PEPD: A Priority Based Packet Discard Scheme to Provide Service Differentiation in Internet

HONGJUN SU
School of Computing
Armstrong Atlantic State University
Savannah, GA 31419, USA.

MOHAMMED ATIQUZZAMAN
School of Computer Science
University of Oklahoma,
Norman, OK 73019-6151, USA.

ABSTRACT
The Internet was initially designed for non real-time data communications, and hence does not provide any Quality of Service (QoS) guarantees. Differentiated Services (DiffServ) is being developed by IETF to provide QoS guarantees by carrying priority information in the packet headers. ATM is expected to coexist with DiffServ in the core of the next generation Internet. In this paper, we *propose the Prioritized Early Packet Discard (PEPD)* scheme which takes into account the priority of packets during discarding of cells in an ATM switch. We developed a Markov chain model of PEPD, and verified the accuracy of the model with simulation results. Results confirm that the proposed PEPD scheme *provides QoS* to end applications by providing differential treatment to packets of different priorities.

KEY WORDS
Differentiated Services, Asynchronous Transfer Mode, Markov Chains, Modeling and Simulation.

1 Introduction

The Internet was initially designed for non real-time data communications, and hence provides only a Best Effort service. It is based on the TCP/IP protocol which uses *large* variable sized packets, and doesn't offer any Quality of Service (QoS) guarantee to applications. On the contrary, Asynchronous Transfer Mode (ATM) is based on *small* fixed sized packets, and has been developed to provide QoS guarantees to real time applications. Because of the QoS offered by ATM, the backbones of many Internet service providers have been implemented using ATM, while TCP/IP is the dominant protocol at the edge of the network, primarily due to the large installed base of TCP/IP-based legacy LANs and applications.

When a TCP/IP network is connected to an ATM network, a large IP packet is broken down into a number of small ATM packets (called cells) by an ATM Adaptation Layer (AAL) at the gateway between the two networks. If a cell is dropped by an ATM switch because of congestion, the destination TCP process drops the entire IP packet to which the dropped cell belongs; the TCP/IP sender has to retransmit the entire packet [1] thereby making the congestion scenario even worse. To solve the above problem, packet based discarding schemes [1, 2, 3, 4, 5, 6, 7, 8] have been proposed to increase end-to-end goodput of TCP/IP applications.

Early Packet Discard (EPD) [1], a packet-based discarding scheme for ATM switches, has been shown to increase the goodput of end-to-end TCP/IP applications by confining cell losses (in ATM switches) to a fewer number of packets, rather than spread out over a large number of packets. EPD drops entire packets that are unlikely to be successfully transmitted prior to buffer overflow. It prevents a congested link from transmitting useless cells, and reduces the total number of incomplete packets at the destination. EPD achieves this by using a threshold in the buffer. Once the queue length in the buffer exceeds this threshold, the buffer will only accept cells that belong to packets that have at least one cell in the queue or have already been transmitted.

Differentiated Services (DiffServ) [9, 10, 11, 12, 13] has been developed by the Internet Engineering Task Force (IETF) to provide QoS guarantees to TCP/IP-based applications. DiffServ is based on service differentiation, and provides aggregate service to the various service classes [10, 14, 15, 16] within the DiffServ framework. The services classes are implemented by classifying and *prioritizing* the packets based on their QoS requirements, such as loss and delay priorities. Priority information is carried in the packet headers.

It is widely believed that the core of the next generation Internet will be composed of a number of technologies such as DiffServ and ATM. It is therefore important to study the interworking issues between DiffServ and ATM. As an example, although the packets from the Assured Service [15] and the Best Effort Service [17] of DiffServ could be stored in the same queue in an ATM switch, they have to be treated with different priorities in the case of network congestion.

Although EPD can improve the performance of TCP applications as mentioned earlier, it can not distinguish between priorities of different applications. *EPD does not consider the priority of packets*, i.e. it treats all packets equally. However, in the case of internetworking between DiffServ and ATM, where DiffServ packets carry priority information, EPD will fail to distinguish between packets of different priorities. The *objective of this paper is to propose a packet-based discarding scheme which will consider the priority of packets* when dropping cells.

Definition of the different service classes of DiffServ and ATM suggest a possible mapping as follows. Premium Service [14, 18] requires delay and loss guarantees, and hence can be mapped to the ATM Constant Bit Rate (CBR) service. Assured Service only requires loss guarantees and hence can be mapped to ATM Unspecified Bit Rate (UBR) service with *CLP bit set to zero*. The Best Effort service does not require any loss or delay guarantee and can be mapped to the ATM UBR service with *CLP bit set to one*. (Mapping to the UBR service is preferred over the ATM Available Bit Rate (ABR) service because it has been found that UBR using EPD can offer a performance similar to ABR, but without having the complexity of ABR flow control [19, 20]). The ATM cell header [21] has a Cell loss Priority (CLP) bit, which can be set to 0 or 1 depending on the loss priority of the cell. The CLP bit can be used to map the priorities of DiffServ packets (as specified in the Differentiated Services Code Point (DSCP) to ATM cells.

In this paper, *we propose the Prioritized EPD (PEPD) scheme which can provide the necessary service differentiation required by future QoS-enabled networks*. In the PEPD scheme, two thresholds are used to provide service differentiation between packets during network congestion. We have developed Markov chain models to study the performance of our proposed scheme in terms of its effectiveness in providing service differentiation to cells of two different priorities. The model has been used to evaluate the effectiveness of PEPD. The accuracy of the model has been validated by simulation. We measured the goodput, cell loss probability, and throughput of the service classes as a function of the load. Given some QoS requirements for the service classes, our model can be used to determine the size of the buffer required at an ATM switch and the value of the two thresholds to achieve the target QoS. The model can be used as a general framework for the analysis of networks carrying packets from applications which require differential treatment in terms of packet losses. The *contributions* of this paper can be summarized as follows:

- We have proposed *a new packet based discarding scheme* which takes into account the priority of packets.
- An *analytical model* has been developed to test the performance of our proposed scheme.
- The model can be used to determine the *buffer size* and drop *thresholds* in an ATM switch buffer which uses priority-aware packet based discarding schemes

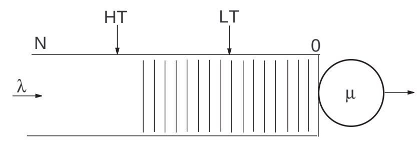

Figure 1: A buffer using the PEPD packet discard scheme.

2 PEPD scheme

As mentioned earlier, our proposed Prioritized Early Packet Discard (PEPD) scheme is a packet based discard scheme which takes into account the priority of packets when discarding cells in an ATM switch. As in EPD, it distinguishes between packets belonging to messages whose packets have not been previously accepted (called *new packets*) and packets belonging to messages whose packets have been previously accepted (called *old packets*) by the buffer. To allow prioritized discarding of cells, PEPD uses two thresholds in the switch buffer (see Fig. 1): a low threshold (LT) and a high threshold (HT), with $0 \leq LT \leq N, LT \leq HT \leq N$. Let QL indicate the instantaneous queue length. The following strategy is used to accept packets in the buffer.

- If $QL < LT$, all packets are accepted in the buffer.
- If $LT \leq QL < HT$, new low priority packets are discarded; only packets belonging to new messages with high priority or packets belonging to messages which have already entered the buffer are accepted.
- If $HT \leq QL < N$, all new packets of both priorities are discarded.
- For $QL \geq N$, all packets are lost because of buffer overflow.

2.1 Proposed Model for the PEPD Scheme

To model the PEPD scheme using Markov chains, we define two modes of the buffer: the *normal mode* in which packets are accepted, and the *discarding mode* in which

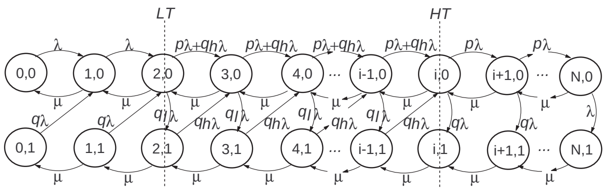

Figure 2: Steady-state transition diagram of a buffer using the PEPD scheme

arriving packets are discarded. The state transition diagram for this policy is shown in Fig. 2. In the diagram, state (i, j) indicates that the buffer has $i, 0 \leq i \leq N$ packets and is in mode $j, 0 \leq j \leq 1$. j =0 and 1 correspond to the normal and discarding modes respectively. We assume that a head-of-message packet arrives with probability q. The probability that an arriving packet is part of the same message as the previous packets is $p = 1-q$, and hence is discarded with that probability in the case that that message is being discarded.

According to PEPD, if a message starts to arrive when the buffer contains more than LT packets, the complete new message is discarded if it is of low priority, while if a new message starts to arrive when the buffer contains more than HT packets, the message is discarded regardless of its priority. Once a packet is discarded, the buffer enters the discarding mode, and discards all packets belonging to this discarded message. The system will remain in discarding mode until another head-of-message packet arrives. If this head-of-message packet arrives when $QL < LT$, it is accepted, and the system enters the normal mode. If this packet arrives when $LT \leq QL \leq HT$, then the system enters the normal mode only if this packet is of high priority. Otherwise, it stays in the discarding mode. Of course, when $QL > HT$, the buffer stays in the discarding mode.

Let's assume that h and $l = 1-h$ be the probabilities of a message being of high and low priority respectively. Also let $P_{i,j}, 0 \leq i \leq N, 0 \leq j \leq 1$ be the steady-state probability of the buffer being in state (i, j). From Fig. 2, we can get the following equations. The solution of these equations will give the steady-state probabilities of the buffer states.

$$
\begin{aligned}
\lambda P_{0,0} &= \mu P_{1,0} \\
q\lambda P_{0,1} &= \mu P_{1,1} \\
(\lambda+\mu)P_{i,0} &= \lambda P_{i-1,0} + \mu P_{i+1,0} + qh\lambda P_{i-1,1} \\
& \quad 1 \leq i \leq LT \\
(\lambda+\mu)P_{i,0} &= (\lambda p + qh\lambda)P_{i-1,0} + \mu P_{i+1,0} + \\
& \quad qh\lambda P_{i-1,1} \quad LT < i \leq HT \\
(\lambda+\mu)P_{i,0} &= p\lambda P_{i-1,0} + \mu P_{i+1,0} \\
& \quad HT < i < N \\
(\lambda+\mu)P_{N,0} &= p\lambda P_{N-1,0} \\
\mu P_{N,1} &= \lambda P_{N,0} \\
\mu P_{i,1} &= q\lambda P_{i,0} + \mu P_{i+1,1} \\
& \quad HT \leq iN \\
(qh\lambda+\mu)P_{i,1} &= q\lambda(1-h)P_{i,0} + \mu P_{i+1,1} \\
& \quad LT \leq i < HT \\
\sum_{i=0}^{N}(P_{i,0}+P_{i,1}) &= 1
\end{aligned}
$$

2.2 Performance of PEPD

Packet based discarding schemes are expected to increase the goodput of end TCP/IP applications. We therefore, measure the effectiveness of the proposed PEPD scheme by the goodput of high and low priority messages. *Goodput* ($\mathcal{G}$) is defined as the ratio between the total number of good packets exiting the buffer and the total number of packets arriving at its input. *Good packets* are those packets that belong to a complete message leaving the buffer. In this paper, we define the *goodput for high (or low) priority packets* as the ratio between total number of high (or low) priority good packets exiting the system and the total number of arriving high (or low) priority packets at the buffer. However, we normalize the goodput to the maximum possible goodput. In additions to goodput, a number of other parameters, such as queue length can be obtained from our model following the approach given in [3].

3 Numerical Results

In this section, we present results from our analytical model to illustrate the performance of PEPD. We also validate the accuracy of our analytical model by comparison with simulation results. We have carried out simulation of a buffer using the PEPD scheme. In our simulation, we set $N = 120, q = 1/6$ which corresponds to the case where the queue size is 20 times the mean message length. The incoming traffic load (ρ) at the input to the buffer was set in the range of $0.8 - 2.2$, where $\rho < 1$ represents moderate load, and $\rho \geq 1$ corresponds to higher load which results in congestion buildup at the buffer. Goodput of the combined low and high priority packets is defined as $G = h * GH + (1-h) * GL$.

Queue occupancy is a critical parameter used for calculating the goodput. We therefore, compare the queue occupancy obtained from the model and computer simulation in Fig. **??**. For $q = 1/6$, it is clear that analytical and simulation results are in close agreement. Even for

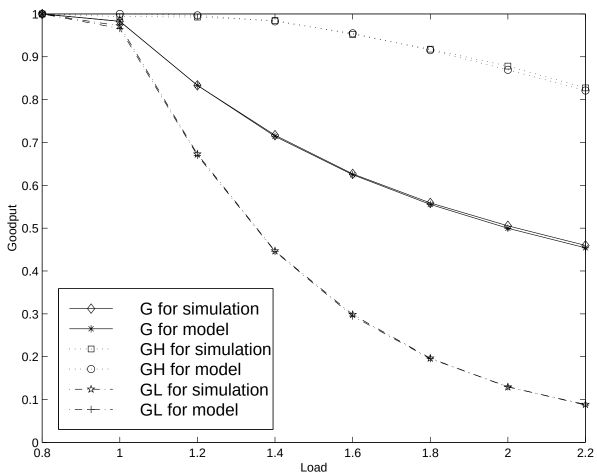

Figure 3: Goodput versus load for $h = 0.5, N = 120, LT = 60, HT = 80, q = 1/6$. G, GH and GL represents average goodput of all, high priority, and low priority packets respectively.

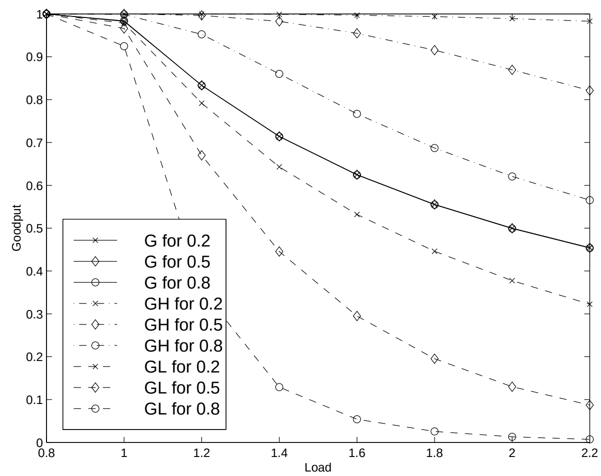

Figure 4: Goodput versus load for $h = 0.2, 0.5$ and 0.8 with $N = 120, LT = 60, HT = 80, q = 1/6$. G, GH and GL represents average goodput of all, high priority, and low priority packets respectively.

high load, our proposed scheme results in the buffer occupancy varying between LT and HT. The exact value depends on the average message length, queue thresholds, etc.

Fig. 3 shows the goodput of the buffer using PEPD for $q = 1/6$ (i.e. mean message length of six) as a function of the offered load. In this figure, the probability that a message is of high priority is 0.5. From Fig. 3, it is clear that the results from our model and computer simulation fit well. So we *conclude that our model can be used to carry out a reasonably accurate analysis of the PEPD policy*. Therefore, in the rest of this section, we will use results from only the model to analyze the performance of the PEPD policy.

Fig. 4 shows the goodput, for $q = 1/6$, as a function of the offered load and for different mix (h) of high & low priority packets. For a particular load, increasing the fraction of High Priority (HP) packets (h) results in a decrease in throughput of both high and Low Priority (LP) packets. The LP throughput decreases because an increase in h results in fewer LP packets at the input to the buffer in addition to LP packets competing with more HP packets in the buffer space (0 to LT). On the other hand, an increase in h results in more HP packets. Since the amount of buffer space (between LT and HT) which is reserved for HP packets is the same, the throughput of HP packets decrease. Note that the decrease in the throughput of LP is much faster than the decrease in throughput of HP. This results in the overall goodput being constant. Our proposed technique allows higher goodput for high priority packets which may be required by real-time applications requiring a preferential treatment over other non-real time applications.

In Fig. 5, we fix LT while varying HT to observe the behavior of the buffer as a function of HT and for various mix of low and high priority packets. It is obvious that for a traffic containing fewer high priority packets (i.e. a low value of h), increasing the HT will increase the performance of the buffer for high priority packets. This is because increasing HT will let the high priority packets get increased benefit by discarding low priority packets, especially for lower values of HT. Increasing HT will result in an initial increase in the goodput for high priority packets followed by a decrease in the goodput. This is obvious, because for a very high value of HT, the behavior of PEPD will approach that of PPD for high priority packets.

Fig. 6 shows the goodput of high priority packets versus the fraction of high priority packets. It is also clear that for a particular load, increasing the high priority traffic will decrease the performance of high priority packets as has been observed in Fig 4.

Finally, in Fig. 7, HT is kept constant while changing LT. For a load of 1.6 and a particular mix of high & low priority packets, we observe that the performance of high priority packets is not very sensitive to a change in LT. However, when LT is set close to HT, the goodput for high priority packets decrease quickly. This is because when the two thresholds are set too close, the high priority packets do not benefit enough from the discarding of low priority packets. *We suggest avoiding this mode of operation because the buffer is not fully utilized.*

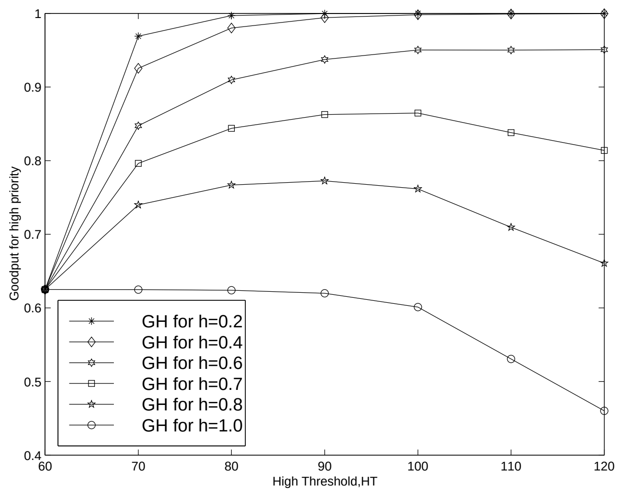

Figure 5: Goodput for high priority messages versus HT for different h with $LT = 60, q = 1/6$ and input load of 1.6

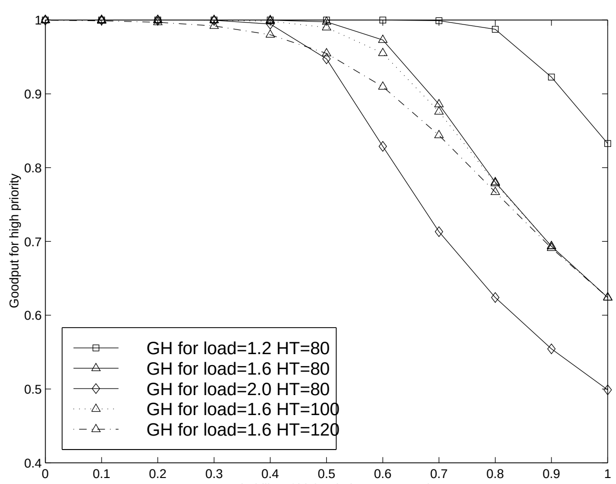

Figure 6: Goodput of high priority messages versus h for different load and HT for $LT = 60$.

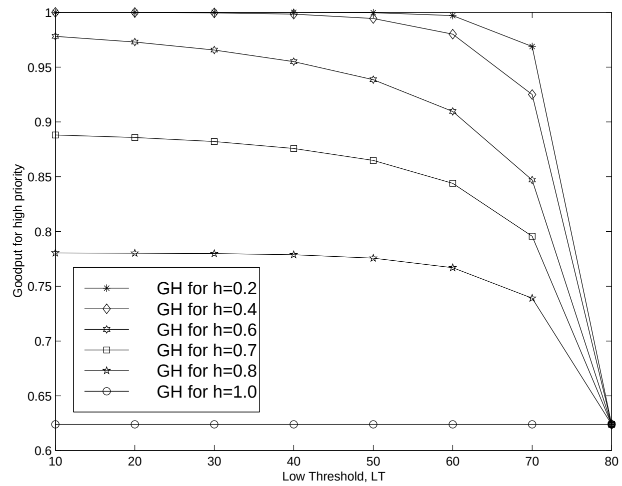

Figure 7: Goodput for high priority versus LT for $N = 120, HT = 80$, load=1.6, $q = 1/6$.

4 Conclusions

In this paper, we have proposed and developed an analytical performance model of our proposed *Priority based Early Packet Discard (PEPD)*, a packet based discard scheme to increase the goodput of end TCP/IP applications. To prove the validity of our proposed analytical model, we compared it with results from computer simulation and found that they are in close agreement. *Numerical results also show that our proposed PEPD scheme can provide differential QoS to low and high priority packets.* Such service differentiation is essential in providing QoS to applications in Differentiated service over ATM networks. Our results show that the performance of PEPD depends on the mix of high & low priority traffic, values of the thresholds, and average message length. Given a certain QoS, the model can be used to *dimension* the size of the buffer and the PEPD thresholds. Our model can serve as a *framework* to study and implement packet based discarding schemes to provide service differentiation in the next generation Internet.

References

[1] A. Romanow and S. Floyd, "Dynamics of TCP traffic over ATM network," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 4, pp. 633–641, May 1995.

[2] K. Kawahara, K. Kitajima, T. Takine, and Y. Oie, "Packet loss performance of the selective cell discard schemes in ATM switches," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 5, pp. 903–913, June 1997.

[3] M. Casoni and J. S. Turner, "On the performance of early packet discard," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 5, pp. 892–902, June 1997.

[4] Y. Lapid, R. Rom, and M. Sidi, "Analysis of discarding policies in high-speed networks," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 764–772, June 1998.

[5] O. Elloumi and H. Afifi, "Improving RED algorithm performance in ATM networks," *GLOBECOM 97: IEEE Global Telecommunications Conference*, Kobe, Japan, pp. 1062–1066, April 7-11, 1997.

[6] N. Tsukutani, K. Kawahara, T. Takine, H. Sunahara, and Y. Oie, "Througput analysis of selective cell discard schemes in transport layer over ATM networks," *GLOBECOM 97: IEEE Global Telecommunications Conference*, Phoenix, USA, pp. 1067–1074, Nov. 3-7, 1997.

[7] W. K. Lai and C. C. Liu, "SWFA: A new buffer management mechanism for TCP over ATM-GFR," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 356–358, March 2003.

[8] L. Kalampoukas, A. Varma, and K.K. Ramakrishnan, "Explicit window adaptation: A method to enhance TCP performance," *IEEE/ACM Transactions on Networking*, vol. 10, no. 3, pp. 338–350, Jun 2002.

[9] S. Blake, D. Black, M. Carlson, Z. Wang, and W. Weiss, "An architecture for differentiated services." RFC 2474, Internet Engineering Task Force, December 1998.

[10] X. Xiao and L. M. Ni, "Internet QoS: A big picture," *IEEE Network*, vol. 13, no. 2, pp. 8–18, April 1999.

[11] K. Nichols, S. Blake, F. Baker, and D.Black, "Definition of the differentiated services field (DS Field) in the IPv4 and IPv6 headers." RFC 2474, Internet Engineering Task Force, December 1998.

[12] Wei Liu, Xiang Chen, Yuguang Fang, and J.M. Shea, "Courtesy piggybacking: Supporting differentiated services in multihop mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 380–393, Oct-Dec 2004.

[13] R.R.-F. Liao and A.T. Campbell, "Dynamic core provisioning for quantitative differentiated services," *IEEE/ACM Transactions on Networking*, vol. 12, no. 3, pp. 429–442, Jun 2004.

[14] V. Jacobson, K. Nichols, and K. Poduri, "An expedited forwording PHB." RFC 2598, Internet Engineering Task Force, June 1999.

[15] J. Heinanen, F. Blaker, W. Weiss, and J. Wroclawski, "Assured forwording PHB group." RFC 2597, Internet Engineering Task Force, June 1999.

[16] G. Karlsson, H. Lundqvist, and I.M. Ivars, "Single-service quality differentiation," *Twelfth IEEE International Workshop on Quality of Service*, pp. 265–272, 7-9 June 2004.

[17] Jian Zhou, P. Martin, and H. Hassanein, "QoS differentiation in switching-based web caching," *2004 IEEE International Conference on Performance, Computing, and Communications*, pp. 453–460, April 2004.

[18] M. Ma, Y. Zhu, and T.H. Cheng, "A bandwidth guaranteed polling MAC protocol for ethernet passive optical networks," *INFOCOM*, pp. 22–31, Mar 30 - April 3, 2003.

[19] T. Ott and N. Aggarwal, "TCP over ATM: ABR or UBR," *Proceedings of the ACM SIGMETRICS '97 Conference*, Seattle, WA, June 1997.

[20] L. Mendes and P.R. Guardieiro, "Proposal and evaluation of ERICA switch algorithm as explicit rate mechanism in GFR backbones with ABR flow control," *5th IEEE International Conference on High Speed Networks and Multimedia Communications*, pp. 152–156, July 3-5, 2002.

[21] ATM Forum, "Traffic management specification version 4.1." AF-TM-0121.000, March 1999.