ANALYSIS OF RATE-BASED CONGESTION CONTROL ALGORITHMS FOR ATM NETWORKS — PART 1: STEADY STATE ANALYSIS —

Hiroyuki Ohsaki, Masayuki Murata, †Hiroshi Suzuki, †Chinatsu Ikeda and Hideo Miyahara

Department of Information and Computer Sciences
Faculty of Engineering Science, Osaka University
Toyonaka, Osaka 560, Japan

†C&C System Research Laboratories
NEC Corporation
Kawasaki, Kanagawa 216, Japan

Abstract — Rate-based congestion control is effective and still simple for traffic management in ATM networks. One of its practical realization schemes is Enhanced Proportional Rate Control Algorithm (EPRCA) which has recently been proposed and adopted as a standard for traffic management scheme in the ATM Forum. In this paper, an analysis for EPRCA is provided for a network with homogeneous source traffic and a single bottleneck ATM link to investigate a dynamic behavior of EPRCA. In addition to three types of switches suggested in EPRCA, a prioritized switch is newly proposed. We exploit the dynamical behavior of EPRCA and obtain the maximum queue length at the switch and the throughput by using the first-order fluid approximation method. Equations for parameter tuning for EPRCA are also provided.

I. INTRODUCTION

Closed-loop rate control is promising for data communications and has been applied to the ABR (Available Bit Rate) service in the ATM Forum. Closed-loop control dynamically regulates the cell emission process of each connection by using feedback information from the network. Among several control schemes for implementation of such a control method, a rate-based congestion control is considered to be effective as a means of controlling connections' flows and of fully utilizing network bandwidth [1]. The function of the rate-based approach is to control the cell emission rate of each source end system by feedback information from the network. If some of the switches in the network becomes congested, it sends congestion indication to the source end systems. Each of source end systems then decreases its cell emission rate to avoid buffer overflow of cells at the congested switch. After congestion is relieved, the transmission rate is increased.

In this paper, we evaluate the rate-based control scheme named Enhanced Proportional Rate Control Algorithm (EPRCA), which has been adopted as a standard traffic management mechanism by the ATM Forum [2]. Refer to [3] for a historical overview of the standardization effort of the rate-based congestion control in the ATM Forum. In the standard, only the behaviors of the source and destination end systems are described, and the implementation issues regarding the ATM switches are left to manufacturers. In [2], however, they suggest three types of switches, EFCI bit setting switch (EFCI), binary enhanced switch (BES), and explicit down switch (EDS), which have different processing capabilities against congestion as will be summarized in Section II.

While a lot of simulation studies have been contributed by researchers at the ATM Forum, our main purpose in the current paper is to give analytic result for the standard mechanism. Regarding the analytic approaches for rate-based control schemes, Blot et al. showed its dynamical behavior by utilizing first-order fluid approximations in [4]. In their model, the switch is assumed to recognize congestion when the cell arriving rate at the switch exceeds its service rate, i.e., the link bandwidth. On the other hand, in the current paper, we will deal with a more realistic situation in which congestion of the switch is detected by its queue length. When the number of cells at the switch buffer exceeds predetermined threshold value, it is recognized as congestion. The same model has been treated in [5] using a similar analytic technique. In their paper, however, the method for notifying congestion from the switch to source end systems is not explicitly modeled; That is, the rate is increased linearly and decreased exponentially, and the time interval to update the increase/decrease rate is assumed to be fixed. This scheme has been found to be of problem in some situations [6].

The rest of this paper is organized as follows. In Section II, we first describe a basic feature of EPRCA. In Section III, our analytic model for EPRCA is described with definition of parameters. In Section IV, steady-state analysis for EPRCA are presented. In Section V, we conclude our paper with some remarks.

II. ENHANCED PROPORTIONAL RATE CONTROL ALGORITHM

We first describe a basic feature of EPRCA, which is based on a positive feedback mechanism. A source end system periodically sends an RM (Resource Management) cell every N_{RM} data cells to check the congestion status of the network. The RM cell received at the destination end system is returned to the source along the backward path if congestion does not occur in the network. The switch can notify its congestion to the destination end system by marking an EFCI bit in the header of data cells. As presented in [5], the basic operation of the rate-based congestion control is that the source end system normally increases its allowable cell transmission rate called ACR (allowed cell rate) while it is decreased when the network falls into congestion. A notable feature of EPRCA is, however, that the source end system always decreases ACR until it receives the RM cell from the network. It can increase the rate only when the RM cell is received. The RM cell is discarded at the destination end system if congestion is indicated by the EFCI bit. It results in that the source end system continues to decrease its ACR. This *positive feedback* mechanism accomplishes a fast rate reduction at the source end system even if RM cells are lost at the switch because of buffer overflow.

In EPRCA, three types of switch architectures with different functions are suggested in the form of pseudo-code [2]. The first one is an EFCI bit setting switch (EFCI) whose operation is described in the above, and it is originally proposed in [7]. However, the EFCI switch has a fault that it produces unfairness among connections [6].

A capability to select source end systems having large ACR is then added in the second switch called Binary Enhanced Switch (BES). More specifically, the BES 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 shared equally and the switch can be fully used without falling into congestion. The key to this is obtaining an accurate $MACR$. The BES switch updates its $MACR$ according to the CCR (Current Cell Rate) field of forward RM cells. For example, $MACR$ is calculated as

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

where AV is used as an averaging factor [2]. When the switch becomes congested, it indicates its congestion to the sources having higher rates. More specifically, the switch marks the CI (Congestion Indication) bit of the backward RM cells if its CCR value exceeds $MACR \times DPF$ (Down Pressure Factor), where a typical value of DPR is 7/8 for safe operation. The switch may remain congested if only the above operation is applied. Therefore when BES switch becomes *very* congested (such that the queue length exceeds DQT), all backward RM cells are marked irrespective of their CCR values. Note that it is evident from the above description that a BECN-like quick congestion notification can be accomplished in BES switches.

A more drastic rate reduction mechanism is adopted by the last one, the Explicit Down Switch (EDS). The EDS switch maintains $MACR$ as the BES switch does, and it controls the transmission rate of sources by setting the ER field of backward RM cells according to degree of congestion. When a backward RM cell with larger CCR than $MACR$ passes through 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 are 7/8 and 1/4 [2].

We note here that since, in the EFCI switch, the congestion occurrence is notified to the destination by marking an EFCI bit in data cells, forward RM cells are not required. However, if at least one BES or EDS switch exists on the path, forward RM cells should be sent by the source end system. From an analytical point of view, however, it is not a problem whether forward RM cells are explicitly sent by sources or not, and both cases of the EFCI switches can be treated in a unified manner as shown in Subsection IV-A.

III. ANALYTIC MODEL

Our model consists of homogeneous traffic sources and a single bottleneck link (see Fig. 1). The number of active connections that share the bottleneck switch is denoted by N_{VC}. We assume that these connections behave identically; that is, they all have the same parameters ICR, PCR, AIR, and MDF. The bandwidth of the bottleneck link (in cells/msec) is denoted by BW, and the propagation delays from the source to the switch and from the switch to the destination are denoted by τ_{sx} and τ_{xd}. 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 denoted 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.

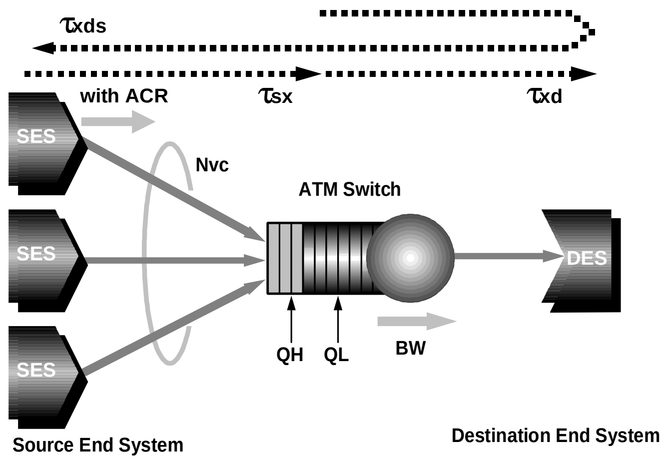

Fig. 1: Analytic Model for EPRCA.

In all three types of switches, congestion is detected by threshold values associated with the queue length at the switch buffer. The EFCI switch has high and low threshold values denoted as Q_H and Q_L. When the queue length at the switch exceeds Q_H, the switch detects its congestion and marks the EFCI bit in the header of data cells. When the queue length goes under Q_L, on the other hand, it is regarded that congestion terminates. The BES and EDS switches have another threshold value DQT and take an action against congestion according to its switch type as described in the previous section. In EPRCA, the source end system sends RM cells proportionally to its rate. Furthermore, the rate is changed in response to backward RM cells returned from the destination end system, and the rate at which the backward RM cells are received 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. We, therefore, require an analytical treatment different from the one presented in [5], where the rate change is performed on a timer basis.

Let us introduce $ACR(t)$ and $Q(t)$, which respectively represent the cell transmission rate ACR of each source end system and the queue length at the switch observed at time t. 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 which is the actual cell transmission rate. For the EDS switch, we will point out its drawback and analyze a modified version of EDS switch (prioritized EDS), which has a capability to provide a simple priority control mechanism. Last, we note that initial transient analyses of these switches are reported in a companion paper [8].

IV. STEADY STATE ANALYSIS

In this section, we analyze performance of EPRCA in steady state with comparison among three types of switches suggested in [2].

A. EFCI Switch

First we focus on the EFCI switch to analyze a dynamical behavior of $ACR(t)$ and $Q(t)$. In the following analysis, forward RM cells are not explicitly considered. Hence this model is equivalent to PRCA proposed in [7]. However, forward RM cells can easily be taken into account by replacing BW in our analysis with BW' defined as

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

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

1) Determination of $ACR(t)$:

Figure 2 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. 2. For simplicity of presentation, we introduce $ACR_i(t)$ and $Q_i(t)$ as

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

where t_i is defined as the time when phase i terminates. Furthermore, 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. 2. $ACR_i(t)$ $(1 \le i \le 4)$ can be determined as follows.

- Phase 1: $ACR_1(t)$

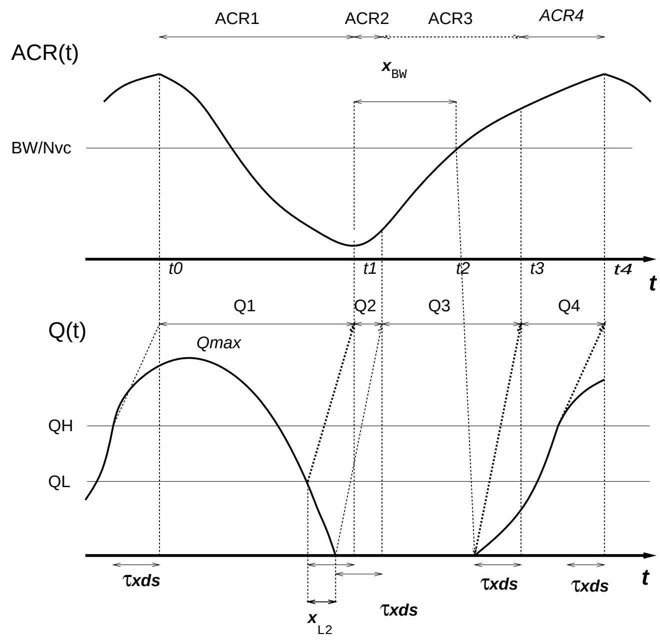

Fig. 2: Pictorial View of $ACR(t)$ and $Q(t)$ in EFCI Switch.

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 $ACR_1(0)/MD$. A differential equation for $ACR_1(t)$ is then obtained as

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

which gives

$$ACR_1(t) = ACR_1(0) e^{-\frac{ACR_1(0)}{MD} t}. \tag{2}$$

Note that MCR is not considered in the above equation; $MCR = 0$ 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, $Q(t - \tau_{xds}) > 0$. 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. We then derive an envelope of ACR in Phase 2 as follows.

$$\begin{aligned} \frac{dACR_2(t)}{dt} &= -\frac{ACR_2(t)^2}{MD} + \frac{BW}{MD\, N_{VC}} ACR_2(t) \\ &\quad + \frac{BW}{N_{VC}} AIR \end{aligned}$$

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

$$ACR_2(t) = \frac{a_1 e^{-a_1 t} + a_2 r e^{-a_2 t}}{c_1 (e^{-a_1 t} + r e^{-a_2 t})}, \tag{3}$$

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

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

The initial transmission rate $ACR_2(0)$ determines r in (3) 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 reaches 0. RM cells arrive 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.

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

Since it is difficult to solve the above equation, 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,

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

which leads to

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

Finally $ACR_3(t)$ is solved as

$$ACR_3(t) \cong ACR_3(0) e^{\beta t}, \tag{4}$$

where β is given as a root of the equation

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

- 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.

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 $t_{i,i+1}$ for each phase. Given initial rates in Phase 1, $Q_1(t)$ is obtained as

$$\begin{aligned} Q_1(t) &= Q_1(\tau_{xds}) \\ &\quad + \int_{x=\tau_{xds}}^{t} (N_{VC}\, ACR_1(x - \tau_{sx}) - BW) dx. \end{aligned} \tag{5}$$

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

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

In other words, x_{L2} is the time for the queue length to reach 0 after it goes below Q_L. Furthermore,

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

which defines the time when the aggregate ACR reaches BW. We should consider the following four cases according to the relation among x_{L2}, x_{BW} and τ_{sx} (Figure 2 corresponds to Case 1).

Case 1: $x_{L2} \leq \tau, \quad x_{L2} < x_{BW} + \tau_{sx}$
Case 2: $x_{L2} \leq \tau, \quad x_{L2} \geq x_{BW} + \tau_{sx}$
Case 3: $x_{L2} > \tau, \quad x_{L2} < x_{BW} + \tau_{sx}$
Case 4: $x_{L2} > \tau, \quad x_{L2} \geq x_{BW} + \tau_{sx}$

Due to lack of space, only Case 1 is 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;

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

Furthermore, $Q(t)$ is determined as follows.

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

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 in steady state.

3) Parameter Tuning:

In this subsection, two suggestive results are presented for parameter tuning: a maximum queue length in steady state and 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 $t \to \infty$.

$$\lim_{t\to\infty} ACR_4(t) = \frac{BW + \sqrt{BW^2 + 4N_{VC}\, MD\, BW\, AIR}}{2N_{VC}}. \quad (6)$$

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.

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

$$Q(\tau_{sx}) = Q_H + (N_{VC}\, ACR_4(\infty) - BW)\tau.$$

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

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

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

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

We then have the maximum queue length Q_{max} as

$$Q_{max} = Q_H + (N_{VC}\, ACR_4(\infty) - BW)\tau \\ + N_{VC} MD(1 - BW/N_{VC}) - \frac{BW\, MD}{ACR_4(\infty)} \log \frac{N_{VC}\, ACR_4(\infty)}{BW} \quad (7)$$

where $ACR_4(\infty)$ is given by (6). We can observe from the above equation that the number of connections 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.

$$x_{L2} > \tau,$$
$$x_{BW} + \tau_{sx} < x_{L2}.$$

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 cell/ms assuming 156Mbit/s ATM link. For other control parameters, the values suggested in [2] are used throughout this paper unless other values are explicitly specified. Equation (1) is used for BW; the overhead for forward RM cells are taken into account to compare with other switches.

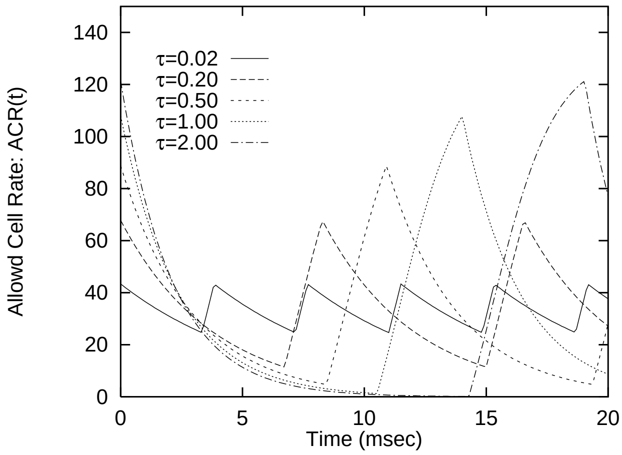

Fig. 3: Effect of Propagation Delay on $ACR(t)$ in EFCI Switch ($N_{VC} = 10$).

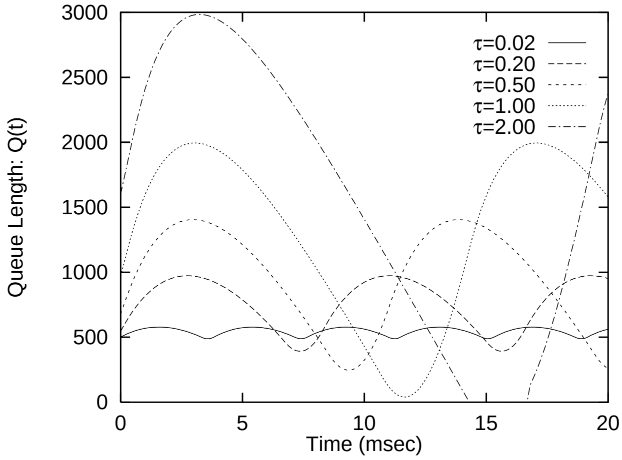

Fig. 4: Effect of Propagation Delay on $Q(t)$ in EFCI Switch ($N_{VC} = 10$).

$ACR(t)$ and $Q(t)$ for different values of propagation delays are compared in Figs. 3 and 4. As can be observed in the figures, the larger τ causes slower congestion notification, and it results in increase of the maximum queue length. Furthermore the under-utilization appears when the propagation delays τ is beyond 2.0 msec (about 400 km). Therefore, from these numerical results, we may conclude that the EFCI switch should be used in rather small networks.

Another problem of the EFCI switch can be found when N_{VC} is set to be large. Figure 5 shows that the maximum queue length

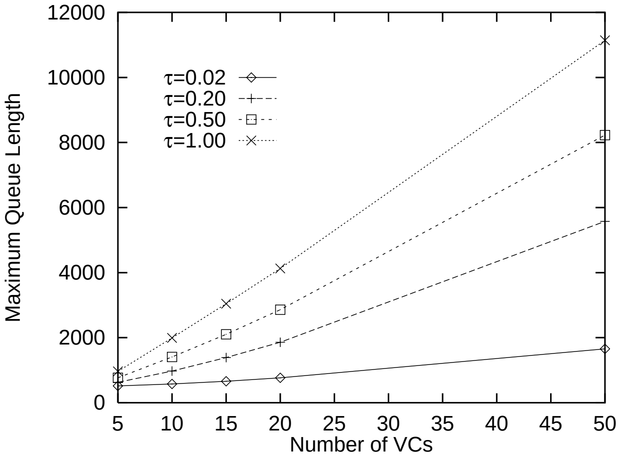

Fig. 5: Effect of N_{VC} on Maximum Queue Length in EFCI Switch.

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.

B. BES Switch

1) Analysis:

As described in Section II, 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 IV-A can directly be applied to the BES switch by (1) replacing the parameter τ_{xds} with τ_{sx}, and (2) 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 (1) in the following analysis.

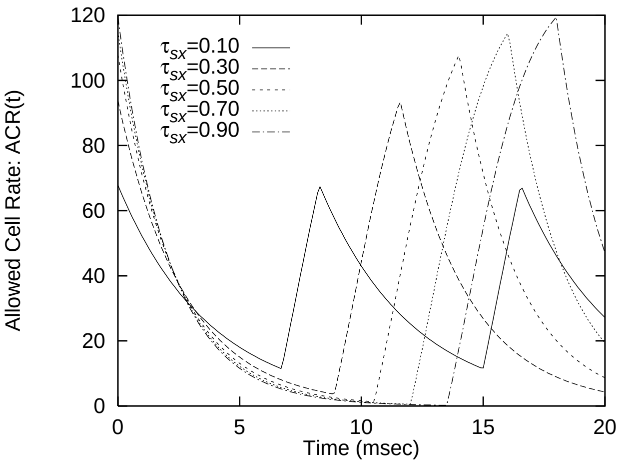

Fig. 6: Effect of Switch Location on $ACR(t)$ in BES Switch ($N_{VC} = 10, \tau_{sx} + \tau_{xd} = 1.0$).

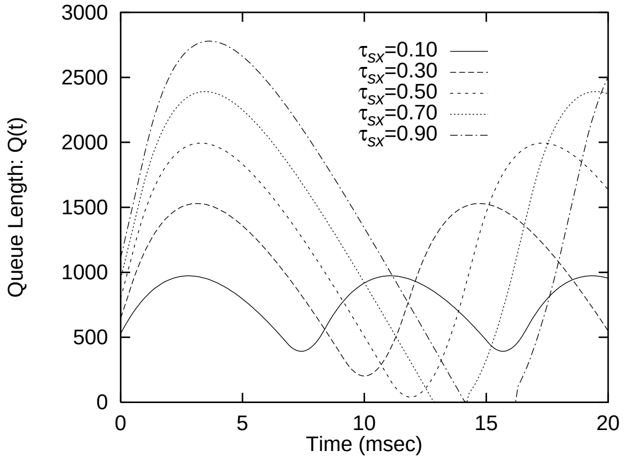

Fig. 7: Effect of Switch Location on $Q(t)$ in BES Switch ($N_{VC} = 10, \tau_{sx} + \tau_{xd} = 1.0$).

With the same approach as in the EFCI switch, the maximum queue length for the BES switch is also obtained in the same form with (7).

2) Numerical Examples:

To see the effect of the BECN-like capability of the BES switch more clearly, the distance between the switch and the source (τ_{sx}) is changed in Figs. 6 and 7. In these 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. 4).

C. EDS and PEDS Switches

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. Typical values for ERF and MRF are 15/16 and 1/4, respectively.

When the queue length becomes beyond Q_H but 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. In our analysis, however, we 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

$$ACR'(t) = ACR_i(t - Q(t)/BW - \tau_{sx}), \quad (8)$$

$$ACR'_i(t) = ACR'(t - t_i), \quad t_{i-1} \le t < t_i. \quad (9)$$

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, Equation (9) becomes

$$ACR'(t) \quad = \quad ACR_i(t - \tau_{sx}). \tag{10}$$

In what follows, analytic results only for the PEDS switch are presented. 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 IV-A.

We divide one cycle into three phases following behaviors of $ACR(t)$ and $Q(t)$ (see Fig. 8). 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 (2) and (3), respectively.

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. 8)

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

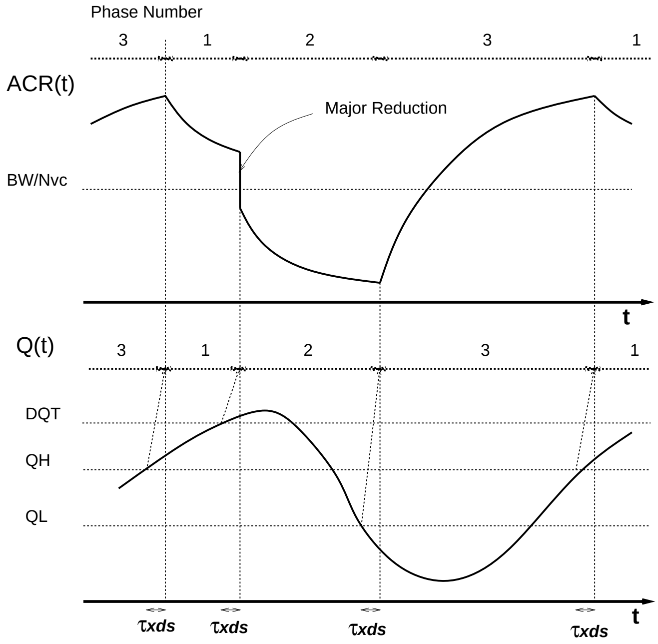

Fig. 8: Pictorial View of EDS Switch with Major Reduction.

- 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.

2) Maximum Queue Length:

The maximum queue length for PEDS switch can be obtained by the same method as in the case of the EFCI switch in Subsection IV-A.

- 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 IV-A:

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

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) Numerical Examples:

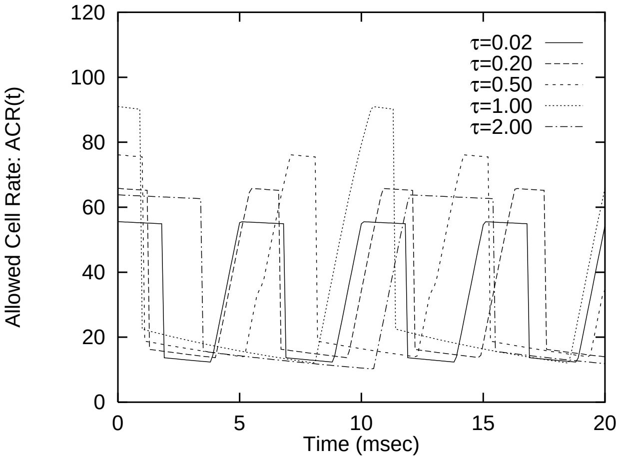

Fig. 9: Effect of Propagation Delay on $ACR(t)$ in PEDS Switch ($N_{VC} = 10$).

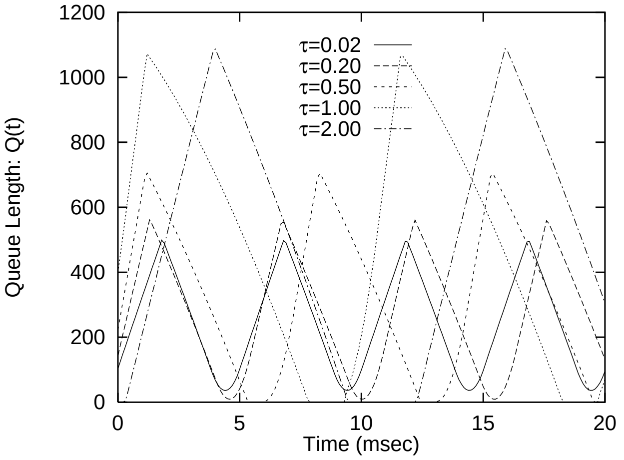

Fig. 10: Effect of Propagation Delay on $Q(t)$ in PEDS Switch ($N_{VC} = 10$).

The effects of propagation delays are depicted in Figs. 9 and 10 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. 10). 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, in the case of $ERF = 1/2$, the maximum queue length is limited under 500 and the network is fully utilized.

V. CONCLUSION

In this paper, we have analyzed a rate-based congestion control mechanism called Enhanced Proportional Rate Control Algorithm (EPRCA), which has been adopted by the ATM Forum as a standard for ABR traffic service. In addition to three types of switches suggested in EPRCA, a prioritized switch in which an RM cell is transmitted with high priority has been considered to analyze the dynamical behavior of ACR and the queue length at the switch in both steady and initial transient state. Furthermore, we have shown the maximum queue length for all switches and conditions that the switch does not fall into underutilized for the EFCI switch.

One problem of the rate-based congestion control exists in the initial transient state. When a number of connections start cell transmission simultaneously, the queue length grows very large. Its analysis is reported in the companion paper [8]. One of the most difficulties in the rate-based congestion control resides in that the active number of VC's (N_{VC}) is not known by the switch. If some appropriate method can be provided, however, the parameter setting at connection setup time would be easier based on the analytic result presented in this paper. For another research topic, fairness achieved by EPRCA should be validated.

REFERENCES

[1] K. K. Ramakrishnan and R. Jain, "A binary feedback scheme for congestion avoidance in computer networks," *ACM Transactions on Computer Systems*, vol. 8, no. 2, pp. 158–181, 1990.

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

[3] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Rate-based congestion control for ATM networks," *ACM SIGCOMM Computer Communication Review*, vol. 25, April 1995.

[4] J.-C. Bolot and A. U. Shankar, "Dynamical behavior of rate-based flow control mechanisms," *Computer Communication Review*, vol. 20, pp. 35–49, 4 1990.

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

[6] J. C. R. Bennett and G. T. D. Jardins, "Failure modes of the baseline rate based congestion control plan," *ATM Forum Contribution 94-0682*, July 1994.

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

[8] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Analysis of rate-based congestion control methods in ATM networks — part 2: Initial transient state analysis —," *to be presented at GLOBECOM '95*, 1995.