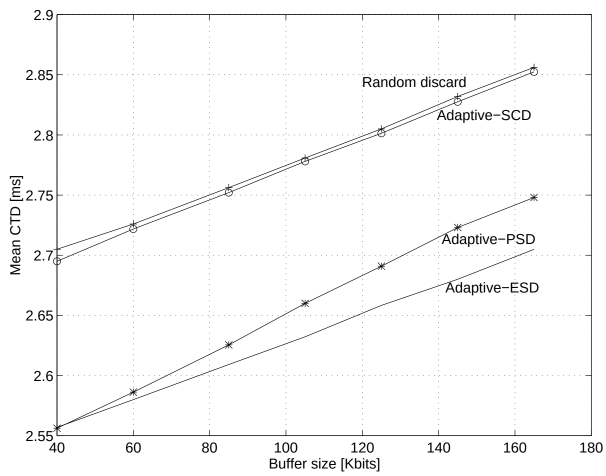

Figure 7 - Mean-Cell Transfer Delay for the aggregate stream

Finally, Figure 7 depicts the mean cell transfer delay *(mean CTD)* versus the buffer size for the aggregate stream. *Mean-CTD* increases in order of magnitude of $Q_{\max}$ regardless to the used drop technique. With limited buffer size (e.g. smaller than 110 Kbits), enh-PBS performs similarly to Random Discard, while the packet video drop schemes show better performance. This is explained by the elimination of groups of cells at the basis of video slices when the current queue length exceeds L_T (i.e. low threshold). The average buffer occupancy rate is reduced which in turn directly reduces the mean cell service delay.

Actually, the transmission delay differences are non-significant since the lowest mean CTD is experienced when running A-ESD with an estimate value in the range of 2.55 and 2.7 ms. For RD and enh-PBS, it linearly increases from 2.7 to 2.85 ms.

V. CONCLUSION

In this paper, we have proposed and compared two intelligent video packet drop schemes, called Adaptive-Partial video Slice Discard and Adaptive-Early video Slice Discard. These schemes aim to improve both transmission performance and picture quality of MPEG-encoded applications over ATM best effort services (e.g. ABR, UBR+, GFR). By combining a preventive, discriminative and grouped packet discard strategy at the basis of video slice rather than cell., the proposed mechanisms substantially increase the number of uncorrupted slices received at destination. In addition, we measured a significant reduction of the bad throughput crossing the network without any bad impact on the mean end-to-end cell transfer delay.

According to the simulation results, the two slice-based discard schemes equally perform in protecting essential video slices (i.e. data issued from Intra-coded and Predictive-coded frames) from loss. The benefit is well measured at the application layer where the effect of error propagation is minimized. Without forward error recovery (FEC) and error concealment mechanisms at the end systems, we recommend A-ESD since he guarantees the highest effective bandwidth utilization rate. Nevertheless, these data protection and recovery techniques are required to reduce the dramatically high video packet loss (i.e. in the order of 10^{-2}) encountered with unreliable best effort services. In such system configuration, A-PSD seems to be more suitable since he permits partial video data recovery and correction at destination.

REFERENCES

[1] ISO/IEC International standard 13818-2, "Generic Coding of Moving Pictures and Associated audio information: Part2 - Video", 1995.

[2] B. DeCleen, P. Pancha, M. El Zarki, ‘Comparison of priority partition methods for VBR MPEG’, IEEE INFOCOM’94, Toronto, Canada, June 1994, pp.689-96.

[3] A. Mehaoua, R. Boutaba and G. Pujolle, ' An Extended Priority Data Partition Scheme for MPEG Video Connections over ATM', IEEE Symposium on Computers and Communications 97, Alexandria, Egypt, July 1997, pp. 62-67.

[4] Y. Wang and Q-F. Zhu, 'Error Control and Concealment for Video Communication: A review', Proceedings of the IEEE, Vol. 86, N.5, May 1998, pp. 974-997.

[5] G. Ramamurthy and D. Raychaudhuri, ‘Performance of packet video with combined error recovery and concealment’, IEEE INFOCOM’95, Boston, April 1995, pp. 753-61.

[6] W. Feng and S. Sechrest, ‘Smoothing and buffering for delivery of prerecorded compressed video’, ACM Computer Communications Review, October 1995, pp. 709-717.

[7] J.D Salehi, Z.-L Zhang, J.F Kurose, and D. Towsley, ‘Supporting stored video : Reducing rate variability and end-to-end resource requirements through optimal smoothing’, ACM SIGMETRICS '96, May 1996, pp. 222-231.

[8] G. Armitage and K. Adams, 'Packet Reassembly during Cell Loss', IEEE Network Magazine, Vol.7, No.5, September 1993, pp. 26-34.

[9] A. Romanov and S. Floyd, 'Dynamics of TCP Traffic over ATM networks', ACM SIGCOMM'94, pp. 79-88, Sept. 1994.

[10] A. Mehaoua and R. Boutaba, 'Performance Analysis of a Partial Slice-based Discard Scheme (Adaptive-PSD) for MPEG Video over UBR+ Service, in 13th IFIP International Conference for Computer Communications (ICCC'97), Cannes, France, 19-21 November 1997.

[11] A. Mehaoua and R. Boutaba, 'Performance Analysis of Cell Discarding Techniques for Best Effort Video Communications over ATM Networks', in Computer Networks and ISDN Systems Journal, Vol.29, issue 17-18, Elsevier Science, North Holland, February 1998, pp. 2021-2037.