On the Effectiveness of Closed-Loop Rate-Based Traffic Management Schemes for ABR Service in ATM Networks *

Yiwei Thomas Hou[†], Nanying Yin[‡], and Shivendra S. Panwar[†]

[†]Department of Electrical Engineering
and Center for Advanced Technology in Telecommunications
Polytechnic University, Brooklyn, NY 11201, USA

[‡]Bay Networks, Inc.
Billerica, MA 01821, USA

February, 1996

Abstract

The Available Bit Rate (ABR) service class has been defined by the ATM Forum. Closed-loop rate-based congestion control has been adopted by the ATM Forum as the standard approach for supporting ABR service in ATM networks. This paper examines congestion control mechanisms for ABR service and presents fundamental performance results for rate-based traffic management schemes under various network operating conditions. We show the frequency range where a rate-based traffic management scheme can operate effectively. Our results contribute to the fundamental understanding of closed-loop traffic management mechanisms for ABR service and provide guidelines for future developments of effective rate-based congestion control algorithms.

Technical Subject Categories: Congestion/Flow Control, ATM Systems and Networks, ABR Service.

For correspondence, contact:

Prof. S. S. Panwar

Dept. of Electrical Engineering, Polytechnic University, Brooklyn, NY 11201 USA
Phone (718) 260-3740; Fax (718) 260-3074; Email: panwar@kanchi.poly.edu

*This work is supported by a NSF Graduate Research Traineeship and by the New York State Center for Advanced Technology in Telecommunications at Polytechnic University, Brooklyn, NY.

1 Introduction

Available Bit Rate (ABR) service has been defined by the ATM Forum [1]. ABR service will support applications that allow the ATM Source End System (SES) to adjust the information transfer rate based on the bandwidth availability in the network. Such applications include LAN interconnect, file transfer, Frame Relay, etc. By definition, on the establishment of an ABR connection, the user shall specify to the network both a maximum bandwidth and a minimum required bandwidth, designated as Peak Cell Rate (PCR) and Minimum Cell Rate (MCR), respectively, for the requested connection. The MCR may be specified as zero. The available bandwidth from the network may vary, but MCR is always guaranteed. A generic closed-loop rate-based traffic management algorithm is shown in Fig. 1 where the control mechanism uses Resources Management (RM) cells with Payload Type Identifier (PTI) "110", to convey network congestion and available bandwidth information to the SES. The SES adjusts the cell generating rate based on the information in the backward RM cells.

There are issues related to this type of closed-loop rate-based traffic management, including link utilization, transient time behavior, buffer requirement, fairness, implementation complexity and cost.

After the traffic management specification [1] is completed, network equipment vendors are working on the implementation of ABR. But it is extremely important to understand how and under what conditions ABR works. The objective of this paper is to investigate fundamental properties of the ABR closed-loop rate-based traffic management schemes. We present performance results under various network operating conditions such as the variation of available bandwidth as a function of time. Both binary and Explicit-Rate (ER) based feedback schemes are examined and defined. Our main objective here is *not* introducing a better traffic management algorithm (even though the algorithms that we define are refined versions of those in the literature), but rather, to show the operating frequency range under which a class of effective traffic management schemes can be designed. Here, "frequency" is a generic term referring the variation of network operating condition, i.e., variation frequency of the available bandwidth or the ABR traffic condition. This paper contributes to the fundamental understanding of closed-loop traffic management mechanisms for ABR service, and provides essential guidelines for the future development of such congestion control algorithms.

The remainder of this paper is organized as follows. Section 2 examines and defines both binary and ER traffic management schemes. The performance results for these algorithms are presented in Section 3 with discussions. Section 4 presents analytical results. Section 5 concludes this paper.

2 Closed-Loop Rate-Based Traffic Management Schemes

There are different approaches and extensive studies on closed-loop rate-based traffic management [3, 4, 6, 8, 9, 11, 12, 13, 15]. These proposals fall into two broad categories, "binary feedback" congestion indication [3, 9, 11, 13, 15] and "explicit rate setting" schemes [4, 6, 8, 12]. For simplicity, we refer them as binary schemes and ER schemes. For binary schemes, a single bit feedback from the network is used to indicate congestion. For ER schemes, rate and/or available bandwidth information is contained in the feedback RM cell to inform the SES. Binary schemes preserve backward compatibility with EFCI-marking switches [2], while the newer ER-based schemes promise higher efficiency and stability with additional implementation cost.

In this paper, we focus on one bottleneck output port of an ATM switch (Fig. 2), also referred to as the bottleneck internodal link. This output port has available bandwidth BW for ABR and is shared by N_{VC} ABR VCs from different SES. The cells from these ABR VCs share the bottleneck internodal link with FIFO queueing discipline.[1] There are extensive control parameters involved in a closed-loop traffic management mechanism. To conserve space here, we list the parameters in the appendix at the end of the paper. Note that among the list, only a subset of parameters will be used for a particular binary or ER scheme.

2.1 Binary-Based Feedback Control

Early binary schemes include Forward Explicit Congestion Notification (FECN) [13] and Backward Explicit Congestion Notification (BECN) [3, 9].

FECN makes use of the Explicit Forward Congestion Indication (EFCI) state in the Payload Type Identifier (PTI) field to convey congestion information in the forward direction. When

[1] Note that the available bandwidth BW for ABR is a variable throughout our paper.

a switch becomes congested, it will mark the EFCI state of all cells being forwarded to the destination. Upon receiving marked cells, the destination returns a RM cell with its Congestion Indication bit, CI, set to "1" to the source along the backward path. Then the SES decreases its ACR multiplicatively as

$$ACR \leftarrow max\{(ACR \times MDF), MCR\}$$

When there is no RM cell returning in a defined time interval, the source recognizes no congestion in the network. Then it increases ACR additively as

$$ACR \leftarrow min\{(ACR + AIR), PCR\}$$

Unlike FECN, which is a rate scheme with an end-to-end control, BECN returns congestion notification to the source directly from the point of congestion. An obvious advantage of BECN over FECN is the faster response to congestion. On the other hand, BECN requires more hardware complexity in the switch to generate backward RM cells.

System Definition

In a manner consistent with the ABR standards, we define our binary scheme as follows.

Source behavior:

1. A source transmits a RM cell for every N_{RM} data cells transmitted. The CI bit in a RM cell is initially set to 0.

2. The initial ACR is set to ICR.

3. Upon receiving a backward RM cell, the new ACR is set to:
   if $CI = 0$, i.e. no congestion,

$$ACR \leftarrow min\{(ACR + N_{RM} \times AIR), PCR\}$$

   else $(CI = 1)$, i.e. congestion,

$$ACR \leftarrow max\{(ACR \times MDF), MCR\}$$

Switch behavior:

1. The forward RM cells carry ACR information and update the ACR table at the switch for each VC.
2. The switch sets the CI bit of backward RM cells according to the following rule.
   When the bottleneck link queue length is over QT, the switch selectively marks $CI = 1$ on those backward RM cells whose VCs correspond to larger ACR.[2]
   When the bottleneck link queue length is over DQT, the switch marks all backward RM cells with $CI = 1$.

Destination behavior:

Upon receiving a forward RM cell from source, the destination simply returns it in the backward direction to the source.

2.2 ER-Based Feedback Control

In ATM networks, CBR, VBR, ABR and UBR share the bandwidth at an internodal link. The available bandwidth for ABR service usually changes with time depending on the variation of other types of traffic in the network. Furthermore, such bandwidth information for ABR service may not be known by the internodal link.

2.2.1 Ideal Case: Bandwidth Information is Known

When the available bandwidth at internodal link to ABR service is known, we can calculate ER for each VC.

In section 3, we will present performance evaluation for the idealized situation under which the internodal link has complete knowledge of available bandwidth and number of active VCs.

[2] This is the so called "relative rate marking" in [1] and helps to solve the "beatdown" problem. Since the focus of this paper is to investigate relative effectiveness of rate-based feedback control schemes under various frequency ranges, which can be clearly demonstrated with the behavior of one VC connection, we will not take the length to define "relative rate marking" mechanism here. In our simulation study for $N_{VC} = 1$ in Section 3.3, CI bit is set to 1 when the bottleneck link queue length is over QT.

Here, a fair bandwidth allocation to each VC may be simply the available bandwidth divided by the number of active VCs [14]. As we will see, the effective operating frequency for an idealized ER scheme is determined by the Round Trip Time (RTT) between SES and bottleneck internodal link. This give us the theoretical operating frequency limit for an optimal ER scheme.

2.2.2 Bandwidth Information is Unknown

This models a practical scenario at an internodal link where available bandwidth for ABR service is an unknown stochastic process.

Here, an accurate "estimation" of available bandwidth is essential for the ER calculation. The "intelligent congestion control" scheme proposed in [12] uses a variable at the bottleneck link, labeled as Mean Allowed Cell Rate ($MACR$), to estimate the optimal cell rate at which a VC can transmit based on the congestion state at the switch. This simple scheme avoids the use of per-VC accounting while achieving satisfactory performance in terms of fairness and link utilization.

In the following, we present our refined version of BW-estimator-based ER scheme based on [12]. In Section 3, we present performance evaluation of our scheme under various bandwidth operating frequencies. As we will see, the performance of such ER scheme is only effective in the "low frequency" range which is determined by the transient time of the switch variable ($MACR$) to reach steady state, a time scale that is much larger than RTT_{SX}.

System Definition

The SES and DES behavior of ER scheme is essentially the same as that in binary scheme, except the use of ER field than the CI bit in binary scheme.

Source behavior:

1. A source transmits a RM cell for every N_{RM} data cells transmitted. The ER field in a RM cell is set initially to PCR.

2. The initial ACR is set to ICR.

3. Upon receiving a backward RM cell, the new ACR is set to:

$$ACR \leftarrow max\{min\{(ACR + N_{RM} \times AIR), ER, PCR\}, MCR\}$$

Switch behavior:

1. The forward RM cells update $MACR$ at the switch according to the flow chart in Fig. 3.[3]
2. The switch sets the ER field of backward RM cells according to the flow chart in Fig. 4.

Destination behavior:

Upon receiving a forward RM cell from source, the destination simply returns it in the backward direction to the source.

Remark:

Our definition for "switch congestion" in Figs. 3 and 4 is a time-based congestion detection algorithm, which is more accurate than a simple threshold-based congestion detection. Here, congestion in the switch is detected by the change of queue length after processing, say, K cells. In our implementation, we use one switch variable that records the queue length seen by the last arriving RM cell. If the queue length seen by a newly arrived RM cell is greater than this variable, i.e. queue length increases between consecutive arriving RM cells, the switch is said to be in congestion. Both this variable and the switch's congestion state are updated by each newly arrived RM cell.

3 Performance Evaluation of Closed-Loop Feedback Control Schemes

In this section, we present simulation results demonstrating the performance of rate-based closed-loop feedback control algorithms defined in Section 2 under various frequency range.

[3]See Section 4 for details of our ER scheme.

Our objective is to identify the frequency range under which a scheme can operate effectively. We choose to use a persistent source (i.e. it always has data to send) under time-varying internodal available bandwidth. Although in practice, most ABR sources are not persistent (but rather, bursty on/off in nature), the operating frequencies can be most clearly illustrated by using a persistent source and frequency-varying bandwidth. Moreover, by studying the performance of a scheme for a persistent source under frequency varying bandwidth, one can imply similar results for bursty source (i.e. source profile characterizable in frequency domain [7]) under constant bandwidth, etc.

3.1 Ideal ER Scheme

This is the ideal case that we discussed in Section 2.2.1. After a SES starts to transmit cells, it takes RTT_{SD} for the first RM cell to return back to the source. This RM cell contains the exact rate at which the source should transmit. After receiving first backward RM cell, the backward RM cells return periodically to the source and the feedback information from the switch is only delayed by the propagation time between SES and switch, τ_{SX} or $\frac{1}{2}RTT_{SX}$ (remember that the switch sets the ER field in the *backward* RM cell).

Fig. 5 shows the low frequency case with $N_{VC} = 1$ and BW variation period $T(= 200$ ms) is much greater than $RTT_{SX}(= 5$ ms). The ACR at source and ACR arriving at the internodal link are delayed waveforms of BW by $\frac{1}{2}RTT_{SX}$ and RTT_{SX}, respectively. Fig. 6 shows the instantaneous load, defined as,

$$Load \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{N_{VC}} ACR_i \, at \, node}{BW}$$

and queue length for the same simulation run. We see that except for short periods (equal to RTT_{SX}) which peak and drop at the time when BW varies, the load is 1 most of time.

Figs. 7 and 8 show the performance of ER scheme when the bandwidth variation period T (= 15 ms) is close to RTT_{SX}. We see ACR arriving at internodal link can no longer keep up with the BW variation and the instantaneous load differs from 1 most of the time. Here, even the ideal ER scheme does not work.

We conclude that under ideal case where the bottleneck internodal link has complete knowledge of BW and the number of active VCs, the operating frequency range for an ER scheme is only limited by RTT_{SX} and should satisfy

$$f \ll \frac{1}{RTT_{SX}} \tag{1}$$

3.2 BW-Estimator-Based ER Scheme

This is the case we discussed in Section 2.2.2. The parameters used in our simulation are listed below.

$PCR = 155$ Mbps	$DQT = 1000$ cells	$AV = 0.25$
$ICR = 10$ Mbps	$QT = 500$ cells	$MRF = 0.5$
$MCR = 0.155$ Mbps	$\tau_{SX} = 1$ ms	$ERF = \frac{31}{32}$
$AIR = 0.03125$ Mbps	$\tau_{XD} = 1$ ms	$N_{RM} = 32$

Fig. 9 shows the ACR arriving at bottleneck internodal link and BW for one VC in low frequency case with $T = 200$ ms (the first variation period of BW actually starts from $t = 100$ ms in our simulation) and $RTT_{SX} = 2$ ms. The ramp up time as illustrated in Fig. 9 is usually a much larger time period than RTT_{SX} or RTT_{SD}. After reaching steady state, the ACR at the node follows BW quite closely. Fig. 10 shows the load and queue length for the same simulation run. We see that except peaks and drops of load at BW transition points (and a few other points), the load is close to 1 most of the time. This shows that the BW-estimator-based ER scheme works fairly well in low frequency case.

Figs. 11 and 12 show the performance of the same ER scheme when the BW variation period decreases to $T = 20$ ms, which is still much greater than $RTT_{SX}(= 2$ ms). We see that such a BW variation is already too fast for our ER scheme to operate effectively. This is due to the transient time required for $MACR$ to reach a steady state whenever BW changes value. This transient time for $MACR$ is the fundamental limitation to the operating frequency of a BW-estimator-based ER scheme.

We conclude that the effective operating frequency range for ER scheme when BW is unknown should satisfy,

$$f \ll \frac{1}{T_{MACR}} \tag{2}$$

where T_{MACR} is the transient time for $MACR$ to reach a new steady state when BW changes between values. In Section 4, we will perform analysis on transient behavior of $MACR$.

3.3 Binary Scheme

We present simulation results for the binary scheme we defined in Section 2.1. The parameters used in our simulation are listed below.

$PCR = 155$ Mbps	$DQT = 1000$ cells	$MDF = 0.99$
$ICR = 10$ Mbps	$QT = 350$ cells	$N_{RM} = 32$
$MCR = 0.155$ Mbps	$\tau_{SX} = 1$ ms	
$AIR = 0.03125$ Mbps	$\tau_{XD} = 1$ ms	

Ignoring those parameters used only for ER scheme, the above parameters are essentially the same as those listed in Section 3.2 for ER scheme.[4]

Figs. 13 and 14 show the performance of our binary scheme in low frequency case with $T = 200$ ms. In comparison with Figs. 9 and 10, it is obvious that the zig-zag nature of ACR under binary scheme makes it less desirable than ER scheme. This illustrates that ER scheme outperforms binary scheme in low frequency case.

Figs. 15 and 16 show the performance of our binary scheme in a higher frequency range with $T = 20$ ms. Here, we have more interesting results. In comparison with Figs. 11 and 12, the obvious advantage of ER scheme over binary scheme in low frequency case disappears. That is, when the link BW's variation frequency becomes sufficiently high, the elaborate BW-estimator-based ER scheme loses its accuracy may not perform better than a simple binary bit setting mechanism.

[4]Since our binary scheme does not employ the elaborate time-based congestion detection algorithm used for ER scheme (see Remark on page 6), Queue Threshold ($QT = 350$ cells) is set to be smaller than that in Section 3.2 (500 cells) to make binary scheme operate properly.

3.4 Discussions

Based on our simulation results, we further make a qualitative plot of "Scheme Effectiveness" vs. "Frequency" in Fig. 17 to give fundamental insights on closed-loop traffic management. The "Scheme Effectiveness" is a measure of: 1. how well the ACR at node can follow the available BW at node; and 2. the fluctuation of buffer occupancy (ideally, we want to keep the buffer content at a steady constant level or load equals to 1).

As shown qualitatively in Fig. 17, at low frequency range ($\ll RTT_{SX}^{-1}$), ideal ER scheme is the most effective closed-loop congestion control. The BW-estimator-based ER scheme is better than binary scheme only at very low frequency range ($\ll T_{MARC}^{-1}$). Once over (T_{MARC}^{-1}), such ER scheme loses its accuracy and may not be better than binary scheme. At higher frequency range ($\sim RTT_{SX}^{-1}$), any feedback scheme does not work well and the ideal ER scheme may be the worst since the ER information being fed back to the SES is always incorrect. At such a frequency range, binary scheme may be the most robust scheme because single bit feedback better reflects the congestion state at bottleneck node.

The operating frequency range for ER and binary feedback mechanisms were clearly demonstrated above by using a persistent source under frequency varying bandwidth. In practice, the number of ABR VCs at a bottleneck internodal link varies with time as well as each ABR VC source profile (usually on/off). A simple expression showing the operating frequency is not obvious. However, it it expected that equivalent or similar results will hold. As an example, for a BW-estimator-based ER scheme, only when the on/off burstiness of each source as well as number of VCs vary on a much larger time scale than T_{MACR}, will such a scheme operate effectively.

4 Analytical Results

In this section, we analyze the transient behavior of ACR and $MACR$ for the BW-estimator-based ER scheme defined in Section 2.2.2.

4.1 Rise Time Analysis

Assuming the buffer is empty at the internodal link, then the node is congestion free when BW changes from BW_l to a higher value BW_h. Suppose a source receives a RM cell at t and the next RM cell at t'. Then we have

$$ACR(t') = ACR(t) + N_{RM} \cdot AIR \tag{3}$$

Rearranging terms and divide both sides by $t' - t$,

$$\frac{ACR(t') - ACR(t)}{t' - t} = \frac{N_{RM} \cdot AIR}{t' - t}$$

We approximate the above by the following continuous-time differential equation:

$$\frac{d}{dt}ACR(t) = \frac{N_{RM} \cdot AIR}{t' - t} \tag{4}$$

Ignoring cell transmission time at internodal link, a RM cell returned to the source at time t must be transmitted at time $t - RTT_{SD}$ under a source cell rate of $ACR(t - RTT_{SD})$. According to the system definition of our ER scheme, N_{RM} cells were transmitted in the interval $[t - RTT_{SD}, t' - RTT_{SD}]$, i.e.

$$N_{RM} = \int_{t-RTT_{SD}}^{t'-RTT_{SD}} ACR(\tau)d\tau \tag{5}$$

Again, we approximate the above by letting

$$N_{RM} \approx ACR(t - RTT_{SD}) \cdot [(t' - RTT_{SD}) - (t - RTT_{SD})]$$

or

$$t' - t = \frac{N_{RM}}{ACR(t - RTT_{SD})} \tag{6}$$

Combining Eqns. (4) and (6), we have

$$\frac{d}{dt}ACR(t) = AIR \cdot ACR(t - RTT_{SD}) \tag{7}$$

Eqn. (7) is a delay-differential equation, whose closed-from solution is difficult to obtain in general. It is possible to solve it using approximate exponentials. Due to the scope of this paper, such analysis will not be given here. Instead, we will give a solution through an iterative numerical procedure.

By letting

$$\frac{d}{dt}ACR(t) \approx \frac{ACR(t+\delta) - ACR(t)}{\delta}$$

Eqn. (7) becomes

$$ACR(t+\delta) = ACR(t) + ACR(t - RTT_{SD}) \cdot (AIR \cdot \delta) \tag{8}$$

Let

$$M = \frac{RTT_{SD}}{\delta}$$

and using the initial conditions

$$ACR(0) = ACR(\delta) = ACR(2\delta) = \cdots = ACR(M\delta) = \frac{BW_l}{N_{VC}} \tag{9}$$

we have

$$\begin{aligned}
ACR[(M+1)\delta] &= ACR(M\delta) + ACR(0) \cdot (AIR \cdot \delta) \\
ACR[(M+2)\delta] &= ACR[(M+1)\delta] + ACR(\delta) \cdot (AIR \cdot \delta) \\
&\vdots \\
ACR[(K+1)\delta] &= ACR(K\delta) + ACR[(K-M)\delta] \cdot (AIR \cdot \delta) \\
&\vdots
\end{aligned}$$

This iterative procedure terminates when

$$ACR[(K+1)\delta] \approx \frac{BW_h}{N_{VC}}$$

and such $(K+1)\delta$ is the transient time for ACR at SES to change from BW_l/N_{VC} to BW_h/N_{VC}.

To validate our analysis, we quantitatively compare our analytical results with simulation in Table 1. The parameter settings are the same as in Section 3.2, with the additional $\frac{BW_l}{N_{VC}} = ICR$, $\delta = \frac{RTT_{SD}}{100}$ and $M = 100$.

BW_h/N_{VC} (Mbps)	Analysis (ms)	Simulation (ms)
20	15.48	14.40
40	27.36	26.49
60	34.32	33.50
80	39.26	38.47
100	43.12	42.32

Table 1: Analytical and simulation results for ACR rise time at SES.

Fig. 18 illustrates the case for BW_h/N_{VC} = 100 Mbps shown in Table 1. Note that it is possible to reduce the transient time for ACR from BW_l/N_{VC} to BW_h/N_{VC} by increasing AIR. But this will results wider oscillation range for ACR once it reaches BW_h/N_{VC}, which would undermine the effectiveness of our ER scheme.

Also shown in Fig. 18 is the $MACR$ which follows closely to the delayed version (by τ_{SX}) of ACR at source. Unlike ACR, which keeps oscillating (in an attempt to increase cell rate should bandwidth becomes available from the network), the $MACR$ stay closely to BW_h/N_{VC} once it reaches there. This is because $MACR$ can be increased only when switch is congestion free *and* ACR is greater than $MACR$ (see protocol in Fig. 3). Or equivalently, $MACR$ can only be increased if

$$MACR < ACR < \frac{BW_h}{N_{VC}} \tag{10}$$

Once $MACR$ reaches BW_h/N_{VC}, Eqn. (10) will no longer be satisfied and $MACR$ will stay at BW_h/N_{VC}, which is the desired optimal cell rate that an ABR VC should set.

4.2 Decay Settling Time

The exact analysis of $MACR$ behavior for our protocol when BW decreases from BW_h to BW_l is very complicated. However, for the purpose of this paper, we will show qualitatively that:

1. $MACR$ will eventually settle down to BW_l/N_{VC}.
2. Such decay settling time is greater than RTT_{SX}.

Claim 2 above is obvious since it will take at least RTT_{SX} to make any necessary change of ACR arriving at node when feedback information is marked on backward RM cells. If the difference between BW_h/N_{VC} and BW_l/N_{VC} is large, it is intuitive that such decay settling time will be much larger than RTT_{SX}.

To show Claim 1 is true, we use similar arguments in Section 4.1. According to our ER protocol (Fig. 3), $MACR$ can only be decreased when switch is in congestion (queue length is found to be increasing by arriving RM cell at node) *and* ACR is less than $MACR$. Or equivalently, $MACR$ can only be decreased if

$$\frac{BW_l}{N_{VC}} < ACR < MACR \tag{11}$$

As an example, Fig. 19 shows the expanded view of part of the simulation results of Figs. 9 and 10. Segments A, B and C show the time intervals where Eqn. (11) is satisfied. Also note that at the end of segments A, B and the entire segment D, there are major rate reductions ($ER \leftarrow min(ER, MRF \cdot MACR)$) when queue length is over DQT (1000 cells). Since ACR adjusts its rate dynamically and based on our protocol, it will attempt to increase its rate beyond BW_l/N_{VC}. Therefore, Eqn. (11) is satisfied from time to time and $MACR$ is decreased eventually down to BW_l/N_{VC}. Once $MACR$ reaches BW_l/N_{VC}, Eqn. (11) will no longer be satisfied. So $MACR$ will settle down at BW_l/N_{VC}, which is again the desired optimal cell rate an ABR VC should set.

5 Concluding Remarks

In this paper, we examined closed-loop rate-based traffic management schemes for ABR service and presented refined versions of binary and ER schemes. More important, we presented performance results for binary/ER schemes under various bandwidth frequency ranges. We obtain the following fundamental principles for designing effective traffic management algorithms for ABR service:

1. Ideally, the ER scheme's operating frequency is constrained by Round Trip time (RTT_{SX}) of feedback loop.
2. In practice, when network bandwidth available to ABR service is unknown, the operating frequency for ER scheme is further limited by the transient response time of the bandwidth estimator (T_{MACR}).

3. BW-estimator-based ER scheme performs better than simple binary-bit-setting scheme in low frequency ($< T_{MACR}^{-1}$). However, such advantage diminishes as internodal link BW variation frequency increases.

4. At higher frequency range ($\sim RTT_{SX}^{-1}$), ideal ER scheme may be worse than binary scheme since the rate information received by the SES is always incorrect. On the other hand, binary scheme is the most robust mechanism because single bit feedback information better reflects the congestion state at node. Thus, the necessity of implementing ER algorithm (with additional cost) over simple binary algorithm at an ATM switch has to be carefully justified according to the actual network environment.

Our conclusions are based on the network model of persistent sources under frequency-varying bandwidth. In practice, the ABR source traffic profile (number of VCs and on/off behavior for a particular VC) is highly bursty and unpredictable, but we expect equivalent or similar results will hold under this general network model.

Appendix: Control Parameters

Source End System Parameters

PCR Peak Cell Rate; a maximum rate which *ACR* can set.

MCR Minimum Cell Rate; a minimum rate of *ACR*.

ICR Initial Cell Rate; an initial value for *ACR*.

AIR Additive Increase Rate; rate increase permitted.

MDF Multiplicative Decrease Factor; rate reduction factor.

N_{RM} Number of Cells/RM; $N_{RM} = 2^N$.

Source End System Variables

ACR Allowed Cell Rate; current cell rate at a source.

Switch Parameter Settings

MACR Modified Allowed Cell Rate; a switch variable used to estimate the optimal *ACR* that an ABR VC should set. Ideally, it should be the available bandwidth divided by the number of active connections.

DQT Down Queue Threshold; the high queue limit to determine very congested.

QT Queue Threshold; the low queue limit to determine congestion.

AV Averaging factor; used by *MACR* to estimate optimal *ACR*.

MRF Major Reduction Factor.

DPF Down Pressure Factor.

ERF Explicit Rate Factor.

RM Cell Fields

ACR Set to Allowed Cell Rate (i.e. current cell rate) by the source when RM cell is generated.

DIR Direction of RM cell; forward or backward.

CI Congestion Indicator; 0 = no congestion, 1 = congestion.

ER Explicit Rate; initially set to PCR, and possibly modified by intermediate nodes along the path.

Other Parameters

τ_{SX} Propagation delay between SES and bottleneck switch.

τ_{XD} Propagation delay between bottleneck switch and DES.

RTT_{SD} Round Trip Time between SES and DES; $RTT_{SD} = 2(\tau_{SX} + \tau_{XD})$.

RTT_{SX} Round Trip Time between SES and bottleneck switch; $RTT_{SX} = 2\tau_{SX}$.

N_{VC} Number of ABR virtual connections.

References

[1] ATM Forum Technical Committee, "Traffic Management Specification - Version 4.0," *ATM Forum/95-0013R10,* Feb. 1996.

[2] ATM Forum Inc., *ATM User Network Interface (UNI) Specification Version 3.1,* First Edition, Prentice-Hall, Inc., New Jersey, 1995.

[3] A. W. Berger, F. Bonomi, and K. W. Fendick, "Proposal for backward congestion feedback at the ATM UNI," *ATM Forum Contribution, 93-839R1,* Sept. 1994.

[4] A. Charny, D. D. Clark, and R. Jain, "Congestion Control with Explicit Rate Indication," *Proc. IEEE ICC '95,* pp.1954-1963, June 1995.

[5] Y. T. Hou, L. Tassiulas, and H. J. Chao, "Overview of Implementing ATM-Based Enterprise Local Area Networks," *IEEE Communications Magazine,* to appear 1996.

[6] R. Jain, S. Kalyanaraman, and R. Viswanathan, "The OSU Scheme for Congestion Avoidance Using Explicit Rate Indication," *ATM Forum Contribution, 94-0884,* Sept. 1994.

[7] S.-Q. Li and C.-L. Hwang, "Queue Response to Input Correlation Functions: Continuous Spectral Analysis," *IEEE/ACM Trans. on Networking,* Vol.1, No.6, Dec. 1993, pp.678-692.

[8] B. Lyles and A. Lin, "Definition and Preliminary Simulation of a Rate-Based Congestion Control Mechanism with Explicit Feedback of Bottleneck Rates," *ATM Forum Contribution, 94-0708,* July 1994.

[9] P. Newman, "Backward Explicit Congestion Notification for ATM Local Area Networks," *Proc. IEEE Globecom '93,* pp.719-723, Dec. 1993.

[10] H. Ohsaki, et al., "Rate-Based Congestion Control for ATM Networks," *ACM Computer Communication Review,* Vol.25, No.2, April 1995, pp.60-72.

[11] K. K. Ramakrishnan and R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks," *ACM Transactions on Computer Systems,* Vol.8, No.2, pp.158-181, May 1990.

[12] K.-Y. Siu and H.-Y. Tzeng, "Intelligent Congestion Control for ABR Service in ATM Networks," *ACM Computer Communication Review,* Vol.24, No.5, Oct. 1994.

[13] N. Yin and M. G. Hluchyj, "On Closed-Loop Rate Control for ATM Cell Relay Networks," *Proc. IEEE Infocom '94*, pp.99-108, June 1994.

[14] N. Yin, "Fairness Definition in ABR Service Model," *ATM Forum Contribution, 94-0928R2,* Nov. 1994.

[15] N. Yin, "Analysis of a Rate-Based Traffic Management Mechanism for ABR Service," *Proc. IEEE Globecom '95*, pp.1076-1082, Nov. 1995.

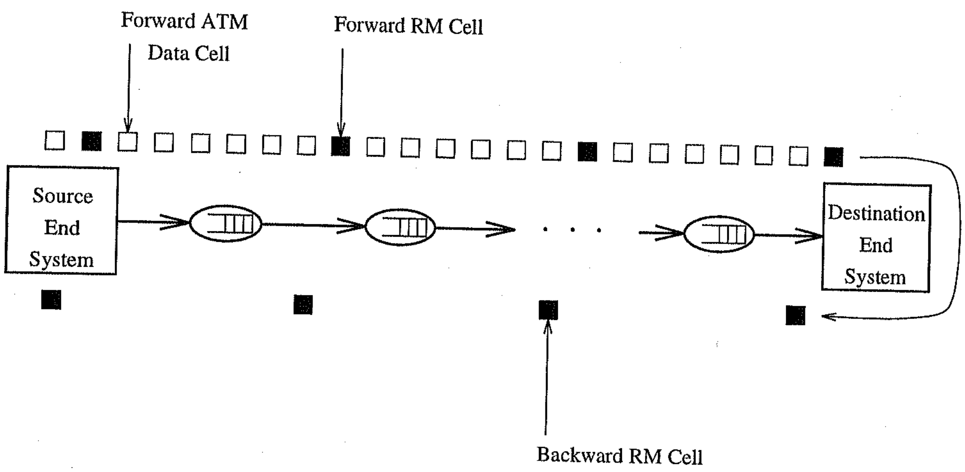

Figure 1: Closed-loop rate-based traffic management for one ABR VC.

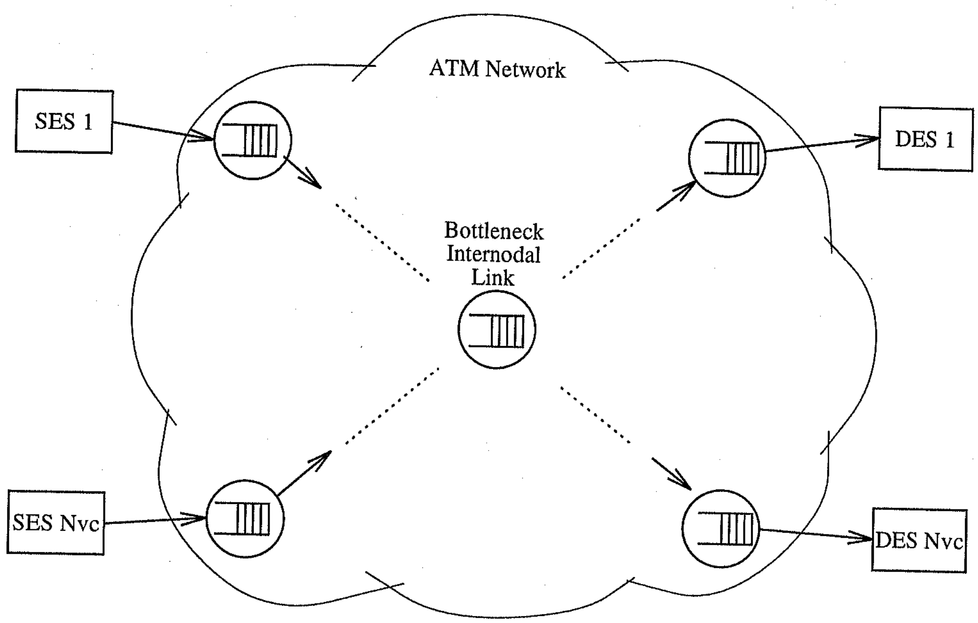

Figure 2: Closed-loop rate-based traffic management for ABR VCs at a bottleneck internodal link.

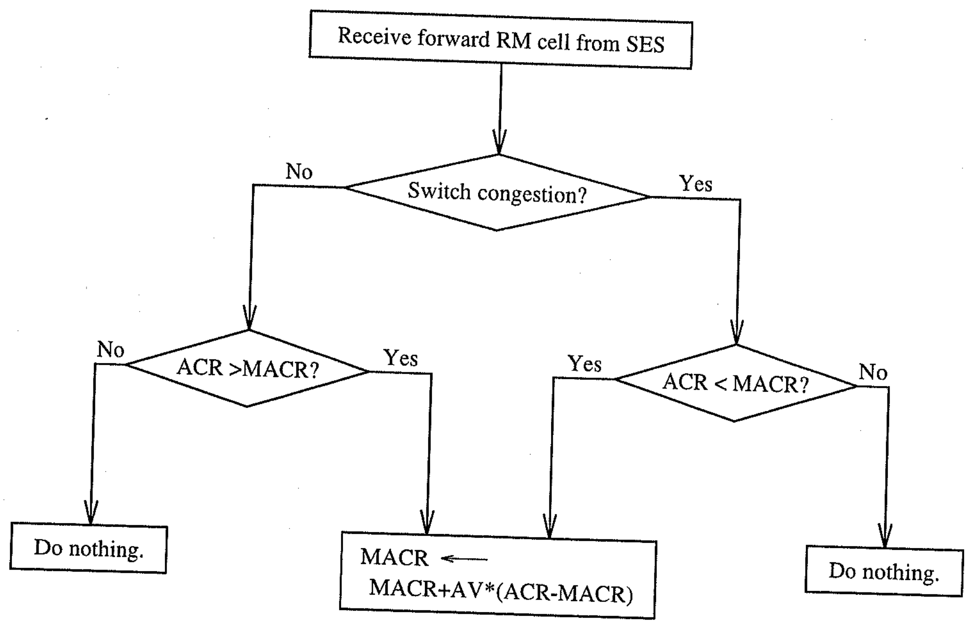

Figure 3: Switch behavior when receiving forward RM cells.

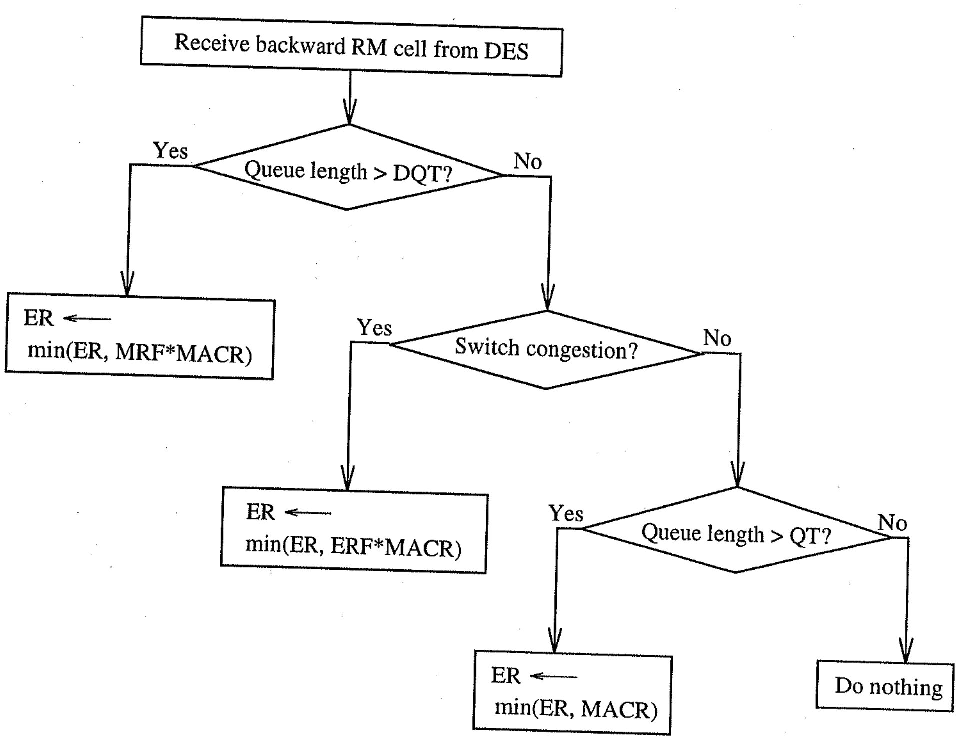

Figure 4: Switch behavior when receiving backward RM cells.

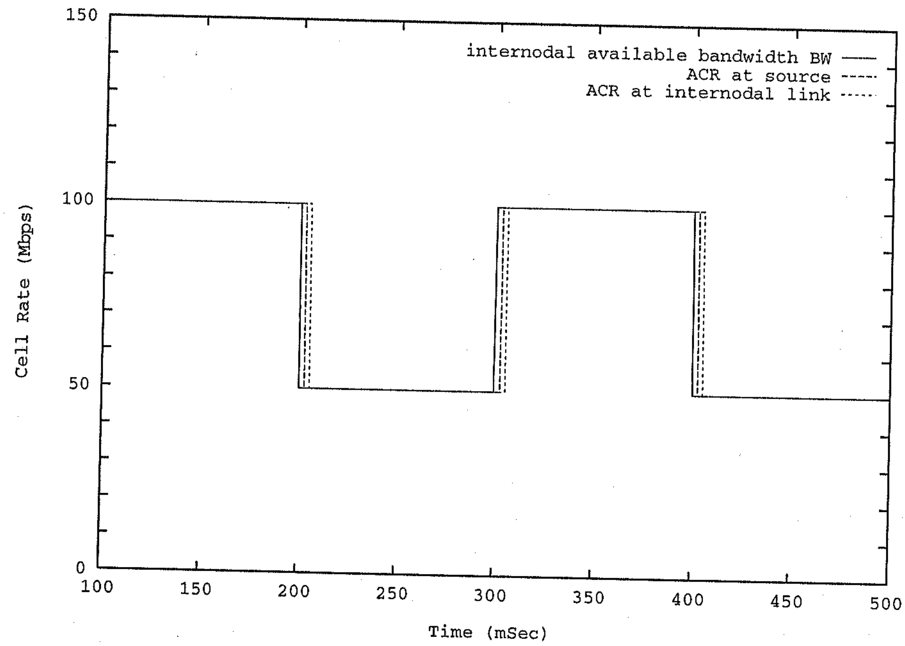

Figure 5: Ideal ER scheme in low frequency case: ACR at source, ACR at node and link BW.

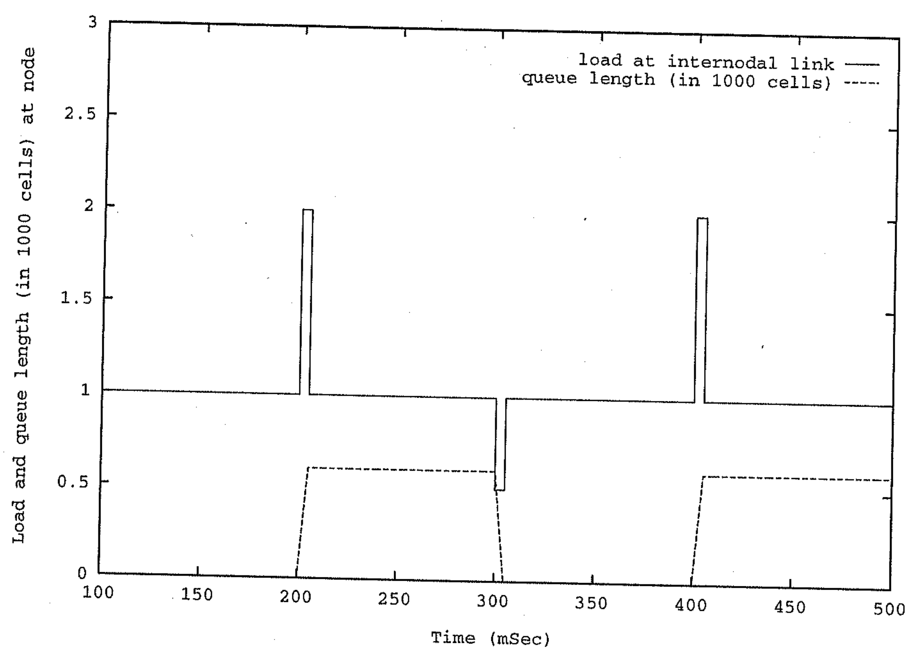

Figure 6: Ideal ER scheme in low frequency case: traffic load and queue length.

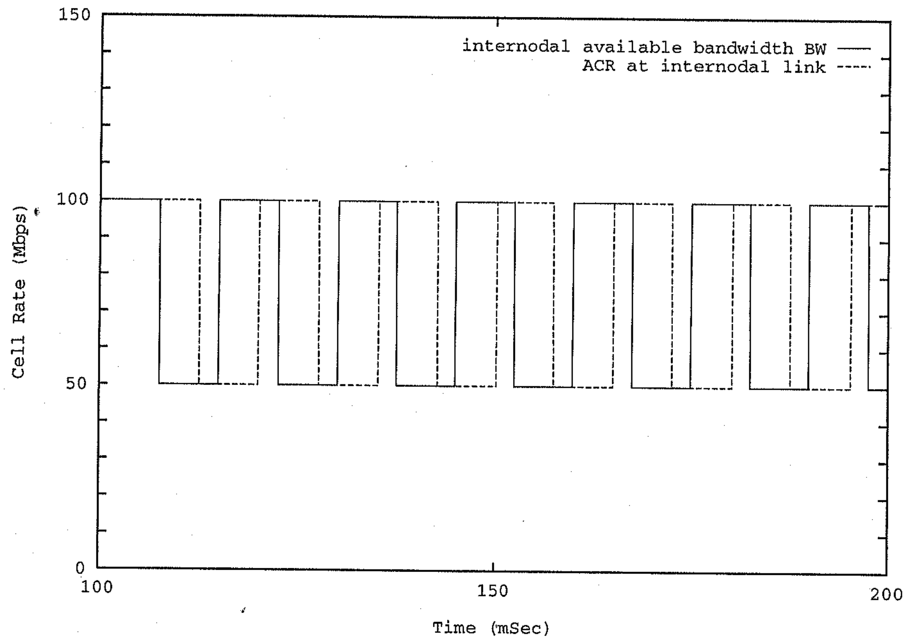

Figure 7: Ideal ER scheme: ACR at node and link BW.

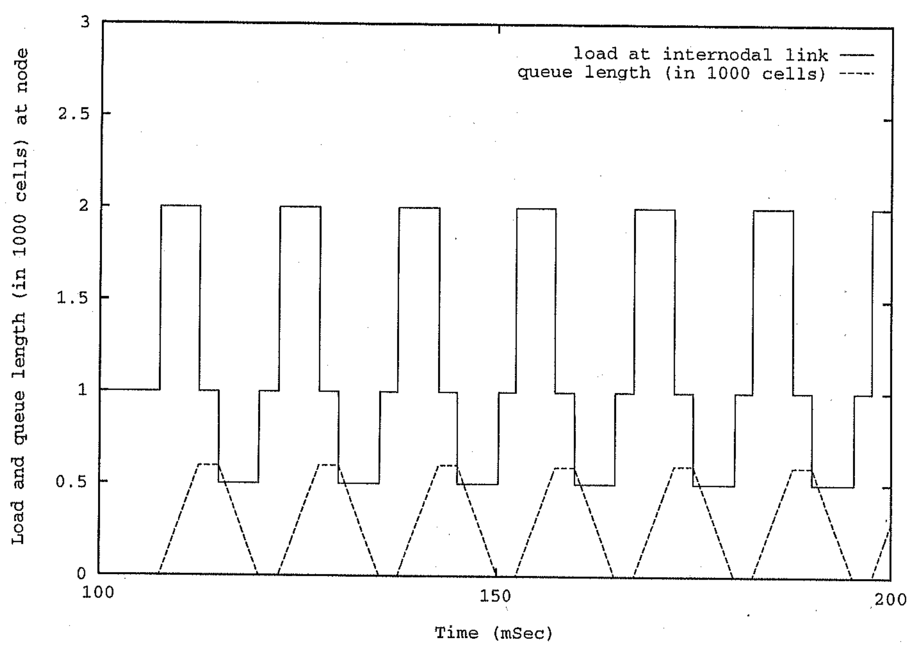

Figure 8: Ideal ER scheme: traffic load and queue length.

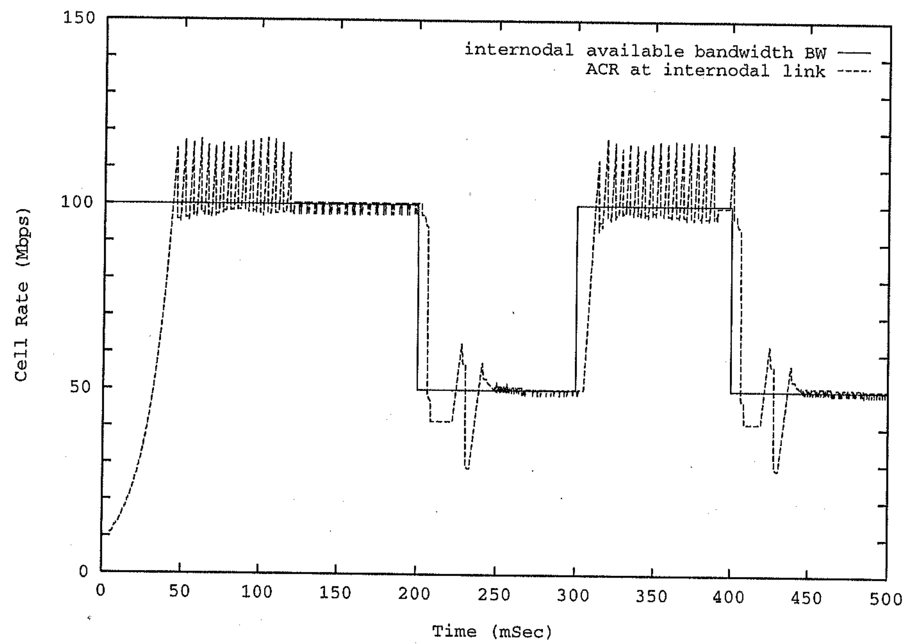

Figure 9: BW-estimator-based ER scheme in low frequency case: ACR at node and link BW.

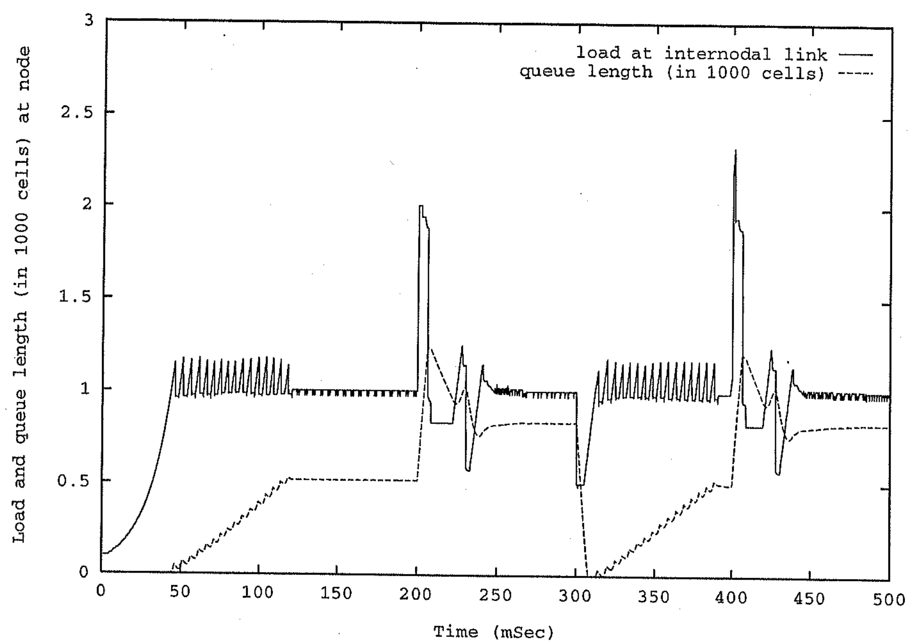

Figure 10: BW-estimator-based ER scheme in low frequency case: traffic load and queue length.

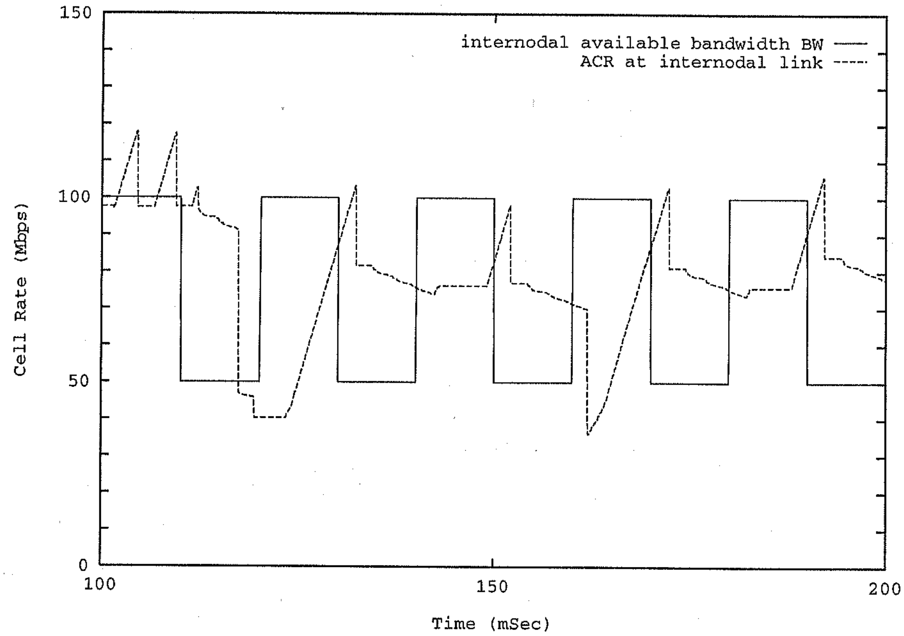

Figure 11: BW-estimator-based ER scheme: ACR at node and link BW.

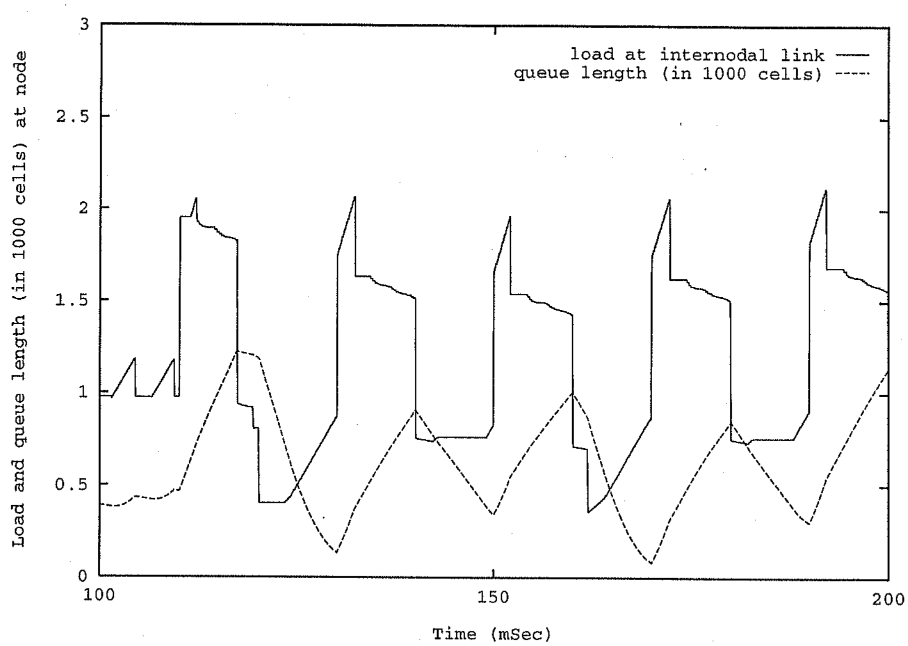

Figure 12: BW-estimator-based ER scheme: traffic load and queue length.

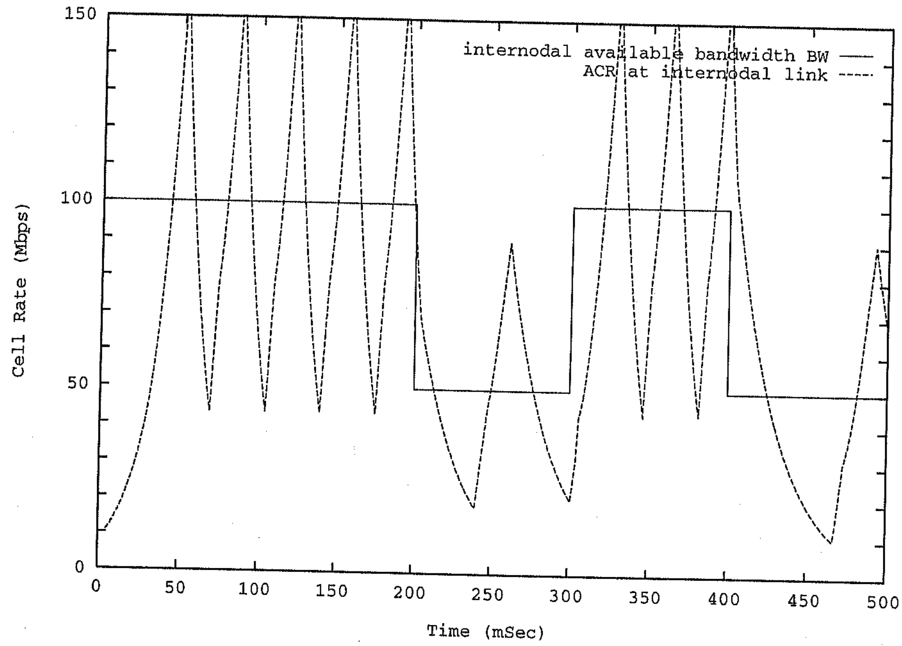

Figure 13: Binary scheme in low frequency case: ACR at node and link BW.

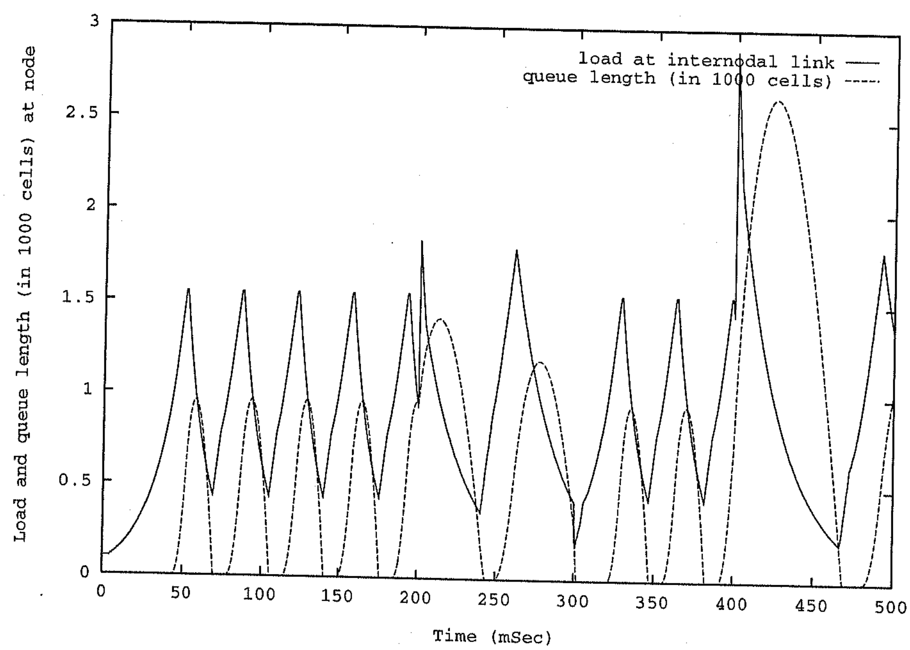

Figure 14: Binary scheme in low frequency case: traffic load and queue length.

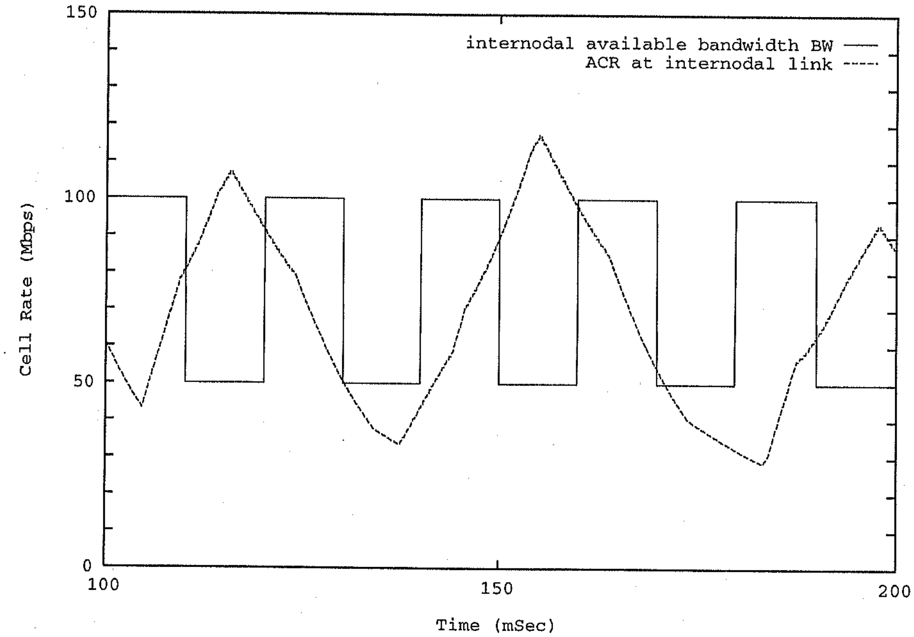

Figure 15: Binary scheme: ACR at node and link BW.

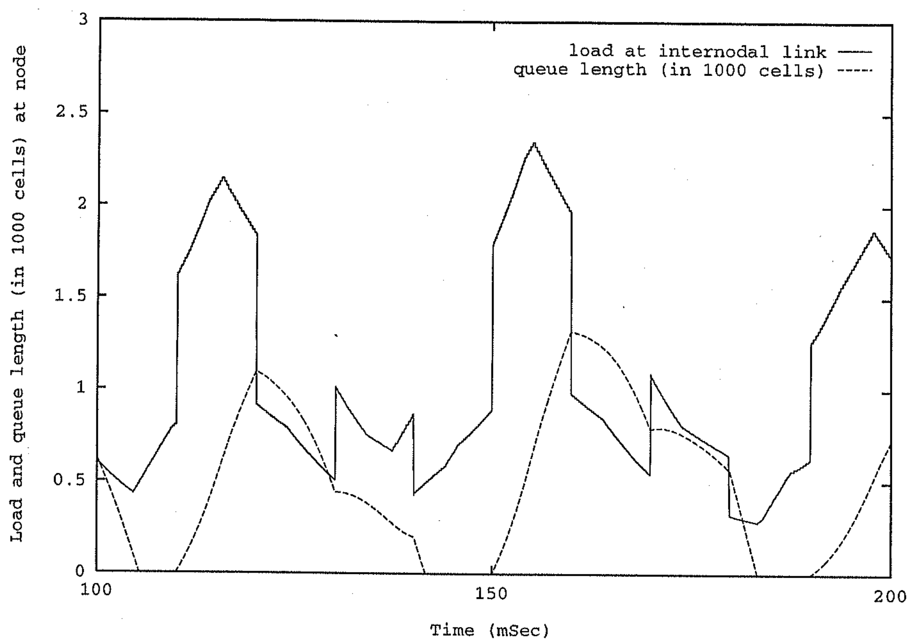

Figure 16: Binary scheme: traffic load and queue length.

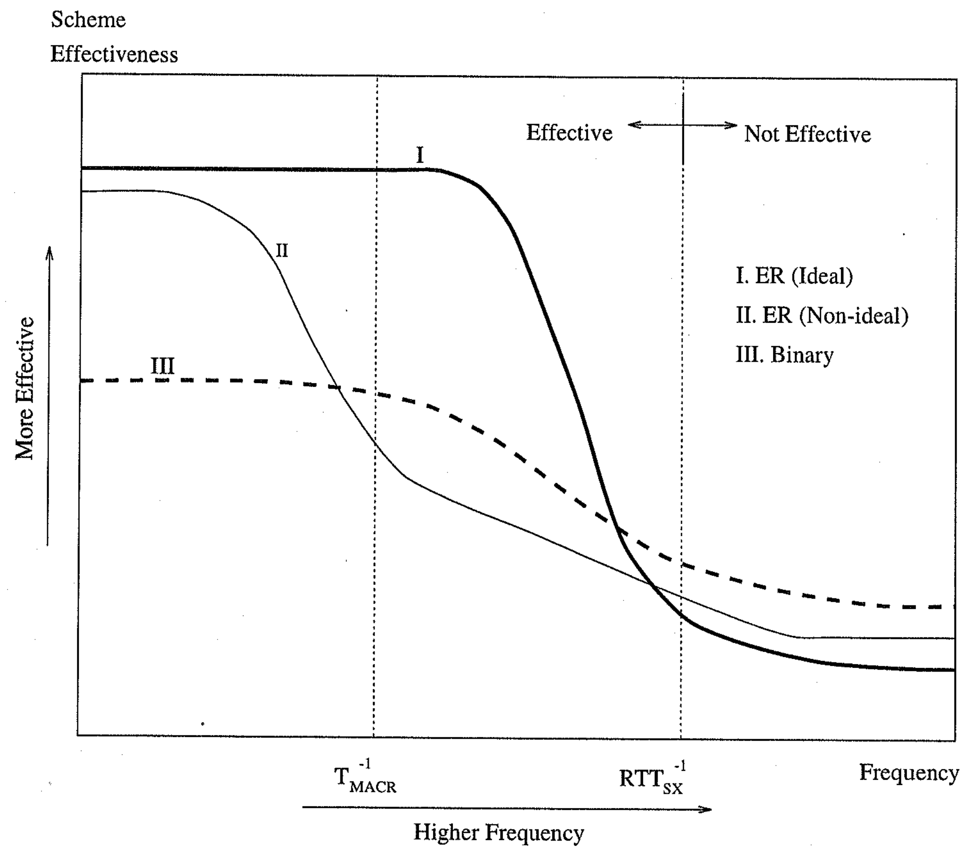

Figure 17: Effectiveness comparison of feedback control schemes.

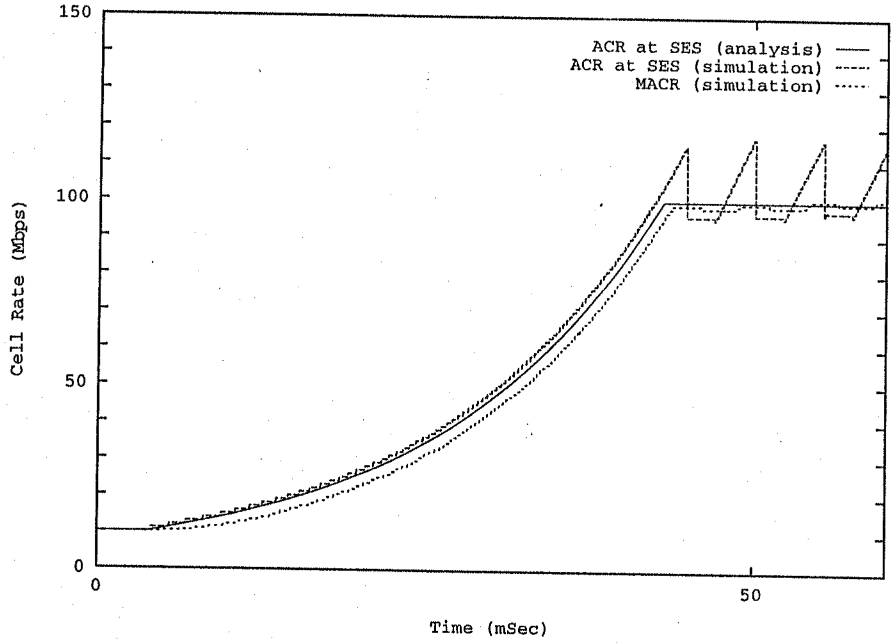

Figure 18: Transient rise behavior of ACR and $MACR$ at SES.

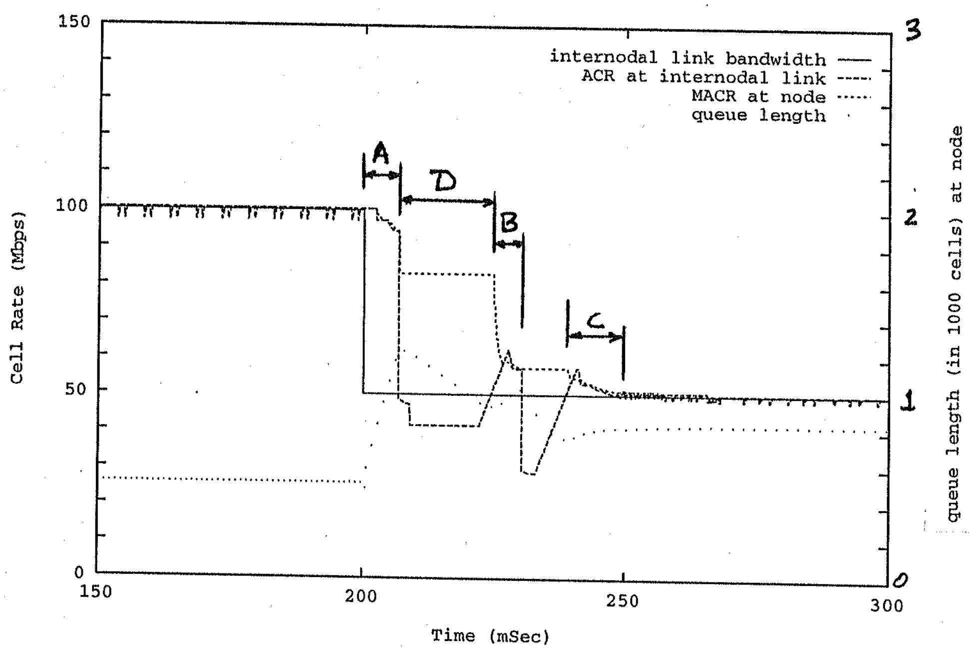

Figure 19: Transient decay behavior of ACR and $MACR$ at SES.