Performance Evaluation of Rate-Based Congestion Control Methods in ATM Networks

Hiroyuki Ohsaki

Department of Information and Computer Sciences
Faculty of Engineering Science
Osaka University

Abstract

An ATM technology realizes B-ISDN by asynchronously treating various multimedia information such as data, voice and video. In ATM networks, traffic management is an important issue to utilize network resources efficiently, and ABR traffic is indispensable for data communication and most of existing applications. Rate-based congestion control is a promising scheme to incorporate ABR traffic into ATM networks because of its simplicity and scalability from local area to wide area ATM networks.

In this paper, we first present a historical overview of rate-based congestion control algorithms developed in the ATM Forum. Scope of this survey includes FECN-based and BECN-based schemes, and binary feedback and explicit rate setting schemes. A summary for fairness definitions for ABR traffic is also presented. In this survey, we show how the current standard mechanism regarding the traffic management is exploited. Then, we quantitatively evaluate the performance of rate-based congestion control method in both initial transient and steady state by using an analytical approach to exhibit its effectiveness. Three types of switch algorithms suggested in the standard are considered. Through numerical examples, we show the effect of the number of active connections and the propagation delays on buffer requirements for ATM switch depending its algorithm. It is also emphasized that appropriate control parameter settings are essential for the rate-based congestion control to offer a proper traffic management tool in an ATM network environment.

Keywords

ATM Networks, Rate-Based Congestion Control, ABR Traffic, EPRCA, Fluid Flow Approximation

Contents

List of Figures

1 Introduction

The capability of ATM (Asynchronous Transfer Mode) networks for providing large bandwidth and handling multiple quality of service (QOS) guarantees can only be realized by preparing effective traffic management mechanisms [1]. Traffic management includes congestion control, call admission control and virtual path (VP)/virtual channel (VC) routing. Especially, an essential issue for stable and efficient operation of ATM networks is congestion control, which is performed between ATM end systems (ES). The ATM end system is a point where an ATM connection is terminated, and the connection goes up to the ATM adaptation layer (AAL). It is also defined as a point where VC connections are multiplexed, demultiplexed or both. Therefore, each ATM segment, a private local area network (LAN) or a public wide area network (WAN), may adopt a different congestion control scheme depending on its characteristics.

Two different congestion control strategies have been discussed for ATM networks; open-loop control and closed-loop control. The open-loop control limits each connection's usable bandwidth based on the notion of a *traffic contract* [2, 3]. To assign bandwidth to the connection, each end system has to declare its traffic parameters to the network prior to connection establishment. These traffic parameters are described in terms of, e.g., peak cell rate, cell delay variation (CDV), sustainable cell rate and/or burst length tolerance. Then, its associated QOS can be guaranteed throughout the session once the connection request is admitted. In other words, a lack of network resources may cause rejection of a newly requested connection in open-loop control. Therefore, open-loop control is sometimes referred to as preventive congestion control by its nature. This sort of congestion control scheme can be applied to real time communications called constant bit rate (CBR) and variable bit rate (VBR) traffic, by which audio and/or motion video can be accommodated.

However, open-loop control becomes insufficient for data communications because each connection can never emit cells exceeding its negotiated rate even when there remains unused bandwidth in the network [1]. Furthermore, data traffic is not likely to have a capability to predict its own bandwidth requirements at connection setup time. Instead, it can adjust the cell transmission rate if the up-to-date congestion status of the network is appropriately informed. These are reasons why closed-loop rate control is promising for data communications, and it is being applied to ABR (Available Bit Rate) service in the ATM Forum. Closed-loop control is sometimes called reactive congestion control, and dynamically regulates the cell emission process of each connection by using feedback information from the network. It is, therefore, essentially suitable to data transfer service because of the incapability to predict QOS required from the network.

For implementation of closed-loop control, two different schemes have been proposed in the ATM Forum; rate-based and credit-based schemes. The credit-based scheme is based on a link-by-link window flow control mechanism [4]. In the credit-based scheme, independent flow controls are performed on each link for different VC's, and each VC must obtain buffer reservations for its cell transmission on each link. This reservation is given in the form of a *credit balance*. As long as each VC gains credit from the next node, it is allowed to continue cell transmission. When the VC is starved of credits, it should wait for the next credits. Owing to this link-by-link fast feedback mechanism, transient congestion can be effectively relieved. In addition, no cell loss occurs since each VC has no opportunity to send cells unless credits are given.

On the other hand, the rate-based scheme controls the cell emission rate of each VC between end systems. It is simpler compared with credit-based flow control schemes in which each switch requires complicated queue management for each VC. Further, it should be appropriate for ATM networks to utilize its large bandwidth. Typical examples of the rate-based approach are forward explicit congestion notification (FECN) and backward explicit congestion notification (BECN) methods [5], which are well-known congestion control strategies in conventional packet switched networks. One of the main purposes of this paper is to discuss the rate-based congestion control schemes proposed and developed so far in the ATM Forum. Implementation aspects of FECN and BECN like methods in the proposed rate-based congestion control method are also described. Further, we will introduce more intelligent schemes in which the switch controls the rate of connections explicitly. Then, quantitative evaluation is provided through an analytical approach to show the effectiveness of the rate-based congestion control method. In numerical examples, it is emphasized that the appropriate control parameter settings play an essential part for the rate-based congestion control to offer a proper traffic management tool in an ATM network environment. Excellent arguments for which methodologies in recent high speed network environments can be found in [6]. Interested readers should refer to it.

The rest of this paper is organized as follows. In Section 2, we introduce several rate-based congestion control schemes proposed in the ATM Forum and qualitatively evaluate these methods in turn. In Section 3, we analyze the Enhanced Proportional Rate Control Algorithm (EPRCA), which is a basis of the rate-based control schemes in the ATM Forum, and illustrate its performance through numerical examples. An analysis of the EPRCA in initial transient state is then presented in Section 4. Simulation studies for rate-based congestin control methods are performed in Section 5. Section 6 includes some concluding remarks and open issues. The control parameters and pseudo-code used in the rate-based congestion control methods are summarized in Appendix A and Appendix B. The current proposal for RM cell format is contained in Appendix C. Appendix D provides the abbreviation list.

2 Rate-Based Congestion Control Schemes in ATM Networks

In this section, we introduce the rate-based congestion control schemes that have been proposed in the ATM Forum. In the ATM Forum, several proposals have been contributed in the rate-based congestion control framework by the end of 1993; the methods based on FECN [7, 8] and the ones based on BECN [5, 9]. Then, the Rate-Based Traffic Management Ad-Hoc working group was established to discuss various aspects of the rate-based congestion control methods. The result was published in the ATM Forum in [10], which will be precisely described in the next subsection. The ATM Forum standard regarding traffic management only specifies the source and destination end systems behaviors, and implementation methods of the switches is left to the manufacturers. We will describe how end systems' behavior is standardized and also how the various switches proposed in the ATM Forum can cooperate with the *standardized* end systems.

2.1 Interval-Based Approach

In this subsection, we explain an original rate-based scheme proposed in [10, 11]. Figure 1 illustrates a basic configuration of the rate-based congestion control scheme in which the ATM connection is terminated at source and destination end systems. A permitted cell transmission rate ACR (Allowed Cell Rate) of the source end system is changed according to the congestion status of the network, which is an essential part of the rate-based congestion control scheme. An initial rate ICR, a maximum allowable rate PCR, and a minimum cell rate MCR are specified by the network at connection setup time. Then, the source is allowed to emit cells at a rate that ranges from 0 to ACR. When compared with later proposals described in the following subsections, a distinguishable point of the original scheme is that both end systems are operated based on interval timers. In addition, the polarity of the feedback information from the network is *negative*, that is, the feedback information is received by the source end system only when the network falls into congestion.

An occurrence of congestion is detected at each intermediate switch by a queue length of the cell buffer (defined as Q_H). Then, it is signaled to the source by a special cell called a RM (Resource Management) cell, of which PTI (Payload Type Identifier) is "110". The FECN like signaling mechanism is defined in this scheme, that is, each switch signals its congestion information to its downstream switches by setting an EFCI (Explicit Forward Congestion Indication) bit contained in the header of passing data cells. When the destination end system receives a data cell in which the EFCI bit is marked, it sends back a RM cell to the source along the backward path. Then, the source end system must decrease its ACR multiplicatively according to this feedback

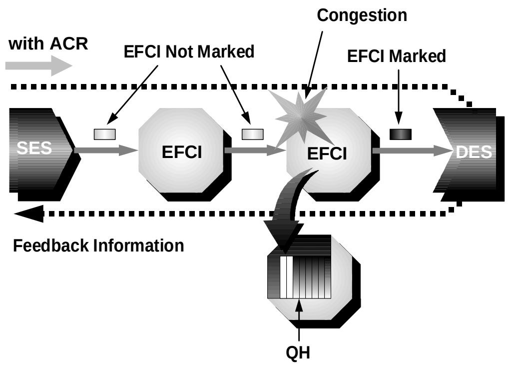

Figure 1: Basic Configuration of the Rate-Based Congestion Control.

information as

$$ACR \quad \leftarrow \quad max(ACR \times RDF, MCR) \tag{1}$$

where RDF is the rate decrease factor and MCR is the minimum cell rate for the ACR. A time interval RMI (RM Interval) is defined at the destination end system, and only one RM cell is allowed to be sent every RMI. The source end system is also provided with an interval timer UI (Update Interval). When the timer expires and no RM cell is received during UI, the source recognizes no congestion occurrence in the network. Then, it increases ACR additively as

$$ACR \quad \leftarrow \quad min(ACR + AIR, PCR),$$

where AIR is the additive increase rate and PCR is the peak cell rate of that connection.

As an implementation option, the network can be divided into two or more segments by introducing *intermediate* networks which should act as *virtual* destination end systems for the source, and as virtual source end systems for the destination (Fig. 2). As a destination end system, an intermediate network has to send back RM cells to the source following the EFCI status of incoming cells. In addition, it is required to regulate the flow of cells destined for the destination end system as a source end system.

Since this approach requires interval timers at both end systems, it increases the complexity of implementation, and could become expensive. In addition, the negative feedback mechanism could cause a collapse of the network under stressful traffic as pointed out in [12]. If the network is heavily congested, RM cells could be delayed or lost due to buffer overflow, with the result that timeliness of the congestion information is lost, or in a more serious situation, the source increases its cell emission rate due to the absence of RM cells in spite of the congestion.

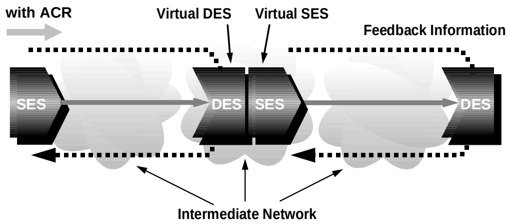

Figure 2: Network Segmentation by a Intermediate Network.

2.2 Counter-Based Approach

The timer-based approach described in the previous subsection is revised due to its drawbacks. In [13], a new scheme called Proportional Rate Control Algorithm (PRCA) is proposed with two major modifications: (1) the polarity of the feedback information is changed to be *positive*, and (2) the need for interval timers is eliminated. The origin of the name PRCA is in that opportunities for rate increases are given in proportion to the current sending rate ACR. In PRCA, the source end system marks the EFCI bit in data cells except for the first cell of every N_{RM} cells. The destination end system instantly returns back an RM cell to the source when it receives the EFCI bit cleared cell. If the EFCI bit is set by the intermediate switches due to its congestion, the destination takes no action. By this mechanism, a positive congestion signaling is established. In other words, receiving the RM cell implies that there is no congestion in the network, and therefore the source end system is given an opportunity to increase its rate.

The source end system sends the cells in the following way. Unless receiving the RM cell, the source determines the next cell transmission time at $1/ACR$ after the current time. It implies that the source continuously decreases its ACR until receiving the RM cell as

$$ACR \leftarrow \max(ACR - ADR, MCR) \tag{2}$$

When the source receives the RM cell, the rate is increased as

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

which compensates the reduced rate since the source received the previous RM cell ($N_{RM}\ ADR$), and increases the rate by $N_{RM}\ AIR$. In an ideal situation with no propagation delay, it should give linear increase of the cell transmission rate. By this rate changing method, the network can be restored even if all RM cells are discarded due to heavy congestion because the rate is always decreased except when the source receives the RM cell. The above operation accomplishes FECN like congestion management. However, if switches have the ability to discard RM cells in the backward direction, BECN like operation can also be achieved, which is one of the notable features of PRCA (Fig. 3).

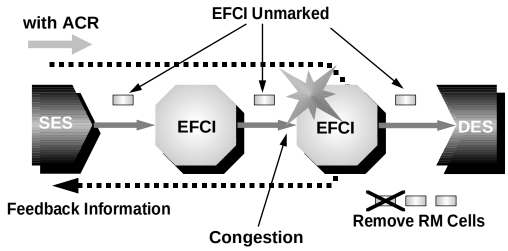

Figure 3: BECN Like Congestion Notification in PRCA.

However, certain problems still remain even in PRCA, which have been revealed in [14]. One of them is referred to as an "ACR beat down" problem, which is explained as follows. Each source of active connections that experiences congestion in several switches has less opportunity to receive positive feedback than other connections with fewer switches. Then, once it decreases its transmission rate to the minimum rate MCR, it is likely to remain at that rate indefinitely in some circumstances. Thus, fairness among connections cannot be achieved. Besides, the problem that PRCA requires a considerable amount of buffers in the existence of a large number of active connections is also pointed out. It is now widely perceived that when the propagation delays are large as in the WAN environment the queue length temporarily grows due to the control information delays. This is an intrinsic and unavoidable feature in recent high speed networks. However, the point above is that such a large queue length takes place even in the LAN environment. During congestion, the rate is decreased by ADR (Eq.(1)). However, when the number of connections is very large, it is too slow for the aggregated input rate to be decreased at the switch. For example, 241,000 cell buffers are required for assuring no cell loss even in the LAN environment for 1000 active connections [14] when the control parameters suggested in [13] are used. Such a large buffer size is unacceptable by the current memory technology.

2.3 Enhanced PRCA Method

Against the problems of PRCA, an improved version called EPRCA (Enhanced Proportional Rate Control Algorithm) is next proposed in [15, 16]. New functionalities are added as an implementation option in two ways. One is a capability to send a congestion indication to particular sources rather than all sources to achieve better fairness among connections, which is referred to as *intelligent marking*. The other is the means for explicitly reducing the rate of each connection, that is, the switch can have a responsibility to determine the cell transmission rate of selected connections. While some modifications are required to incorporate these new features into original PRCA, EPRCA preserves a backward compatibility with PRCA. Namely, a switch supporting only PRCA can also be employed in EPRCA. For distinguishing this switch from other new switches, it is called an EFCI bit setting switch.

EPRCA requires forward RM cells as well as backward RM cells. RM cells contain a CI (Congestion Indication) bit, which is used to notify congestion information to the source. (See Appendix C for the latest proposal for RM cell format.) The source end system periodically sends a forward RM cell every N_{RM} data cells instead of unmarking an EFCI bit of data cells as PRCA does. When the destination end system receives the forward RM cell, it returns back the RM cell to the source as a backward RM cell. In doing so, the CI bit of the backward RM cell is set following the EFCI status of the last incoming data cell. Therefore, the congestion detected at the intermediate switches can be notified by marking the EFCI bit of data cells in the forward path. By this mechanism, the source knows the congestion status of the network when it receives the backward RM cell. It is a FECN like implementation of the congestion notification. Further, the switch can be allowed to set a CI bit of backward RM cells as a BECN implementation.

Two major enhancements of EPRCA – *intelligent marking* and *explicit rate setting* – require additional information in each RM cell: CCR (Current Cell Rate) and ER (Explicit Rate) fields. An ER element is used to decrease the source rate explicitly, and is initially set to PCR by the source. One or more intermediate switches in congestion may change it to a lower value so that the rate of the source end system is rapidly decreased for quick congestion relief. The CCR element is set to the current ACR of the source in effect. A fair distribution of the bandwidth can be achieved with these two values. In congestion, the intermediate switch selectively signals a congestion indication to the sources with higher transmission rate. This intelligent marking mechanism can be used in conjunction with the explicit rate setting mechanism to achieve more impartial sharing of the available bandwidth.

Three types of switch architectures with different functions are suggested in the form of pseudo code in [16]; EFCI bit setting switches (EFCI), Binary Enhanced Switches (BES) and Explicit Down Switches (EDS). EFCI switches, which are already existing in the market, are just the same as switches supporting PRCA, and are expected to be least expensive.

In BES switches, two threshold values for indicating congestion are defined; QT and DQT. When a BES switch is congested, i.e., when the queue length in the cell buffer of a BES switch exceeds QT, the switch performs intelligent marking. It selectively reduces the rate of sources with larger ACR (Fig. 4). For this purpose, the switch maintains a control parameter $MACR$ (Mean ACR) that should ideally be the mean of the

ACR's of all active connections. When the rate of all connections is equal to $MACR$, the bandwidth is equally shared and the switch can be fully utilizing without falling into congestion. The key to realize this feature is how to obtain an accurate $MACR$. The BES switch updates its $MACR$ based on the CCR field of forward RM cells. For example, $MACR$ is calculated as

$$MACR \leftarrow MACR(1 - AV) + CCR \times AV$$

where AV is used as an averaging factor [16]. When the switch becomes congested, it indicates its congestion occurrence to the sources having higher rates. More closely, the switch marks the CI bit of the backward RM cells if its ACR value is beyond $MACR \times DPF$ (Down Pressure Factor) where a typical value of DPF is 7/8 for safety. However, the switch may still remain congested if only intelligent marking is employed. Therefore, when a BES switch becomes *very* congested such that the queue length is beyond DQT, all backward RM cells are marked irrespective of their ACR values. By this mechanism, relief from severe congestion can be accomplished. Note that it is evident from the above description that a BECN like quick congestion notification can be accomplished in BES switches.

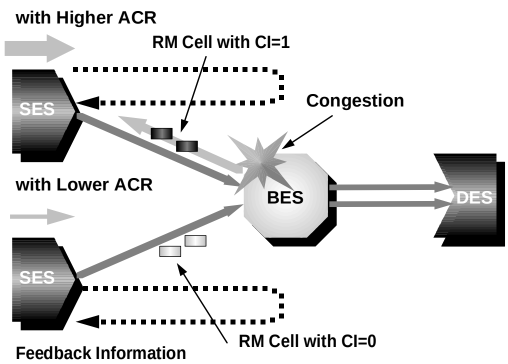

Figure 4: Intelligent Marking in BES Switch.

EDS switches are provided with an *explicit rate setting* capability in addition to intelligent marking (Fig. 5). These maintain $MACR$ as BES switches do, and control the transmission rate of a sources by setting the ER field of backward RM cells following a degree of congestion. When a backward RM cell with larger CCR than $MACR$ passes by the congested EDS switch, the value of its ER element is set to $MACR \times ERF$ (Explicit Reduction Factor). If the switch becomes very congested, $MACR \times MRF$ (Major Reduction Factor) is set in all backward RM cells to achieve quick congestion relief, which is called *major reduction*. Typical values of ERF and MRF shown in [16] are 7/8 and 1/4.

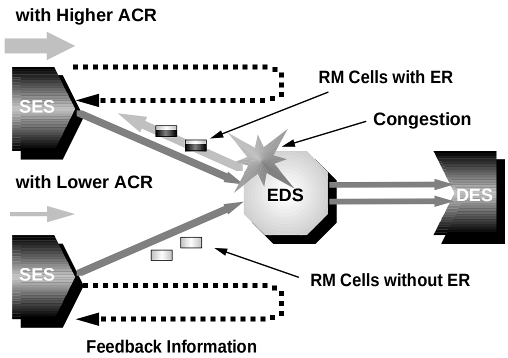

Figure 5: Intelligent Marking and Explicit Rate Setting in EDS Switch.

As described before, three types of switches, EFCI, BES and EDS switches, can co-exist in EPRCA. Furthermore, the operation of EFCI switches can be enhanced by BES or EDS switches located downstream. BES and EDS switches interpret the EFCI bit of forward data cells. When a BES or EDS switch is congested, entries in the VC table are marked following the EFCI status of forward RM cells, and the EFCI bit is cleared. Then, the CI bit of backward RM cells is set to notify the source of its congestion if its associated entry in the VC table is marked. This mechanism enables a BECN like quick congestion notification even if EFCI switches exist in the network. Namely, the enhanced switches can behave as virtual destination end systems to the EFCI switch. A quantitative evaluation of the above three types of switches will be presented in the next section.

The behavior of the source end system is simplified in [17], which will be included as an example source code in the standard. In this proposal, the source end system decreases its ACR by ACR/RDF every N_{RM} cells sent until reaching MCR. While this modification on EPRCA would degrade its performance to some extent, the complexity at the source end system can be considerably decreased.

2.4 Recent Proposals for Enhancements of EPRCA

In this subsection, several proposals which enhance the switch capabilities are introduced. As described before, since the standard does not specify the switch's behavior, the methods shown in this subsection will not be reflected in the standard. However, we believe that those methods would help understanding how the current EPRCA can be improved for effective congestion control.

2.4.1 Adaptive Proportional Rate Control

EPRCA has superiority over PRCA in fair share and high utilization of link bandwidths. However, it still contains a certain fault that fairness of EPRCA is not fulfilled in some configurations [18]. When a BES or EDS switch is very congested, it enforces all connections to decrease their rates equally but not selectively. Hence, when the switch is very congested for a rather long duration, *intelligent marking* does not work well. While it has been suggested to eliminate this operation in *very congested* states to avoid this problem, it results in an excessive queue length instead of preserving fairness [18].

Adaptive Proportional Rate Control (APRC) gives a solution to this problem [19]. Congestion in the switch is detected based on a change of its queue length in a fixed time interval rather than by a threshold value. If the queue length increases in N cell times, the switch is expected to fall into congestion. Then, each connection that has a higher rate than $MACR$ is enforced to a lower value by setting $MACR$ to the ER field of the backward RM cell. When the number of cells in the buffer exceeds DQT, the switch selectively set $MACR \times DPF$ to ER only for connections with higher rate. This modification improves the responsiveness to congestion, and therefore it can reduce the maximum queue size. The fairness among connections is also improved by this method. This *intelligent marking* capability of APRC was incorporated into EPRCA.

A new scheme called APRC2 with an improved *intelligent marking* capability is introduced in [20]. In EPRCA, ACR in the forward RM cell is used to compute $MACR$. However, at the switch, the effective rate of some connections that experience congestion at other switch may be quite different from the CCR values contained in the RM cells, which lead to a misbehavior of the explicit rate control. This problem can be avoided by virtue of UCR to establish stable operation [20]. In APRC2, UCR is defined as a mean of ACR's only for connections with larger ACR in addition to $MACR$. The value of UCR is updated as

$$UCR \leftarrow UCR + a(ACR - UCR)$$

only when ACR is greater than $MACR$. Then, UCR is used to effectively determine explicit rate ER.

Furthermore, it enhances operation of the source end system and the switch of EPRCA to shorten *ramp-up* time, which is the duration between the time when a connection begins/resumes its transmission and the time when the network settles into steady state. The ramp-up time is of importance because (1) since each connection starts its transmission with rate ICR regardless of the network status, it might cause a large queue buildup or under-utilization of the link unless ICR is set properly, and (2) most networks operate in a transient state since many connections are established but idle due to the bursty nature of the ABR traffic. Simulation results in [21] show improvements in ramp-up time, link utilization and maximum queue length when compared with previous schemes such as EPRCA and APRC.

2.4.2 EPRCA+ and EPRCA++ Methods

For implementing *explicit rate setting*, it is expected that the number of active connections is known by the switch. This is achieved by, for example, *per-VC accounting*, which can be implemented in several ways with additional hardware complexity. For instance, each switch may have a VC table to capture the number of active connections. Each VC entry is marked or unmarked according to the status of its corresponding VC (active or inactive), and the number of marked entries represents the number of active connections. In this way, rates of all sources are adjusted through RM cells in one round-trip time when there is one congested switch in the network. EPRCA+ proposed in [22] also employs this kind of scheme, and its simplified version can be found in [23].

In EPRCA+, congestion detection is performed by estimating the traffic load at the switch rather than by using a threshold value in the cell buffer. For this purpose, the switch should be provided with an interval timer, and the switch counts the number of cells received during this fixed time interval. The source end system is also equipped with a (more expensive) interval timer instead of a counter for sending RM cells. The rate of the source is kept unchanged until receiving a backward RM cell in which the explicit rate ER determined by the switch is contained. The older value of ACR is also carried in the CCR field of the backward RM cell. Then, the source changes its ACR following the relation between the current rate ACR and the ER value in the backward RM cell. If the older ACR (CCR) and the current ACR are not far from the ER value, i.e., if

$$CCR \leq ER \leq ACR$$

or

$$ACR \leq ER \leq CCR,$$

then ACR is unchanged. Otherwise, the value ER specified by the network is used for the new ACR.

One attractive feature of EPRCA+ is its small number of control parameters, which can easily be set by a network manager. Many control parameters required in EPRCA (see Appendix A) are eliminated in EPRCA+ in spite of its explicit rate setting capability. Further, in EPRCA+, the target utilization band (TUB), around which the switch is utilized, can be set freely. One may set the TUB of the switch under 100% link utilization, and then a smaller queue size and shorter cell delays at the switch can be achieved. Although an additional expense for timers and the VC table is required, EPRCA+ can provide better fairness and responsiveness than EPRCA [22].

Redundant complexities in the latest EPRCA are pointed out in [24]. For example, in the current proposal, an active source end system should decrease its rate by ACR/RDF for every N_{RM} cell times until reaching MCR. However, this excessive function makes it difficult to implement the source end system and to police the cell flow even though its necessity is not well justified, and it might be unnecessary under stable environments. A new scheme called EPRCA++, which is an enhancement of EPRCA+, is proposed in [25]. EPRCA++ adopts a counter at the source end system for forward RM cells instead of a timer as in EPRCA+. Furthermore, the source end system decreases its ACR only if no backward RM cell is received in $k \times N_{RM}$ cell times where k is set to a rather large value. By these modifications, EPRCA++ achieves higher performance even compared with EPRCA+ especially in transient state.

2.5 Fairness Definitions

As described so far, rate-based congestion control schemes have been improved to achieve higher link utilization and to require less buffer size at the switch as well as to obtain better fairness among connections. However, a clear definition for the fairness measure is necessary when concerned with the fairness degree of control methods. Since there is no absolute definition of the fairness measure in the literature, the ATM forum is also discussing fairness criteria. The fairness for the ABR service is defined by a max-min criteria in [26]. In the max-min criteria, all active connections are served *fairly* if the following two conditions are preserved [27]: (1) each connection must pass through at least one bottleneck switch along its path, and (2) it should be fairly shared when the available bandwidth is assigned to connections which do not experience the non-bottlenecked switch. However, the max-min criteria becomes inadequate for ABR traffic where MCR, the minimum usable rate for each connection, is incorporated. In what follows, we introduce four definitions of the fairness measure proposed in [28, 29].

1. MCR plus equal share

   The first definition is that each active source is allocated its MCR at first, then the rest of the available bandwidth is fairly assigned by the max-min criteria. That is, the nth active connection's rate B_n is given by

   $$B_n = MCR_n + \frac{C - \sum_{i=1}^{N_{VC}} MCR_i}{N_{VC}}, \quad 1 \le n \le N_{VC}, \tag{3}$$

   where C and N_{VC} are the available bandwidth and the number of active connections at the switch, respectively. Further, MCR_n represents MCR of the nth connection. It is known that this definition can be achieved not only by enhanced switches like BES or EDS but also by EFCI switches [28].

2. Maximum of MCR or Max-Min share

   In the second definition, each connection acquires bandwidth between MCR and the bandwidth equally divided by the number of connections:

   $$B_n = max(\frac{C}{N_{VC}}, MACR_n), \quad 1 \leq n \leq N_{VC}. \tag{4}$$

   This assignment needs an iteration for the sum of B_n's to be settled at the available bandwidth C, and the required number of iterations cannot be estimated. However, larger bandwidth can be allocated than in the previous case for connections with larger MCR.

3. Allocation proportional to MCR

   The third one assigns the available bandwidth to unconstrained connections in a weighted manner as;

   $$B_n = C \frac{MCR_n}{\sum_{i=1}^{N_{VC}} MCR_i}, \quad 1 \leq n \leq N_{VC}. \tag{5}$$

   This scheme cannot be applied to connections with $MCR = 0$. The bandwidth allocated to each connection is proportional to its MCR.

4. Weighted allocation

   The fourth one is a hybrid of the first and third definitions:

   $$B_n = MCR_n + F_n(C - \sum_{i=1}^{N_{VC}} MCR_i), \quad 1 \leq n \leq N_{VC}, \tag{6}$$

   where F_n is a weight for the nth connection defined as

   $$F_n = \frac{b}{N_{VC}} + (1-b)\frac{MCR_n}{\sum_{i=1}^{N_{VC}} MCR_i}, \quad 0 \leq b \leq 1.$$

   Note that it becomes identical to the first and third definitions with $b = 0$ and $b = 1$, respectively.

An appropriate definition of fairness depends on the environment to which it is applied. For example, requirements for charges may become an important factor in deciding which one is appropriate. In the public WAN environment, if the tariff is determined in proportion to MCR, the third definition would be suitable. On the other hand, in private networks, the first definition may be meaningful since it is less expensive in implementation. In the above definitions, however, the different values of PCR_n for connections are not taken into account. Since various speeds of interfaces are currently defined, PCR may also have to be incorporated.

3 Steady State Analysis of Enhanced PRCA

In this section, we analyze the performance of EPRCA in steady state with comparison among three types of suggested switches — EFCI, BES and EDS switches, which have been described in Subsection 2.3.

3.1 Analytic Model

First, we provide our analytical model with definitions of parameters. Our model is rather simple and consists of homogeneous traffic sources and a single bottleneck link (see Fig. 6). Nevertheless, it is important to explore the performance of rate-based congestion control method and to give a basis for comparisons with other proposed methods presented in the previous section. The active number of VC's that share the bottleneck switch is denoted by N_{VC}. We assume that these VC's behave identically, that is, all VC's the have same parameters, ICR, PCR, AIR and RDF (for these parameters, see Appendix A). The bandwidth of the bottleneck link is denoted by BW in cells/msec. We represent propagation delays from the source to the switch and from the switch to the destination by τ_{sx} and τ_{xd}, respectively. These parameters τ_{sx} and τ_{xd} are given according to the network configuration, i.e., LAN or WAN. The round-trip propagation delay from the source to the destination is defined by τ, and the following relation holds: $\tau = 2(\tau_{sx} + \tau_{xd})$. We further introduce $\tau_{xds}(= 2\tau_{xd} + \tau_{sx})$ which is the propagation delay of congestion indication from the switch to the source end system via the destination end system.

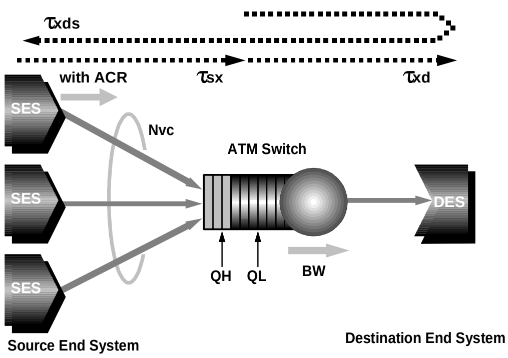

Figure 6: Analytic Model.

In all three types of switches, its congestion is detected by threshold values associated with the queue length at switch buffers. EFCI switches have high and low threshold values denoted as Q_H and Q_L, respectively. When the queue length at a switch exceeds Q_H, the switch detects its congestion and marks the EFCI bit in the header of data cells. On the other hand, it is regarded that congestion terminates when the queue length goes under Q_L. BES and EDS switches have another threshold value DQT to indicate *very* congested status and take distinct actions against congestion as has been described in Subsection 2.3.

In EPRCA, the source end system sends RM cells proportionally to its rate. Furthermore, the rate change is performed by backward RM cells returned from the destination end system, and the received rate of the backward RM cells is bounded by $BW/(N_{VC}\,N_{RM})$ when the switch is congested. Otherwise, it is identical to the transmission rate of the source end system. Therefore, we require a different analytical treatment from the one presented in [7] in which the rate change is performed on a timer basis.

Let us introduce $ACR(t)$ and $Q(t)$, which represent the cell transmission rate ACR of each source end system and the queue length at the switch observed at time t, respectively. In what follows, evolutions of $ACR(t)$ and $Q(t)$ in steady state are analyzed assuming that (1) the switch has infinite capacity of the buffer, and that (2) the source end system always has cells to be sent. Therefore, $ACR(t)$ becomes equivalent to CCR (Current Cell Rate), which is the actual cell transmission rate.

Furthermore, we will point out drawbacks of the EDS switch and present its improved version, called a PEDS (Prioritized EDS) switch, which has a capability to provide simple priority control for RM cells over data cells. As will be shown later, the maximum queue length can be decreased by the PEDS switch when compared with other types of switches.

3.2 EFCI Switch

In this subsection, we focus on the EFCI switch to analyze a dynamical behavior of $ACR(t)$ and $Q(t)$. In our analysis, forward RM cells are not considered explicitly. Hence, this model is equivalent to PRCA in Subsection 2.2. However, forward RM cells can easily be taken into account by replacing BW in our analysis with BW', which is defined as

$$BW' = BW\frac{N_{RM}}{N_{RM}+1}. \tag{7}$$

In numerical examples, BW' will be used to compare with other switches which require forward RM cells.

3.2.1 Determination of $ACR(t)$

Figure 7 shows a pictorial view of $ACR(t)$ and $Q(t)$ which have a periodicity in steady state. An initial point of one cycle is defined at the time when the congestion indication is received at the source end system. In the EFCI switch, it takes τ_{xds} for the congestion indication to reach the source end system after the queue length becomes Q_H at the switch. We divide one cycle into four phases following behaviors of $ACR(t)$ and $Q(t)$ as depicted in Fig. 7. For simplicity of presentation, we introduce $ACR_i(t)$ and $Q_i(t)$ as

$$\begin{aligned} ACR_i(t) &= ACR(t - t_{i-1}), \quad 0 \leq t < t_i, \\ Q_i(t) &= Q(t - t_{i-1}), \quad 0 \leq t < t_i, \end{aligned}$$

where t_i is defined as the time when phase i terminates. Further, the length of phase i is represented by

$$t_{i-1,i} = t_i - t_{i-1}.$$

We note here that a more strict treatment is required for representing system behaviors dependent on system parameters as will be presented in the next subsection. For a meanwhile, however, we assume that the system behaves as in Fig. 7. $ACR_i(t)$ $(1 \leq i \leq 4)$ can be determined as follows.

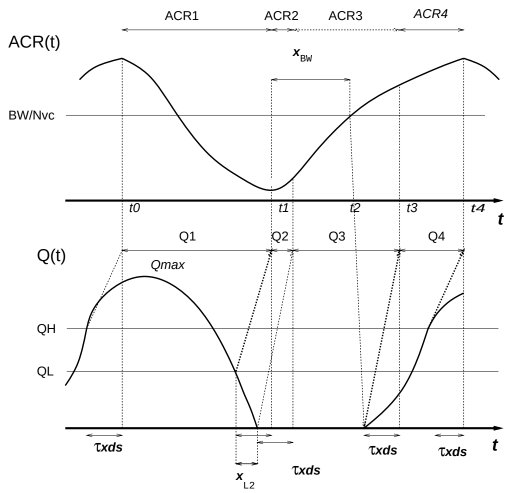

Figure 7: Pictorial View of $ACR(t)$ and $Q(t)$ in EFCI Switch.

- Phase 1: $ACR_1(t)$

  At time $t = 0$, a congestion indication from the switch is received at the source end system, and Phase 1 starts with an initial value $ACR_1(0)$. During this phase, the next cell emission time is determined by an inverse of the current $ACR_1(t)$, and the new cell emission rate is determined by subtracting ADR from $ACR_1(t)$, where ADR equals to $ACR_1(0)/RD$. A differential equation for $ACR_1(t)$ is then obtained as

$$\frac{dACR_1(t)}{dt} = -\frac{ACR_1(0)}{RD} ACR_1(t),$$

which gives

$$ACR_1(t) = ACR_1(0) e^{-\frac{ACR_1(0)}{RD} t}. \tag{8}$$

Note that MCR is not considered in the above equation, i.e., $MCR = 0$ is assumed in our analysis.

- Phase 2: $ACR_2(t)$

  In this phase, the source end system receives the backward RM cells at a constant rate since the switch is fully utilized, that is, $0 < Q(t - \tau_{xds})$. By letting the interarrival time of the backward RM cells be x, we have

$$\frac{1}{x} = \frac{BW}{N_{VC}\, N_{RM}}.$$

  Actually, ACR is increased only when the backward RM cell is received, and otherwise decreased continuously like Phase 1. However, we consider the values of ACR at the time when ACR is increased. Then, we derive an envelop of ACR in Phase 2 as follows.

$$\begin{aligned}
\frac{dACR_2(t)}{dt} &= -\frac{ACR_2(t)ADR + N_{RM}\, ADR + N_{RM}\, AIR}{x} \\
&= -\frac{N_{RM}}{BW/N_{VC}} ACR_2(t) \frac{ACR_2(t)}{RD} + N_{RM} \frac{ACR_2(t)}{RD} + N_{RM} AIR \\
&= -\frac{ACR_2(t)^2}{RD} + \frac{BW}{RD\, N_{VC}} ACR_2(t) + \frac{BW}{N_{VC}} AIR
\end{aligned}$$

  By solving the above equation for $ACR_2(t)$, we obtain

$$ACR_2(t) = \frac{a_1 e^{-a_1 t} + a_2 r e^{-a_2 t}}{c_1 (e^{-a_1 t} + r e^{-a_2 t})}, \tag{9}$$

  where a_1 and a_2 are roots of the equation

$$a^2 + c_2 a + c_1 c_3 = 0,$$

  and c_1, c_2 and c_3 are given by

$$c_1 = -\frac{1}{RD}; \qquad c_2 = \frac{BW}{RD\, N_{VC}}; \qquad c_3 = \frac{BW\, AIR}{N_{VC}}.$$

  The initial transmission rate $ACR_2(0)$ determines r in Eq.(9) as

$$r = \frac{a_1 - c_1\, ACR_2(0)}{c_1\, ACR_2(0) - a_2}.$$

- Phase 3: $ACR_3(t)$

  Phase 3 continues during τ_{xds} after the queue length at the switch becomes 0. In this phase, the RM cell arrives at the source end system depending on its own rate $ACR_3(t-\tau)$. By letting the interarrival time of two successive backward RM cells be x, a differential equation for $ACR_3(t)$ should satisfy the following equation.

$$\frac{dACR_3(t)}{dt} = \frac{N_{RM}\,AIR + N_{RM}\frac{ACR_3(t)}{RD} + ACR_3(t)\{1 - e^{-\frac{ACR_3(t)}{RD}x}\}}{x}.$$

  However, it is difficult to solve the above equation, then we approximately use the following equation by neglecting the second and third terms in the numerator.

$$\frac{dACR_3(t)}{dt} \cong \frac{N_{RM}\,AIR}{x}$$

  Recalling that ACR is decreased during not receiving RM cells, the interarrival times of two successive RM cells x satisfies,

$$\int_0^x ACR_3(t-\tau)e^{-\frac{ACR_3(t-\tau)}{RD}y}dy = N_{RM},$$

  which leads to

$$x = \frac{RD\log(\frac{RD}{RD-N_{RM}})}{ACR_3(t-\tau)}.$$

  Finally, $ACR_3(t)$ is solved as

$$ACR_3(t) \cong ACR_3(0)e^{\beta t}, \tag{10}$$

  where β is given as a root of the equation

$$\beta = \frac{N_{RM}\,AIR}{RD\,\log(\frac{RD}{RD-N_{RM}})}e^{-\tau\beta}.$$

- Phase 4: $ACR_4(t)$

  $ACR_4(t)$ is given in the equivalent form to $ACR_2(t)$ since the receiving rate of RM cells at the source end system is just same as in the case of Phase 2,

3.2.2 Evolution of $ACR(t)$ and $Q(t)$

In what follows, we present the evolution of $ACR(t)$ and $Q(t)$. For this purpose, we should determine the initial values $ACR_i(0)$ and the length of each phase $t_{i,i+1}$. Given initial rates in Phase 1, $Q_1(t)$ is obtained as

$$Q_1(t) = Q_1(\tau_{xds}) + \int_{x=\tau_{xds}}^{t} (N_{VC}\,ACR_1(x-\tau_{sx}) - BW)dx. \tag{11}$$

The length of Phase 1 $t_{12}(= t_1)$ is given as

$$t_{12} = x_{L1} + \tau_{xds},$$

where x_{L1} is obtained by solving the equation $Q_1(x_{L1}) = Q_L$. In what follows, we will use a convention $x_{L1} = Q_1^{-1}(Q_L)$ for brevity.

For Phase 2 and later phases, we need a careful treatment since some of phases do not appear dependent on the parameters. First, let us introduce x_{L2} as

$$x_{L2} = Q_2^{-1}(0) - Q_2^{-1}(Q_L).$$

i.e., x_{L2} is the time for the queue length to reach 0 after it goes below Q_L. Further,

$$x_{BW} = ACR_2^{-1}(\frac{BW}{N_{VC}}) - t_1,$$

which defines the time when the aggregate ACR reaches BW. We should consider the following four cases depending on x_{L2}, x_{BW} and τ_{sx}. See Fig. 7 which corresponds to Case 1 in the below.

Case 1: $x_{L2} \leq \tau, x_{L2} < x_{BW} + \tau_{sx}$

Case 2: $x_{L2} \leq \tau, x_{L2} \geq x_{BW} + \tau_{sx}$

Case 3: $x_{L2} > \tau, x_{L2} < x_{BW} + \tau_{sx}$

Case 4: $x_{L2} > \tau, x_{L2} \geq x_{BW} + \tau_{sx}$

Due to lack of space, only Case 1 and Case 4 are explained. In Case 1, we have

$$\begin{aligned} t_2 &= t_1 + x_{L2} \\ Q(t) &= 0, \quad t_2 + \tau_{sx} < t \leq t_3 + \tau_{sx}. \end{aligned}$$

In the above equation, t_3 is given by

$$t_3 = t_2 + x'_{BW} + \tau,$$

where x'_{BW} is the time when the aggregate ACR reaches BW, that is,

$$x'_{BW} = ACR_3^{-1}(\frac{BW}{N_{VC}}) - t_2.$$

Further,

$$Q(t) = \begin{cases} 0, & t_2 + \tau_{sx} < t \leq t_2 + \tau_{sx} + x'_{BW} \\ \int_{t_2+\tau_{sx}+x'_{BW}}^{t} (N_{VC}\, ACR_3(x - \tau_{sx}) - BW)dx, & t_2 + \tau_{sx} + x'_{BW} < t \leq t_3 + \tau_{sx} \end{cases}.$$

Finally, we obtain equations for Phase 4 as

$$\begin{aligned} t_4 &= t_3 + t_{H4} + \tau_{xds} \\ Q(t) &= \int_{t_3+\tau_{sx}}^{t} (N_{VC}\, ACR_4(x-\tau_{sx}) - BW)dx, \quad t_3+\tau_{sx} < t \leq t_4+\tau_{sx}, \end{aligned}$$

where t_{H4} is given by

$$t_{H4} = Q_4^{-1}(Q_H).$$

In Case 4, the queue length never reaches 0. Namely, the switch is always fully utilized. Since neither Phase 2 nor 3 appears in this case, we have

$$t_4 = Q_4^{-1}(Q_H) + \tau_{xds}.$$

$Q_4(t)$ is then obtained from Eq.(11).

Finally, by setting

$$\begin{aligned} ACR_1(0) &= ACR_4(t_4) \\ Q_1(0) &= Q_4(t_4+\tau_{xds}), \end{aligned}$$

and calculating the above equations iteratively, we can obtain the dynamical behavior of EPRCA with the EFCI switch in steady state.

3.2.3 Parameter Tuning

In this subsection, two suggestive results are presented for parameter tuning: a maximum queue length in steady state and a condition that the link is never under-utilized. The former is required for dimensioning the appropriate buffer size of the switch with cell-loss free. The latter is necessary to obtain high throughput.

Maximum Queue Length

To obtain the maximum queue length, we first obtain a maximum rate of $ACR_4(t)$ by letting $t \to \infty$.

$$\begin{aligned} \lim_{t\to\infty} ACR_4(t) &= \left.\frac{a_1 e^{(a_2-a_1)t} + a_2 r}{c_1(e^{(a_2-a_1)t}+r)}\right|_{t\to\infty} \\ &= \frac{BW + \sqrt{BW^2 + 4N_{VC}\, RD\, BW\, AIR}}{2N_{VC}}. \end{aligned} \tag{12}$$

Noting that $ACR_4(t)$ may terminate before reaching its maximum value dependent on system parameters, we can obtain an upper bound of the maximum queue length from this equation as follows.

Suppose now that the source end system receives congestion indication at time $t = 0$, and that its ACR equals to its maximum value, i.e.,

$$ACR_1(0) = ACR_4(\infty).$$

$Q_1(t)$ starts at time $t = \tau_{sx}$ with its initial value $Q(\tau_{sx})$ given as

$$\begin{aligned} Q(\tau_{sx}) &= Q_H + \int_0^{\tau_{xds}+\tau_{sx}} (N_{VC}\, ACR_4(\infty) - BW)dt \\ &= Q_H + (N_{VC}\, ACR_4(\infty) - BW)\tau. \end{aligned}$$

The queue length begins to decrease when the aggregate cell arrival rate at the switch is below the link bandwidth BW at time t_{max}, which is given by

$$N_{VC}\, ACR_4(\infty) e^{-\frac{ACR_4(\infty)}{RD} t_{max}} = BW,$$

By solving the above equation for t_{max}, we have

$$t_{max} = \frac{RD}{ACR_4(\infty)} \log \frac{N_{VC}\, ACR_4(\infty)}{BW}.$$

Then, from Eq.(11), we have the maximum queue length Q_{max} as

$$\begin{aligned} Q_{max} &= Q_1(t_{max}) \\ &= Q_H + (N_{VC}\, ACR_4(\infty) - BW)\tau + N_{VC} RD(1 - \frac{BW}{N_{VC}}) \\ &\quad - \frac{BW\, RD}{ACR_4(\infty)} \log \frac{N_{VC}\, ACR_4(\infty)}{BW}, \end{aligned}$$

where $ACR_4(\infty)$ is given by Eq.(12). We can observe from the above equation that the number of VC's has a serious impact on the maximum queue length especially when the propagation delay between the source and destination end system is large.

Conditions for Avoiding Under-Utilization

As previously noted, a full link utilization is accomplished by the following conditions, which corresponds to Case 4 in the previous subsection.

$$\begin{aligned} x_{L2} &> \tau, \\ x_{BW} + \tau_{sx} &< x_{L2}. \end{aligned}$$

3.2.4 Numerical Examples

In this subsection, we provide numerical examples for the EFCI switch. Threshold values Q_H and Q_L are identically set to 500, and the bandwidth of bottleneck link is set to 353.208 cells/msec assuming 156Mbps ATM link. For other control parameters, the values suggested in [16] (see also Appendix A) are used throughout this paper unless other values are explicitly specified. Eq.(7) is used for BW, that is, the overhead for forward RM cells are taken into account to compare with other switches.

Figures 8 and 9 show the effect of the number of VC's, N_{VC}, on $ACR(t)$ and $Q(t)$, respectively. We choose the propagation delay $\tau_{sx} = \tau_{xd} = 0.005$ msec (2km between source and destination) as a typical value for a LAN environment. In the figures, it is easily observed that the maximum queue length becomes large and the cycle is lengthened as N_{VC} increases. However, the maximum queue length can be limited to an acceptable value in the LAN environment when the number of VC's is fairly small.

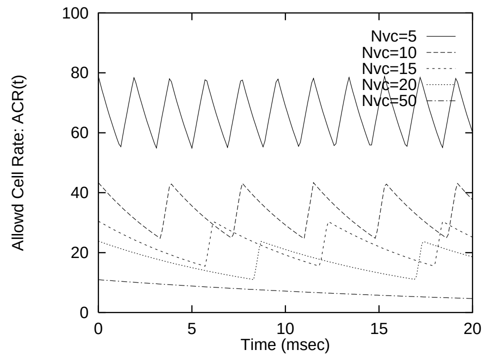

Figure 8: Effect of N_{VC} on $ACR(t)$ in EFCI Switch ($\tau = 0.02$msec).

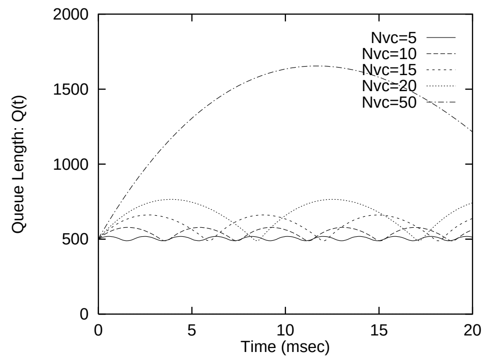

Figure 9: Effect of N_{VC} on $Q(t)$ in EFCI Switch ($\tau = 0.02$msec).

$ACR(t)$ and $Q(t)$ for different values of propagation delays are next compared in Figs. 10 and 11. As can be observed in the figures, the larger τ causes slower congestion notification, and it results in increase of the maximum queue length. Further, the under-utilization appears when the propagation delays τ is beyond 2.0 msec (about 400km). Therefore, from these numerical results, we may conclude that the EFCI switch should be used in rather small networks.

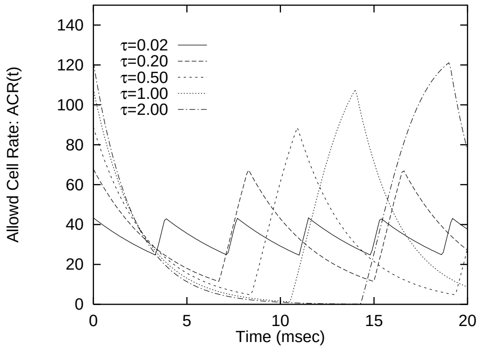

Figure 10: Effect of Propagation Delay on $ACR(t)$ in EFCI Switch ($N_{VC} = 10$).

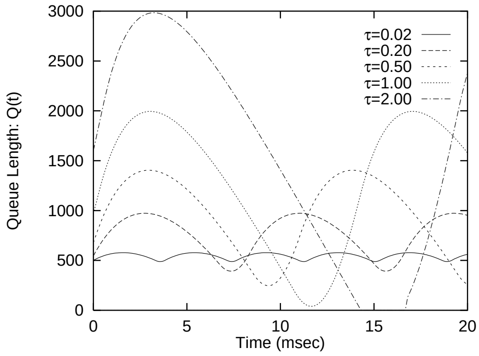

Figure 11: Effect of Propagation Delay on $Q(t)$ in EFCI Switch ($N_{VC} = 10$).

Another problem of the EFCI switch can be found by larger N_{VC} as has been described in Subsection 2.2. Figure 12 shows that the maximum queue length grows almost linearly. As will be shown in the following subsections, the maximum queue length can be decreased considerably by introducing BES or EDS switches. However, when we consider an initial transient state, that is, when all connections start its cell transmission at same time, the growth of the maximum queue length becomes unacceptable even with BES or EDS switch as has been described in Subsection 2.3. This is extensively investigated in [30]. In [30], the effect obtained by introducing the PEDS switch is also demonstrated.

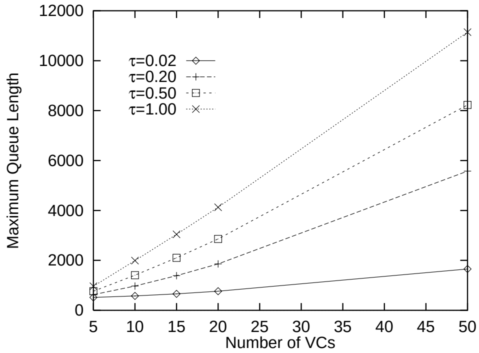

Figure 12: Effect of N_{VC} on Maximum Queue Length in EFCI Switch.

3.3 BES Switch

3.3.1 Analysis

As described in Subsection 2.3, the BES switch informs its congestion occurrence to the source end system by setting the CI bit in the backward RM cell. While the main purpose of the BES switch is to achieve fairness among connections by intelligent marking, the above mechanism enables a faster congestion signaling to the source than the EFCI switch. In our current model, the analysis of the EFCI switch obtained in Subsection 3.2 can directly be applied to the BES switch (1) by replacing the parameter τ_{xds} with τ_{sx}, and (2) by assuming that CCR used for calculating $MACR$ is not an outdated value. By this assumption, CCR is likely to be equal to $MACR$. Noting that the forward RM cells are required in the BES switch, BW should be replaced by BW' given in Eq. (7) in the following analysis.

With the same approach as in the EFCI switch, the maximum queue length for the BES switch is obtained as

$$\begin{aligned} Q_{max} &= Q_H + (N_{VC}\, ACR_4(\infty) - BW')\tau + N_{VC}\, RD(1 - BW'/N_{VC}) \\ &\quad - \frac{BW'\, RD}{ACR_4(\infty)} \log\left(N_{VC}\, ACR_4(\infty)/BW'\right). \end{aligned}$$

3.3.2 Numerical Examples

Figures 13 and 14 show the effect of Nvc on $ACR(t)$ and $Q(t)$ for $\tau_{sx} = \tau_{xd} = 0.005$. The effects of the propagation delays are also shown in Figs. 15 and 16 for $N_{VC} = 10$. To see the effect of the BECN like capability of the BES switch more clearly, the distance of the switch from the source (τ_{sx}) is changed in Figs. 17 and 18. In those figures, $\tau_{sx} + \tau_{xd}$ is fixed at 1.0 and N_{VC} is 10. We can easily observe the effect of the BECN like quick congestion notification of the BES switch. The maximum queue length can be reduced from 3,000 to 2,000 in the case of $\tau = 2.0$ (see Fig. 11).

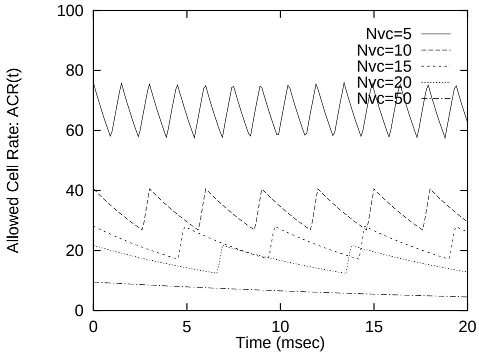

Figure 13: Effect of N_{VC} on $ACR(t)$ in BES Switch ($\tau = 0.02$msec).

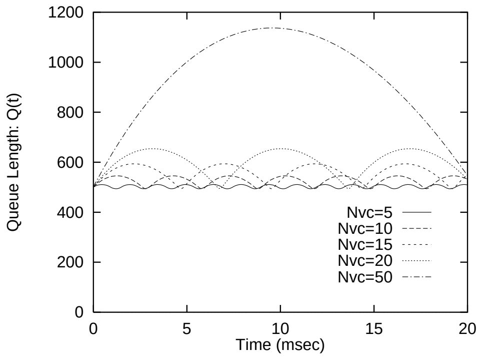

Figure 14: Effect of N_{VC} on $Q(t)$ in BES Switch ($\tau = 0.02$msec).

3.4 EDS and PEDS Switches

3.4.1 Analysis

The EDS switch has an explicit rate reduction mechanism. When the queue length becomes Q_H, the switch forces sources with larger ACR to gradually decrease its rate using parameter ERF. If the queue length still grows and reaches DQT, the major rate reduction is performed. In this case, ACR's of all sources are reduced by using parameter MRF. As shown in Appendix A, typical values for ERF and MRF are 15/16 and 1/4, respectively.

When the queue length becomes beyond Q_H and still below DQT, $MACR$ is used to determine the sources of which ACR is larger than other sources. This selection capability is used for achieving fairness among connections. However, in our analysis, we

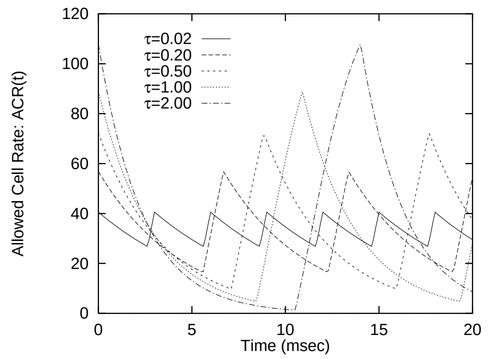

Figure 15: Effect of Propagation Delay on $ACR(t)$ in BES Switch ($N_{VC} = 10$).

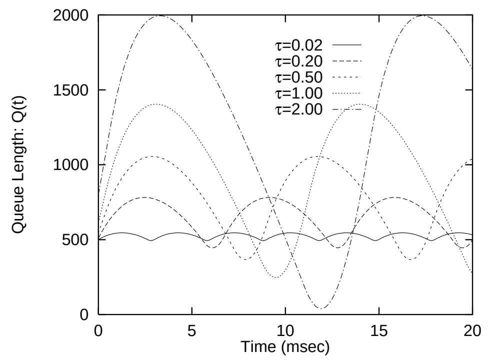

Figure 16: Effect of Propagation Delay on $Q(t)$ in BES Switch ($N_{VC} = 10$).

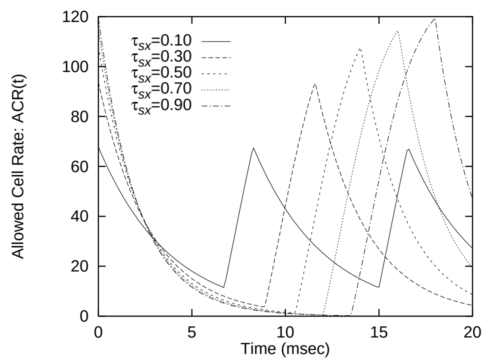

Figure 17: Effect of Switch Location on $ACR(t)$ in BES Switch ($N_{VC} = 10, \tau_{sx} + \tau_{xd} = 1$).

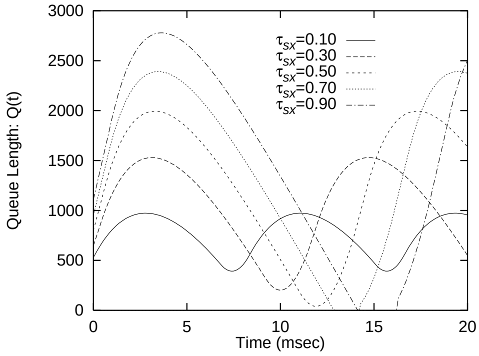

Figure 18: Effect of Switch Location on $Q(t)$ in BES Switch ($N_{VC} = 10, \tau_{sx} + \tau_{xd} = 1$).

assume that all connections behave identically. Therefore, when the queue length exceeds Q_H, all source end systems are enforced to reduce its ACR if $MACR$ is calculated properly according to the algorithm, which is our main assumption in this section. Let $ACR'(t)$ be ACR recognized by the EDS switch at time t. We further introduce $ACR_i'(t)$ which is $ACR'(t)$ in Phase i. According to the above assumption, we have

$$ACR'(t) = ACR_i(t - Q(t)/BW - \tau_{sx}), \tag{13}$$

$$ACR_i'(t) = ACR'(t - t_i), \quad t_{i-1} \le t < t_i. \tag{14}$$

In the EDS switch, $MACR$ is used to determine new ACR for each source end system when the switch is congested. $MACR$ should remain fixed in an ideal situation. However, as indicated in the above equations, $ACR'(t)$ contains $Q(t)$ because the switch recognizes ACR of the sources when it processes the forward RM cells. In congestion, the queue length tends to become large, which shows that ACR contained in the RM cell becomes too old to be used for estimating appropriate $MACR$. Therefore, we introduce the prioritized EDS (PEDS) switch in which the RM cells are processed with high priority over the data cells. By this mechanism, Eq. (14) becomes

$$ACR'(t) = ACR_i(t - \tau_{sx}). \tag{15}$$

In what follows, analytic results only for the PEDS switch are presented. Further, due to a lack of space, we show the analysis only for the case where the network is fully utilized. However, other cases can also be treated in a similar manner as in Subsection 3.2.

We divide one cycle into three phases following behaviors of $ACR(t)$ and $Q(t)$ (see Fig. 3.4.1). First, $ACR_i(t)$ is determined. During Phase 1, the rate is gradually decreased with parameter ERF. A differential equation for $ACR_1(t)$ can be written as

$$\frac{dACR_1(t)}{dt} = -\frac{MACR(1-ERF)}{x} ACR_1(t),$$

where x is a fixed interval between two consecutive RM cells and given by

$$\frac{1}{x} = \frac{BW/N_{VC}}{N_{RM}}.$$

Therefore,

$$ACR_1(t) = ACR_1(0)e^{-\frac{N_{VC}\, MACR\,(1-ERF)}{BW\, N_{RM}} t}.$$

$ACR_2(t)$ and $ACR_3(t)$ can be obtained by Eqs. (8) and (9), respectively.

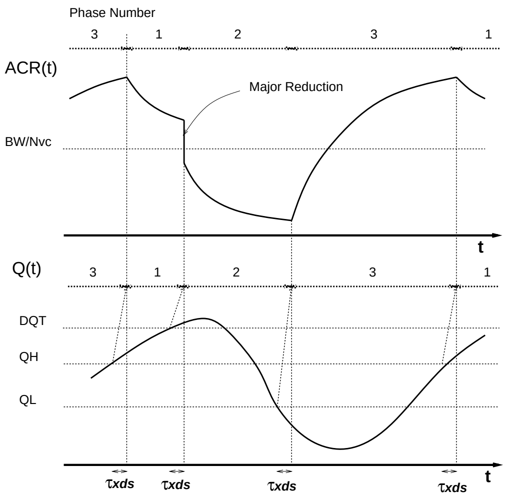

Figure 19: Pictorial View of EDS Switch with Major Reduction.

To consider $Q(t)$, let us introduce t_{DQT}, t_{Q_H} and x_1 defined as

$$\begin{aligned} t_{DQT} &= Q_3^{-1}(DQT), \\ t_{Q_H} &= Q_3^{-1}(Q_H), \\ x_1 &= t_{DQT} - t_{Q_H}, \end{aligned}$$

where x_1 is the time duration that queue length reaches DQT since it is beyond Q_H. Based on the relation between x_1 and τ, the following three cases should be considered.

- Case 1: $x_1 < \tau$ (Fig. 3.4.1)

  The length of Phase 1 is given by

$$t_{0,1} = x_1 + \tau_{xds}.$$

Then, initial values for $ACR_2(t)$ and $Q_2(t)$ in Phase 2 are derived as

$$\begin{aligned} ACR_2(0) &= MRF\, ACR_1(t_1 - \tau_{sx}), \\ Q_2(0) &= Q_1(t_1 + \tau_{sx}). \end{aligned}$$

The length of Phase 2 is obtained as

$$t_{1,2} = x_2 + \tau_{xds},$$

where $x_2 = Q_2^{-1}(Q_L)$. In the same way, initial values for $ACR_3(t)$ and $Q_3(t)$ and the length of Phase 3 are obtained as

$$\begin{aligned} ACR_3(0) &= ACR_2(t_2), \\ Q_3(0) &= Q_2(t_2 + \tau_{sx}), \\ t_{2,3} &= x_3 + \tau, \end{aligned}$$

where $x_3 = Q_3^{-1}(Q_H)$.

- Case 2: $x_1 > \tau$ and $max(Q_1(t)) \geq DQT$

  The length of Phase 1 is given by

  $$\begin{aligned} t_{0,1} &= \tau_{xds} + t_{DQT} + \tau_{xds}, \\ t_{DQT} &= Q_1^{-1}(DQT). \end{aligned}$$

  Accordingly, the initial values and the length of Phase 2 are derived as

  $$\begin{aligned} ACR_2(0) &= MRF\, ACR_1(t_1 - \tau_{sx}), \\ Q_2(0) &= Q_1(t_1 + \tau_{sx}), \\ t_{1,2} &= x_2 + \tau_{xds}, \end{aligned}$$

  where $x_2 = Q_2^{-1}(Q_L)$. Phase 3 is equivalent to that of Case 1.

- Case 3: $x_1 > \tau$ and $max(Q_1(t)) < DQT$

  The behavior of this case is obtained by eliminating Phase 2 of Case 1.

3.4.2 Maximum Queue Length

The maximum queue length for the EDS switch can be obtained by the same method as in the case of the EFCI switch in Subsection 3.2.

- Cases 1 and 2: $max(Q_1(t)) > DQT$

  In this case, the queue length reaches its maximum value τ_{sx} after when major reduction occurs at $t = t_1$. Hence, the maximum queue length can be obtained as

  $$Q_{max} = Q(t_1 + \tau_{sx}).$$

- Case 3: $max(Q_1(t)) < DQT$

  In this case, no major reduction occurs. It means that the maximum queue length Q_{max} can be obtained with the same method in the case of the EFCI switch in Subsection 3.2, that is,

$$Q_{max} = Q_1(t_{max}).$$

  where t_{max} is given as

$$t_{max} = \frac{BW\,N_{RM}}{N_{VC}\,MACR(1 - ERF)} \log \frac{N_{VC}\,ACR_3(\infty)}{BW}.$$

3.4.3 Numerical Examples

In Figs. 20 and 21, the effect of N_{VC} on $ACR(t)$ and $Q(t)$ in the PEDS switch are illustrated for $\tau = 0.02$. The effects of propagation delays are depicted in Figs. 22 and 23 for $N_{VC} = 10$. In obtaining these figures, DQT is set to 500 and both of Q_H and Q_L are set to 100. In all cases, the effect of major rate reduction is apparent because the queue length can drastically be decreased when compared with other EFCI and BES switches. However, we notice that the link is not fully utilized when τ becomes large (Fig. 23). This is due to inappropriate control parameter settings $ERF(= 15/16)$ and $MRF(= 1/4)$. We need set control parameters carefully to achieve better performance in the EDS switch. For example, the case of $ERF = 1/2$ is shown in Figs. 24 and 25. In those figures, we can see that the maximum queue length is limited under 500 and that the network is fully utilized.

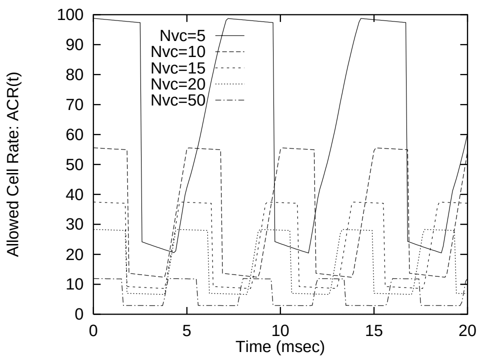

Figure 20: Effect of N_{VC} on $ACR(t)$ in PEDS Switch ($\tau = 0.02$msec, $ERF = 15/16$).

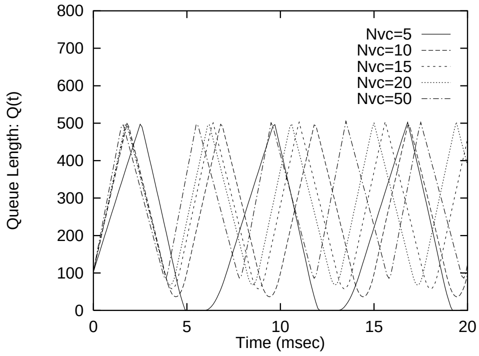

Figure 21: Effect of N_{VC} on $Q(t)$ in PEDS Switch ($\tau = 0.02$msec, $ERF = 15/16$).

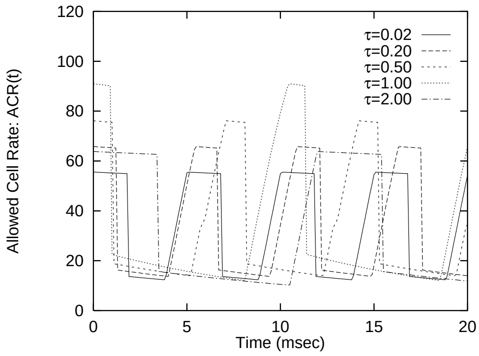

Figure 22: Effect of Propagation Delay on $ACR(t)$ in PEDS Switch ($N_{VC} = 10$).

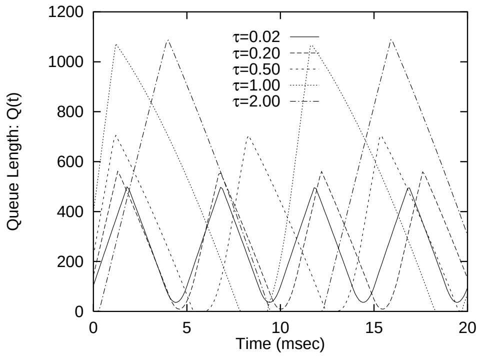

Figure 23: Effect of Propagation Delay on $Q(t)$ in PEDS Switch ($N_{VC} = 10$).

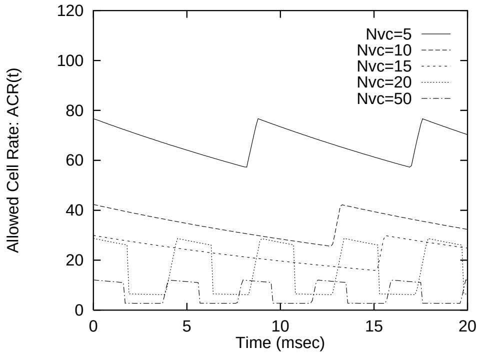

Figure 24: Effect of N_{VC} on $ACR(t)$ in PEDS Switch ($\tau = 0.02$msec, $ERF = 1/2$).

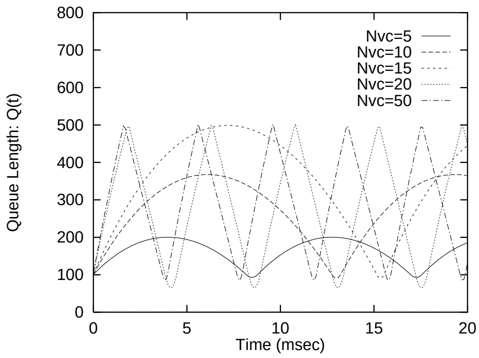

Figure 25: Effect of N_{VC} on $Q(t)$ in PEDS Switch ($\tau = 0.02$msec, $ERF = 1/2$).

4 Initial Transient State Analysis of Enhanced PRCA

When one or more connections start their cell transmission at once, queue length at each switch grows considerably unless control parameters such as ICR (Initial Cell Rate) are set properly. In this section, we analyze the EPRCA in initial transient state using the same analytic model presented in Subsection 3.1.

4.1 EFCI Switch

In this subsection, we analyze the initial transient behavior of the allowed cell rate $ACR(t)$ and the queue length $Q(t)$ of the EFCI switch for a given ICR to show that an appropriate choice of ICR plays an important role for achieving effective control while supressing the maximum queue length to an appropriate value. The latter is important for buffer-dimensioning in designing ATM switches.

4.1.1 Derivation of $ACR(t)$

Figures 26 and 27 show pictorial views of $ACR(t)$ and $Q(t)$ of the EFCI switch in the initial transient state for different ICR's. As illustrated, the evolution of $ACR(t)$ and $Q(t)$ is classified into two categories according to the following relations.

$$N_{VC}\, ICR > BW \tag{16}$$

or

$$N_{VC}\, ICR < BW \tag{17}$$

To see this, we will derive $ACR(t)$ and $Q(t)$ from now on. Here, we should note that a more rigorous treatment is required regarding the above classification because it should also be affected by other parameters as will be shown in the below.

We first divide $ACR(t)$ into phases, each of which has a different form dependent on the congestion status of the switch. There are three types of $ACR(t)$.

Type I: $ACR_d(t)$

$ACR(t)$ is decreased exponentially.

Type II: $ACR_{iu}(t)$

$ACR(t)$ is increased and the offered load to the link is below its capacity BW.

Type III: $ACR_{io}(t)$

$ACR(t)$ is increased and the offered load to the link is beyond its capacity BW.

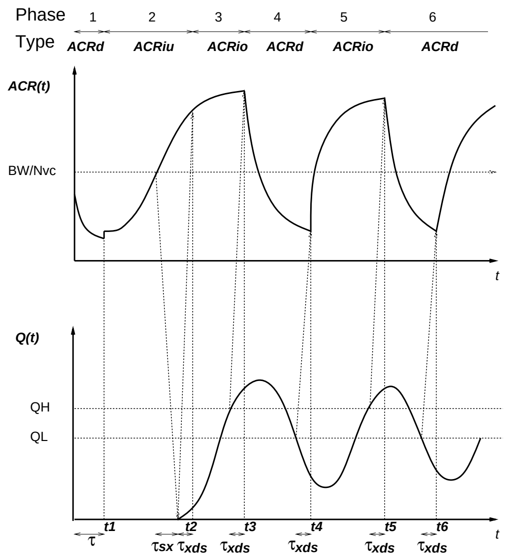

Figure 26: Pictorial View of EFCI Switch ($ICR < BW/N_{VC}$).

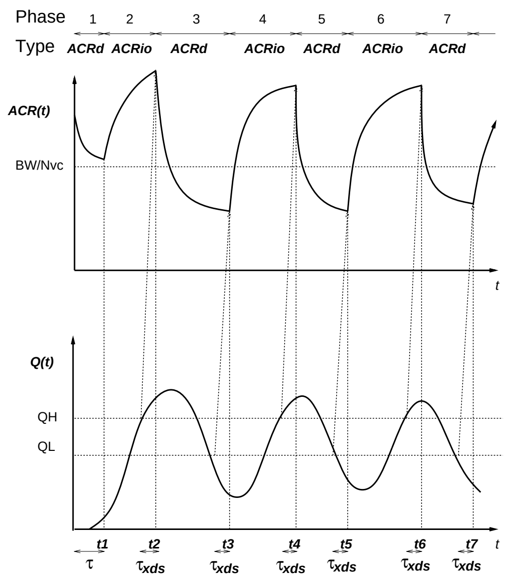

Figure 27: Pictorial View of EFCI Switch ($ICR > BW/N_{VC}$).

In the case of $N_{VC}\,ICR > BW$, $ACR_d(t)$, $ACR_{iu}(t)$, $ACR_{io}(t)$ appears in that order (Fig. 26). Then, $ACR_d(t)$ and $ACR_{io}(t)$ are repeated. On the other hand, a cycle consisting of $ACR_d(t)$ and ACR_{io} is repeated in the case of $N_{VC}\,ICR < BW$ (Fig. 27). Let us denote $ACR_i(t)$ and the corresponding $Q_i(t)$ as $ACR(t)$ and $Q(t)$ of Phase i, respectively, that is,

$$\begin{aligned} ACR_i(t) &= ACR(t - t_{i-1}), \quad 0 \le t < t_i, &(18)\\ Q_i(t) &= Q(t - t_{i-1}), \quad 0 \le t < t_i, &(19) \end{aligned}$$

where t_i is the time when Phase i terminates. Further, the length of Phase i is defined by

$$t_{i-1,i} = t_i - t_{i-1}. \qquad (20)$$

Each source end system transmits an RM cell followed by data cells at time 0. Until it receives the first RM cell, its ACR is decreased exponentially. At time τ_{sx}, the RM cells arrive at the switch. At this time, the switch is not congested, and the RM cells will be returned to the source. In actual, it takes one cell time for the switch to handle each RM cell. Then, it may be queued up during processing RM cells when the number of VC's is large. However, we assume that all RM cells from sources are transferred simultaneously by the switch. Since the queue length is zero at time τ_{sx}, all sources will increase ACR at time $\tau_{sx} + \tau_{xds} = \tau$ by receiving the first backward RM cell. The next phase is determined by the relation between ICR and BW/N_{VC}. Type II ($ACR_{io}(t)$) appears when $N_{VC}ICR < BW$ (Fig. 26), and Type III if $N_{VC}ICR < BW$ (Fig. 27). Note that the derivation of $ACR(t)$ and $Q(t)$ has already been shown in Subsection 3.2, that is, $ACR_d(t)$, $ACR_{io}(t)$ and $ACR_{uo}(t)$ correspond to $ACR_1(t)$, $ACR_2(t)$ and $ACR_3(t)$, respectively. Then, evolution of $ACR(t)$ and $Q(t)$ is shown by determining the length and the initial value of each phase.

4.1.2 Evolution of $ACR(t)$ and $Q(t)$

In this subsection, $ACR(t)$ and $Q(t)$ are derived by taking the same approach presented in Subsection 3.2.

- $ICR < BW/N_{VC}$

  In this case, ACR is decreased until the first RM cell is returned to the source end system at $t = \tau$. Therefore,

$$\begin{aligned} t_1 &= \tau &(21)\\ ACR_1(t) &= ACR_d(t), \quad 0 \le t \le t_1, &(22)\\ Q_1(t) &= 0, \quad 0 \le t \le t_1 + \tau_{sx}, &(23) \end{aligned}$$

where the initial value of $ACR_1(t)$ is ICR. ACR is then increased until reaching at the value BW/N_{VC}.

$$ACR_2(t) = ACR_{iu}(t), \quad 0 \leq t \leq t_{12}, \tag{24}$$
$$Q_2(t) = 0, \quad \tau_{sx} \leq t \leq t_{12} + \tau_{sx}. \tag{25}$$

The time t_{12} is given by as

$$t_{12} = ACR_2^{-1}(BW/N_{VC}) + \tau \tag{26}$$

During Phase 3, the queue length grows and RM cells are returned with a fixed interval $N_{VC}N_{RM}/BW$ and ACR is increased according to $ACR_{io}(t)$.

$$ACR_3(t) = ACR_{io}(t), \quad 0 \leq t \leq t_{23} \tag{27}$$
$$Q_3(t) = \int_{x=\tau_{sx}}^{t} (N_{VC}\, ACR_3(x - \tau_{sx}) - BW)dx, \quad \tau_{sx} \leq t \leq t_{23} + \tau_{sx} \tag{28}$$

where t_{23} is a solution of

$$Q_3(t_{23}) = Q_H. \tag{29}$$

In what follows, we will use the convention $t_{23} = Q_3^{-1}(Q_H)$ for brevity. After the queue length reaches the threshold value Q_H, ACR is again decreased. ACR is then increased when Q becomes the lower threshold value Q_L.

- $ICR > BW/N_{VC}$

  In this case, ACR is decreased until the first RM cell is returned to the source end system at $t = \tau$ as in the above case, i.e.,

$$t_1 = \tau \tag{30}$$
$$ACR_1(t) = ACR_d(t), \quad 0 \leq t \leq t_1, \tag{31}$$
$$Q_1(t) = \int_{x=tsx}^{t} (N_{VC}\, ACR_1(x - \tau_{sx}) - BW)dx, \quad \tau_{sx} \leq t \leq t_1 + \tau_{sx} \tag{32}$$

  If $ACR_1(\tau)$ is still beyond BW/N_{VC}, ACR is increased and the RM cells are returned with a fixed interval $N_{VC}\, N_{RM}/BW$ during Phase 2.

$$ACR_2(t) = ACR_{io}(t), \quad 0 \leq t \leq t_{12}, \tag{33}$$
$$Q_2(t) = Q_1(t_1 + \tau_{sx}) + \int_{x=tsx}^{t} (N_{VC}\, ACR_2(x - \tau_{sx}) - BW)dx, \quad \tau_{sx} \leq t \leq t_{12} + \tau_{sx} \tag{34}$$

  where t_{12} is a solution of

$$t_{12} = Q_2^{-1}(Q_H) + \tau_{xds} \tag{35}$$

  Otherwise, Phase 2 according to $ACR_{iu}(t)$ begins as in the above case. When the queue length reaches the high threshold value Q_H, ACR is again decreased.

Last, we note that for later phases, the steady state analysis presented in Subsection 3.2 can be applied.

Maximum Queue Length

In this subsection, we show the maximum queue length for two cases.

- $ICR < BW/N_{VC}$

  As shown in Fig. 26, $Q(t)$ reaches at its maximum value during Phase 4. Let the queue length take its maximum value Q_{max} at t_{max}. Then, we have a relation

$$Q_{max} = Q^{-1}(t_{max}) \tag{36}$$

$$t_{max} = ACR^{-1}(BW/N_{VC}) + \tau_{sx} \tag{37}$$

- $ICR > BW/N_{VC}$

  As shown in Fig. 27, $Q(t)$ reaches at its maximum value during Phase 3 as

$$Q_{max} = Q^{-1}(t_{max}), \tag{38}$$

$$t_{max} = ACR^{-1}(BW/N_{VC}) + \tau_{sx}. \tag{39}$$

4.1.3 Numerical Examples

In this subsection, some numerical examples for the EFCI switch are provided. In these examples, both Q_H and Q_L are identically set to 500, and other control parameters are set to the suggested values shown in [16] (see Appendix A).

The effects of the propagation delays on $ACR(t)$ and $Q(t)$ are displayed in Figs. 28 and 29 for $N_{VC} = 10$. Since a larger value of τ causes slower congestion notification, the queue length is built up initially. Therefore, it is hard to directly apply EFCI switches in the case where we interconnect LANs located in the long distance. After then, the queue length is cyclically fluctuated. For example, the maximum queue length is 2,000 in the case of $\tau = 1.0$ in steady state (see Subsection 3.2) while the queue length becomes about 1,240 initially.

Figs. 30 and 31 show $ACR(t)$ and $Q(t)$, respectively, for different values of N_{VC}. The propagation delay between the source and destination end systems are set to be 0.05 msec (around 2km) as a typical value for a LAN environment. It is obvious that the large N_{VC} causes an increase of the maximum queue length even in the case of short propagation delays. In Fig. 30, $ACR(t)$ is growing in spite of the switch congestion when $N_{VC} = 50$. It is due to the form of $ACR(t)$ given by Eq. (9). Therefore, the queue length becomes large instead of congestion relief.

A possible solution for decreasing the maximum queue length is to set ICR properly. In Fig 32, the different values of ICR are used in the case where $N_{VC} = 50$ and $\tau = 0.02$. As can be seen in the figure, appropriate ICR can decrease the maximum queue length to some extent. However, it requires to know the active number of connections, N_{VC}, in advance. If it is difficult to estimate it, we need (1) faster congestion

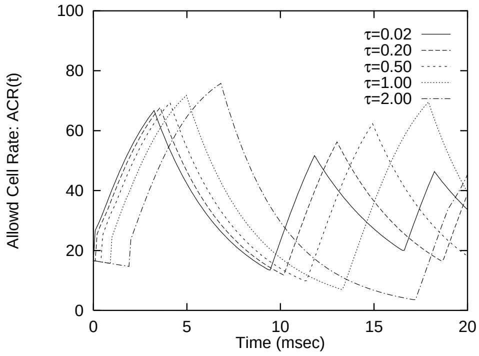

Figure 28: Effect of Propagation Delay on $ACR(t)$ in EFCI Switch ($N_{VC} = 10$).

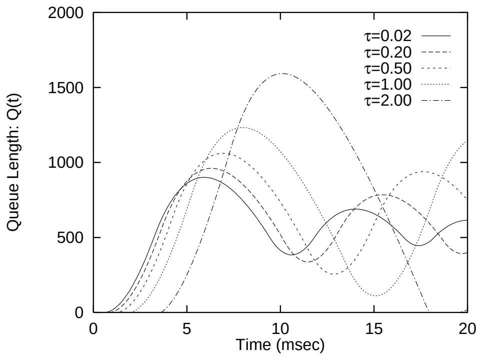

Figure 29: Effect of Propagation Delay on $Q(t)$ in EFCI Switch ($N_{VC} = 10$).

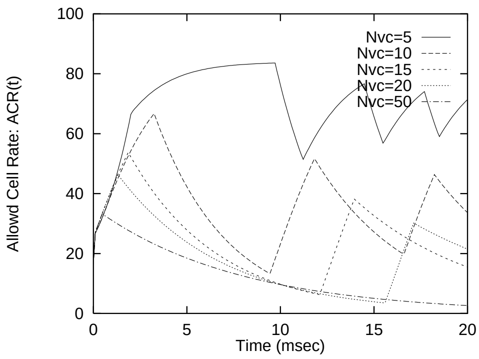

Figure 30: Effect of N_{VC} on $ACR(t)$ in EFCI Switch ($\tau = 0.02$ msec).

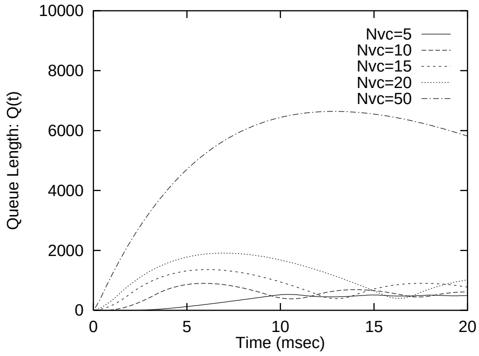

Figure 31: Effect of N_{VC} on $Q(t)$ in EFCI Switch ($\tau = 0.02$ msec).

notification achieved by the BES switch which will be presented in the next section, or (2) explicit rate reduction by EDS switch in Section 4.3.

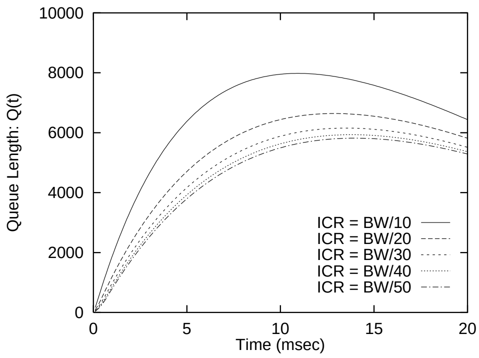

Figure 32: Effect of Initial Transmission Rate on $Q(t)$ in EFCI Switch ($N_{VC} = 50$).

4.2 BES Switch

4.2.1 Analysis

In our current model, the analysis for EFCI switches obtained in Section 4.1 can be directly applied to the BES switch by replacing the parameter τ_{xds} with τ_{sx}. In the BES switch, when the queue length is between Q_H and DQT, the congestion notification is selectively performed using $MACR$. Therefore, if $MACR$ is started with a small value at time 0, congestion is notified to the source end system if the queue length exceeds Q_H according to the control algorithm. However, in the current study, we want to estimate the upper bound of maximum queue length in the initial transient state. Thus, we assume that $MACR$ is started with a rather large value so that the source end system is not notified of congestion even when the queue length becomes Q_H. This situation takes place when a number of connections simultaneously starts cell transmission after a single connection has been occupied the link. To avoid this, some mechanism might be required for the BES switch so that if the link is not used during a some period, $MACR$ is reset to its initial value IMR.

4.2.2 Numerical Examples

To see the effect of fast congestion notification of the BES switch, the different values of the propagation delays are used to illustrate Figs. 33 and 34. Here, the propagation delays τ_{sx} (from the source end system to the switch) and τ_{xd} (from the switch to the destination end system) are set to be identical. The initial value of $MACR$, IMR, is set to BW. As described in the previous subsection, the threshold values Q_H is not meaningful in the current analysis, and DQT is set to be 500 which is identical to Q_H of the EFCI switch. By comparing with Figs. 28 and 29 for the EFCI switch, superiority of the BES switch is obvious. For example, the maximum queue length can be decreased from 1700 to 1200 when $\tau = 2.0$msec.

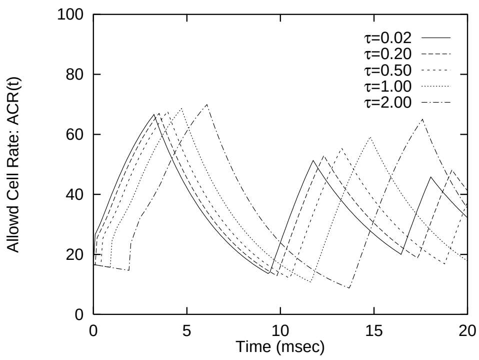

Figure 33: Effect of Propagation Delay on $ACR(t)$ in BES Switch ($N_{VC} = 10$).

The location of BES switch affects its performance. Fig. 35 shows evolutions of $Q(t)$ for different τ_{sx}'s while the number of VC's and the distance between the source and destination end systems are fixed ($N_{VC} = 10$ and $\tau_{sx} + \tau_{xd} = 1.0$). This result indicates that as the switch is located near the source, the maximum queue length is decreased because of faster congestion notification to the source. However, the BES switch is helpless regarding the large VC's problem. The maximum queue length cannot be decreased even if the BES switch is introduced when τ_{sx} is large relative to the end-to-end propagation delay.

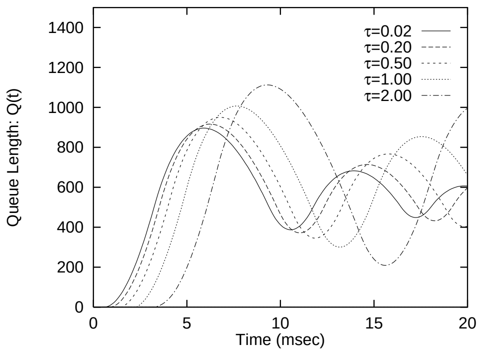

Figure 34: Effect of Propagation Delay on $Q(t)$ in BES Switch ($N_{VC} = 10$).

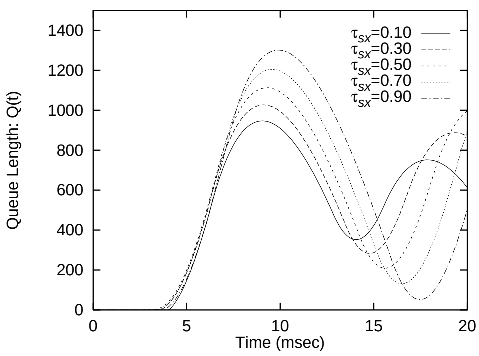

Figure 35: Effect of Switch Location on $Q(t)$ in BES Switch ($N_{VC} = 10$).

4.3 EDS Switch

4.3.1 Assumptions

In addition to the assumptions introduced in Section 3.1, the following assumptions are made in analyzing the EDS switch.

1. The number of VC's is large such that $N_{VC} > BW/ICR$. The condition $N_{VC} \leq BW/ICR$ is ideal in the sense that the number of cells initially accumulated at the switch buffer can be eliminated, as has already been shown in EFCI switches. Therefore, we only consider the former condition to investigate the worse case.

2. Each connection sends a first RM cell at $t = 0$. Then, it is followed by data cells. Let t'_n be the time when the switch processes n-th forward RM cell sent by the source and carries out EDS operations, i.e., an update of MACR.

$$t'_n = (n-1)\frac{N_{VC}(N_{RM}+1)}{BW} + \tau_{sx}. \tag{40}$$

   Here, we assume that RM cells from all sources are processed at time t'_n while, in actual, it takes one cell time to handle each RM cell. If the queue length is above Q_H, the estimation using MACR is performed according to the algorithm. As in the case of the BES switch, MACR is assumed to begin with a large value. Further, we introduce t_n which is the time when the source end system possibly receives n-th RM cell for major rate reduction. Such a case takes place when the switch revives the backward RM cell and the queue length exceeds DQT.

$$t_n = (n-1)\frac{N_{VC}(N_{RM}+1)}{BW} + \tau, \tag{41}$$

   where t_0 is defined as 0.

3. The rate reduction to $ER = MACR \times ERF)$ does not take place because an initial value of MACR is large by our assumption.

4.3.2 EDS Switch without Priority Control

In this subsection, we treat the EDS switch without priority control in which the RM cells and data cells are emitted in a FIFO order at the switch.

Before deriving $ACR(t)$ and $Q(t)$, we introduce two additional functions; $ACR'(t)$ and $MACR(t)$. Let $ACR'(t)$ be ACR recognized at the switch. This is required because $ACR(t)$ is the rate of the source at time t, but due to the queueing delay at the switch, ACR which the switch recognizes at time t is not $ACR(t - \tau_{sx})$. Actually, $ACR'(t)$ is given as;

$$ACR'(t) = ACR\left(\frac{BW}{N_{VC}(N_{RM}+1)}(t - \tau_{sx})\right). \tag{42}$$

Inversely, ACR of the source end system at time t is recognized at the switch at time

$$\frac{N_{VC}(N_{RM}+1)}{BW}(t+\tau_{sx}). \tag{43}$$

At time t'_n, if the queue length is above Q_H, $MACR(t)$ begins to be updated using ACR contained in the RM cell. We assume that RM cells from the number N_{VC} of sources are processed at the switch at same time, i.e., the switch calculates MACR through ACR's contained in RM cells from N_{VC} sources immediately without delay. This assumption gives $MACR(t)$ as

$$\begin{aligned} MACR(t'_n) &= ACR'(t'_n)\times AV\sum_{i=0}^{N_{VC}-1}(1-AV)^i+(1-AV)^{N_{VC}}MACR(t'_n-) \\ &= ACR'(t'_n)\times(1-(1-AV)^{N_{VC}-1})+(1-AV)^{N_{VC}}MACR(t'_n-), \end{aligned} \tag{44}$$

where AV is an average factor (see Appendix). Note that $MACR(t'_n)$ becomes identical to $ACR'(t'_n)$ when $N_{VC}=\infty$. We have $MACR_n(t)$, which is defined as $MACR(t)$ for $t_n-\tau_{sx}-2\tau_{xd}\le t<t_{n+1}-\tau_{sx}-2\tau_{xd}$ as

$$\begin{aligned} MACR_1(t) &= IMR, \qquad 0\le t<t_2-\tau_{sx}-2\tau_{xd}, \\ MACR_n(t) &= ACR'(t_n-\tau_{sx}-2\tau_{xd})\times(1-(1-AV)^{N_{VC}-1}), \\ &\qquad +(1-AV)^{N_{VC}}MACR(t_n-\tau_{sx}-2\tau_{xd}) \\ &\qquad n\ge 2,\quad t_n-\tau_{sx}-2\tau_{xd}\le t<t_{n+1}-\tau_{sx}-2\tau_{xd}. \end{aligned} \tag{45}$$

Using the above functions, we now obtain $ACR(t)$ and $Q(t)$. Let $ACR_n(t)$ be $ACR(t)$ for $t_n\le t\le t_{n+1}(n\ge 0)$. Then, we have (Fig.3.4.1)

$$ACR_n(t)=ACR_n(0)e^{-\frac{ACR_n(0)}{RD}(t-t_n)},\quad t_n\le t<t_{n+1} \tag{46}$$

where the initial values $ACR_n(0)$ are given as

$$ACR_0(0) = ICR, \tag{47}$$

$$ACR_1(0) = \begin{cases} MACR_1(t_n-\tau_{sx})\times MRF, & \text{if } Q_0(t_1-\tau_{sx})\ge DQT \\ ACR(2\tau_{sx}+\tau_{xd})+N_{RM}(AIR+ADR), & \text{otherwise} \end{cases} \tag{48}$$

$$ACR_n(0) = \begin{cases} MACR_n(t_n-\tau_{sx})\times MRF, & \text{if } Q_{n-1}(t_n-\tau_{sx})\ge DQT \\ ACR_{n-1}(t_n), & \text{otherwise} \end{cases} \tag{49}$$

Here, we should take account of possibilities of major rate reduction.

The corresponding queue length is determined as

$$Q_0(t) = \int_{\tau_{sx}}^{t}\left\{\frac{N_{RM}+1}{N_{RM}}N_{VC}\times ACR_0(x-\tau_{sx})-BW\right\}dx,\quad \tau_{sx}\le t<t_1+\tau_{sx}, \tag{50}$$

$$\begin{aligned} Q_n(t) &= Q_{n-1}(t_n+\tau_{sx})+\int_{t_n+\tau_{sx}}^{t}\left\{\frac{N_{RM}+1}{N_{RM}}N_{VC}\times ACR_n(x-\tau_{sx})-BW\right\}dx, \\ &\qquad t_n+\tau_{sx}\le t<t_{n+1}+\tau_{sx}. \end{aligned} \tag{51}$$

In the below, we will derive an formula for the queue length just before the first major reduction, which is likely to happen at time t_2 in the case of the large number of N_{VC}. First, the queue length at time $t_1 + \tau_{sx}$ is given by

$$\begin{aligned} Q_0(t_1+\tau_{sx}) &= \int_{\tau_{sx}}^{t_1+\tau_{sx}} \left\{ \frac{N_{RM}+1}{N_{RM}} N_{VC}\, ACR_0(t-\tau_{sx}) - BW \right\} dt \\ &= \frac{N_{RM}+1}{N_{RM}} N_{VC}\, RD\,(1-e^{-\frac{ICR}{RD}\tau}) - BW\,\tau \end{aligned} \tag{52}$$

Then, the queue length at time $t_2 + \tau_{sx}$ is given by

$$\begin{aligned} & Q_1(t_2+\tau_{sx}) \\ &\quad = Q_0(t_1+\tau_{sx}) + \int_{t_1+\tau_{sx}}^{t_2+\tau_{sx}} \left\{ \frac{N_{RM}+1}{N_{RM}} N_{VC}\, ACR_1(t-\tau_{sx}) - BW \right\} dt \\ &\quad = Q_0(t_1+\tau_{sx}) + \frac{N_{RM}+1}{N_{RM}} N_{VC}\, RD\,(1-e^{-\frac{ACR_1(0)}{MD}\frac{N_{VC}(N_{RM}+1)}{BW}}) - N_{VC}(N_{RM}+1), \end{aligned} \tag{53}$$

where $ACR_1(0)$ was given in eq. (48).

Here, we have assumed that the queue length at time $t_2 - \tau_{sx} - 2\tau_{xd}$ is beyond Q_H, and that the queue length at time $t_2 - \tau_{sx}$ exceeds DQT. Due to the lack of space, we omit the condition that this case takes place, but it is easily obtained. The condition that the queue length does not increase after t_2 is given by

$$N_{VC} \times ACR_2(0) \le BW$$

where

$$\begin{aligned} ACR_2(0) &= MRF \times MACR_2(t_2-\tau_{sx}) \\ &\cong MRF \times ACR'(t_2-\tau_{sx}-2\tau_{xd}) \\ &= MRF \times ICR \times e^{-\frac{ICR}{RD}\left[1+\frac{BW\tau_{sx}}{N_{VC}(N_{RM}+1)}\right]}. \end{aligned} \tag{54}$$

As will be shown in the numerical examples, the above queue length does not always give the maximum value at the initial transient state. However, in eq.(53), we have a free design parameter MRF, and it can prevent the queue length to unacceptably become large.

4.3.3 EDS Switch with Priority Control

In this subsection, the prioritized EDS switch is evaluated where the RM cells are given preferential service at the switch via Head-of-Line priority discipline, i.e., the RM cells are always served if any at the switch buffer. The additional assumptions for the analysis to be tractable are described below.

1. The bandwidth available to the data cells is given by subtracting the bandwidth for high priority RM cells from the total bandwidth, i.e.,

$$BW \frac{N_{RM}}{N_{RM}+1}.$$

2. $ACR'(t)$ equals to $ACR(t)$ at time $t-\tau_{sx}$ because of preferential service to RM cells, that is, the RM cell is processed immediately on its arrival at the switch.

$$ACR'(t) = ACR(t-\tau_{sx}) \tag{55}$$

   Therefore, we neglect the delay due to contention among RM cells.

3. n-th backward RM cell arrives at the switch at time $t_n = s_n + \tau$ where s_n satisfies

$$\int_0^{s_n} ACR(t)dt = (n-1)\, N_{RM}. \tag{56}$$

   That is, we need not take account of queueing delay of RM cells due to its priority service.

The definition of $ACR(t)$ is just same as in the case of the EDS switch without priority control. The queue length at the switch is then obtained as

$$\begin{aligned} Q_0(t) &= \int_{\tau_{sx}}^{t} \left\{ N_{VC} \times ACR_0(x-\tau_{sx}) - BW \frac{N_{RM}+1}{N_{RM}} \right\} dx, \quad \tau_{sx} \le t < t_1 + \tau_{sx} \qquad (57)\\ Q_n(t) &= Q_{n-1}(t_n+\tau_{sx}) + \int_{t_n+\tau_{sx}}^{t} \left\{ N_{VC} \times ACR_n(x-\tau_{sx}) - BW \frac{N_{RM}+1}{N_{RM}} \right\} dx, \\ & \qquad n \ge 1, \quad t_n + \tau_{sx} \le t < t_{n+1} + \tau_{sx}. \qquad (58) \end{aligned}$$

If the first major reduction occurs at time t_2, the queue length at time $t_2 + \tau_{sx}$ is given as

$$\begin{aligned} Q_1(t_2+\tau_{sx}) &= Q_0(t_1+\tau_{sx}) + \int_{t_1+\tau_{sx}}^{t_2+\tau_{sx}} \left\{ N_{VC} \times ACR_1(t-\tau_{sx}) - BW \frac{N_{RM}+1}{N_{RM}} \right\} dt \\ &= Q_0(t_1+\tau_{sx}) + RD\, N_{VC}(1 - e^{-\frac{ACR_1(0)\, t_2}{RD}}) - BW\,(t_2-t_1) \frac{N_{RM}}{1+N_{RM}} \qquad (59) \end{aligned}$$

where

$$\begin{aligned} Q_0(t_1+\tau_{sx}) &= \int_{\tau_{sx}}^{t_1+\tau_{sx}} \left\{ N_{VC} \times ACR_0(t-\tau_{sx}) - BW \frac{N_{RM}+1}{N_{RM}} \right\} dt \\ &= RD\, N_{VC}(1 - e^{-\frac{ICR\, t_1}{RD}}) - BW\, t_1 \frac{N_{RM}}{1+N_{RM}}. \qquad (60) \end{aligned}$$

In the above equation, t_2 is obtained as follows. When $s_2 \le \tau$, the following inequality holds.

$$\int_0^{2\tau_{sx}+\tau_{xd}} ACR_0(t) < N_{RM} + 1.$$

Then s_2 is obtained from the following equation,

$$\int_0^{s_2} ACR_0(t) = RD(1 - e^{-\frac{ICR}{RD}t}). \tag{61}$$

That is,

$$s_2 = \frac{RD}{ICR} \log \frac{RD}{RD - (N_{RM} + 1)}, \tag{62}$$

which determines t_2 from eq. (56).

On the other hand, when $s_2 > \tau$, s_2 is given by

$$s_2 = -\frac{RD}{ICR} \log \left\{ \frac{N_{RM} + 1}{RD} - (1 - e^{\frac{ICR}{RD}\tau} + e^{\frac{ACR_1(0)}{RD}\tau}) \right\}. \tag{63}$$

If ACR is not decreased enough to sustain the queue length, the queue length would still grow. In that case, a smaller value of MRF should be used to decrease the maximum queue length.

4.3.4 Numerical Examples

In this subsection, we provide numerical examples for EDS and PEDS switches. The suggested control parameters in Appendix are used except DQT, Q_H and Q_L. DQT is set to 500, and both Q_H and Q_L are set to 100 for comparison purposes with EFCI and BES switches.

Initial evolutions of $ACR(t)$ and $Q(t)$ are plotted in Figs. 36 and 37, respectively. In these figures, N_{VC} is set to 50 and τ is varied. The case of the EFCI switch is also shown for comparison purposes. The result of the BES switch is not shown because it is almost same as the one of the EFCI switch. It is remarkable that giving priority to RM cells in the PEDS switch can drastically decrease the maximum queue length.

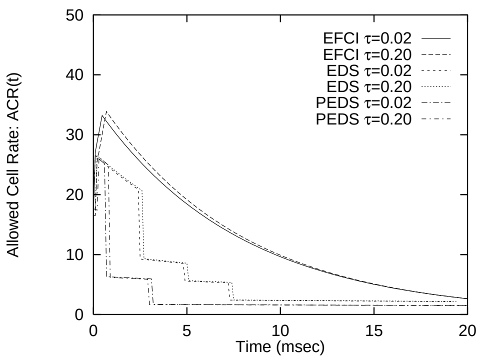

Figure 36: Initial Evolution of $ACR(t)$ ($N_{VC} = 50$).

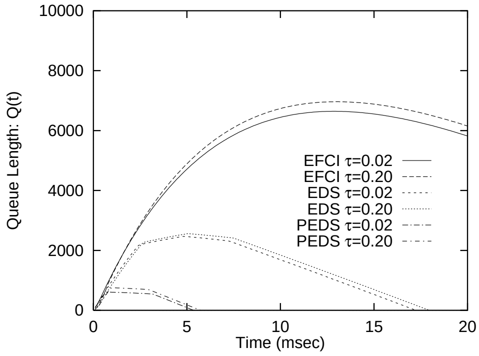

Figure 37: Initial Evolution of $Q(t)$ ($N_{VC} = 50$).

In Fig. 38, the maximum queue length in the EDS switch for different values of N_{VC}'s are plotted. The increase of N_{VC} affects the maximum queue length directly while the propagation delay has little influence on it. The number of connections, N_{VC}, is still a major dominant factor for the maximum queue length even in the case of the EDS switch. On the contrary, the influence of N_{VC} on the maximum queue length becomes less in the PEDS switch. For example, the required buffer for PEDS switch becomes less than 3000 cells even if $\tau = 2.0$ (about 200km) and $N_{VC} = 100$. In the case of EFCI and EDS switches, the corresponding values are 18,000 and 14,000, respectively. Consequently, the PEDS switch can be valuable to be implemented even though it requires an additional control complexity.

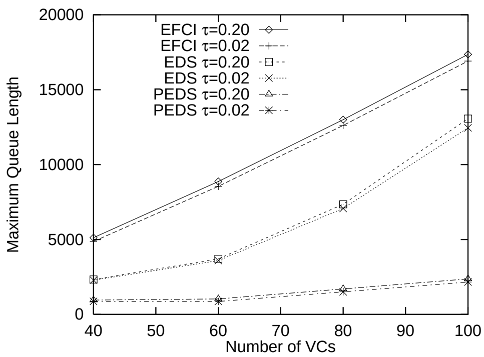

Figure 38: Effect of N_{VC} on Maximum Queue Length.

5 Simulation Study for Rate-Based Congestion Control Methods

In the previous sections, we have evaluated EPRCA analytically. As described in Section 2, other methods are proposed in the ATM Forum. In this section, we evaluate one of new methods called EPRCA++ which has been introduced in Subsection 2.4.2 by using a simulation technique. The effect of VBR traffic on ABR traffic is also investigated for EPRCA and EPRCA++ when VBR connections are established as well as ABR connections.

5.1 Comparison of EPRCA and EPRCA++

In this subsection, performance of EPRCA and EPRCA++ are evaluated for the network model, which is same as the one used in our analysis presented in Subsection 3.1. The propagation delay between the source and destination end systems τ are set at 0.01 msec, 0.1 msec, and 1.00 msec as typical values for LAN, MAN and WAN environments, respectively. The link speed at the switch is set to 150M bits/sec. Each connection establishment is staggered by 5 msec, that is, the nth connection starts its cell cell transmission at $(n-1)\times 5$ msec. For control parameters of EPRCA and EPRCA++, the values suggested in [16, 25] are used. Each simulation run is started at time 0 msec, when no connections exist in the network, to time 200 msec.

Figs. 39 through 41 show the allowed cell rate ACR of each connection and the queue length at the switch for EPRCA with EFCI, BES and EDS switches, The number of connections N_{VC} and the the propagation delay τ are set to at 5 and 0.01 msec, respectively. The case of EPRCA++ is shown in Figure 42. From figures, one can easily find that EPRCA++ gives extremely stable operation compared with EPRCA since the queue length is very small in the case of EPRCA++. It is owing to EPRCA++'s precise explicit rate setting capability, i.e., an explicit rate ER in RM cells is directly computed by the switch based on the available bandwidth, the input rate and the number of active VC's although it requires additional mechanisms at the switch, for example, per-VC accounting and a counter for incoming cells.

However, EPRCA++ causes a serious problem as the propagation delay becomes larger. In Figs. 43 through 46 and Figs. 46 through 50, simulation results for EPRCA and EPRCA++ are plotted for $\tau = 0.10$ and $\tau = 1.00$, respectively. It can be observed that the maximum queue length of EPRCA++ grows up to 5,500 cells, that is much larger than the values of EPRCA. It is true that EPRCA++ adopts the FECN like operation. However, the maximum queue length of EPRCA++ is about twenty times larger than the one of the EFCI switch case, which is also based on FECN. It is because EPRCA++ adopts quick start mechanism, i.e., AIR is initially set to PCR. Therefore,

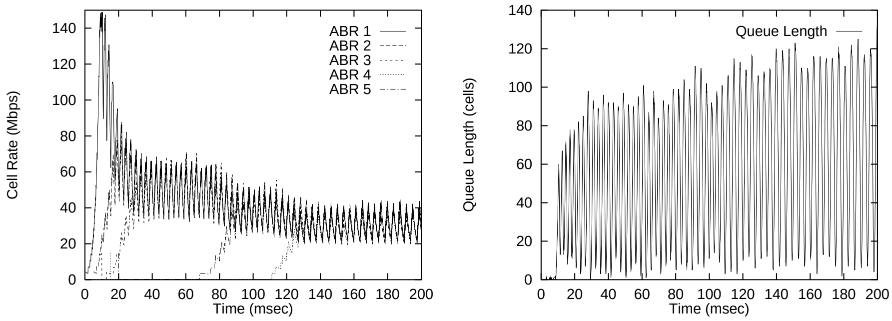

Figure 39: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 0.01$).

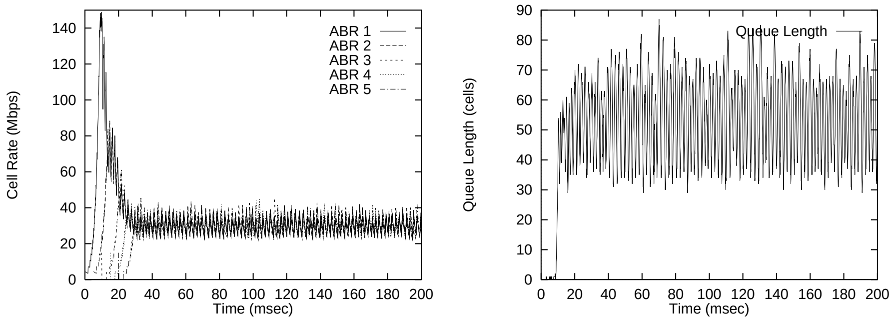

Figure 40: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 0.01$).

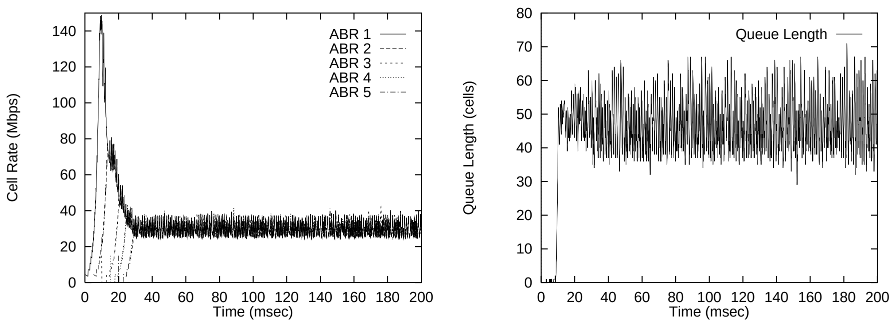

Figure 41: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 0.01$).

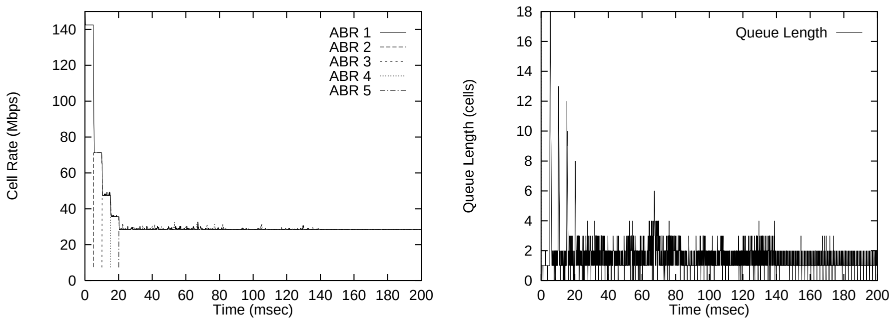

Figure 42: EPRCA++ ($N_{VC} = 5$, $\tau = 0.01$).

EPRCA++ could become unacceptable for rather large networks. However, the performance of EPRCA++ might be improved by, for example, changing the target utilization band (TUB) according to τ and N_{VC}, which requires a more parameter tuning.

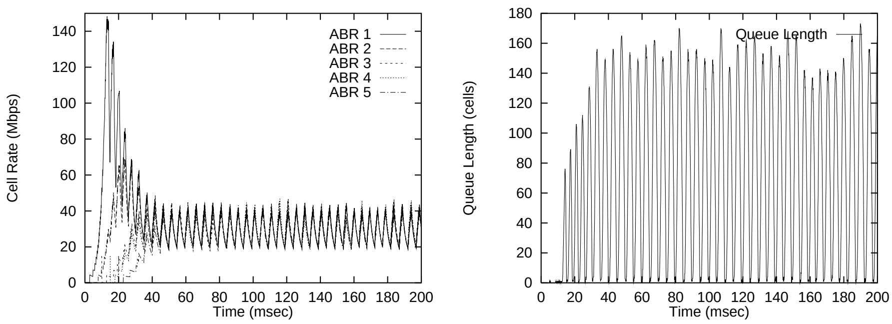

Figure 43: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 0.10$).

5.2 Effect of VBR traffic

In this subsection, we evaluate the performance of EPRCA and EPRCA++ for the case of a mixture of ABR and VBR traffic to examine their robustness in the multimedia traffic environment.

One of attractive capabilities of ATM networks is that it can support multiple QOS's following requirements of each traffic. As described before, rate-based congestion control has been developed for ABR service class, which is suitable for data communication and/or existing applications. On the other hand, real-time traffic such as voice and motion video is accommodated into CBR or VBR service class. In spite of this fact,

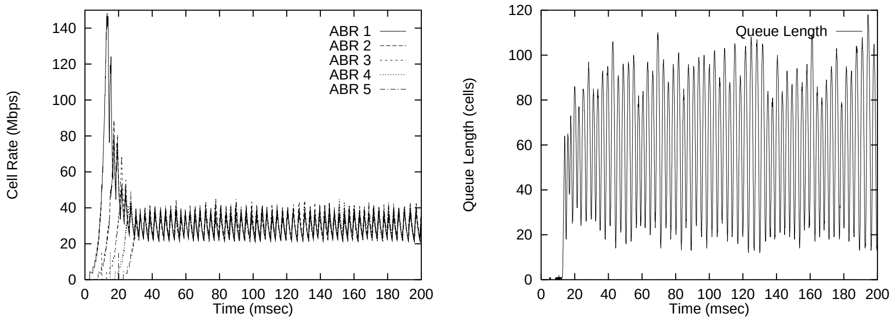

Figure 44: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 0.10$).

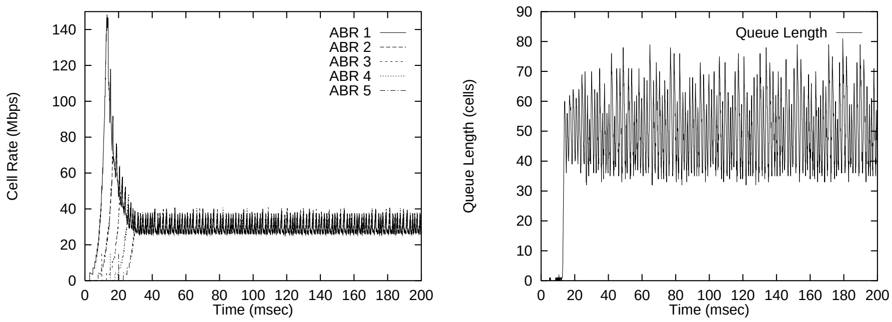

Figure 45: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 0.10$).

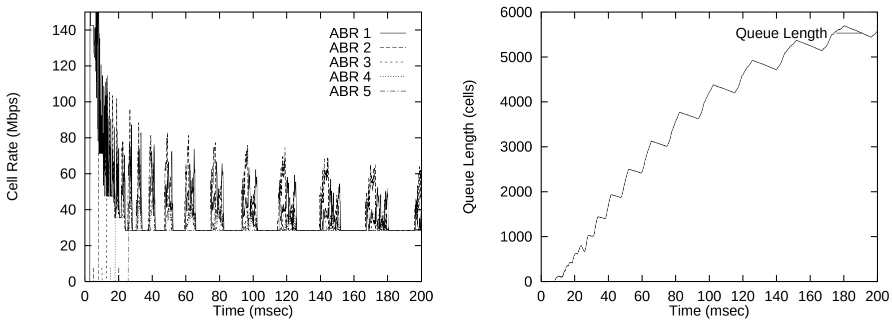

Figure 46: EPRCA++ ($N_{VC} = 5$, $\tau = 0.10$).

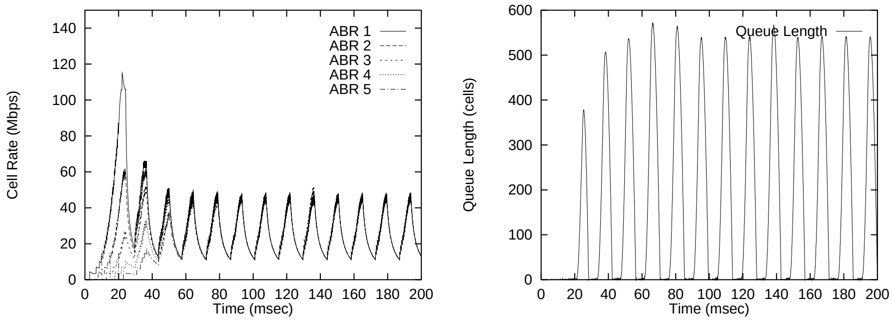

Figure 47: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 1.00$).

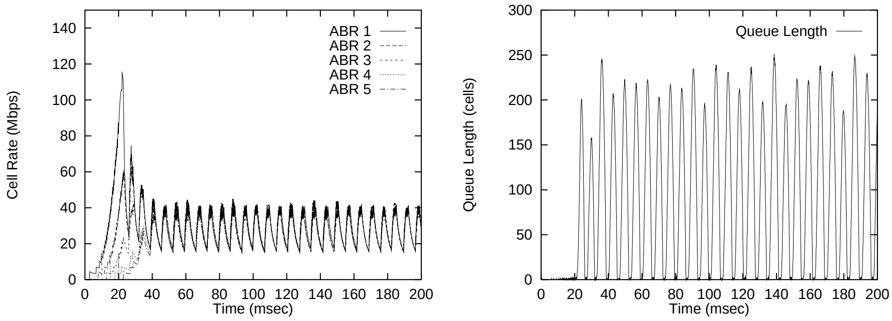

Figure 48: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 1.00$).

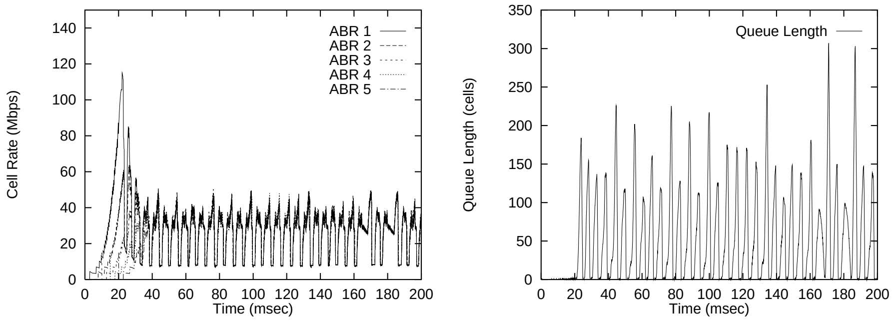

Figure 49: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 1.00$).

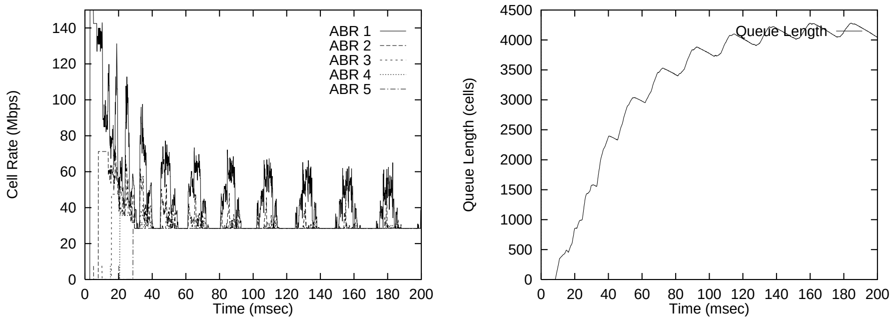

Figure 50: EPRCA++ ($N_{VC} = 5$, $\tau = 1.00$).

rate-based congestion control methods have been developed without considering those real time traffic classes.

In this paper, the effect of VBR traffic on ABR traffic is evaluated by a simulation technique using the same model presented in Subsection 5.1. It is assumed that VBR traffic is assigned high priority than ABR traffic, that is, VBR traffic cells are transmitted prior to ABR traffic cells at the switch if VBR traffic cells exist in its buffer. The difficulty exists in that the cell generation rate of VBR traffic is varied according to the time. Therefore, the bandwidth available to the ABR traffic is also changed. As a typical application of VBR traffic, we adopt MPEG-1 encoded video stream of 30 frames/sec, 352 $\times$ 240 pixels with average rate 4.5 Mbits/sec. In our simulation, 10 identical VBR sources are multiplexed with different starting points.

In EPRCA++, the switch requires information about the available bandwidth for ABR traffic. If we consider only the ABR traffic, it is identical to the VP capacity, which is fixed at 150 Mbps physical capacity in many cases. However, when the VBR traffic is also accommodated on the link, some method to measure the available bandwidth for the ABR traffic should be provided. Since it is not described in the original EPRCA++ method, we assume that the switch counts incoming VBR cells, and estimates available bandwidth BW' for ABR traffic as

$$BW' = BW \times (T - N_{VBR})/T,$$

where T and N_{VBR} denote an averaging interval (in cells) and the number of VBR cells in this interval, respectively.

Simulation results of EPRCA and EPRCA++ for $N_{VC} = 5$ and $\tau = 0.01$ are presented in Figs. 51 through 54. For the propagation delay τ, 0.01 msec, 0.1 msec and 1.00 msec are used. Simulation results in these cases are shown in Figs. 57 through 62. When the propagation delay is small (0.01 msec.), EPRCA++ method outperforms

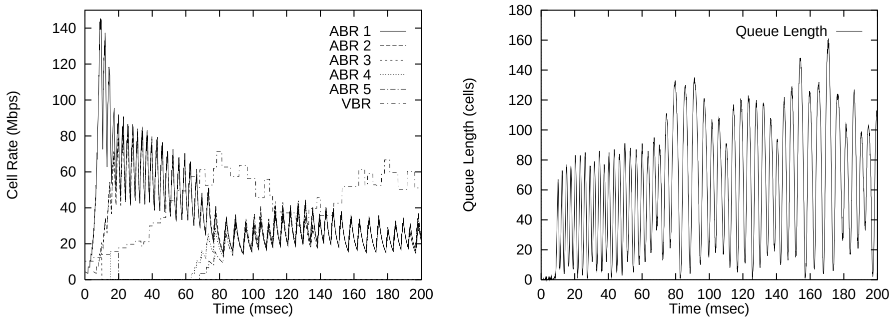

Figure 51: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 0.01$).

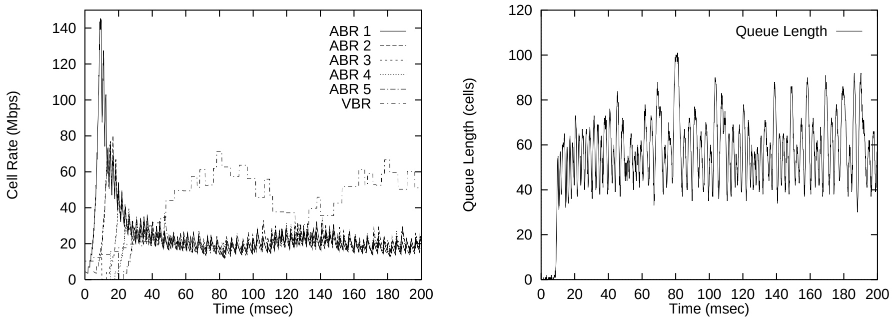

Figure 52: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 0.01$).

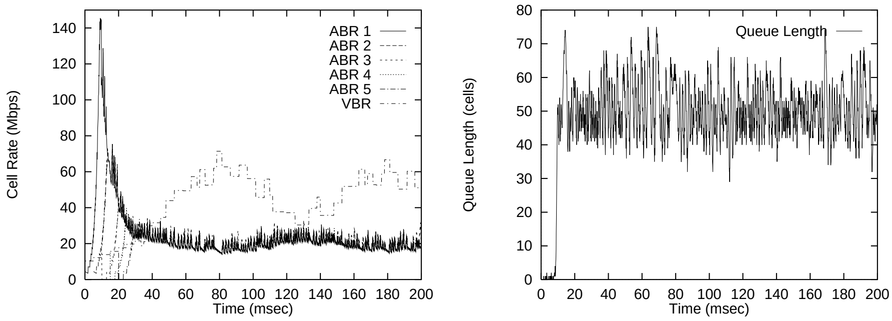

Figure 53: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 0.01$).

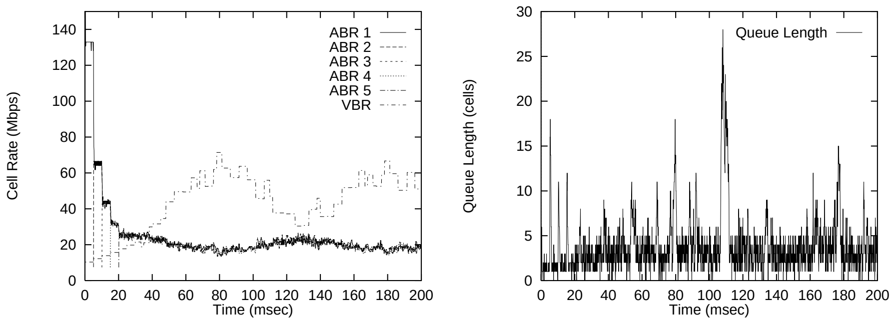

Figure 54: EPRCA++ ($N_{VC} = 5$, $\tau = 0.01$).

other EPRCA methods. However, the larger propagation delays becomes, the performance degradation in terms of the queue length is significant in the case of EPRCA++. Therefore, we may conclude that EPRCA is more robust against the VBR traffic than EPRCA++. The reason why EPRCA++ shows worse performance is explained as follows. EPRCA++ tries to utilize the bottlenecked link at target load, which is set at, for example, 0.95. However, the switch is overloaded (or underloaded) when the rate of VBR traffic increases (decreases) since an observed available bandwidth for ABR traffic is too old when τ is large. Remind that EPRCA++ uses FECN. Hence, the larger τ becomes, more overloaded the switch becomes.

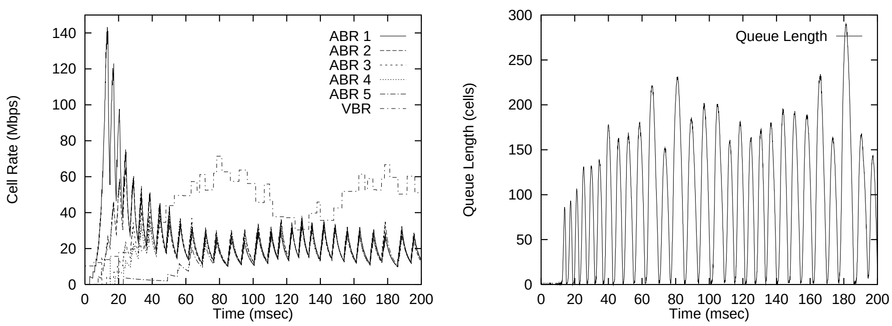

Figure 55: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 0.10$).

6 Conclusion

As a congestion control mechanism, rate-based control is a promising approach for incorporating ABR traffic into high speed networks based on ATM technology. Many

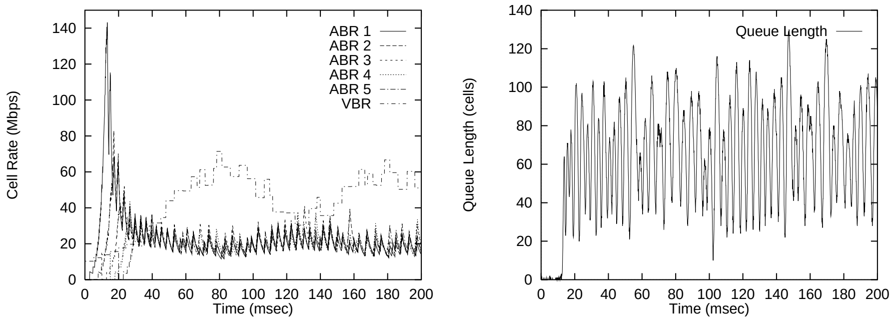

Figure 56: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 0.10$).

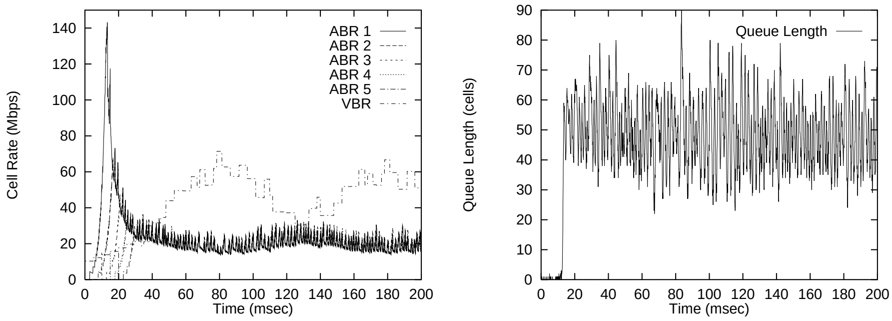

Figure 57: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 0.10$).

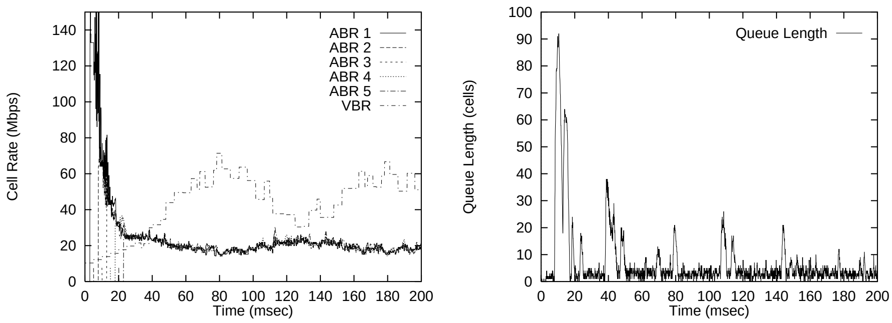

Figure 58: EPRCA++ ($N_{VC} = 5$, $\tau = 0.10$).

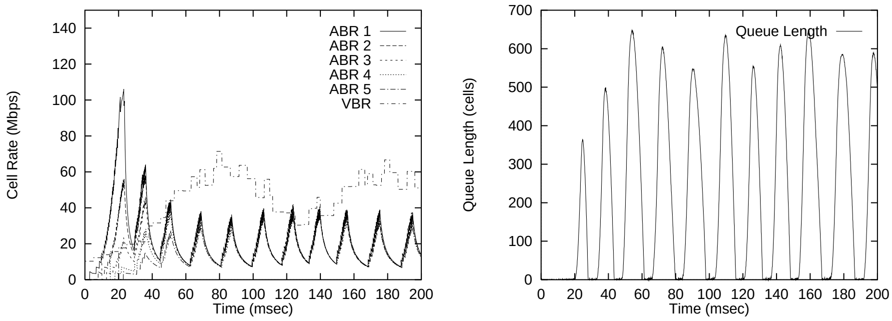

Figure 59: EPRCA with EFCI Switch ($N_{VC} = 5$, $\tau = 1.00$).

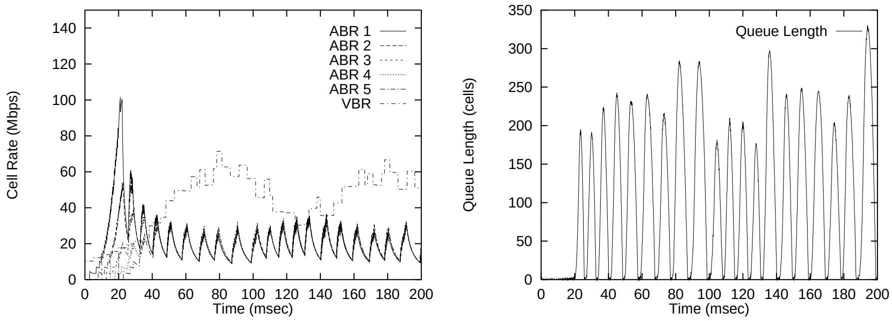

Figure 60: EPRCA with BES Switch ($N_{VC} = 5$, $\tau = 1.00$).

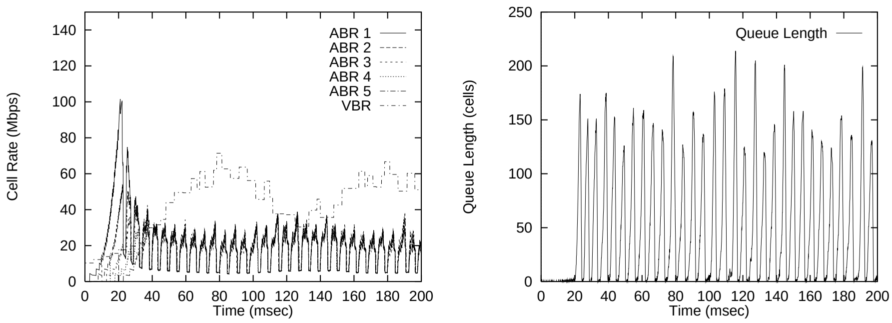

Figure 61: EPRCA with EDS Switch ($N_{VC} = 5$, $\tau = 1.00$).

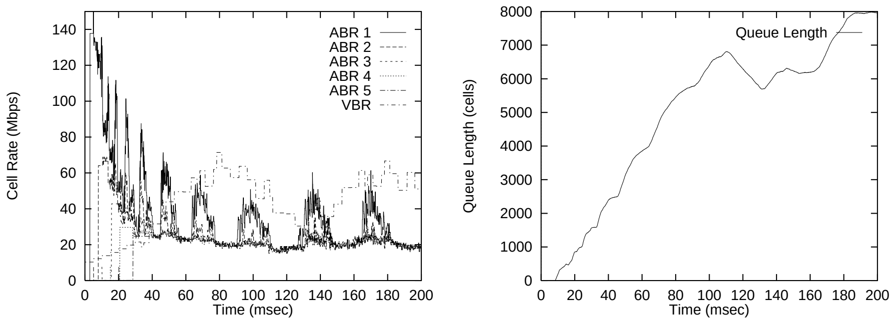

Figure 62: EPRCA++ ($N_{VC} = 5$, $\tau = 1.00$).

schemes for rate-based control approaches are proposed and studied in this area from a "binary feedback" congestion indication scheme to an "explicit rate setting" scheme.

In the first half of this paper, we gave a historical overview for rate-based schemes proposed in the ATM Forum so far with their advantages and disadvantages. One important feature of the rate-based algorithms is in its flexibility, that is, backward compatibility is always preserved in these schemes. One may choose cheaper end systems and simple switches such as the EFCI switches proposed in PRCA or EPRCA because of their cost-effectiveness. On the other hand, those who prefer stability and efficiency may use more expensive but more superior switch such as the EDS switch in EPRCA. Even if different switch architectures are adopted, all switches (and networks) should interwork with each other. Rate-based congestion control methods are still actively discussed. The final voting for standard documentation, to be included in User-Network Interface (UNI) Specification Version 4.0, is planned for the June $\sim$ August 1995 meeting.

In the second half of this paper, we analyzed the performance of EPRCA, a representative scheme in rate-based approaches, in both of initial transient and steady states. We provide analyses for three types of switches — EFCI, BES and EDS switches — suggested in EPRCA. EFCI switches are simple by means of a "binary feedback congestion indication" capability. Through numerical examples, it has been shown that EFCI switches work well in the LAN environment if the number of active VC's is limited. Next, BES switches are analyzed and the effect of their BECN like capability is illustrated depending on the distance between the source and the switch. The last one, EDS switches can dramatically decrease the maximum queue length by its explicit rate setting capability.

Acknowledgment

I would like to express my sincere appreciation to Professor Hideo Miyahara of Osaka University, who introduced me to the area of broadband telecommunication networks including the subjects in this thesis.

All works of this thesis would not have been possible without the support of Associate Professor Masayuki Murata of Osaka University. It gives me great pleasure to acknowledge his assistance. He has been constant sources of encouragement and advice through my studies and preparation of this manuscript.

I would like to thank Dr. Hiroshi Suzuki and Dr. Chinatsu Ikeda with NEC Corporation, C&C System Laboratories, for their invaluable suggestions.

Thanks are also due to many friends and colleagues in the Department of Information and Computer Science of Osaka University who gave me useful suggestions.

Appendix A Control Parameter Definitions

In this appendix, control parameters and variables used by rate-based control algorithms are listed. These parameters are summarized from [10, 13, 15, 16, 17].

- Source End System Parameters

  — General

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/reset value for ACR.
AIR	Additive Increase Rate; rate increase permitted.
RDF	Rate Decrease Factor; $RDF = 2^{RD}$.

  — Timer-based Only

UI	Update Interval; a period for which ACR is re-evaluated.

  — After PRCA

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

- Source End System Variables

  — General

ACR	Allowed Cell Rate, current maximum for CCR.
CCR	Current Cell Rate, current rate of cell transmission.

  — After PRCA

ADR	Additive Decrease Rate (cells/unit time).

- Destination End System Parameters

  — Timer-based Only

RMI	Minimum Interval of RM cells transmitted.

- Switch Parameters and Suggested Settings [16]

— BES and EDS Switches

$MACR$	N/A	Congestion point rate computed by switch per queue, Ideally, should be the available bandwidth divided by the number of active connections
DQT	100	High queue limit to determine very congested.

— EDS Only

VCS	7/8	VC Separator.
AV	1/16	Exponential Averaging Factor; for averaging ACR's
MRF	1/4	Major Reduction Factor; for major reduction.
DPF	7/8	Down Pressure Factor.
ERF	15/16	Explicit Reduction Factor

- RM Cell Fields

— General

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

— After PRCA

DIR	Direction of RM cell; forward or backward.

— After EPRCA

ACR	Allowed Cell Rate in effect when forward RM cell is generated.
ER	Explicit Rate; initially set to PCR, and possibly modified by intermediate nodes along the path.

Appendix B Pseudo Codes for End Systems

In this appendix, the pseudo-code for the source and destination end systems are presented. These codes are extracted from [17].

Source End System Behavior

```
Initialization
    Call setup assigns values to PCR, MCR, ICR, Nrm, AIR, and RDF.
    ACR =    ICR
    ADR =    ShiftR(ACR, RD)
    count =  0
    active = 0



While VC-on-line do

if  cell-to-send                                    ! starting event
    if  not active                                  ! active==0
        next-cell-time = now                        ! now is from real time clo

if  now >= next-cell-time                           ! scheduled event
    if cell-to-send                                 ! queue not empty
        active = 1
        if count>0
            send data cell
        else
            count = ShiftL( 1, N)                           ! count = 2^N
            send RM(ACR, DIR=forward, ER=PCR, CI=0)         ! every Nrm cells
            decrease = 1
        count =  count - 1

    if  not active
            count = 0
            decrease = 1
    if decrease
            decrease = 0
            ACR = max( ACR - ADR, MCR)
            interval =  1/ACR
```

```
    if  not cell-to-send
        interval = ShiftL(interval, N)
        active = 0

    if active or ACR>ICR
         next-cell-time = next-cell-time + interval     ! schedule cell


if receive RM(ACR, DIR=backward, ER, CI)
    if(CI = 0)
         ACR = ACR + AIR + ADR
    ACR = min(ACR, ER, PCR)
    ACR = max(ACR, MCR)
    interval =  1/ACR
    ADR = ShiftR(ACR, RD)
```

Destination End System Behavior

```
if receive data cell
    VC-CI = EFCI state of cell

if receive RM(ACR, DIR = forward, ER, CI)
    if(VC-CI=1)
        CI = 1
    send RM(ACR, DIR = bsckward, ER, CI)
```

Appendix C Current RM Cell Format Proposal

In this appendix, the current proposal for RM cell format is shown, which is extracted from [31].

FIELD	BYTES	BITS	NAME	SOURCE SET TO	NETWORK SET TO	DESTINATION SET TO
Header	1–5		ATM Header	Standard Header $PT = 6$		
ID	6	Protocol ID	0			
DIR	7	1	Direction	0=Forward	See Below	1=Backward
CI	7	2	Congestion Indication	0=Increase OK	1 if very congested	may set=1 if $EFCI = 1$
BS	7	3	Block Start	1 if start block		
Reserved	7	4	Reserved			
N	7	5–8	log_2 Num. Cells/RM	$log_2 N_{RM}$		
Reserved	8–9		Reserved			
CCR	10–11		Current Cell Rate	ACR		
MCR	12–13		Minimum Cell Rate			
ER	14–15		Explicit Cell Rate	PCR	Explicit Rate	may be set if $EFCI = 1$
CRC	52–53		Cyclic Check	CRC16	CRC16	CRC16

- The network should not modify the direction bit DIR except when it is segmenting the network or multicasting a copy back to the source.

Appendix D Abbreviation List

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
APRC	Adaptive Proportional Rate Control
BECN	Backward Congestion Notification
BES	Binary Enhanced Switch
CCR	Current Cell Rate
CDV	Cell Delay Variation
CBR	Constant Bit Rate
DES	Destination End System
EDS	Explicit Down Switch
EFCI	Explicit Forward Congestion Indication
EPRCA	Enhanced Proportional Rate Control Algorithm
ES	End System
FECN	Forward Explicit Congestion Notification
PRCA	Proportional Rate Control Algorithm
PTI	Payload Type Identifier
RM	Resource Management
QOS	Quality Of Service
SES	Source End System
UNI	User–Network Interface
VBR	Variable Bit Rate
VC	Virtual Channel
VP	Virtual Path

References

[1] P. Newman, “Traffic management for ATM local area networks,” *IEEE Communications Magazine*, pp. 44–50, August 1994.

[2] G. Woodruff and R. Ksitpaiboon, “Multimedia traffic management principles for guaranteed ATM network performance,” *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 437–446, April 1990.

[3] “Special issue on “Congestion Control in High Speed Networks”,” *IEEE Communications Magazine*, vol. 29, October 1991.

[4] H. T. Kung and A. Chapman, “Credit-based flow control for ATM networks: Credit update protocol, adaptive credit allocation, and statistical multiplexing,” *The Proceedings of the SIGCOMM ’94*, vol. 24, pp. 101–114, October 1994.

[5] P. Newman, “Backward explicit congestion notification for ATM local area netwoks,” *IEEE GLOBECOM ’93*, pp. 719–723, December 1993.

[6] R. Jain, “Myths about congestion management in high-speed networks,” in *Information Network and Data Communications, IV* (M. Tienari and D. Khakhar, eds.), pp. 55–70, Elsevier Science Publishers B.V. (North-Holland), 1992.

[7] N. Yin and M. G. Hluchyj, “On closed-loop rate control for ATM cell relay networks,” *IEEE INFOCOM ’94*, pp. 99–109, June 1994.

[8] C. Ikeda and H. Suzuki, “Adaptive congestion control schemes for ATM LANs,” *IEEE INFOCOM ’94*, pp. 829–838, June 1994.

[9] A. Berger, F. Bonomi, K. Fendick, and J. Swenson, “Control of multiplexed connections using backward congestion control,” *ATM Forum Contribution 94-064*, January 1994.

[10] B. Makrucki et al., “Closed-loop rate-based traffic management,” *ATM Forum Contribution 94-438*, May 1994.

[11] A. W. Barnhart, “Baseline model for rate control simulations,” *ATM Forum Contribution 94-0399*, May 1994.

[12] Jon C. R. Bennett and G. Tom Des Jardins, “Failure modes of the baseline rate based congestion control plan,” *ATM Forum Contribution 94-0682*, July 1994.

[13] L. Roberts et al., “Closed-loop rate-based traffic management,” *ATM Forum Contribution 94-0438R1*, June 1994.

[14] Jon C. R. Bennett and G. Tom Des Jardins, "Comments on the July PRCA rate control baseline," *ATM Forum Contribution 94-0682*, July 1994.

[15] H. Hsiaw et al., "Closed-loop rate-based traffic management," *ATM Forum Contribution 94-0438R1*, September 1994.

[16] L. Roberts, "Enhanced PRCA (proportional rate-control algorithm)," *ATM Forum Contribution 94-0735R1*, August 1994.

[17] L. Roberts and A. W. Barnhart, "New pseudocode for explicit rate plus EFCI support," *ATm Forum Contribution 93-0974*, October 1994.

[18] Y. Chang, N. Golmie, L. Benmohamed, and D. Su, "Simulation study of the new rate-based EPRCA traffic management mechanism," *ATM Forum Contribution 94-0809*, September 1994.

[19] H.-Y. Tzeng and K.-Y. Siu, "A class of proportional rate control scheme and siulation results," *ATM Forum Contribution 94-0888*, September 1994.

[20] K.-Y. Siu and H.-Y. Tzeng, "Limits of performance in rate-based control scheme," *ATM Forum Contribution 94-1077*, November 1994.

[21] H.-Y. Tzeng and K.-Y. Siu, "Comparison of performance among existing rate control schemes," *ATM Forum Contribution 94-1077*, November 1994.

[22] R. Jain, S. Kalyanaraman, and R. Viswanathan, "Simulation results: The EPRCA+ scheme," *ATM Forum Contribution 94-0988*, October 1994.

[23] R. Jain, S. Kalyanaraman, and R. Viswanathan, "The OSU scheme for congestion avoidance using explicit rate indication," *ATM Forum Contribution 94-0883*, September 1994.

[24] R. Jain, S. Kalyanaraman, and R. Viswanathan, "Current default proposal: Unresolved issues," *ATM Forum Contribution 94-1175R1*, November–December 1994.

[25] R. Jain, S. Kalyanaraman, and R. Viswanathan, "Transient performance of EPRCA and EPRCA++," *ATM Forum Contribution 94-1173*, November–December 1994.

[26] S. Sathaye, "ATM Forum traffic management specification," *ATM Forum Contribution*, December 1994.

[27] D. Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, New Jersey: Prentice-Hall, 1987.

[28] N. Yin, “Fairness definition in ABR service model,” *ATM Forum Contribution 94-0928R2*, November 1994.

[29] D. Hughes, “Fairness share in the context of MCR,” *ATM Forum Contribution 94-0977*, October 1994.

[30] H. Ohsaki, M. Murata, H. Suzuki, and H. Miyahara, “Analysis of rate–based control algorithm for ATM LANs, part 2: Initial transient state analysis,” *submitted to ICC ’95*, 1994.

[31] L. Roberts, “RM cell format proposal,” *ATM Forum Contribution 94-0973R1*, November 1994.