INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

On the Need for Frame Discard in ATM Networks

Nelson L. S. Fonseca *Sergio A. Yunes*

Technical Report - IC-03-008 - Relatório Técnico

April - 2003 - Abril

On the Need for Frame Discard in ATM Networks

Nelson L. S. da Fonseca and Sergio A. Yunes
State University of Campinas
Institute of Computing
Campinas, Brazil
e-mail:nfonseca@ic.unicamp.br

Abstract - This paper introduces a novel policing mechanism, called Packet Leaky Bucket which marks all cells of a frame with the same level of priority. Moreover, it discusses the need for having frame discard mechanisms in ATM networks when frame-oriented policing mechanisms are used.

I. INTRODUCTION

The ATM network technology was conceived in the middle 1980's to be the technology for a universal multimedia network. Although such goal has never been achieved, a mature transfer technology was established and, currently, several backbones use ATM. Besides that, the TCP/IP stack has been widely disseminated. Since a considerable amount of TCP flow are carried over ATM, it is of paramount importance to optimize such transport.

In ATM networks, IP datagrams are encapsulated in AAL-5 PDU's which are, then, divided into several ATM cells. The ATM Unspecified Bit Rate (UBR) service category provides best effort services with no Quality of Service (QoS) guarantee. UBR was created for transporting IP flows. When UBR is used by upper layer protocols which do not implement congestion control, such as UDP, there is a high risk of network collapse due to congestion. Such collapses can be avoided if the Available Bit Rate Service category is used. Nonetheless, the use of ABR implies in implementing complex shapers inside the end system [1].

The Guarantee Frame Rate (GFR) category is designed to furnish better support to the TCP/IP protocol stack than the one provided by UBR. GFR provides bandwidth guarantee and performs congestion control at the frame (datagram) level, rather than at the cell level as is done in UBR. Furthermore, for the end system, it is as easy to use as is the UBR service category.

In the GFR service category, the cells entitled to the guaranteed bandwidth are the ones conforming to the Frame Generalized Cell Rate Algorithm (F-GCRA). This algorithmis a variation of the Generalized Cell Rate Algorithm (GCRA), it marks all cells of a frame with the same priority value in an attempt to avoid the corruption of a large number of frames when congestion occurs. Nonetheless, F-GCRA may lead to situations where fewer cells are given service than would be eligible for the service guarantee. In other words, a lower number of cells are marked as high priority than it would be if GCRA were in place.

Cell discard can be pursued either at the cell level or at the frame level. Discarding at the frame level considers the frame boundaries but does not consider the cell priority level. In this paper, a new policing mechanism which takes frame boundaries into account without the same deficiencies presented by F-GCRA is introduced. This new mechanism is not restricted to ATM networks and can be used in any frame (packet / datagram) oriented networks such as IP DiffServ networks. Moreover, the need to adopt frame discard mechanisms in networks using frame-oriented policing mechanisms is also discussed.

This paper is organized as follow. Section 2 introduces the Packet Leaky Bucket algorithm for policing flows at the frame level. Section 3 gives a brief overview of information discard mechanisms. Section 4 presents numerical results about the interaction between policing and discarding mechanisms. Finally, in Section 5 conclusions are drawn.

II. THE PACKET LEAKY BUCKET

In GFR, there are two modes of operation, which differ by the interpretation of the Cell Loss Priority (CLP) bit value. In GFR.1, the priority value of a cell is related to all the other cells in the same connection. In GFR.1, there is no way to differentiate eligible frames from non-eligible frames for the guaranteed bandwidth based on the CLP value. The network is not allowed to change the CLP bit set by the end system. The bandwidth consumption of each connection has to be accounted at each switch in order to guarantee the minimum bandwidth to each connection.

In GFR.2, the CLP bit value indicates those cells conforming to the contract and are, thus, eligible for the guaranteed bandwidth. In the latter, the CLP bit value is determined by the F-GCRA algorithm, a variation for frames of the GCRA algorithm. Actually, F-GCRA assigns the same level of priority to all cells in a frame sets all cells of a frame with the same priority level which is determined in conformity with the level of priority of the first cell subject to a GCRA (T, L) algorithm, where T is the allowed inter arrival time and L is the burst tolerance. T and L are given, respectively, by 1 / MCR and (MBS - 1) x (1/MCR - 1/PCR), where MCR is the Minimum Cell Rate, MBS is the Maximum Burst Size and PCR is the Peak Cell Rate. The Simple Frame GCRA (SF-GCRA), a simplified version of the F-GCRA, is easier to implement than F-GCRA is.

One problem with F-GCRA is the drawback that fewer cells may receive guaranteed service (bandwidth) than are eligible. In networks using GCRA, a connection policed by GCRA (T', L') has at least as many cells marked as having high priority as when one is policed by GCRA (T, L), for $T' \leq T$ and $L'/T' \geq L/T$. However, this does not always happen when the F-GCRA algorithm is used. [1]. A novel policing mechanism, called Packet Leaky Bucket (PLB), introduced here, tries to overcome this shortcoming. PLB also marks all cells of a frame with the same priority value. The PLB is a Leaky Bucket (LB) coupled with a counter which accounts for the number of cells marked with a different priority value than it would be marked if the LB mechanism were used. The Packet Leaky Bucket algorithm is given in Figure 1.

If the counter value is positive, the number of cells marked as high priority has exceeded the number of high priority cells the flow should have. Conversely, if the counter value is negative, the number of cells given low priority is higher than the number of low priority cells the flow should have.

If at the arrival of the first cell of a frame the counter value is either zero or has a positive value then the frame is set as low priority, but if the counter value is negative, the frame will be assigned high priority.

Whenever an arriving high priority cell finds the bucket empty the counter is incremented. Otherwise, a token is consumed. If an arriving low priority cell finds at least a token in the bucket, a token is consumed and the counter is decremented. Otherwise, no action is taken.

This counter is used to reduce the discrepancy between the number of cells marked by the PLB as having a certain priority level and the corresponding number which would have been marked if LB were used. This reduction is performed frame by frame, i.e., any discrepancy arising in one frame is compensated in the next frame. Note that the maximum difference between the number of cells marked by PLB with a given priority value and that with the same priority value marked by LB is at most the maximum frame size minus one since the priority value of the first cell of a frame, by definition, conforms to the LB algorithm.

The PLB algorithm was validated by comparing the violation probability produced by itself to the violation probability produced by LB. Trace driven as well as synthetic-driven simulation were used. Real data used in the experiments are detailed in the next section. In synthetic-driven simulation, ON-OFF heavy tail with Pareto distributed ON and OFF periods were employed. PLB produces the same violation probability produced by the Leaky Bucket mechanism which was already expected given that the maximum possible difference between the number of cells marked with low priority (maximum frame size minus one) by the two mechanisms is much smaller than the trace length.

The Packet Leaky Bucket Algorithm

```
C - Counter
priority - priority value of a cell
first_priority - priority value of first celkl of a frame

if (arriving cell is the first cell of a frame)
    if ( $B = 0$ ) then
        if ( $C < 0$ )
            first_priority = high
        else
            first_priority = low
    else
        if ( $C <= 0$ )
            first_priority = high
        else
            first_priority = low
    priority = first_priority

else    # cell is not the first of a frame #
    if (priority = high)
        if ( $B > 0$ ) then
            $B = B - 1$
        else
            $C = C + 1$
    else
        if ( $B > 0$ ) then
        begin
            $B = B - 1$
            $C = C - 1$
        end {then}
```

Figure 1: The Packet leaky Bucket Algorithm

The use of PLB is not restricted to ATM networks. It can be incorporated in any network with variable-size PDU such as IP DiffServ networks. Actually, it is recommended for DiffServ networks since the described discrepancy can happen in any network based on variable size PDU, when the leaky bucket (GCRA) is adopted to policy flows at the packet (datagram) level [1].

III. INFORMATION DISCARD

The value of the CLP bit determines which cells should be discarded when congestion occurs. High priority cells have preference to stay in buffer. Discarding at the cell level does not take into account the boundaries of higher layer PDU's and, consequently, the lost cells may be spread among several frames. At the frame level, cells belonging to the same frame are firstly discarded. Frame boundaries and not the value of the CLP bit is the information considered for discarding. Since a corrupted frame is useless for data applications, retransmiting corrupted frames increases congestion. Several policies have been proposed for discarding cells at the frame level, i.e., policies which consider frames as unit of discard.

A cell discard mechanism is completely specified by a buffer organization policy and by a push-out policy. A buffer organization policy defines which buffer slot may be occupied by which cell. Push-out policies choose a cell to be discarded among the low priority cells. In the Complete Sharing with push-out buffer organization (CS), any buffer slot can be occupied by any cell irrespective of its priority level. Cells are dropped if and only if the buffer is full. When a high priority cell finds the buffer full and there is an enqueued low priority cell, then the low priority cells is dropped. Complete Sharing is loss-conserving since cells are lost only when necessary. Loss conserving policies maximizes the cell goodput [2]. Push-out policies differ by the age of the cell selected to be drop. The First-in-First-Drop chooses the oldest low priority in queue whereas the Last-In-First-Drop chooses the most recently arrived cell.

Tail Drop and Early Packet Discard are frame discarded policies widely implemented in ATM switches. In Tral Drop if a cell of frame is discarded, all cells of the same frame which arrive subsequently to the dropped one are also discarded. In Early Packet Discard, a threshold queueing position is defined. If a frame arrives and the queue length has passed the threshold position the whole frame is discarded.

IV THE NEED FOR FRAME DISCARD

Setting all the cells of a frame with the same level of priority in networks with discarding at the cell level increases the chances that losses are concentrated in a smaller number of frames than when they have different priority values. Therefore, there is a greater chance of having a smaller number of corrupted frames when frame-based policing mechanisms are used than when policing mechanisms that do not consider frame boundaries are employed. One of the cental questions investigated in this paper is thus whether the use of policing mechanisms at the frame level used in conjunction with cell discard mechanisms produces the same effect that frame discard mechanisms do Trace drive simulation was used to investigate this question.

In such simulation experiments, a set of sources feeds a single server employing a cell discarding mechanism (Figure 2). In some experiments, all sources are policed by LB while in others sources are policed by PLB. The metrics of interest is the frame goodput which is the fraction of non-corrupted frames that leaves the queue. The buffer organization is the complete sharing with push-out (CS), since such a policy it maximizes cell goodput, and, consequently, has the potential to maximize frame goodput. The push-out policy used was the First-In-First-Drop. FIFD gives the highest high priority cell loss rate [2], thus, tending to maximize the goodput of high priority frames, which should be the majority of the conforming frames in the flows. FIFD gives the highest overall frame goodput, as confirmed in the simulation experiments.

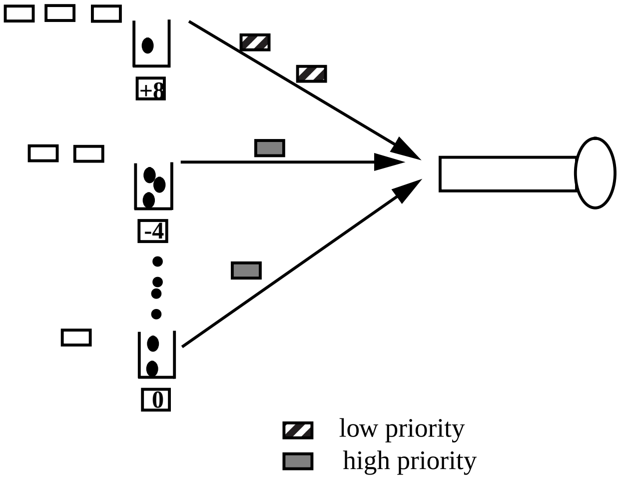

Figure 2: The Simulation Experiment

Similar experiments were performed using frame discard policies without policing. The frame discard policies used were Tail Drop (TD) and Early Packet Discard (EPD). These policies do not work towards the provision of guaranteeing bandwidth which is an added value offered by GFR. However, the present paper aims at investigating the need to have discard mechanisms at the frame level, and, thus, it is desirable to remove the effect of bandwidth guarantee in the simulation results. It is believed that GFR provides better services than UBR even when there is no bandwidth guarantee [1].

Traces of real data collected in IP over ATM networks were used for these simulations. Traces contain the size of IP datagrams encapsulated in AAL-5 PDU’s as well as the time stamp of their arrival time. The traces are publicly available at the National Laboratory for Applied Research (NLAR) site (http://www.nlanr.net). Table 1 shows characteristics of the traces used. Data shown in this paper corresponds to the experiments using the University of San Diego trace since they are representative of the results derived with the other traces. The San Diego trace has a Hurst parameter value of 0.91.

A single trace of datagrams (frames) was divided among N sources with consecutive datagrams being assigned to sources in a circular fashion. The same trace for each source is used in various experiments. The service rate, and consequently, the load, was varied by changing the duration of a time slot, i.e, the time required to transmit a cell. The arrival time of a datagram in the time stamp field of the trace was

Table 1: Description of Traces used in the Simulation Experiments

Network	Link Speed	Duration	# of IP Packets
San Diego Super Computer Center	OC12c	94 sec.	327.000,00
Michigan U.	OC3c	89sec.	434.000,00
Columbia U.	OC3c	83sec	325.000,00
Florida U.	OC3c	82 sec.	301.000,00
Indiana U.	OC12c	85 sec.	270.000,00

then translated into time slots. The buffer size is large enough to accommodate a cell of an average frame size for each source. The leaky rate of each LB was set to the mean arrival rate of each source, and the bucket size was set to the maximum frame size present in the corresponding trace.

In Figure 3, the packet goodput is plotted as a function of the load for the various simulations. Results for Tail Drop, EPD with threshold at 80% of the buffer size, CS used jointly with LB and CS jointly used with PLB are displayed. Comparing these results shows the advantage of ultilizing policing at the frame level. The frame goodput produced when PLB is used is 0.225 greater than that using LB. Moreover, the frame goodput decay is greater with the LB mechanism than when PLB is in employed. Tail Drop produces the highest goodput, since it does not discard frames unnecessary as EPD does. The difference between the goodput produced by CS jointly used with PLB and the goodput produced by Taildrop is approximately 0.02, which is not significant at all. Therefore, the same effect achieved by frame discard mechanism can be achieved with the emplyment of cell discard mechanisms in conjunction with policing mechanisms marking all cells of a frame with the same level of priority.

Another question of interest in the present investigation is whether it is advantageous to discard cells at both the cell level and the frame level. In such an integrated scheme, before discarding a frame, low priority cells will be dropped for releasing buffer slots to the incoming frame. Simulation experiments were pursued to assess the effectiveness of the integration of discard mechanism. Figure 4 illustrates the results of such experiments which uses the same scenario of Figure 3. CS and PLB are jointly used with EPD and also with Taildrop. It can be observed that the integration of both mechanisms does not bring any advantage at all. Discarding low priority cells produces corrupted frames without leading to the acceptance of a greater number of frames.

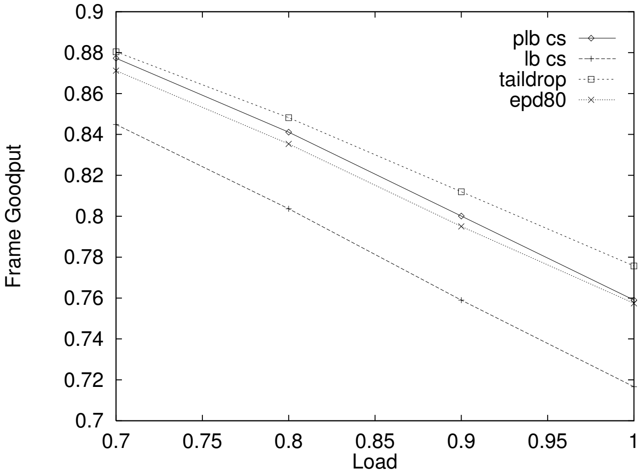

Figure 3: A Comparison Between Frame Discard Mechanisms and Cell Discard Mechanisms Jointly used with Frame-oriented Policing Mechanisms

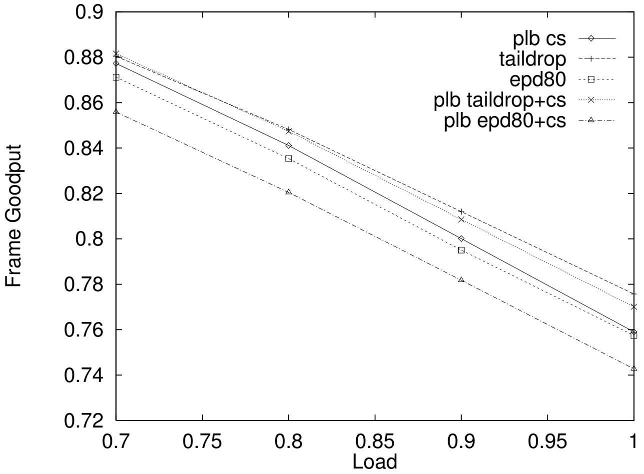

Figure 4: An Evaluation of the Effectiveness of the Integration of Frame and Cell Discard Mechanisms

V. CONCLUSIONS

This paper has introduced a new policing mechanism, the Packet Leaky Bucket, which marks all cells in a frame with the same level of priority. It is composed of a traditional Leaky Bucket accompanied by a counter which monitors the discrepancy between the cells marked by the Packet Leaky Bucket mechanisms and the cells that would be marked by the Leaky Bucket mechanism. This counter is used to control such discrepancies.

When all cells of a frame have the same priority value losses are clustered in a lower number of frames than when cells belonging to the same frame have different priority value. It was shown that the frame goodput produced when PLB is employed is much higher than when LB is used.

Moreover, it was shown that cell discard mechanisms used in conjunction with PLB produces a frame goodput produced equivalent to that rsulting from frame discard policies making the track of frame boundaries at the switches unnecessary at the line speed. Besides that, it was pointed out that it is not worthwhile integrating cell and frame discard.

The PLB policing mechanism can be easily extended to operate within the DiffServ framework for provisioning QoS in the Internet by counting the number of bytes in a datagram rather than the number of cells in a frame.

ACKNOWLEDGMENT

This work was partially sponsored by CNPq and by FAPESP

REFERENCES

[1] The ATM Forum Technical Committee, Traffic Management Working Group. Traffic Management Specification, Version 4.1, Technical Report, ATM Forum, March, 1999

[2] R. Cohen, N.L.S. Fonseca and M. Zukerman, “Traffic management and Network Dimensioning”, in Multimedia Communications Networks: Technologies and Services, M Tatipamula and B. Kasnabish Editors, pp. 133-173,Artech House, 1998.