14

Quantitative study of uncontrolled transients on ABR congestion control

Eduardo S. C. Takahashi[1]
Department of ECE, Carnegie Mellon University
Pittsburgh, PA 15213, USA, takahasi@cs.cmu.edu

Abstract

The ABR service category of ATM architecture is a best effort service which aims to provide rapid access to unused network bandwidth at up to the link rate whenever the network bandwidth is available. As a result, the transfer characteristics provided by the network may change subsequent to connection establishment, and a congestion control mechanism is required. The ATM Forum has specified a congestion control infrastructure from where congestion control mechanisms like EFCI, EPRCA, ERICA and others have been proposed. The analysis of the underlaying infrastructure shows that, for the present switch and end-system behavior definitions, situations can arise when sources are allowed to transmit without control by the implemented congestion control mechanism. This study presents a quantitative analysis of the expected effects.

Keywords

Asynchronous transfer mode; available bit rate service; buffer storage; telecommunication congestion control; cell loss probability; bursty traffic; transient analysis; closed loop systems; ATM; ABR; ATM Forum; EFCI; EPRCA; ERICA

1 INTRODUCTION

ATM technology is still under development with standards being specified and updated. As part of this process, the ATM Forum has defined guidelines for several service categories on its ATM architecture (ATM Forum, 1996). While CBR and VBR aim to provide specific quality of service (QoS), ABR and UBR work as best effort services, without specific guarantees. The difference between ABR and UBR is that in ABR specification a loose concept of QoS still remains, ABR should provide "low" cell loss ratio target while UBR is purely best effort. Although there is no specific definition of "low" cell loss ratio in the ATM Forum specifications, the adoption of that as a target has deep implications, separating ABR from UBR. ABR implementations should provide more reliable connections than UBR. The cost is a more complex algorithm, employing congestion control. An unexpected side effect though is the vulnerability of ABR to uncontrolled transients, what can compromise the "low" cell loss ratio objective.

---

[1] Research at CMU was supported through a scholarship from CNPq - Brasilia/Brazil.

Management of Multimedia Networks and Services R. Boutaba & A. Hafid (Eds.)

The coexistence of guaranteed and non-guaranteed services sharing resources in a ATM network is a major advance that allows a compromise between efficient usage of the resources and user satisfaction (Clark *et al.*, 1992). While some applications have strict requirements, like bandwidth and delay, imposing the necessity of guaranteed services, other applications are not so demanding, fitting well on best-effort services that use whatever fraction of the resources is still available at that moment. Those are the extremes of the spectrum, with a variety of applications fitting in the between.

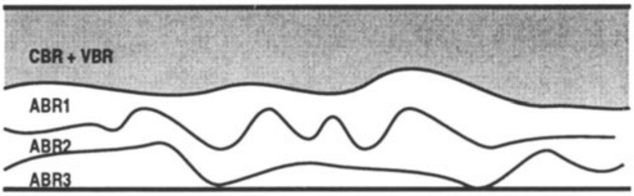


Figure 1. Link sharing.

Figure 1 presents a typical scenario where the aggregate bandwidth used by some number of guaranteed services (CBR and VBR) fluctuates with time. VBR connections fluctuate by definition, while the aggregate CBR connections may fluctuate when sessions are established and closed. The remaining bandwidth is divided between the best-effort connections, in this example three ABR connections. To maximize resource usage, the best-effort connections compete for that available bandwidth, using as much as they can. In order to accommodate changes in the fraction of the bandwidth used by guaranteed services and still prevent high cell loss ratio, ABR implements congestion control mechanisms.

The congestion control mechanism adopted for ABR receives feedback information from the intermediate switching elements and destination in order to determine the available capacity. A closed loop mechanism is responsible to keep track of the changing conditions and maximize resource usage while minimizing losses. The problem of simultaneously achieving both those goals is not trivial and another constraint that must be considered is implementation complexity. Several mechanisms were proposed generating intense debate (Chen *et al.*, 1996, Kung and Morris, 1995, Siu and Tzeng, 1994). Under all those considerations, the ATM Forum specified an ABR congestion control mechanism with a control loop associated with each information flow. Resource Management (RM) cells are used in the control loop, being inserted in the information flow and carrying information from source to switching elements and destination and then bringing feedback to the source. Instead of traffic descriptors, ABR employs procedural definitions that specify source, destination, and switch behaviors in response to the data carried on RM cells. More detailed description of the ATM Forum ABR congestion control specification can be found in ATM Forum (1996) and Jain *et al.*(1996c).

While end-point behaviors are fully determined, in the ATM Forum specification the switch behavior is an infrastructure with minimum requirements for operation and compatibility, allowing a great degree of freedom to implementations. The basic functions performed by switches on the control loop are congestion evaluation and feedback calculation. For congestion evaluation, developers are free to use whatever information they have available at the switch, and two approaches currently being implemented are queue threshold based and incoming rate based. Feedback calculation requires standard interfaces with the fixed end-point behaviors, so two formats were specified, not mutually exclusive, binary and explicit rate feedback. A possible classification of switch behaviors for ABR congestion control mechanisms would include those two dimensions. Table 1 summarizes the congestion control mechanisms covered by this paper, showing that they are representative of the possible combinations.

TABLE 1. Classification of ABR mechanisms.

	Binary Feedback	*Explicit Rate Feedback*
Threshold Based	EFCI	EPRCA
Rate Based		ERICA

Some work has been done in the direction of describing ABR congestion control behavior quantitatively (Kolarov and Ramamurthy, 1994, Walthall and Clement, 1996, Siu and Tzeng, 1994, Johansson and Karlsson, 1995). In those references simulations and analytical models were employed to characterize normal operation conditions of some of the switch algorithms. This paper describes the quantitative study of one particular aspect of ABR congestion control not foreseen in the previous references, the uncontrolled transients resulting from ACR accumulation. Section 2 introduces a qualitative view of the potential vulnerability of ABR congestion control. Section 3 presents the quantitative analysis and comparisons between ABR mechanisms. Section 4 checks the analytical study with simulation results. Finally, Section 5 reviews the results and conclusions of this work.

2 CRITICAL CASES: A QUALITATIVE VIEW

Cell loss happens whenever the input rate to a specific output port exceeds its maximum output rate for a period long enough to overflow the output buffer. Under normal operation (Jain *et al.*, 1996c), once the congestion control mechanism detects congestion (or the imminence of it) the feedback information indicates to the sources the necessity to slow down their transmission rates, preventing or reducing losses. When the congestion evaluation mechanism determines the end of congestion state, sources are allowed to increase their transmission rates, if they can, in order to use any unused bandwidth increasing efficiency of resource utilization. Of course, different algorithms perform differently, as well as different scenarios would produce different results. In any case, the source behavior specifies the maximum transmission rate by ACR based in its present state and the feedback received.

One important distinction that has to be made is between ACR and the actual transmission rate of a source. ACR is a state variable and holds the maximum rate that a source can transmit at that moment, according to the congestion control mechanism. The actual transmission rate of a source can be any value between zero and ACR, and a specific value cannot be imposed by the network since it is up to the end user to decide transmission. Also the end system itself may have limitations, permanent or temporary, that prevent it from transmitting at ACR.

The ATM Forum traffic management specification (ATM Forum, 1996) address briefly the problem of disparity between ACR and actual transmission rate in the context of ACR *retention* (Jain *et al.*, 1996c, Jain *et al.*, 1995, Jain *et al.*, 1996a). ACR retention occurs when a RM cell going back to the source encounters only uncongested nodes, allowing the source to increase ACR, as shown in Figure 2. If, by any reason, that increase is not reflected on an increase of actual transmission rate, the difference may account as a credit retained by the source. Until a

RM cell indicating congestion arrives, the source with credit may use it at any time creating sudden bursts. The purpose of this work is twofold: first extend the concept of ACR retention to a more general concept of ACR *accumulation*, and then analyze quantitatively the effects of it.

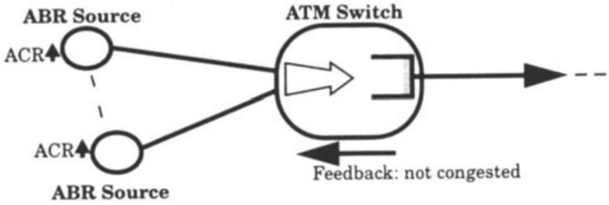


Figure 2. ACR retention.

ACR accumulation is the process of increasing ACR values with time. As in ACR retention, RM cells going back to their sources may encounter only uncongested nodes, signaling the sources to increase their ACR, as shown in Figure 3. The rate at which ACR is increased depends on several factors including negotiated parameters and switch algorithm used. The final ACR value may exceed by far the actual available rate at the bottleneck node of a connection, creating opportunity for sudden rate jumps at the sources what may cause severe congestion. Basically, after ACR accumulation the source is free from control by the congestion control mechanism, being able to generate *uncontrolled transients*. Therefore, the time required for ACR to differ considerably from the actual transmission rate is an important variable that must be considered.

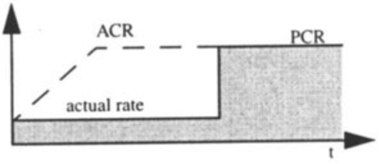


Figure 3. ACR accumulation and sudden rate jump.

We notice, however, that the degree of danger created by uncontrolled transients depends on the ability of the uncontrolled sources to generate sudden bursts of intense traffic. The objective of this work is to translate that qualitative statement into a quantitative one, being able to understand if ACR accumulation and uncontrolled transients are really a potential problem for sources transmitting with certain characteristics and switches operating with certain mechanism and with a determined buffer size.

We adopted in our work connections involving OC-3 links (155.52 Mbit/s) but the results can be scaled up and down. Looking through the universe of possible users and applications that may request networking services one can see that most of the applications that are bandwidth intensive concentrate around multimedia applications (Little and Ghafoor, 1990). Although the delay requirements of live multimedia applications in general make CBR and VBR better options than ABR, the last is not necessarily ruled out. Also, some data applications are able to generate high intensity traffic. Another point to consider is the capacity of endpoints to drive high intensity traffic, requiring powerful hardware. All the above suggests that, at least in the beginning, the majority of traffic sources would be constituted of low intensity traffic sources, with few high intensity traffic connections present, as shown in Figure 4. The low intensity sources are unable to generate major traffic transients, and the statistical independence allows us to approximate the aggregate of those sources as a constant background traffic. A not so good approximation, but very helpful, the major traffic sources can be added and considered as one high intensity bursty source.

Figure 4. ABR traffic distribution.

3 QUANTITATIVE STUDY

3.1 Analytical Models

Before going inside the quantitative analysis, targets must be specified. Since the major concern about congestion on best-effort services is cell loss, the measurement of queue sizes is fundamental. In a closed loop control the response is delayed by the feedback time and the worst case scenario would involve uncontrolled transients with high intensity bursts of long duration. Relevant parameters to be analyzed are the ramp-up time required for a source to accumulate high ACR and the response of the system to bursts (step response). The real effect of uncontrolled transients can be derived from the step response as the actual burst characteristics are imposed on the step function. The duration and intensity of the burst can be used to determine the cell loss ratio for a given buffer size. Also, the long term impact of uncontrolled transients is highly dependent on the repetition of bursts in the traffic pattern. If intense bursts are frequent a sustained high cell loss ratio is possible, as shown in Figure 5.

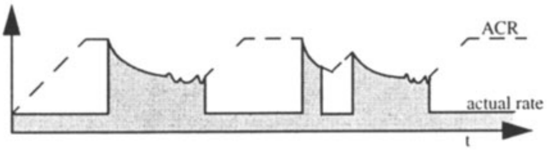


Figure 5. Bursty source.

The procedural description specified by the ATM Forum together with various switch behavior implementations make analytical modeling a hard task. One useful tool employed in this work was the fluid flow approximation. A similar approach was adopted in Johansson and Karlsson (1995) and Siu and Tzeng (1994). Assuming the flow of data as a function of continuous time a simple expression can be used to determine queue length as function of time:

$$Q(t) = \int^{t} (ACR(t) + BGR(t) - PCR)dt \qquad \text{(EQ 1)}$$

where ACR(t) is the allowed cell rate of a source, BGR(t) is the aggregate background traffic, and PCR is the peak cell rate. Assuming that the buffer is large enough to avoid losses, the maximum queue size condition will be when the incoming rate matches the output capacity:

$$ACR(t) + BGR(t) = PCR \tag{EQ 2}$$

Instead of trying to describe this complex system analytically as a whole, a close look inside the components shows inherently simple behavior. The ABR source in particular has the property that ACR(t) is not continuous, but all transitions are steps because the source state changes only on the reception of RM cells. Therefore, the integral in Equation 1 can be easily calculated if the interval between RM cells and their effects are known.

The way to calculate those parameters depends on the specific switch behavior implemented. Figure 6 shows the basic idea employed with EFCI switches. Under congestion feedback, the source will reduce its ACR by a factor of (1 - RDF) for each backward RM cell received. Although the time between RM cells is variable, for short intervals it is a good approximation to consider it constant.

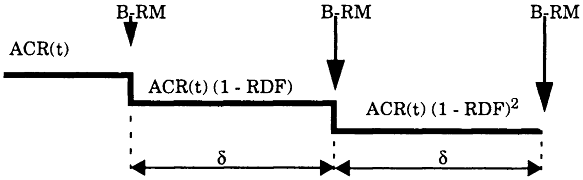


Figure 6. EFCI model.

For EPRCA the analysis is more complex, the behavior of EPRCA switches is quite different for short and long distances. Figure 7 presents the model for LAN (short distances), when the propagation delay between source and switch is too small when compared to the period between RM cells. In the presence of congestion cells start to get backlogged in the switch, what makes the output rate at the switch constant at its link rate. Therefore, the rate at which backward RM cells (B-RM) are received at the source becomes approximately constant. The rate at which the source send forward RM (F-RM) cells however is tied with its ACR and thus is variable and higher than the rate of B-RM cells. Since the source is responding to the congestion indicated by the switch, the number N of F-RM cells sent between the arrival of two B-RM cells is monotonically decreasing. Starting from any initial condition, we are able to evaluate the compound effects of F-RM cells at the switch and B-RM cells at the source.

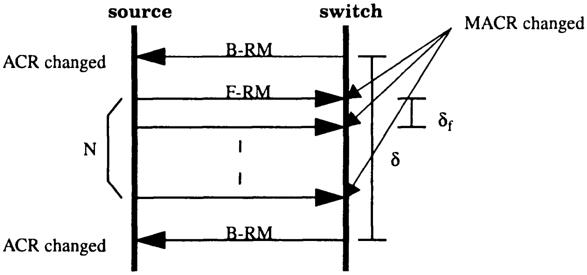


Figure 7. EPRCA LAN model.

When long distances are involved and the propagation delay between source and switch is considerably larger than the interval between F-RM cells, EPRCA behavior can be better described by Figure 8. On the onset of congestion the switch sends a B-RM cell that reduces ACR at the source just like in the LAN model. The difference is that the following B-RM cells sent will carry the same feedback information based on the MACR value hold by the switch, and therefore will have no effect. Only after the next F-RM cell sent by the source reaches the switch MACR will be modified, so state changes on both sides will approximate a staircase.

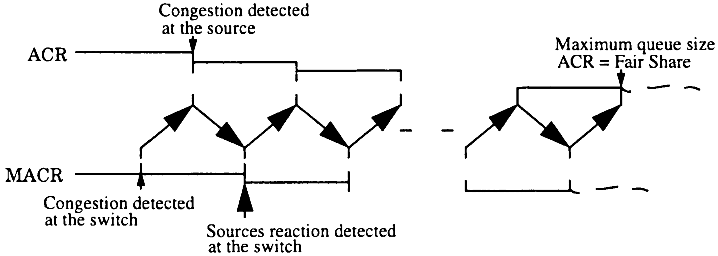


Figure 8. EPRCA WAN model.

ERICA works in a very different way (Jain *et al.*, 1996b). ERICA switches are responsible for keeping track of the input rate so that they can calculate the load factor z as an average value for certain switch interval *SWI*. After the source jumps from an initially low transmission rate to the maximum, it may take one SWI until z accurately indicates congestion plus twice the propagation delay until the rate reduction reaches the switch. If congestion is still present, new cycles may be needed.

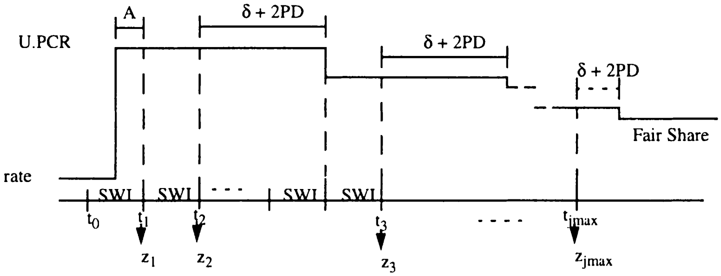


Figure 9. ERICA model.

Considering now the ACR ramp-up time, EFCI and EPRCA behaviors are identical since under uncongested conditions the switches do not change the RM cells going back to the source. Starting from an equilibrium initial condition with sources using their fair shares of the bandwidth and oscillating lightly between congested and uncongested states, if from a given instant on one source reduces its rate and the others do not use the freed resource, all sources will start accumulating ACR. As Figure 10 shows, for each B-RM cell received at the source ACR will be increased by RIF times PCR.

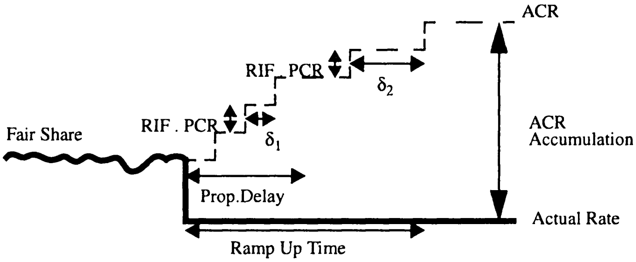


Figure 10. EFCI and EPRCA ramp up time.

ERICA again follows a more complex behavior since even when switches are uncongested they monitor B-RM cells and may change the explicit rate feedback (ER). Starting from a stable condition with z = 1, if one source reduces its transmission rate and the others do not use the freed bandwidth after around a propagation delay time the switch will recognize $z < 1$. ER sent by the switch will allow the source to increase its ACR until the maximum value ACR/z. At the source point of view, that change in ACR will take twice the switch-source propagation delay to result in change of ER, resulting in a staircase increase of ACR. Note that ACR accumulation is possible only because sources are not increasing their actual transmission rates, making 1/z constant.

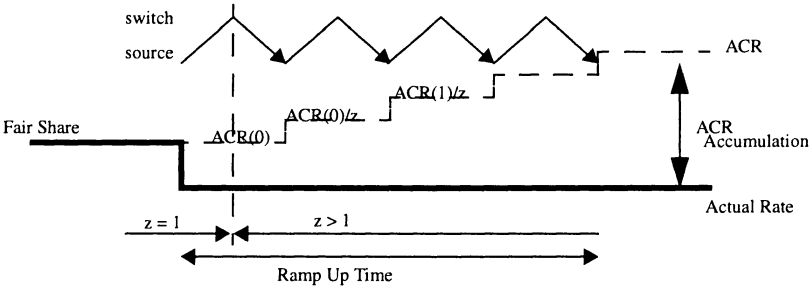
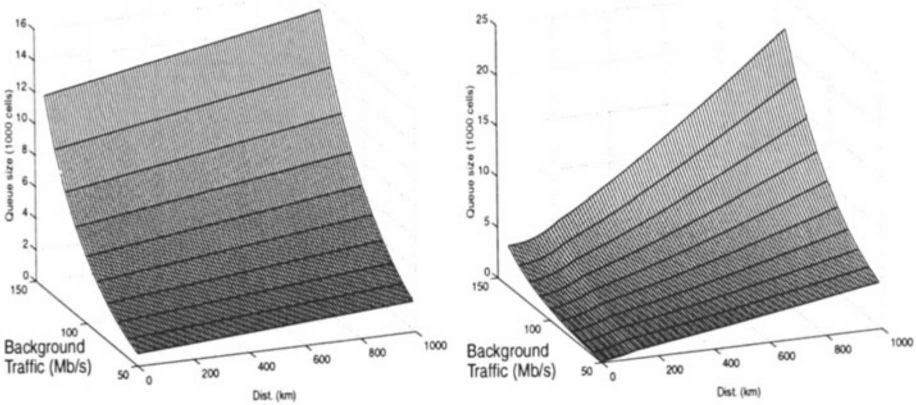


Figure 11. ERICA ramp up time.

3.2 Results

Using the models developed, expressions were derived allowing a quantitative analysis of the different switch implementations. Although the expressions were functions of several variables, to make visualization easier and staying close to possible scenarios some variables were fixed in the following examples. We consider first the maximum queue size required to achieve zero cell loss as a function of the background traffic present and the delay between source and bottleneck switch. Knowing that maximum queue size, the expected cell loss can be calculated for a given buffer size, and based on burst probabilities the expected sustained cell loss ratio can be obtained.

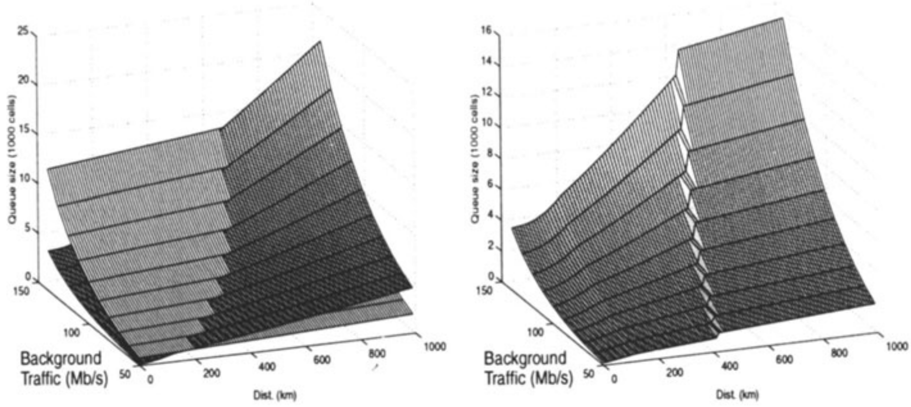


a. EFCI

b. EPRCA

Figure 12. EFCI and EPRCA.

Figure 12 shows the behaviors of EFCI and EPRCA mechanisms. Although it may be unfair to compare different algorithms based only on this results because different parameter sets yield different results, some characteristics are intrinsic to each algorithm. EFCI usually employs conservative bandwidth allocation parameters, what allows very large queues in the presence of uncontrolled transients. As the delay in the control loop increases, however, the increase in queue size is slow. EPRCA starts with smaller queue sizes due to its fast bandwidth deallocation, but as the delay increases it suffers a fast decrease in responsiveness, resulting in fast queue size growth. At some point, EFCI and EPRCA behaviors intercept, making EPRCA queues longer than EFCI, as shown in Figure 13. Since EFCI and EPRCA implementations are very close, it seems reasonable that all EPRCA implementations also implement EFCI without incurring in increased complexity, producing a more stable queue size.



a. EFCI x EPRCA

b. EFCI + EPRCA

Figure 13. EFCI x EPRCA.

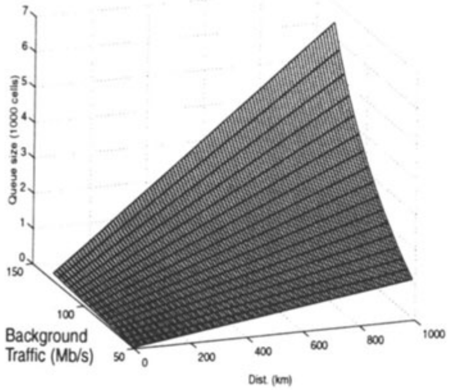
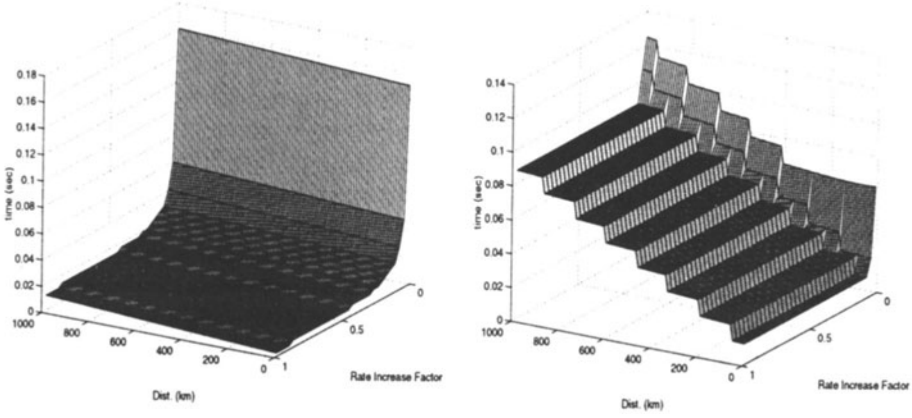


Figure 14. ERICA.

The ERICA model presents better results due to its very fast bandwidth deallocation. As Figure 14 shows, the queue size starts at least at an order of magnitude smaller than the previous mechanisms. As the distance increases, this difference decreases to a factor of two, but another parameter must be considered, the ACR ramp-up time.

Figure 15 shows the ACR ramp-up times as a function of source-switch distance and RIF. Although EFCI and EPRCA share the same model, their actual behaviors will probably differ in the sense that while EFCI uses conservative RIF (much smaller that one) EPRCA uses aggressive RIF (close or equal to one). Thus, for EPRCA the ramp-up time will be mostly dependent on the propagation delay of the first B-RM cell that carries uncongested information. ERICA is also likely to employ aggressive RIF (equal to one) but since it controls also the increase of ACR the ramp-up time will have a very different characteristic. As the delay to get an updated maximum value for ACR increases, the ramp-up time will also increase, with some other complicating effects due to the correlation of intervals between RM cells.



a. EFCI and EPRCA

b. ERICA

Figure 15. Ramp up time.

One effect of ramp-up time is that while for EPRCA the uncontrolled transient can happen after a small window of opportunity, with ERICA that window must be much larger in order to allow considerable ACR accumulation. In that way, ERICA is less likely to produce long queues than EPRCA since it has less chances of waiting long enough to accumulate ACR, what makes the graph shown in Figure 14 to be considered very conservative.

4 SIMULATION RESULTS

Simulation studies performed with a modified version of the NIST ATM network simulator (Golmie *et al.*, 1995) confirmed the previous analytical results with errors around 10%. An application example is shown in Figure 16. That example involves four independent simulation runs and the cumulative percentile cell loss ratio observed in the long term. There is a bursty ABR connection and background traffic sharing a bottleneck OC-3 link 400 km away from the ABR source. The switch output buffer is set to 1000 cells. The interval between bursts is either a constant or exponentially distributed with same mean as in the constant case.

With exponentially distributed interarrival cell loss rates are higher because bursts can arrive before the queue is reduced considerably, while with constant interarrival the flush interval is guaranteed. Notice also that EPRCA produces consistently higher cell loss ratios. The cell loss rates expected from the previous analysis are presented on the right of Figure 16, showing relatively small errors.

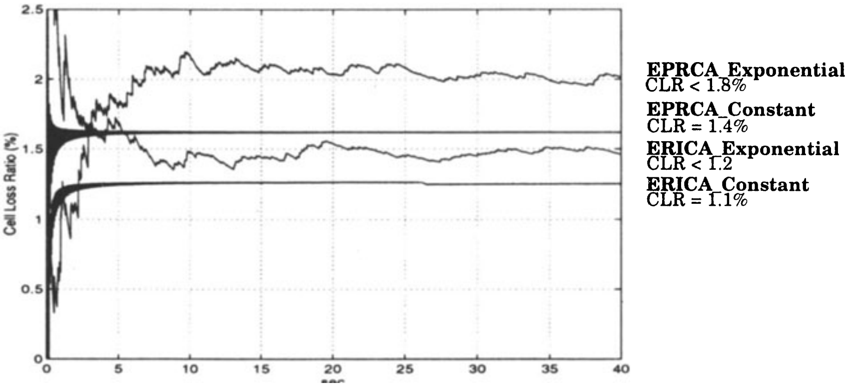


Figure 16. Simulation results.

5 CONCLUSION

This work showed that ABR service is potentially vulnerable to uncontrolled transients and contributed with a quantitative analysis of that behavior. Even though previous papers stated concern about ACR retention, an analytical study of the real impact of that was not available. The extension to the concept of ACR accumulation and the analysis of several representative switch behaviors allowed a definition of scenarios where uncontrolled transients can and cannot be harmful to ABR services. As a result, it was observed that ERICA algorithm, which control ACR even when switches are not congested, provides higher protection against uncontrolled transients. Also, the potential danger of ACR accumulation increases with the delay between source and bottleneck switch, being small for LAN but considerable for WAN.

It is important to observe, however, that ACR accumulation is not a flaw of the ATM Forum ABR specification, but a side effect. It is ultimately the ACR accumulation that allows efficient resource usage when sources are well behaved.

6 REFERENCES

ATM Forum (1996) Traffic Management Specification Version 4.0.

Chen, T. Liu, S. and Samalam, V. (1996) The Available Bit Rate Service for Data in ATM Networks. *IEEE Communication Magazine,* (34) **5**, 56-8, 63-71.

Clark, D. Shenker, S. and Zhang, L. (1992) Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism. *Proceedings of ACM SIGCOMM '92*, 14-26.

Golmie, N. Koenig, A. and Su, D. (1995) The NIST ATM Network Simulator. *NISTIR 5703.*

Jain, R. Kalyanaraman, S. Goyal, R. Fahmy, S. and Lu, F. (1995) A Fix for Source End System Rule 5. *AF-TM 95-1660.*

Jain, R. Kalyanaraman, S. Goyal, R. Fahmy, S. and Lu, F. (1996) Comments on Use-it or Lose-it. *AF-TM 96-0178.*

Jain, R. Kalyanaraman, S. Goyal, R. Fahmy, S. and Viswanathan, R. (1996) ERICA Switch Algorithm: a Complete Description. *AF-TM 96-1172.*

Jain, R. Kalyanaraman, S. Fahmy, S. and Goyal, R. (1996) Tutorial Paper on ABR Source Behavior. *AF-TM 96-1270.*

Johansson, P. and Karlsson, J. (1995) Transient Analysis of a Closed Loop Rate Control Algorithm for ATM. *Proceedings of the Twelfth Nordic Teletraffic Seminar*, 33-46.

Kolarov, A. and Ramamurthy, G. (1994) Comparison of Congestion Control Schemes for ABR Service in ATM Local Area Networks. *1994 IEEE GLOBECOM*, 913-8.

Kung, H. and Morris, R. (1995) Credit-Based Flow Control for ATM Networks. *IEEE Network*, (9) **2**, 40-8.

Little, T. and Ghafoor, A. (1990) Network Considerations for Distributed Multimedia Object Composition and Communication. *IEEE Network*, (4) **6**, 32-40, 45-49.

Siu, K. and Tzeng, H. (1994) Intelligent Congestion Control for ABR Service in ATM Networks. *Computer Communication Review*, (24) **5**, 81-106.

Walthall, R. and Clement, M. (1996) Simulation and Analysis of the Performance of EPRCA in a Wide Area ATM Network Consisting of Both ER and EFCI Switch Mechanisms. *4th International Conf. on Telecommunications Systems, Modeling and Analysis*, 330-40.

7 BIOGRAPHY

Eduardo S. C. Takahashi received his B.S. degree in Electrical Engineering from Instituto Tecnologico de Aeronautica, Brazil, in 1989, M.S. degree in Electrical Engineering from the Universidade de Sao Paulo, Brazil, in 1995, and is a Ph.D. candidate in the Department of Electrical and Computer Engineering at Carnegie Mellon University. His areas of research interests include high-performance networking, QoS provisioning, resource management, and application aware networks.