An Adaptive and Selective Cell Drop Policy with Dynamic Data Partitioning for Best Effort Video over ATM

Ahmed Mehaoua[1,2], Raouf Boutaba[1] and Guy Pujolle[2]

[1] Computer Research Institute of Montreal,
1801, McGill College, Montreal (Qc), Canada
Email: {amehaoua, rboutaba}@crim.ca

[2] University of Versailles,
45 Av. des Etats-Unis, 78000 Versailles, France
Email: Guy.Pujolle@prism.uvsq.fr

Abstract

In this paper, we propose and evaluate a new MPEG-based video delivery framework for use with ATM best effort services (e.g. Available Bit Rate and Unspecified Bit Rate). The presented framework relies on three components: a video-oriented cell discarding scheme, which adaptively and selectively adjusts drop level to switch buffer occupancy and video cell payload types; a dynamic frame-level priority data partition mechanism based on MPEG data structure and feedback from the network; and an enhanced ATM Adaptation Layer type 5 associated with a new slice-based MPEG2 encapsulation strategy. This best-effort video delivery framework is evaluated using simulation and real MPEG video data.

Keywords : ATM, AAL5, MPEG, Best Effort, Quality of Service, Discarding.

1. Introduction

To address the problem of transmission of compressed variable bit rate video over lossy networks, several protection and recovery techniques have been proposed to minimize video quality degradation due to cell loss. Layered coding with prioritization is one of the most popular approach [1][2][3]. Forward error recovery schemes with destination concealment have also been designed to cope with this problem [4][5][6][7]. Another approach consists to reduce the burstiness and the peak bandwidth requirements of video by applying complex smoothing and buffering techniques at the source [8][9].

With bursty packet-oriented services (TCP, LANs) over ATM, various techniques have also been developed to preserve packet integrity and achieve higher effective throughput during overloads. These mechanisms may be classified in respect to the four ATM service classes.

Variable Bit Rate (VBR) service, with policing and smoothing techniques, requires accurate traffic characterization with source parameters declaration and is complex to manage for unpredictable connections [10].

ATM Block Transfer (ABT) service, with Fast Reservation Protocols (FRP), is shown to maintain high link utilization during network overloads with small buffer requirements [11][12].

Available Bit Rate (ABR) service, with end-to-end rate-based flow control, has been standardized by ATM Forum in 1995 to manage congestion of bursty data traffics [10]. However, like adaptive windowing mechanisms, the rate-based mechanism require one or more network round-trip times before it effectively react to congestion [13].

Unspecified Bit Rate (UBR) service, with no flow control and no loss guarantees, is the true best effort service and provides the least expensive service for the transport of packet-based applications. However, because of its simplicity, plain UBR with inadequate buffer sizes performs poorly in a congested network. Packet Tail Discard or Partial Packet discard (PPD) has been proposed to address this problem [14]. If a cell is dropped from a switch, the subsequent cells of the higher layer protocol data unit are also discarded. 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 [15] that achieves better throughput performance but does not guarantee fairness among the connections [16]. When the switch buffer queues reach a threshold level, entire higher level data units (e.g. AAL5 Protocol Data Unit) are preventively dropped. To improve its fairness, selective packet drops based on per-Virtual Circuit accounting have been introduced by Heinanen and Kilkki and referenced as Fair Buffer Allocation (FBA) [17].

However, none of the mentioned congestion control and quality of service (QoS) management schemes are focusing on the transmission of MPEG-encoded video packets over non-guaranteed ATM best-effort services. Thus in this paper, we are proposing a new delivery framework which targets to gracefully degrade picture quality during congestion period.

The balance of this paper is organized as follows. Section 2 is devoted to the description of the different components of the proposed best effort video delivery framework. In Section 3, we introduce the network model, the investigated performance parameters and we discuss the simulation results. Finally, we, conclude and present directions for future works in Section 4.

2. A Best Effort Delivery Framework for Video over ATM Networks

2.1 MPEG Video Coding

Next generation video applications will widely use Motion Picture Expert Group compression standards to save network resources. MPEG coding algorithm structures video data in a hierarchical format in order of increasing spatial size : 8x8 pixels Block, 16x16 pixels Macroblock, Slice, Frame, Group of Pictures and Sequence [18]. Two of them have significant impact on decompression and displaying process.

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, coding of a slice is done independently from its adjacent slices, making it the smallest autonomous unit. Consequently, slices serve as resynchronization decoding units in case of long error bursts due to transmission problems.

Frame or picture is the basic unit of display. Three picture types may be present in MPEG streams. They differ from the applied coding method : Intra-coded (I) picture, Predictive-coded (P) picture and Bidirectionally predictive-coded (B) picture. Intra- and Predictive-coded pictures are essential and have to be preserved from corruption during transmission. Indeed, Due to error propagation 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.

From the presented video properties, three obvious remarks stand out. First, the smallest transmission data unit is rather slice than cell, ATM Adaptation Layer (AAL) or MPEG system (e.g. Transport Stream or Program Stream) packet. Secondly, in situation of congestion, dropping video cells indiscriminately can cause serious degradation in picture quality. I- and P frames have to be better protected during transmission. Finally, unlikely to delay insensitive applications, error recovery techniques based on retransmission are useless to recover corrupted data. Therefore, best effort video delivery services should be evaluated in their ability to provide an efficient message-based service. In this paper, the term message is interpreted specifically in reference to MPEG video. A message may be an encoded block, macro-block, slice or an entire picture.

2.2 A Video-oriented Dynamic Priority Assignation Scheme (Dynamic-PAS)

Since human perception is less sensitive to low frequency components of a video signal, subsequent blocks are transformed into the frequency domain using Discrete Cosine Transform (DCT). Each transformed block may then be partitioned into an essential layer (comprising the lowest frequency DC coefficient), and an enhancement layer (consisting of the set of high frequency AC coefficients). The information contained in the essential layer is packetized and transmitted at high priority, which ensures a guaranteed quality of service. Information in the enhancement layer is transmitted at a low priority, which provides only a best effort service. The Cell Loss Priority (CLP) mechanism is usually used to provide a two-level priority service within a single ATM channel.

In [19], this previous technique is applied at the macroblock layer. Similarly to the block-based approach, the DC value for each of the 8x8 block are assigned to the high priority (HP) stream. The macroblock header, and the motion vector in case of P-frames are also included in the HP stream. For the remaining 63 DCT coefficients of each block, the authors define a parameter β which specifies the number of AC coefficients that are to be placed in the HP stream. The remaining (63-β) coefficients are transmitted in a low priority (LP) stream. To allow the regeneration of the original bit stream by the destination, the macroblock address is joined to the LP information.

To overcome the limited priority capability of the CLP mechanism, a connection-level prioritization approach is evaluated in [20] to transport a layered MPEG-2 video application over different ATM service classes. The scheme uses static data partition between two connections by means of a Load Balancing factor (LBF). The virtual connections are associated with a guaranteed service class (i.e VBR-rt) and a best effort service class (e.g. ABR) to respectively carry the base layer and the enhancement layer.

The drawbacks of these frequency domain techniques are the added complexity and the special devices required at the destination to synchronize and recover the original video stream.

Consequently, data partition with priority assignation can simply be implemented at the frame level. The cells belonging to following frames are set to different priorities. For instance I-frame cells may have the highest priority over P and B-frame cells. In [21], two static priority partition strategies are proposed which also use the CLP-bit :

1. *Static I/PB priority partition* : in this method, I-frame cells are considered with a high priority and have their CLP bit set to '0', while P- and B-frame cells are assigned a lower priority with CLP flag set to '0'. If a congestion occurs, cells from P- and B-frames are discarded first.
2. *Static IP/B priority partition* : in this variant, I- and P-frames cells are both considered with high priority and are better preserved from elimination. Only B-frames cells are assigned a low priority.

The main drawback of these methods is that they can not dynamically adapt to network load changes.

Therefore in this section, we propose a new video partition and prioritization scheme, named Dynamic Priority Assignation Scheme (Dynamic-PAS). The scheme is simple to implement and sufficiently generic to be performed at any MPEG data layer. In this paper, the emphasis is on the slice layer. The scheme uses the classical CLP bit and dynamically assigns cell priorities according to the current MPEG frame type (I, P or B) and the reception of backward congestion signals from the network.

Cells belonging to Intra-coded frame have the highest priority and their CLP-bit is set to '0'. B-frames have the lowest priority and the associated cells have a CLP value of '1'. Regarding to P frames, they are alternatively assigned high and low priority depending on the network load. At the beginning of the transmission, P-cells are initialized with a high priority (e.g. IP/B partitioning mode).

When the buffer queue length (QL) exceeds an upper threshold, an early congestion is detected and the ATM switch emits a signal to the source which in turn, adjusts P-cells priority level to low (i.e. I/PB partitioning mode). In our implementation we use forward Resource Management (RM) cells, with Congestion Indication flag (CI) tagged, to notify the destination and afterward the source . These cells are called I/PB-RM cells (see Figure 1). We choose RM cells rather than EFCI mechanism to notify the sources for reliability consideration. Indeed, the switch is highly congestion, all cells are dropped and none are then available for carrying the indication signal to the source.

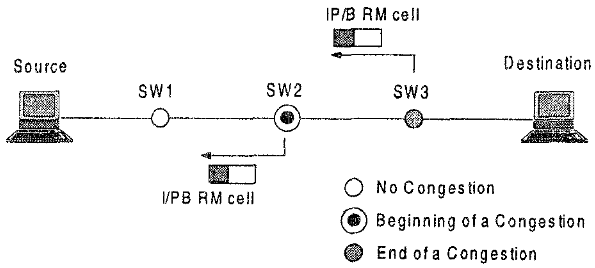

Figure 1 - Transmission of I/PB RM cells

When the queue length decreases below a lower threshold, a new signal (i.e. using IP/B RM cells) is transmitted to the source to modify its data partitioning mode. P-cells priority are switched back to a high priority (e.g. IP/B mode).

The PTI ATM-user-to-ATM-user bit (AUU flag), is employed to indicate whether it is the last cell of an upper message. We propose to use this flag to distinguish between successive slices. In this paper, the cell having its PTI-AUU flag set to '1' is termed End Of Slice (EOS) cell. Using Dynamic-PAS, each CLP and PTI bit singly assume its original meaning according to the standard. This dynamic priority assignation mechanism takes the advantages of both Static I/PB and Static IP/B priority partition. The main drawback of the scheme is that is stringently dependent to the round trip time delay, and thus to the network topology and link distances.

However, to support this slice-based mechanism, enhancements to AAL5 and a new MPEG2 video stream encapsulation strategy are required.

2.3 A Slice-based MPEG2 Video Stream Encapsulation Strategy

The key factor that controls the end-to-end performance is the ATM adaptation layer. AAL type 5 is currently the most commonly used. It is used for encapsulating User-Network Interface (UNI) 4.0 signaling messages [22], Private Network-Network Interface (PNNI) topology information exchange [23] and, in most cases for transporting non-real time data traffic. In 1995, the ATM Forum has recommended the carrying of constant bit rate MPEG-2 streams over

AAL5 with a null convergence sub-layer [24]. However, AAL5 is inadequate for transmission of variable bit rate video and required extended features. Forward error correction, jitter removal, multiplexing and multi-priority support are some of them. AAL type 2 is supposed to address these requirements but is not yet standardized.

As illustrated in Figure 2., uncompressed video frames, called presentation units, are individually encoded according to the MPEG standard and are referenced as access units. The stream produced by these access units is then named elementary stream. The next step is its packetization. The resulting stream is called Packetized Elementary Stream (PES).

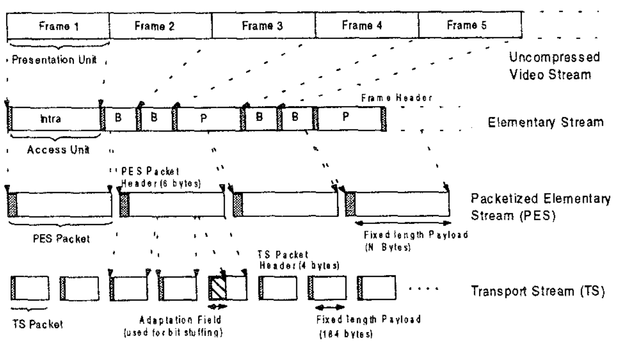

Figure 2 - PES Encapsulation using fixed length packet

There is no specific requirements for encapsulating encoded data in a Packet Elementary Stream (PES). This means that an access unit may start at any point within a PES packet. In addition, more than one access unit may be present in one PES packet. Nevertheless, the way this packetization is done can significantly affect the performance of the decoding process and the quality of the service provided by the network. For instance, if each PES packet contains exactly one MPEG data message (e.g. frame, slice, macroblock or bloc) the decoder can easily determine the start and end of the message. Similarly network transport and control policies can take benefit of this structure to offer a guaranteed packet-oriented service. However, this approach requires use of variable size packet, which induce a slight added complexity for the encoding process.

Instead of encapsulating MPEG video data at the macroblock level, as in the method by Ghanbari and Hugues [25], we propose that each PES is built from a single encoded video slice (see Figure 3). Indeed as mentioned in the introductory section slice, is the main coding processing unit and the smallest autonomous unit. Coding and decoding of blocks and macroblocks are feasible only when all the data of a slice are available.

At the next step, we propose to segment the PES packet into a number of 188-byte fixed length Transport Stream (TS) packets. In respect to the MPEG system multiplex layer [26], every TS packet embeds data from only one PES packet. In our case, the last Transport Packet may not be completely full since it is unlikely that a variable PES packet will fit exactly into an integer number of transport packets. Thus as shown in Figure 3., stuffing bytes are placed in the adaptation field to complete the payload.

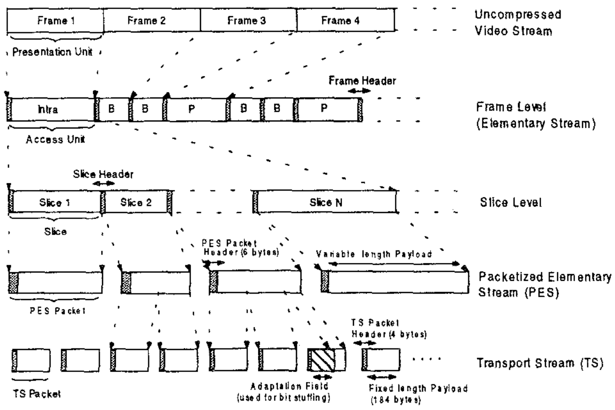

Figure 3 - Slice-based PES Encapsulation using variable length packet

In the worst case, e.g. padding 183 bytes, we evaluate the average overhead to 0.2 % with a single slice per frame, to 3.5 % with 15 slices per frame. We assume a NTSC TV broadcast quality of 512x480 picture resolution. The introduced overhead is highly dependent on the distribution and the number of slice per frame. However, it can be avoided using Program Stream (PS) packets.

2.4 An AAL5 Service Specific Convergence Sublayer (SSCS)

At the ATM Adaptation Layer Service Access Point (AAL SAP), the transport layer passes the TS packets to the Service Specific Convergence Sublayer (SSCS) using message mode service with blocking/deblocking internal function [27]. The following AAL_UNIDATA_REQUEST (ID, M, SLP, CI) primitive is used. The 'Interface Data' (ID) parameter specifies the exchange TS packet. The 'More' (M) parameter indicates if it is the last AAL Service Data Unit (SDU) of the upper message (e.g. end of the slice). The 'Submitted Loss Priority' (SLP) parameter gives the priority level of the TS packet.

In this paper, we propose to extend its range from two to three possible values and to initialize it in respect to the 'picture_coding_type' field located in the MPEG frame header [26]. This field specifies for each frame, the used coding mode (e.g. Intra, Predictive or Bi-directional Predictive). Consequently, three different types of SSCS-PDUs are newly defined : high-priority (I-frames), medium-priority (P-frames) and low priority (B-frame). This parameter also indicates how the 'SLP' parameter of the CPCS_UNIDATA_invoke primitive shall be set. Finally, the 'Congestion Indication' indicates how the 'CI' parameter of the CPCS_UNIDATA_invoke primitive shall also be set.

As illustrated in Figure 4., SSCS groups every three (3) TS packets and adds a header and a trailer information. The header is composed of a 4-bit Sequence Number (SN), and a 4-bit SN Protection (SNP). The trailer consists of a 3-byte Forward Error Correction (FEC).

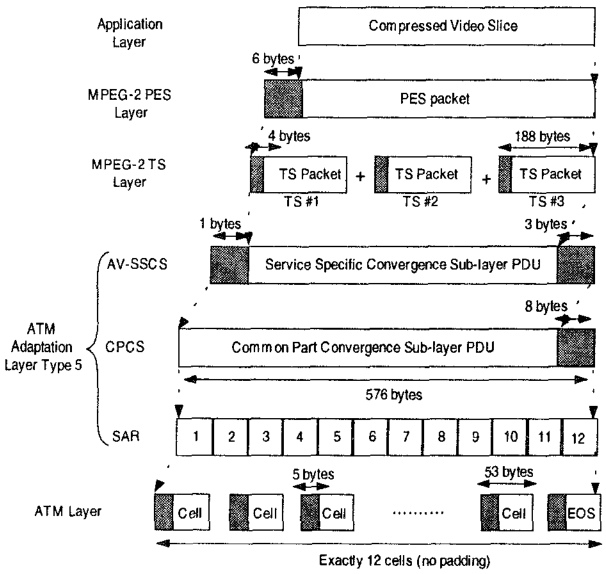

Figure 4 - New Mapping of MPEG Slices on AAL5

The FEC scheme uses a Reed-Solomon (RS) code [28], which enables the correction of up to 2 loss bytes in each block of 564 bytes (e.g. 3x188). The addition of a sequence number modulo-16 of 4 bits enables the AAL5 receiver to detect and locate up to 15 consecutive SSCS PDU losses. When losses are detected, dummy bytes are inserted in order to preserve the bit count integrity at the receiver. The SNP contains a 3-bit Cyclic Redundant Code (CRC) generated using the generator polynomial $g(X)=X^3+X+1$, and the result is protected by an even parity check bit. The SNP field is then capable of correcting single bit errors and detecting multiple bit errors.

The overhead introduced by the SSCS is 0.7 % and the delay is three TS packets (about 12 cells) at the transmitter and the receiver.

The SSCS-PDU are then transmitted to the Common Part Convergence Sublayer (CPCS) using the CPCS_UNIDATA_Invoke primitive. The 8-byte CPCS trailer information is appended to the CPCS SDU and no byte padding is required. The resulting Convergence Sublayer PDU is passed to the Segmentation And Reassembly (SAR) layer using the SAR_UNIDATA_Invoke primitive.

The under-laying SAR protocol will subsequently segment the CS-PDU into exactly twelve (12) 48-byte ATM SDUs. The ATM layer will then marked the CLP field of every cell using the 'AUU' and the 'SCLP' parameters of the AAL_DATA_Request [25].

2.5 An Adaptive and Selective Cell Discard Scheme (Adaptive-SCD)

One of the simplest switch buffer scheduling algorithm is to serve cells in First-In First-Out (FIFO) order. If buffer congestion occurs, the incoming cells are dropped regardless to their importance. This random discard strategy is not suitable for video transmission.

In this section we propose a variant of the Selective Cell Discard (SDC) scheme [11], which provides better performance for carrying video streams over lossy environment. The proposed Adaptive-SDC scheme [8] is associated with Dynamic-PAS and the Extended AAL5 to form a quality of picture (QoP) control framework. The aim of this framework is to ensure graceful picture degradation during overload periods. It allows accurate video cell discrimination and progressive drop by adjusting dynamically Adaptive-SCD mode in respect with cell payload types and switch buffer occupancy.

During light congestion, we propose to drop a lower priority cell first rather than delayed it and give its buffer space to a higher priority cell. This proactive strategy is performed gradually by including medium and high priority cells .

As depicted in Figure 5., four buffer thresholds are used : Stop_Threshold (ST), Low_Threshold (LT), Medium_Threshold (MT) and High_Threshold (HT). These thresholds define three operation modes. The utilization of four thresholds instead of two reduces the speed of oscillation for the Dynamic-PAS Resource Management (RM) notification mechanism and have shown better performance.

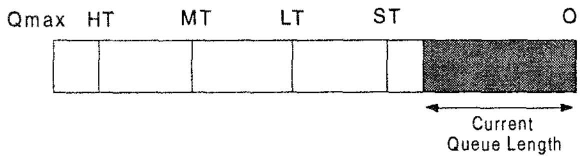

Figure 5 - Switch Buffer thresholds

- Mode {idle} : If the buffer queue length (QL) is lower than Low_Threshold no cells are discarded.
- Mode {1} : If the total number of cells in the buffer exceeds Low_Threshold but is still below High_Threshold, a medium congestion is detected and only the low priority cells are eligible for elimination. This mode stops when QL falls down to Stop_Threshold.
- Mode {2} : When QL exceeds High_Threshold, all the incoming cells are eliminated until the queue length drops below Medium_Threshold.

Queue Length	*Increase direction*	*Decrease direction*
$QL \leq ST$	Idle	Idle
$LT \leq QL < ST$	Idle	{1}
$MT \leq QL < LT$	{1}	{1}
$HT \leq QL < MT$	{1}	{2}
$QL \geq HT$	{2}	{2}

Table 1 - Adaptive-SCD Operation Modes

The I/PB RM cells are transmitted to all the video sources when MT is exceeded, while IP/B RM cells are send only when QL drops below LT. At the reception of feedback signals, the sources immediately change their operation mode. Consequently, a single P-frame may be transmitted with cells of different priority.

Using this adaptive discarding technique low priority cells are firstly dropped to quickly reduce buffer occupancy during light congestion, while higher priority cells are preserved from elimination. If the congestion worsens, all the cells are progressively candidate to discarding. This approach allows a graceful and controllable picture degradation with low operation complexity.

3. Performance evaluation

3.1 Network Model

As illustrated in Figure 6., we consider a network simulation model composed with an ATM switch, an OC-1 (e.g. 55.1 Mbps) bottleneck link (L) and three VBR MPEG connections. The distance between the sources and the switch are constant and set to 0.2 km (e.g. 0.125 miles). To emulate a Local Area Network (LAN), 'L' is initialized to 2.5 Km (e.g. 1.56 miles). 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.

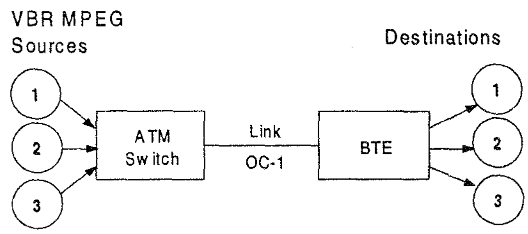

Figure 6 - Network Model

The VBR video connections are generated using three MPEG-1 frame traces : 'Star-Wars', 'Tennis' and 'Soccer' [29]. The main statistics of the MPEG sequences are summarized in Table 2. and 3.

	Star Wars	Tennis	Soccer
Compression Rate (X:1)	130	121	106
Quantizer scale	I : 10	P : 14	B : 18
GOP pattern (N=12 M=2)	IBBPBBPBBPBB		

Table 2 - Coding parameters

	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 3 - Statistics of the video sources

Every connection starts transmitting at different times in the range [0, 83.3 ms.]. Indeed, we have noticed that the deterministic Group Of Picture (GOP) pattern of the video sequences have an important impact on switch buffer occupancy and confirms the conclusions in [30]. Thus, the first frame (e.g. I-frame) of the video sequences are desynchronized to ovoid I-frames overlapping and therefore periodic buffer congestion.

In this paper, we assume 15 uniformly distributed slices per frame and a simulation time of 1.43 min, which represents 37080 slice transmissions. We also assume a shared output FIFO buffer and a constant decoder output rate during transmission of a frame. The level of congestion is measured by means of thresholds.

Seven switch buffer configurations are investigated. For each of them, the same method is applied to determine the values of the four thresholds. HT, MT, LT and ST are respectively set to 1.0, 0.9, 0.8 and 0.7 fraction of the maximum buffer size (Qmax). Qmax is varying in the range of 40 to 165 Kbits.

3.2 Performance Parameters

Let us define a lost cell as a cell discarded by a switch.

Let us define the aggregate cell loss ratio, as the number of lost cells from the three video connections vs. the total transmitted cells.

Let us define the I-frame cell loss ratio *CLR-i,* as the number of lost cells belonging to I-frames from the three connections vs. the total number of transmitted I-cells. The same metric is applied for P- and B-frames and are referenced as *CLR_p* and *CLR_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_{0K}): $D_K = t_{0K} - t_{iK}$.

During simulations, the processing delays at the ATM and AAL5 layers are not explicitly modeled. We assume that their contribution to the end-to-end delay experienced by the cell is relatively constant, and thus it can be omitted. In this paper, emphasis is on the variation of the mean-CTD for the aggregate stream.

3.3 Simulation Results

In this paper, we compare the performance of Dynamic-PAS with the three following static CLP-based partition techniques.

- CLP_1 : Static IPB/- frame partition (No Priority)
- CLP_2 : Static IP/B frame partition
- CLP_3 : Static I/PB frame partition

For all these prioritization schemes, the switch runs the proposed Adaptive and Selective Cell Discard (Adaptive-SCD) scheme. Actually, exception is on 'CLP_1' where random discard is performed when High Threshold is exceeded. This hypothesis allows us to simulate a non selective discard mechanism.

Figures 7., 8., 9., and 10. show the cell loss ratio for respectively the aggregate, I-, P- and B- cell flows. From Figure 7., we can notice no significant difference between the aggregate cell loss curves for the different priority assignation techniques. Since the same dropping mechanism is applied, e.g. Adaptive-SCD, we obtain approximately the same loss ratio. This ratio decreases from about 3.43 % to 2.6% while the buffer size increases by a factor of four.

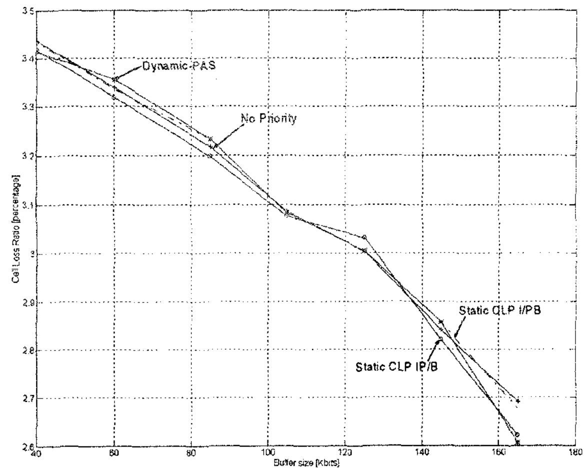

Figure 7 - Aggregate Cell Loss Ratio

As illustrated in Figures 8., 9. and 10., Dynamic-PAS concentrates the loss within the B-frames and protects efficiently the reference I- and P-frames. Indeed, using automatic switching between two operation modes, it takes advantages of the both static partition methods. For the protection of I-frames, it performs as good as 'Static CLP I/PB' and much better than the two others. For the protection of P-frames, 'Static CLP IP/B' scheme gives the best results. This is due to the fact that it maintains the highest priority level to P-frames during whole communication time. However he and 'Static CLP IPB/-' experience the worst I-cell loss ratio and they are not recommended to carry encoded video streams. This is due to the GOP structure of the sequence and the great sizes of the Intra-frames. Indeed, when an I-frame is transmitted, the buffer queue length rapidly increases to accommodate the burst. The two schemes start to drop the I-cells immediately and stop when the queue length decreases below ST. The same phenomena is repeated with the following P-frame. Indeed, 'Soccer' sequence has several scene changes during the simulation period. This yields to code some of the frames as intra-coded picture rather than predictive coded. To illustrate this characteristic, the mean size of the I and P-frames are respectively 65.9 Kbits and 38.2 Kbits for 'Soccer', and 57.4 Kbits and only 16.3 Kbits for 'Star-wars'.

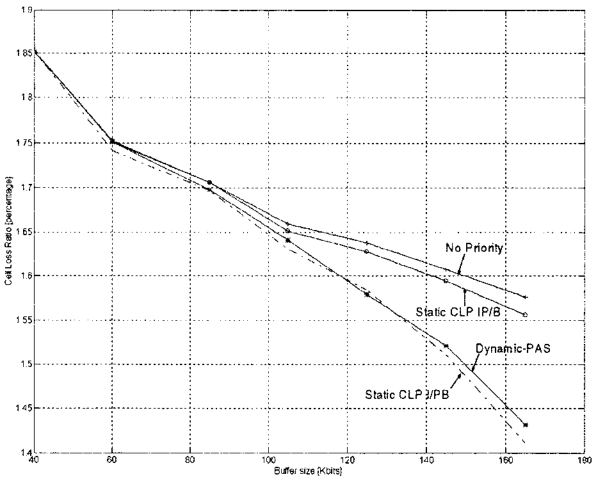

Figure 8 - (I)ntra-frame Cell Loss Ratio

When the buffer size is small (e.g. Qmax lower than 120 Kbits), Dynamic-PAS performs like the other schemes. This can be explained by the fact that the buffer can not simultaneously accommodate cells from different connections. Thus when the lower and medium thresholds are crossed due to the arrivals of an I-frame burst, few B- or P-cells are available for discard. The queue keeps on rising and Adaptive-SCD switches to Mode {2}.

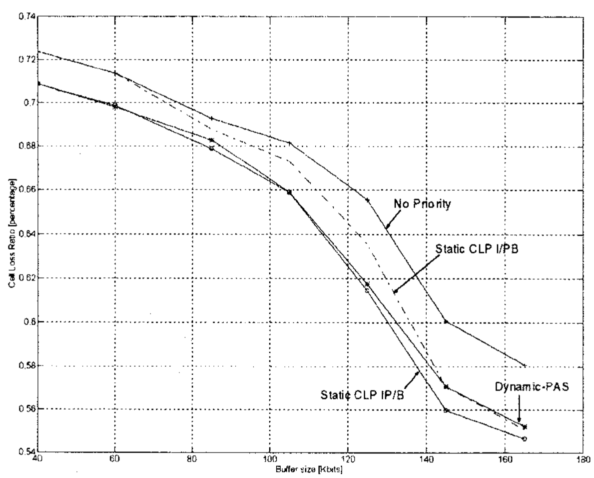

Figure 9 - (P)redictive-frame Cell Loss Ratio

It is interesting to notice that, even though 'Static CLP IPB/-' assigns a high priority to I-frame cells, it performs less efficiently than 'Static CLP I/PB' for preserving I-cells from loss. This is due to the random discard effect when High Threshold is reached.

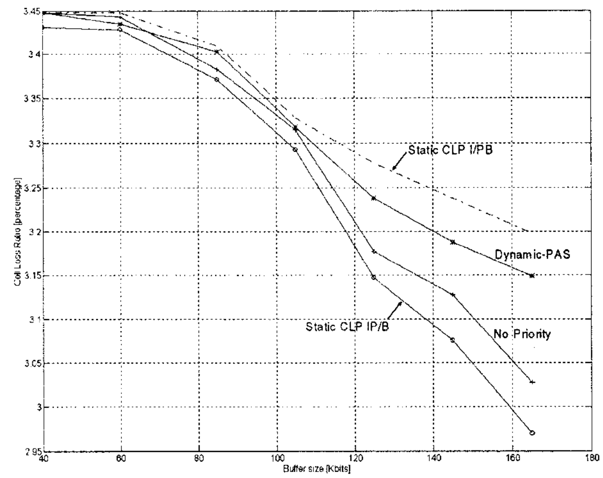

Figure 10 - (B)idirectional Predictive-frame Cell Loss Ratio

B-frames are the most concern by loss and contribute largely to the overall loss ratio. Afterwards, I- and P-cells are, in this order, the most subject of drop. This is explained by the frequency of B-frames in video sequences, as well as the drop policy of Adaptive-SCD. Indeed, in our MPEG video samples, the proportion of I, P and B data are respectively to 53%, 24% and 23 % of the aggregate stream. Due to the GOP pattern and the multiplexing process, B-frames occurs more often and are more likely to be discarded, e.g. 48 B-frame occurrences per second, for only 2 and 6 for I and P-frames.

From Figure 11., the mean-cell transfer delay increase in order of magnitude of the buffer size. We may notice that Dynamic-PAS has not much effect on the variation of the mean-CTD. With limited buffer size it performs similarly to the Static CLP-based schemes. A mean ratio varying from 2.7 and 2.9 ms. He even shows better results than 'Static CLP IP/B' and 'static CLP IPB/-' partition schemes while 'Qmax' rises. This is because 'Dynamic-PAS' starts to drop P-cells earlier than these methods and thus reduce the buffer occupancy much faster. When Queue length exceeds the medium threshold (MT), both B- and P-cells are eligible for elimination with Dynamic-PAS, which represent approximately fifty percent of the whole stream. With 'Static CLP IP/B' and 'Static CLP IPB/-', the buffer must be filled until High threshold to start P-cells elimination. The impact of this strategy on buffer occupancy lead on a greater mean CTD.

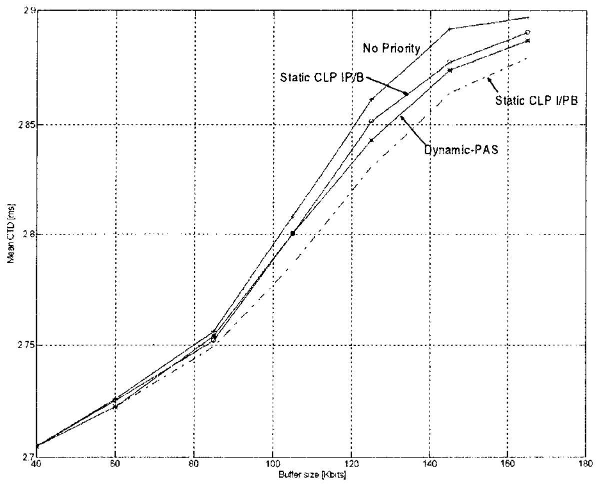

Figure 11 - Mean-CTD for the aggregate stream

4. Conclusion

In this paper we have shown that a dynamic priority assignation strategy can significantly improve the quality of service provided to video encoded applications. By automatically adjusting its partition mode to the network load, the proposed Dynamic Priority Assignation scheme (Dynamic-PAS) efficiently prevent critical video data (e.g. I and P- frames) from loss.

In association with an intelligent video-oriented discard scheme and an enhanced ATM Adaptation Layer type 5, Dynamic-PAS better exploits the structure of the MPEG video traffic and overcomes the difficulty imposed by random cell discarding. Together, they allow accurate cell discrimination and progressive cell group discard at the slice level, which lead to graceful picture quality degradation during congestion periods.

The proposed MPEG-based discard scheme, named Adaptive and Selective Cell Discard (Adaptive-SCD), selects the cell to be dropped with respect to MPEG data hierarchy and the current switch load. Simulation results show that the number of critical Intra-coded and Predictive-coded cells arriving at destination significantly increases using both Dynamic-PAS and Adaptive-SCD.

Future work will involve studying performance evaluation of Dynamic-PAS in WAN configurations and the adaptation of Partial and Early Packet Discard [15][31] mechanism to better suit with MPEG video transmission over ATM best effort services.

Acknowledgments

The authors are grateful to Oliver Rose, at University of Wüezrburg (Germany), for providing them with the MPEG video traces.

5. References

[1] P. Pancha and M. El Zarki, 'MPEG coding for variable bit rate video transmission', IEEE Communication magazine, pp.54-66, May 1994.

[2] B. DeCleen, P. Pancha, M. El Zarki, 'Comparison of priority partition methods for VBR MPEG', IEEE INFOOM'94, 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. To be published.

[4] W. Luo and M. El Zarki, 'Analysis of error concealment schemes for MPEG2 video transmission over ATM based network', SPIE'95, vol. 2501, pp. 1358-68, 1995.

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

[6] S. Lee, 'Cell loss and error recovery in variable rate video', Journal of Visual communication and image representation, pp.39-45, March 1993.

[7] A. Mehaoua, R. Boutaba, and G. Pujolle", A Picture Quality Control Framework for MPEG video over ATM", IFIP/IEEE Workshop on Protocols for High Speed Networks'96, Sophia-Antipolis, France, October 1996.

[8] W. Feng and S. Sechrest, 'Smoothing and buffering for delivery of prerecorded compressed video', Computer Communications, October 1995, pp. 709-717.

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

[10] ATM Forum, TM-SWG, 'Traffic Management Specification 4.0', at-tm-0056.000, April 1996.

[11] ITU Telecommunication Standardization Sector, recommendation I.371, "Traffic Management and Congestion Control in B-ISDN", Perth, November 1995.

[12] J. Turner, 'Managing bandwidth in ATM Networks with bursty Traffic', IEEE Network, vol6, n5, Sept. 1992, pp.50-58.

[13] F. Bonomi and K.W Fendick, 'The rate-based flow control framework for the Available Bit Rate ATM service', IEEE Network, pp. 25-39, March 1995.

[14] G. Armitage and K. Adams, 'Packet Reassembly during Cell Loss', IEEE Network Magazine, vol7, n5, Sept. 1993, pp. 26-34.

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

[16] R. Jain and al., 'Buffer requirements for TCP over UBR', ATM Forum 96-0518, April 1996.

[17] J. Heinanen and K. Kilkki, 'A Fair Buffer Allocation scheme', unpublished manuscript.

[18] D. LeGall, 'MPEG : a video compression standard for multimedia application', Communications of ACM, vol. 34, April 1991, pp.47-58.

[19] P. Pancha and M. El Zarki, 'Prioritized Transmission of Variable Bit Rate MPEG Video', GLOBECOM'92, pp. 1135-39, 1992.

[20] A. Mehaoua and R. Boutaba, 'Layered transmission of MPEG over VBR and ABR connections', Fifth International conference on Telecommunication systems, Modeling and Analysis, Nashville, TN, March 1997, pp. 425-430.

[21] T. Han and L. Orozco-Barbosa, 'Performance requirements for the transport of MPEG video streams over ATM networks', ICC'95, Washington, June 1995, pp. 221-225.

[22] ATM Forum, 'ATM User-Network Interface (UNI) Signalling specification Version 4.0', af-sig-0061.000, July 1996, pp.35-37.

[23] ATM Forum, 'Private Network-Network Interface (PNNI) Specification Version 1.0', af-pnni-0055.000, March 1996, pp. 54-62.

[24] ATM Forum SAA SWG, "Audiovisual Multimedia Services : Video on Demand", Specification 1.0, ATM_FORUM/af-saa-0049.000, December 1995.

[25] M. Ghanbari and C. J. Hugues, 'Packing Coded Video Signals into ATM cells', IEEE/ACM Transactions on Networking, vol1, n5, pp.505-509, October 1993.

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

[27] ITU-T I.363, "B-ISDN Adaptation Layer (AAL), Perth, November 1995.

[28] S. B. Wicker and V. K. Bhargava, 'Reed-Solomon Codes and their Applications', IEEE Press, 1994.

[29] O. Rose, 'Statistical properties of MPEG video traffic and their impact on traffic modeling in ATM systems', 20th Conference on Local Computer Networks (LCN'95), Minneapolis, MN, October 1995.

[30] O. Rose and M. R. Frater, 'Impact of MPEG Video Traffic on an ATM multiplexer', IFIP High Performance Networking '95, Palma, Sept. 1995, pp. 157-168.

[31] H.Li, K.Y. Siu, and al., TCP Performance over ABR and UBR services in ATM, IPCCC'96, March 1996.