Performance of TCP over ATM for Various ABR Control Policies

Carlos M. D. Pazos†‡
Dirceu Cavendish Jr†‡
† UCLA – Computer Science Department
405 Hilgard Ave., Los Angeles, CA 90024

Vincenzo A. Signore†
Mario Gerla‡
‡ Politecnico di Bari – DEE
Via Orabona 4, 70125, Bari Italy

Abstract

Connectionless traffic originating from LANs is carried as Available Bit Rate (ABR) traffic in the ATM network and is subjected to ATM level flow control, either rate-based or credit-based. In addition, transport layer flow control (e.g. TCP window control) is applied at the end points. In this paper, we study the interaction between TCP and ATM flow controls. We use simulation to compare the “goodput” of different ABR flow control schemes and to evaluate the sensitivity to different parameter selections (e.g., TCP Maximum Segment Size (MSS), round trip delay, etc). The study uncovers an unfairness situation caused by the use of non uniform MSS values.

1 Introduction

ATM networks must be able to support connectionless traffic. This turns out to be a challenge, since ATM requires that sources provide traffic descriptors at call set up time for bandwidth negotiation and assignment. However, data sources cannot predict their traffic behavior. The ATM network will thus carry connectionless traffic on an Available Bit Rate (ABR) basis, with minimum or no bandwidth assignment or QoS guarantee. Flow control must also be exercised to prevent ABR traffic from congesting the ATM network.

In addition to the ATM flow control mechanisms, connectionless traffic is also controlled by Transport layer flow control mechanisms, for example, the TCP dynamic window scheme. Hence, in this paper we investigate the interaction between the TCP flow control and various ABR controls implemented at the ATM layer, namely, FCVC, PRCA, EPRCA and SP-EPRCA. A more extended discussion of the results presented here can be found in [1].

2 ABR Flow Control Schemes

Since no UPC (User Parameter Control) can be applied to ABR traffic, flow control has to be implemented through feedback from the network congestion points to the traffic sources, allowing the ABR traffic to use the bandwidth left over by CBR and VBR traffic. Two ways of doing it have received considerable attention at the ATM Forum. In one of them, buffers are reserved at intermediate nodes along a VC and credit information is exchanged between consecutive intermediate nodes, from the destination all the way back to the source. One such scheme is the FCVC [2]. In the other approach, the source transmission rate is controlled based on congestion feedback information from the network. Examples of rate-based solutions are PRCA, EPRCA [3] and SP-EPRCA [4].

3 The Simulation Model

We have developed a simulation model to evaluate the system shown in Figure 1. The congested link is shared by ABR and higher priority VBR traffic.

We consider three different sets of experiments. First we compare the performance of different congestion control schemes and evaluate the effect of the Maximum Segment Size (MSS), with all stations restricted to use the same MSS. We study the case of a campus ATM with only one switch and contrast it with the case of an ATM MAN when we have seven switches. In both cases, each link is considered to be 7Km long. The link speeds are always 150Mpbs.

In the second experiment, we study the effect of having the ATM network carry traffic from different types of LANs which put different restrictions on the maximum PDU size (MSS) that can be used by TCP entities. So, we divide the stations into three classes of sources as shown in Figure 1. Each class is restricted to a different MSS.

In the third experiment, we study the effects of having longer links, which translates into larger buffers.
Namely, we assume a cross-country distance separating the sources from the destination. We have the same configuration of seven switches interconnected by 600 Km links.

Our TCP implementation is based on the TCP description found in [5] with the slow-start feature. The VBR generator uses an MMDP model that describes the aggregated traffic of multiple VBR sources. The VBR traffic uses most of the bandwidth, leaving just a little unused bandwidth to be shared among the ABR (TCP) sources.

All the ABR traffic is generated by a set of 15 workstations, connected possibly to different LAN/MANs. Each workstation is running an application that is generating traffic simulated by an ON-OFF model. The aggregate TCP traffic is such that it requires more bandwidth than what is left over by the VBR traffic, thus saturating the trunk.

For the ATM switch, we have implemented priority queueing for VBR and ABR cells, with ABR queue serviced only when no cells are available on the VBR queue. Furthermore, for the FCVC scheme, the ABR queue is subdivided into different VC queues. The buffer allocation for each queue is derived as per [2]. For the short links the VC buffer accommodates 11 cells, while in the longer links (as in the cross-country case), the buffers can hold 81 cells. This is also the buffer allocation used for the SP scheme. Since in our study we have 15 ABR sources, in order to carry out a fair comparison we dimensioned the total PRCA and EPRCA buffers to be 155 cells for the short links and 1,215 cells for the longer links.

4 Results

Simulation run duration was 8.5 s (simulated time) for the one switch case, and 11 s for the seven switch case. These simulated times were long enough to produce 95% confidence intervals, not shown here to make the figures more readable.

4.1 The TCP Performance

In Figures 2 and 3 we present the TCP goodput as a function of the MSS (for MSS values: 256, 1024, 2048, 8192) when the TCP traffic is sent over connections crossing one and seven switches, respectively. Here we investigate the effect of different round trip delays (RTDs) (without modifying buffer requirements) by comparing the results in both figures. We also investigate the performance of the different ABR controls when the ABR traffic is generated by homogeneous (same MSS) TCP sources. Hence, we evaluate the performance (goodput) as a function of TCP MSS and RTD.

Figures 2 and 3 show the effect of the MSS on performance. Since the amount of protocol overhead is the same regardless of MSS, the larger the MSS the better the efficiency in the transport of data bytes. However, larger MSS means that more data has to be retransmitted in case of cell loss. For FCVC and SP, no cells are lost and a larger MSS only improves performance. For PRCA and EPRCA, the loss of cells and consequent retransmission of segments defies the overhead efficiency of large MSS values.

Comparing the results for FCVC on Figures 2 and 3, we see that this scheme is unaffected by the difference in RTD. This is because FCVC employs backpressure and buffer reservations which prevent cell losses that would trigger TCP retransmissions. Furthermore, the goodput experienced with FCVC is very close to the theoretical bound discussed in [6].
uses for the $MDF$ (smaller MDF) would unnecessarily slow down the sources even further. An $MDF$ value of 6 was also tested but with no performance gains. The reduction in the level of goodput due to the increased distance results from the out of date congestion feedback information, leading to higher chances of cell losses and TCP retransmissions. The larger the MSS, the worst the goodput since more data must be retransmitted. Furthermore, the TCP slow start forces sources affected by losses to wait for an acknowledgement before transmitting another segment. Since there is an increase in the RTD with seven switches, the sources waiting for the acknowledgement cannot use their share of bandwidth which accounts for the reduced aggregated goodput.

For the EPRCA scheme we simulated the cases with $MDF = 5$ and $MDF = 6$. Since the EPRCA uses the Explicit Rate ($ER$) to limit the increase in the $ACR$, $MDF = 5$ proved to be too conservative as this value implies a fast decrease of the $ACR$ even when the network agreed to support the $ER$. Hence, the sources do not transmit at the $ER$ for very long. With the $MDF = 6$, the sources reduce their $ACR$ slower and can use more of the available bandwidth. With longer distances, we notice the same effect of goodput reduction due to the out dated congestion notifications already observed with PRCA.

For the SP scheme, there is no parameter tuning involved and only one curve was plotted. Since this scheme implements a rate based control which guarantees no loss, no bandwidth is wasted with retransmitted segments. However, the scheme relies on feedback from the network, thus it depends on RTD. For these reasons, the SP performs better than the EPRCA or PRCA schemes in the topology with only one switch. As the distance is increased, the loss free property is offset by the dependence on RTD.

To sum up, for the two network configurations considered in this section, we notice that FCVC always outperforms PRCA, EPRCA and SP. Despite its implementation complexity, FCVC does provide an upper bound on performance. EPRCA can provide a better performance than PRCA. However, the performance improvement is rather small.

### 4.2 Non Uniform Sources

In the case study of non uniform TCP sources, the aggregate goodput experienced by the sources in each class is plotted in Figure 4, with the MSS values in each class been 512, 1024 and 2048. As it can be seen, the FCVC, EPRCA and SP schemes allow the different classes to experience an almost identical goodput. Indeed, the small difference is due to the efficiency of the protocol stack, which varies slightly with the MSS.

Unfairness, on the other hand, manifests itself with PRCA, as can be observed. Furthermore, the unfairness varies with distance, see Figure 5. Two factors are responsible for the unfairness, and have an impact which depends on the RTD. First of all, the larger the MSS the more resource is wasted on retransmissions. Secondly, when congestion occurs in the network, cells are dropped and the TCP flow control moves into a slow start phase. During this phase, with a larger MSS more RM cells are sent to probe the network. Consequently, the corresponding sources can reclaim bandwidth faster once congestion subsides.

With long RTDs, more losses occur and the first factor prevails, penalizing the sources using larger MSS. On the other hand, with short RTDs, the second factor prevails because longer congestion recovering periods are more likely.

### 4.3 Cross-country Distances

In the cross-country case study, larger buffers are required which gives the switch more capability to handle less conservative ABR controls, as observed in Figure 6. For the EPRCA case, the results for
$MDF = 6$ is very poor compared to the results in Figure 3 in which $MDF = 6$ was the best we could get with this scheme. Because we have a very large RTD, the $ER$ advertised back to the sources is much out dated. However, the large buffers allow for the accommodation of the rate mismatch between the sending rate and the actual available bandwidth. This is so much the case that even for larger $MDF$ values, say 7 and 8, we still get an improvement in performance over the $MDF = 6$ case.

For PRCA the case is exactly the same. When an RM cell is received with no congestion indication, the $ACR$ can be set to the PCR regardless of whether that rate can be supported by the network. However, PRCA with an $MDF = 5$ reduces the rates very quickly in between arrival of RM cells. So, the PRCA aggressiveness, coupled with large buffers to accommodate the rate mismatches, is responsible for the PRCA better performance in this scenario. We also notice that increasing the aggressiveness of PRCA by using an $MDF = 6$ did not prove fruitful.

![Figure 6: TCP goodput for a cross-country topology.](image)

The SP scheme is heavily penalized by the large RTD value involved in this experiment. Since SP does not provide a minimum cell transmission rate, the sources are forced to stop transmissions while they wait for network feedback information.

The results for the FCVC scheme have presented a reduction in performance for such large RTD. This happens due to the TCP/ATM interaction. Since sources always start transmitting in the slow start phase, the window takes time to expand. At the ATM level, the sources are allowed to transmit their segments at full link speeds given they have ATM credits for that. However, the very large RTD requires the TCP entity to wait for an entire RTD before submitting a new set of segments, defying the FCVC ability to send more data.

5 Conclusions

In this paper we have compared the performance of TCP traffic carried as ABR traffic over ATM networks, with several different ABR flow control schemes (FCVC, PRCA, EPRCA and SP). We have also investigated the effects of TCP maximum segment size, RTD, and buffering at the ATM switches on the TCP/ATM interaction. Unquestionably, the FCVC scheme provides for the best performance for all configurations considered. However, even if stations are backlogged, the FCVC potential of cannot be fully utilized if the distances involved are very long. Thus, the cost effectiveness of very robust and complex schemes can very poor.

The ATM Forum has already ruled that rate-based control should be used for ABR traffic flow control [7]. However there is still plenty of room for improvement. Suggesting that the switches should only provide cell marking for congestion indication, allows switch manufacturers not to implement the explicit rate computation for all VCs traversing the switches. In such case, the sources cannot implement the EPRCA scheme, leaving PRCA as the only alternative. This is not a good idea since in this paper we showed that PRCA can be very unfair in a realistic non homogeneous TCP environment.

Acknowledgments

We would like to thank Abra Davini for her contributions to the simulation study presented here.

References


