Parameter tuning of rate-based congestion control algorithm for ABR service class in ATM Networks

Hiroyuki Ohsaki, Masayuki Murata and Hideo Miyahara

Department of Information and Computer Sciences
Faculty of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka, Osaka 560, Japan

Phone: +81-6-850-6588
Fax: +81-6-850-6589
E-mail: oosaki@ics.es.osaka-u.ac.jp

Abstract

A rate-based congestion control algorithm has been developed and standardized in the ATM forum for ABR service class. In the standard, the behavior of source and destination end systems is specified by several control parameters such as RIF (Rate Increase Factor) and RDF (Rate Decrease Factor). In spite of the fact that the performance of the rate-based congestion control algorithm heavily depends on the selection of these control parameters, the selection method of parameters is not shown in the standard. In this paper, by extending our previous work, appropriate settings of rate-control parameters in the various circumstances are investigated. We first analyze the dynamical behavior of the rate-based congestion control for multiple groups of ABR connections with different propagation delays. Next, we evaluate the effect of CBR traffic on ABR connections. Simulation results for a multi-hop network configuration is also presented to exhibit the tradeoff relationship among cell loss probability, link utilization and fairness. Finally, the selection method of control parameters in the multi-hop network is proposed based on our analytic methods and simulation results.

1 Introduction

A rate-based congestion control algorithm has been standardized for ABR (Available Bit Rate) service class by the ATM forum [1, 2, 3, 4]. The target of the standard is an operation algorithm of both source and destination end systems. Although some example behaviors of intermediate switches are introduced in the standard activities [5, 6, 7], implementation issues regarding intermediate switches are left to manufactures. In this paper, we will focus on the simplest switch among them, which is referred to as an "EFCI bit setting switch" or a "binary switch". In the binary switch, the congestion is detected by a predefined threshold in the switch buffer. If the number of cells in the buffer exceeds this threshold value, it is recognized as congestion. See Subsection 2.1 for a more precise description of the switch behavior.

In the standard document [2], several control parameters are defined for controlling cell transmission at the source end system. These include RIF (Rate Increase Factor) and RDF (Rate Decrease Factor) that control rate increase and decrease envelopes. During a connection establishment process, the source end system negotiates these control parameters with the network. In [8, 9], we have shown that effectiveness of the rate-based congestion control is heavily dependent on a choice of control parameters. If control parameters are configured properly, the rate-based congestion control can achieve high performance (i.e., no buffer overflow, high link utilization and small cell delay). However, a selection method of control parameters has not been specified in the standard, and parameters should be determined intuitively unless a proper tool is provided.

In [10, 11, 12], we have shown the analytic method to determine an appropriate setting of control parameters including RIF, RDF and ICR (Initial Cell Rate) for a single-hop network configuration. In the analysis, we have assumed that all source end systems behave identically, and that they always have cells to transmit. Under these assumptions, we have derived conditions that control parameters should satisfy to achieve two main objectives: preventing cell loss and achieving full link utilization. Based on these results, we have proposed a simple guideline for parameter tuning at the ATM Forum [13]. In addition to obtain high performance (in terms of cell loss and link utilization), fairness among connections is also an important issue. Actually, each connection may have a different round-trip delay according to the network configuration. In such a case, fairness among connections may be degraded due to the different feedback delays. When another ABR connection is newly established in the network, the ramp-up time of this connection is also important. Furthermore, we need further investigation to apply our analysis to more general network configurations with multiple switches where each connection has the different number of hops and the different propagation delay.

We further need to consider existence of real-time applications such as audio and video in a multimedia network environment. Since these applications use CBR (Constant Bit Rate) or VBR (Variable Bit Rate) service class, multiple service classes co-exist in the network. For ABR service class to utilize the available bandwidth unused by CBR/VBR service class, CBR/VBR traffic should be given higher priority than ABR traffic at the switch to guarantee QoS (Quality of Service) requirements of CBR/VBR traffic. Namely, cells of ABR traffic are awaited in the switch buffer if a CBR/VBR cell exists in the switch buffer in the case that the switch has two logically independent buffers — one for CBR/VBR service class and the other for ABR service class [14]. In other words, the bandwidth available to the ABR service class is limited by CBR/VBR traffic. Therefore, when a CBR/VBR connection is newly added into the network, the bandwidth available to the ABR service class is suddenly decreased, which would give a serious effect on the performance of the ABR connections; the switch buffer for ABR cells may become overloaded for a while leading to a large queue buildup and eventually cell losses due to the buffer overflow.

In this paper, we focus on the two subjects. We first analyze the behavior of the rate-based congestion control for a single-hop network but each group of connections is allowed to have the different propagation delay. In [15], Blot et al. have analyzed a dynamical behavior of a rate-based congestion control for connections with different

propagation delays. However, their analytic model was quite simple and different from the rate-based congestion control standard [2]. Through numerical examples, we show the effect of control parameters on the ramp-up time of a new ABR connection. We also derive the maximum queue length at the switch after a new CBR connection is established in the network. In [16], we have investigated some aspects of the performance of the rate-based congestion control algorithm in the multimedia network environment through simulation experiments. The author in [17] has analyzed the effect of CBR traffic on the ABR service class. However, he has considered only the case where the CBR connection requires the bandwidth close to the link capacity, and his analytic model was different from the rate-based congestion control standard [2]. On the contrary, in the current paper, we treat a more general and realistic model where the CBR connection requires any portion of the link bandwidth, and derive the maximum queue length via the worst case analysis.

We next investigate an appropriate setting of control parameters for a multi-hop network configuration by simulation. In the simulation, we use the model with multiple connections with different numbers of hops. The main purpose is to evaluate the effect of two rate-control parameters (RIF and RDF) on the performance. In [18, 19], the authors have provided simulation results for several combinations of control parameters. In this paper, control parameters are chosen based on our analytic results. As performance measures, cell loss possibility, link utilization and fairness among connections are considered. We also validate how our analytic results of the single-hop model can be applied to generic network models.

The rest of this paper is organized as follows. Section 2 is devoted to the analysis of the rate-based congestion control algorithm, and it contains two parts: the analysis for multiple groups of ABR connections where each group has the different propagation delay, and the derivation of the maximum queue length at the switch resulting from a new CBR connection. In Section 3, we present several simulation results for a multi-hop network configuration with discussion of the applicability of our analysis to generic network configurations. In Section 4, we present some concluding remarks, and summarize a detailed guideline for determining control parameters for the ABR service class with a binary-mode switch.

2 Parameter Tuning for Single Binary-Mode Switch

In what follows, we first summarize the analytic results of dynamical behavior of the rate-based congestion control obtained in [10, 11, 12] in Subsection 2.1. In Subsection 2.2, we then extend our analysis to the model in which multiple groups of connections with different propagation delays are allowed, and focus on a fairness problem by adding a newly established ABR connection. In Subsection 2.3, we analyze the effect of an additional CBR connection on the behavior of ABR connections in terms of the maximum queue length.

2.1 Analysis for The Homogeneous Connections: Summary

The analytic model consists of homogeneous traffic sources and a single bottleneck link as shown in Fig. 1. The number of active connections that share the bottleneck link is denoted by N_{VC}. The bandwidth of the bottleneck link is denoted by BW, and propagation delays between the source and the switch, and between the switch and the destination are defined by τ_{sx} and τ_{xd}, respectively. The round-trip propagation delay at the switch is denoted by $\tau(= 2\tau_{sx} + 2\tau_{xd})$. We further introduce $\tau_{xds}(= 2\tau_{xd} + \tau_{sx})$ as the propagation delay of congestion indication from the switch to the source via the destination. Propagation delays τ_{sx} and τ_{xd} are assumed to be identical for all source and destination pairs. In the analysis, it is also assumed that each source end system has infinite cells to transmit. Namely the permitted cell rate ACR (Allowed Cell Rate) is always equal to the actual cell transmission

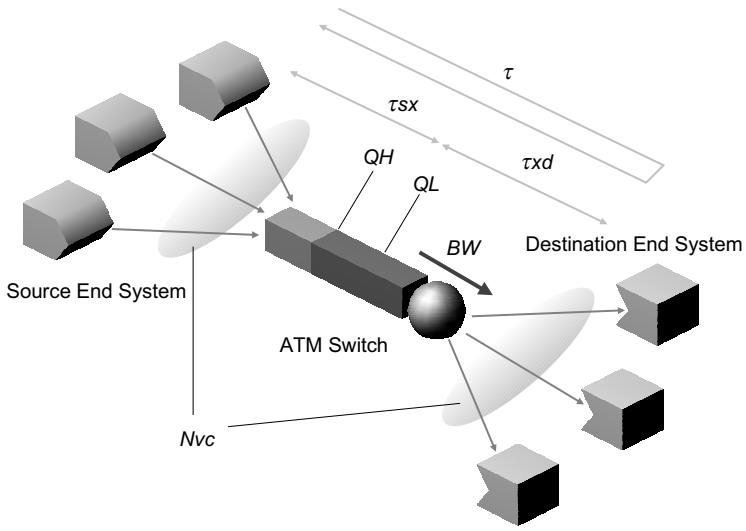

Figure 1: Analytic Model.

rate CCR (Current Cell Rate). Furthermore, all connections are assumed to start cell transmission simultaneously with the same rate-control parameters. We also assume that propagation delays τ_{sx} and τ_{xd} are the same for all connections. Thus, all connections behave identically. This assumption is, however, relaxed in Subsection 2.2 such that several groups of connections can have different propagation delays.

At the switch, congestion occurrence is detected by high and low threshold values associated with the queue length of the buffer. These threshold values are denoted by Q_H and Q_L, respectively. When the queue length exceeds the high threshold value Q_H, the switch detects congestion and notifies the source end system via the destination end system by marking an EFCI (Explicit Forward Congestion Indication) bit in the data cell, or by marking the CI (Congestion Indication) bit in the forward RM (Resource Management) cell. By receiving congestion information from the network, each source end system decreases its rate. After the queue length goes under the low threshold value Q_L, it is regarded as congestion termination. The EFCI bit in the data cell or the CI bit in the forward RM cell is then cleared to indicate congestion relief to source end systems via the destination. This sort of switch operation is often referred to as an "EFCI bit marking switch", or simply as a "binary switch" [3].

Behavior of source and destination end systems is standardized in the ATM forum, and is specified in [2]. Each source end system periodically generates a forward RM cell in proportion to its rate; that is, it sends one forward RM cell per $(N_{RM} - 1)$ data cells. On the receipt of the forward RM cell, the corresponding destination end system sends it back as the backward RM cell to the corresponding source end system. The allowed cell rate ACR of the source is then changed in response to the backward RM cell. When it receives the CI bit cleared cell, rate increase is performed as

$$ACR \leftarrow \min(ACR + RIF \times PCR, PCR), \tag{1}$$

where PCR (Peak Cell Rate) is a maximum rate for this connection negotiated at its connection setup, and RIF (Rate Increase Factor) is a ratio for proportional rate increase. The CI bit marked cell decreases ACR as

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

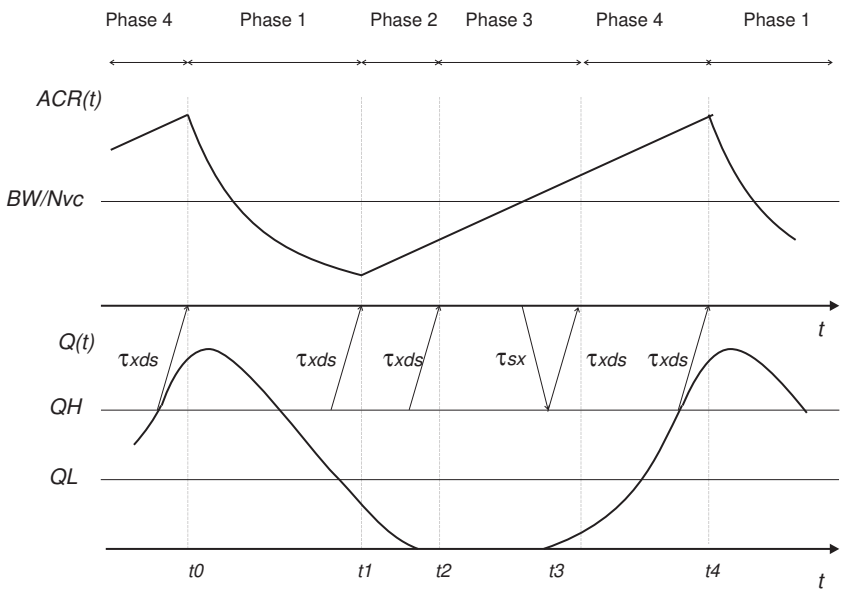

Figure 2: Pictorial View of $ACR(t)$ and $Q(t)$.

where MCR (Minimum Cell Rate) is a minimum rate (or guaranteed rate) for this connection, but in usual case MCR may be set to zero. RDF (Rate Decrease Factor) is a ratio for exponential rate decrease.

In the following analysis, the forward RM cell is not explicitly considered; congestion indication is performed by the EFCI bit of the data cell. However, the forward RM cell can easily be taken into account by replacing BW in our analysis with BW', which is defined as

$$BW' = BW\,\frac{N_{RM}-1}{N_{RM}}. \tag{3}$$

It is noted that we will use BW' instead of BW in numerical examples presented in Subsections 2.2 and 2.3.

Let us introduce $ACR(t)$ and $Q(t)$ as the allowed cell rate ACR of each source and the queue length at the switch observed at time t, respectively. In what follows, we present analytic results for $ACR(t)$ and $Q(t)$ that are taken from our previous work [10, 11, 12].

As shown in Fig. 2, $ACR(t)$ and $Q(t)$ have periodicity in steady state. The behavior of $ACR(t)$ and $Q(t)$ are divided into four phases called Phase 1 through Phase 4. We introduce $ACR_i(t)$ and $Q_i(t)$ as $ACR(t)$ and $Q(t)$ in Phase i, which are defined as

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

where t_i is the time when Phase i terminates. For simplicity, we further introduce $t_{i-1,i}$ as the interval of Phase i, which is defined as $t_i - t_{i-1}$. $ACR_i(t)$ is given by the following equations.

$$ACR_1(t) = ACR_1(0)e^{-\frac{BW\,RDF}{N_{VC}\,N_{RM}}t} \tag{4}$$

$$ACR_2(t) = ACR_2(0) + \frac{BW\,RIF\,PCR}{N_{VC}\,N_{RM}}t \tag{5}$$

$$
\begin{aligned}
ACR_3(t) &\cong ACR_3(0)e^{\frac{RIF\,PCR}{N_{RM}}t} && (6)\\
ACR_4(t) &= ACR_4(0) + \frac{BW\,RIF\,PCR}{N_{VC}\,N_{RM}}t && (7)\\
0 \leq \quad t &< t_{i-1,i}
\end{aligned}
$$

The evolution of $ACR(t)$ and $Q(t)$ are finally obtained by determining the initial value $ACR_i(0)$ and the duration of Phase i $t_{i,i+1}$. Given initial rates of Phase i, $Q_i(t)$ is obtained as

$$
Q_i(t) = Q_i(\tau_{sx}) + \int_{\tau_{sx}}^{t} \max(N_{VC}\,ACR_i(x-\tau_{sx}) - BW, 0)dx, \quad \tau_{sx} \leq t < \tau_{sx} + t_{i-1,i}.
$$

The duration of Phase i, $t_{i-1,i}$, is obtained as

$$
t_{i-1,i} = \begin{cases} Q_1^{-1}(Q_L) + \tau_{xds} & i = 1 \\ \min(Q_2^{-1}(Q_H) + \tau_{xds}, Q_2^{-1}(0) + \tau_{xds}) & i = 2, 4 \\ ACR_3^{-1}(BW/N_{VC}) + \tau & i = 3 \end{cases} \tag{8}
$$

where $ACR_i^{-1}(t)$ and $Q_i^{-1}(t)$ are defined as the inverse representations of $ACR_i(t)$ and $Q_i(t)$, respectively.

2.2 Analysis for Multiple Groups of Connections

In this subsection, we derive the dynamical behavior of the rate-based congestion control for N groups of connections with different propagation delays. Through numerical examples, we show the importance of parameter tuning for achieving good fairness and the short ramp-up time for an additional ABR connection.

2.2.1 Analysis

We divide ABR connections into N groups with different propagation delays. Within a group, connections have identical propagation delays. Figure 3 depicts our analytic model in the case of $N = 2$. Propagation delays from each source to the switch, and from the switch to each destination of group n ($1 \leq n \leq N$) are denoted by τ_{sxn} and τ_{xdn}, respectively. For brevity, we introduce $\tau_n (= 2\tau_{sxn} + 2\tau_{xdn})$ and $\tau_{dsn} (= \tau_{sxn} + 2\tau_{xdn})$. The number of connections in group n is denoted by N_{VCn}. Thus, we have a relation:

$$
N_{VC} = \sum_{n=1}^{N} N_{VCn}
$$

We assume that all connections in each group behave identically. Namely, all connections in each group have the same control parameters. Let us introduce RIF_n, RDF_n and N_{RMn} as RIF, RDF and N_{RM} of group n, respectively. We also assume $\tau_{sxi} \leq \tau_{sxj}$ and $\tau_{xdi} \leq \tau_{xdj}$ for any i and j ($i < j$) without loss of generality.

Let us introduce $ACR^n(t)$ and $Q(t)$ that represent ACR of the source end system in group n and the queue length at the switch observed at time t, respectively. As with the case in Subsection 2.1, $ACR^n(t)$ and $Q(t)$ have periodicity (see Fig. 4 for a pictorial view of $ACR^n(t)$ and $Q(t)$ for $N = 2$). We further introduce $ACR_i^n(t)$ and $Q_i(t)$ as the $ACR^n(t)$ and $Q(t)$ in Phase i, which are defined as

$$
\begin{aligned}
ACR_i^n(t) &= ACR^n(t - t_{i-1}),\\
Q_i(t) &= Q(t - t_{i-1}).
\end{aligned}
$$

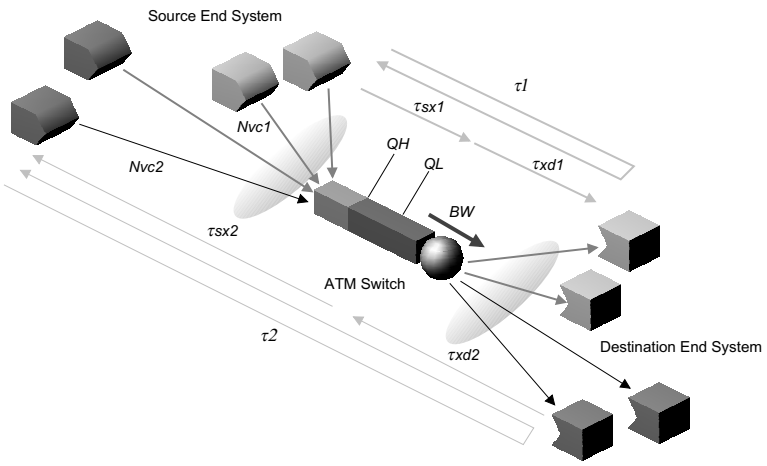

Figure 3: Analytic Model for Multiple Groups for $N = 2$.

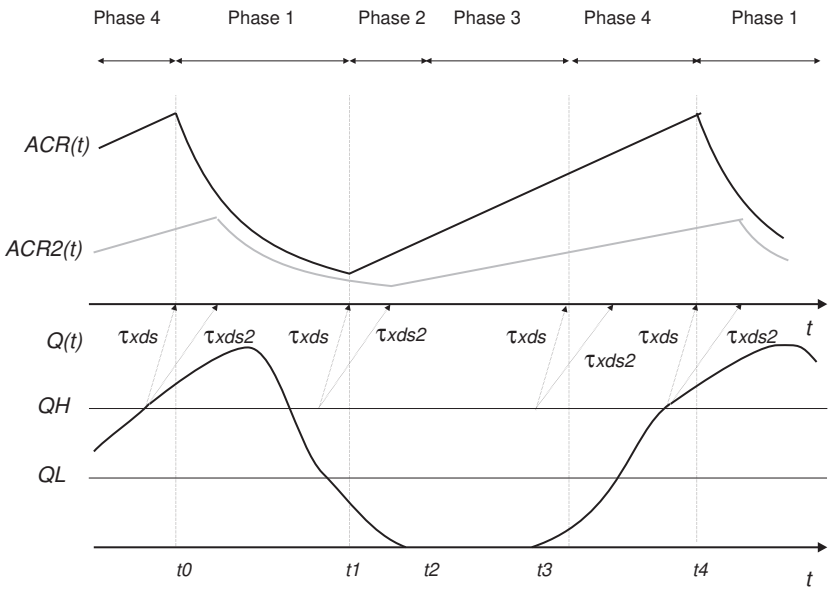

Figure 4: Pictorial View of $ACR^n(t)$ and $Q(t)$.

Because of the difference in propagation delays between the switch and the source via the destination (τ_{xdsn}), congestion information from the switch arrives at the sources at different time. Hence, $ACR_i^n(t)$ is obtained as follows (see Eqs.(4)–(7)).

$$\begin{aligned}
ACR_1^n(t) &= ACR_1^n(\tau_{xdsn} - \tau_{xds1})e^{-\frac{BW\,RDF_n}{N_{VC}\,N_{RMn}}(t-(\tau_{xdsn}-\tau_{xds1}))} \\
ACR_2^n(t) &= ACR_2^n(\tau_{xdsn} - \tau_{xds1}) + \frac{BW\,RIF_n\,PCR}{N_{VC}\,N_{RMn}}(t - (\tau_{xdsn} - \tau_{xds1})) \\
ACR_3^n(t) &\cong ACR_3^n(\tau_{xdsn} - \tau_{xds1})e^{\frac{RIF_n\,PCR}{N_{RMn}}(t-(\tau_{xdsn}-\tau_{xds1}))} \\
ACR_4^n(t) &= ACR_4^n(\tau_{xdsn} - \tau_{xds1}) + \frac{BW\,RIF_n\,PCR}{N_{VC}\,N_{RMn}}(t - (\tau_{xdsn} - \tau_{xds1}))
\end{aligned}$$

for

$$\tau_{xdsn} - \tau_{xds1} \leq \quad t \quad \leq \tau_{xdsn} - \tau_{xds1} + t_{i-1,i}.$$

At time t, the switch observes $ACR^n(t - \tau_{sxn})$ for group n because of the propagation delay from the source to the switch, τ_{sxn}. Therefore, $Q_i(t)$ in Phase i is obtained as

$$Q_i(t) = \max(Q_i(\tau_{sx1}) + \int_{\tau_{sx1}}^{t} (\sum_{n=1}^{N} N_{VCn}\,ACR_i^n(x - \tau_{sxn}) - BW), 0), \quad \tau_{sx1} \leq t < \tau_{sx1} + t_{i-1,i}.$$

The duration of Phase i, $t_{i-1,i}$, is simply obtained by replacing τ_{xds} in Eq. (8) with τ_{xds1}.

2.2.2 Numerical Examples

In this subsection, we provide several numerical examples. To exhibit the effect of the rate-control parameters on the ramp-up time of an additional ABR connection, we first add connections of group 1 in the network. After these connections are stabilized, another connection of group 2 with $PCR = BW$ and $ICR = PCR/20$ is established. The number of connections for each group is set to $N_{VC1} = 10$ for group 1 and $N_{VC2} = 1$ for group 2. We fixed the bandwidth of bottleneck link BW at 353.7 cell/ms assuming 150 Mbit/s ATM link. At the switch, its buffer size BL is assumed to be infinite for the purpose of obtaining the maximum queue length. Both high and low threshold values Q_H and Q_L are fixed at 150Kbyte. At each source end system, N_{RM} is set to 32.

We first examine the effect of the propagation delay on the ramp-up time. In Figure 5, we plot $ACR^n(t)$ and $Q(t)$ for $\tau_1 = \tau_2 = 0.02$ ms. In this figure, $RIF = 1/64$ and $RDF = 1/16$ (i.e., $RIF_n = 1/64$ and $RDF_n = 1/16$) are chosen to satisfy two objectives — preventing cell loss and achieving full link utilization — for connections of group 1 [12]. We add group 2 to the network when group 1 is at the beginning of Phase 1. In Figure 6, we change only the round-trip delay of group 2, τ_2, from 0.02 ms to 2.00 ms. In Table 1, we also show effective throughput normalized by the link capacity for connections in each group where τ_1 is fixed at 0.02 ms but τ_2 is varied as 0.02 ms, 0.20 ms and 2.00 ms. From these results, one can find that the difference in round-trip delays of group 2 has little effect on fairness and the rump-up time. For example, the ramp-up time in Fig. 6 is almost equivalent to Fig. 5. To show validity of our analysis, we show simulation results for the same parameters with Fig. 6 in Fig. 7. It can be found that our analytic results are quite close to simulation results.

The effect of RIF and RDF on the additional ABR connection is next investigated. Figure 8 shows the case where a larger value of RDF is used; that is, the rate decrease is faster than the case of Fig. 5. Here, $RDF = 1/4$ is used instead of 1/16 while $RIF = 1/64$ is unchanged. On the other hand, slower rate increase is considered in

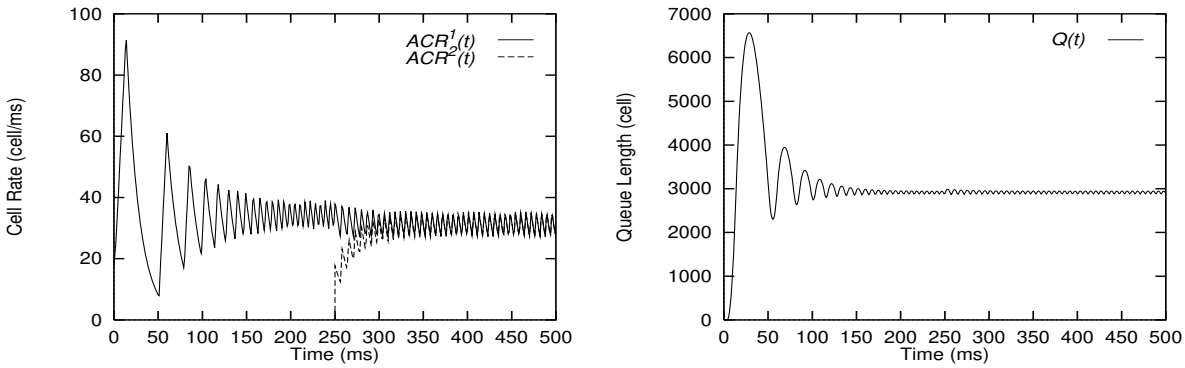

Figure 5: Effect of Propagation Delay for $\tau_1 = 0.02$ ms and $\tau_2 = 0.02$ ms.

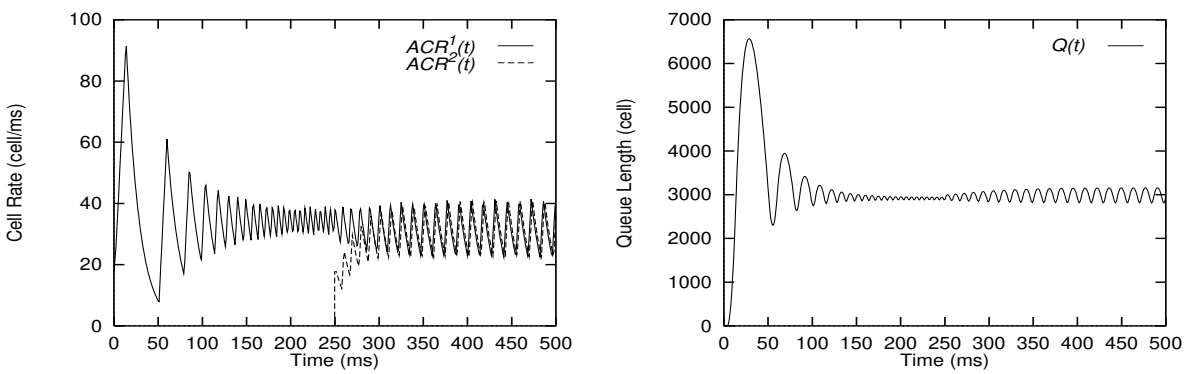

Figure 6: Effect of Propagation Delay for $\tau_1 = 0.02$ ms and $\tau_2 = 2.00$ ms.

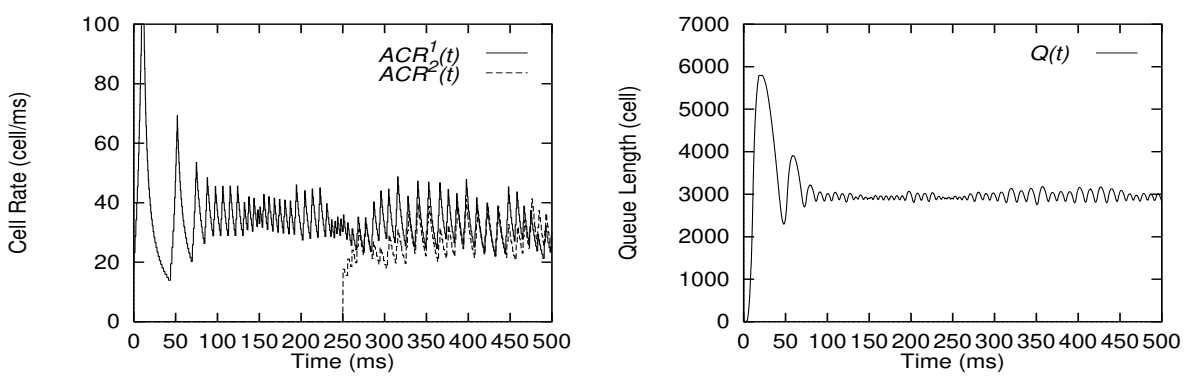

Figure 7: Simulation Results for $\tau_1 = 0.02$ ms and $\tau_2 = 2.00$ ms.

Table 1: Effective Throughput for Each Group.

Round-Trip Delay of Group 2 (τ_2)	Group 1	Group 2
0.02 ms	0.0880	0.0880
0.20 ms	0.0880	0.0880
2.00 ms	0.0882	0.0875

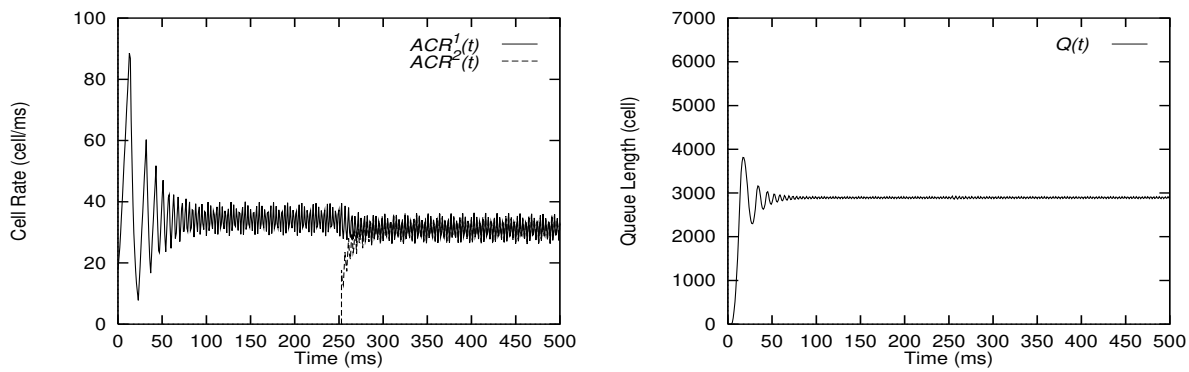

Figure 8: Effect of Control Parameters for $RIF = 1/64$ and $RDF = 1/4$.

Fig. 9 where we use $RIF = 1/256$ and $RDF = 1/16$. These parameter sets also prevent cell loss and achieve full link utilization. It can be found that the ramp-up time of group 2 is considerably affected by the setting of RIF and RDF. Namely, the ramp-up time becomes shorter by increasing RDF, and longer by decreasing RIF. Especially, the small value of RIF leads to much larger ramp-up time as can be observed in Fig. 9. Therefore, for fulfilling good responsiveness, RIF and RDF should be set to large values as long as no cell loss and full link utilization can be satisfied.

2.3 Effect of CBR Traffic

In this subsection, by extending analytic results provided in Subsection 2.1, we derive the maximum queue length at the switch when a CBR connection is newly established.

2.3.1 Analysis

We add a CBR connection to the model presented in Subsection 2.1 (see Fig. 1) at time t' with a fixed bandwidth $p \times BW (0 \leq p \leq 1)$. The available bandwidth to ABR traffic is therefore suddenly changed from BW to

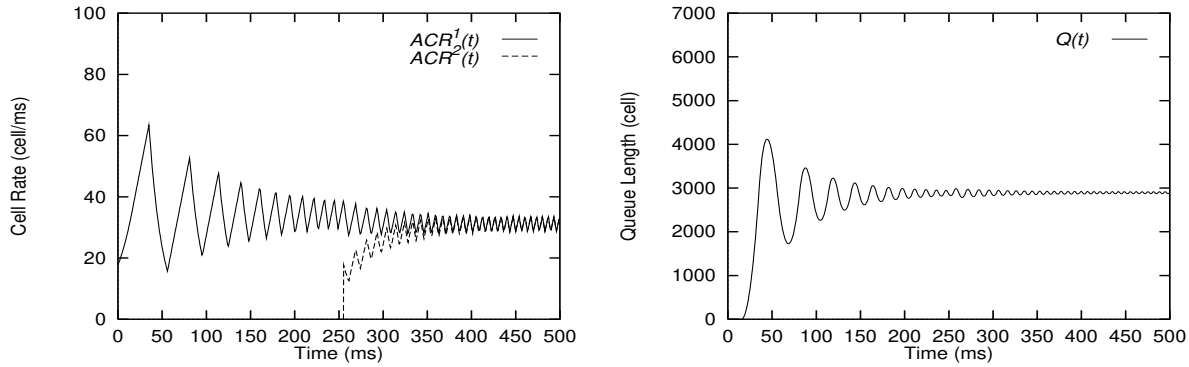

Figure 9: Effect of Control Parameters for $RIF = 1/256$ and $RDF = 1/16$.

$(1-p)BW$ at the time t'. Let us introduce Q_{max} as the maximum queue length after the establishment of the CBR connection at the time t'. First, Q_{max} is given by

$$Q_{max} = Q(t' + \tau_{sx}) + \int_{t'+\tau_{sx}}^{t'_{max}} \max(N_{VC} ACR'(x - \tau_{sx}) - (1-p)BW, 0)\, dx, \tag{9}$$

where $ACR'(t)$ is defined as the allowed cell rate ACR at time $t(\geq t')$, and t'_{max} is the time when $Q(t)$ takes its maximum value (see Fig. 10). Since $Q(t)$ starts to decrease again after τ_{sx} from when the aggregate cell rate of ABR connections is decreased to $(1-p)BW$, t'_{max} is obtained as

$$t'_{max} = ACR'^{-1}\left[\frac{(1-p)BW}{N_{VC}}\right] + \tau_{sx},$$

where $ACR'^{-1}(x)$ is the inverse representation of $ACR'(t)$.

After the time t', each source receives backward RM cells with a fixed interval since the switch has always cells in the buffer. By letting T_{RDF} be the interval of two successively received backward RM cells at the source end system, T_{RDF} is given by

$$T_{RDF} = \frac{N_{RM}\, N_{VC}}{(1-p)BW}.$$

However, when the arrival rate of the backward RM cell is too slow, each source end system decreases its rate by CDF (Cutoff Decrease Factor). In particular, when it receives no backward RM cell after transmitting the number C_{RM} of forward RM cells, it begins to reduce its ACR at each forward RM cell transmission as

$$ACR \leftarrow \max(ACR - ACR \times CDF, MCR). \tag{10}$$

The main purpose of the rate reduction mechanism introduced by C_{RM} and CDF is to allow the source end system to emit cells before receiving the first backward RM cell in its initial transient state [2]. Thus, C_{RM} may be set to a rather large value. However, as will be shown in numerical examples, this mechanism is also helpful to avoid cell loss for ABR connections caused by background traffic such as CBR traffic.

By letting T_{CDF} denote the duration of transmitting C_{RM} forward RM cells without receipt of backward RM cells, T_{CDF} is given by

$$T_{CDF} = \frac{N_{RM}\, C_{RM}}{ACR}.$$

According to the relation between T_{RDF} and T_{CDF}, $ACR'(t)$ is obtained as follows.

1. $T_{RDF} \leq T_{CDF}$; the source end system receives one or more backward RM cells before transmitting C_{RM} forward RM cells.

   In this case, $ACR'(t)$ is equivalent to $ACR_1(t)$ in Phase 1. Therefore, we have

   $$ACR'(t) = ACR(t')e^{-\frac{(1-p)BW\ RDF}{N_{VC}\ N_{RM}}(t-t')}.$$

2. $T_{RDF} > T_{CDF}$; no backward RM cell is received by the source end system before transmitting C_{RM} forward RM cells.

   After the time $(t' + T_{CDF})$, the source end system decreases its rate according to Eq.(10) for each forward RM cell transmission. Thus, we have a differential equation as

   $$\frac{dACR'(t)}{dt} = -\frac{(ACR'(t))^2\, CDF}{N_{RM}\, C_{RM}}.$$

   By solving this equation, we have

   $$ACR'(t) = \begin{cases} ACR(t'), & t' \leq t < t' + T_{CDF} \\ \left[\frac{CDF}{N_{RM}}(t-t') + \frac{1}{ACR(t')}\right]^{-1}, & t' + T_{CDF} \leq t \end{cases}$$

   Actually, the backward RM cell arrives at the source end system at $t = t' + T_{RDF}$, and it decreases ACR by RDF. In the above analysis, we ignored the rate reduction by receiving backward RM cells at the source end system since the arrival rate of backward RM cells is slow enough, and RDF is usually smaller than CDF. Furthermore, even in the case where RDF is not small compared with CDF, our analysis gives the upper-bound of the maximum queue length.

As can be found from Eq. (9), Q_{max} depends on the initial values such as $Q(t'+\tau_{sx})$ and $ACR'(t')$ that further depends on time t'. In what follows, we derive the maximum of Q_{max} for any t', which is defined as

$$Q'_{max} = \max_{t'}(Q_{max}). \tag{11}$$

As shown in Fig. 2, ACR takes its maximum value at the end of Phase 4 (at the beginning of Phase 1). In addition, $ACR(t')$ is maximized when the switch is not fully utilized since the large amplitude of $Q(t)$ means the large amplitude of $ACR(t)$. Therefore, Q'_{max} is obtained by setting $t' = t_4$, and by giving initial values of Phase 4 as

$$\begin{aligned} ACR(t_3) &= \frac{BW}{N_{VC}}, \\ Q(t_3 + \tau_{sx}) &= 0. \end{aligned}$$

At last, we note that the maximum queue length Q'_{max} is given by a closed-form equation.

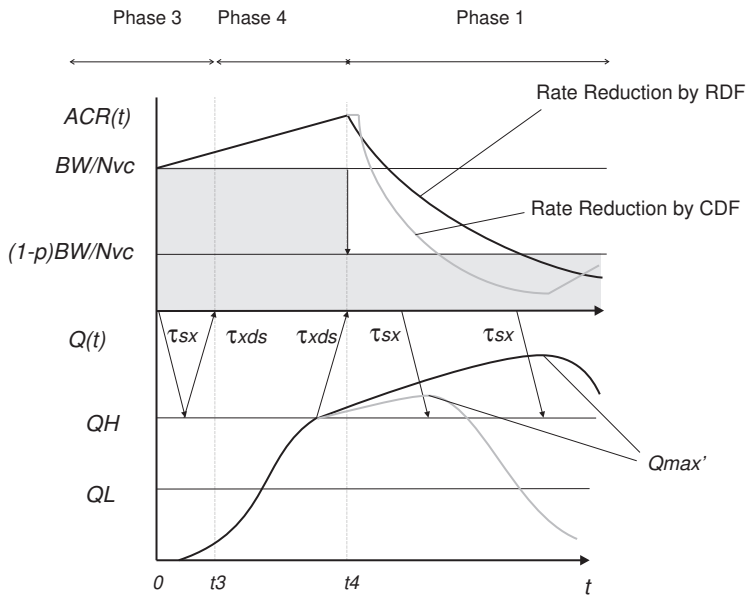

Figure 10: Pictorial View of $ACR(t)$ and $Q(t)$ with CBR Traffic.

2.3.2 Numerical Examples

In the following numerical examples, both τ_{sx} and τ_{xd} are fixed at 0.005 ms (about 1 km) as a typical value of the LAN environment. Furthermore, the number of ABR connections N_{VC} is set to 10. For other control parameters except RIF and RDF, we use the same values employed in Subsection 2.2.

We first show the maximum queue length Q'_{max} obtained by Eq.(11) as a function of p in Fig. 11. In this figure, RIF is fixed at 1/64, and C_{RM} and CDF is at 32 and 1/2, respectively, while RDF is varied as 1/4, 1/16 and 1/64. It can be found that Q'_{max} increases as p increases at first. For example, once a CBR connection that requires a half of the link bandwidth (75Mbit/s, in this case) is added, the switch should have 17,000 cells of buffer capacity to avoid cell loss of ABR connections with $RDF = 1/16$. Then, Q'_{max} is suddenly reduced around $p = 0.9$. It is because the source end system decreases its rate by CDF rather than RDF when the available bandwidth for ABR connections becomes too small. Moreover, one can find that the maximum queue length can be reduced by setting RDF to a large value (i.e., faster rate decrease). In Fig. 12, RIF is changed from 1/64 to 1/1024, which means slower rate increase. In this figure, the maximum queue length is decreased to some extent when compared with Fig. 11. However, a large amount of buffer capacity is still required to prevent cell loss if p is large.

By setting C_{RM} properly, cell loss can be prevented even when the CBR connection reserves the bandwidth close to the link capacity as shown in Figs. 13 and 14. In these figures, as with the previous examples, RIF is set to 1/64 and 1/1024, respectively. However, C_{RM} that decides the duration to rate reduction by CDF is changed from 32 to 4. These figures show that the maximum queue length can be limited even when p becomes large. For example, 12,000 cells of the buffer capacity is sufficient for preventing cell loss with $RDF = 1/16$ even when the CBR connection requires the entire bandwidth.

We plot Q'_{max} as the functions of C_{RM} and p in Figs. 15 and 16. In these figures, RIF is set to 1/64 and 1/1024, respectively, while RDF is fixed at 16 in both cases. The z-axis is ranged from 0 to 20,000 cells. As can be found from these figures, C_{RM} should be set to be a smaller value to avoid cell losses completely for any traffic

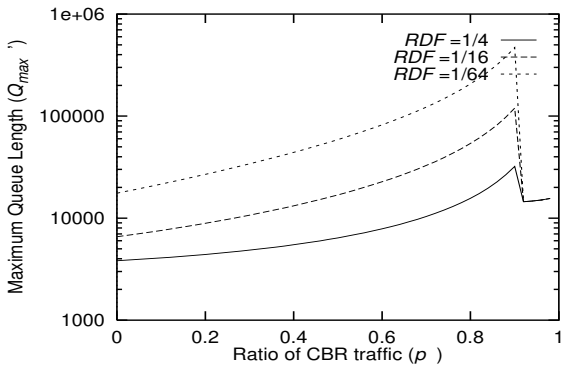

Figure 11: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/64$ and $C_{RM} = 32$.

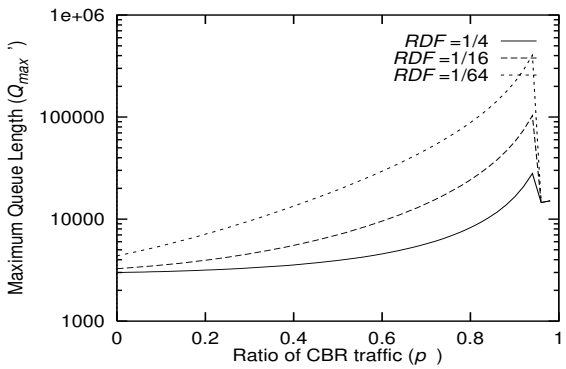

Figure 12: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/1024$ and $C_{RM} = 32$.

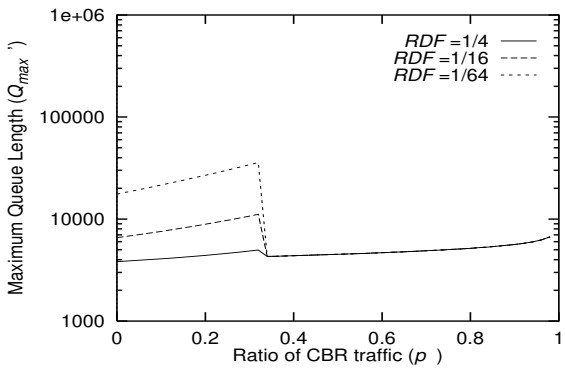

Figure 13: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/64$ and $C_{RM} = 4$.

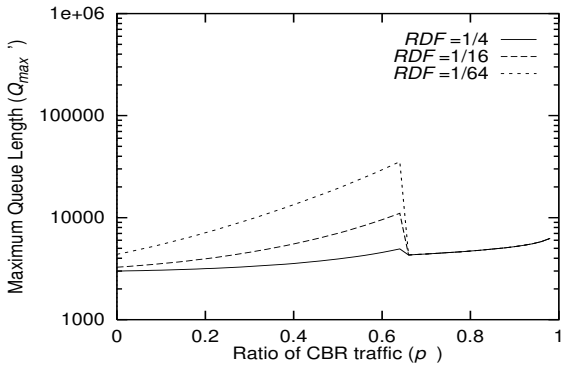

Figure 14: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/1024$ and $C_{RM} = 4$.

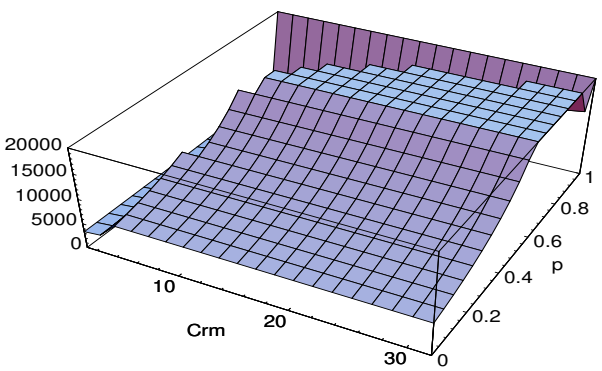

Figure 15: The Maximum Queue Length for $RIF = 1/64$ and $RDF = 1/16$.

load of the CBR connection. However, by setting RIF to a smaller value such as 1/1024, the queue buildup may be limited to some degree in the region where p is not large. Therefore, we can conclude that to limit the queue buildup by a new CBR connection, each of RIF and RDF should be small and large, respectively. Moreover, a smaller value of C_{RM} is helpful to prevent cell loss.

3 Simulation Results of Multi-Hop Network Configuration

In this section, we investigate an appropriate setting of rate-control parameters for a generic network configuration by simulation. In Subsection 3.1, we first introduce our simulation model. In Subsection 3.2, we then show simulation results to discuss the robustness of the rate-based congestion control in terms of cell loss, link utilization and fairness.

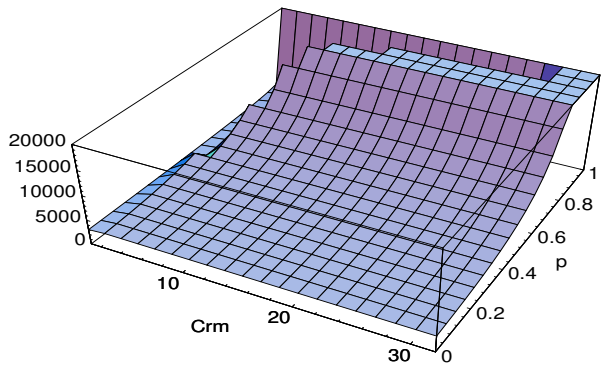

Figure 16: The Maximum Queue Length for $RIF = 1/1024$ and $RDF = 1/16$.

3.1 Simulation Model

Figure 17 illustrates our simulation model that is commonly referred to as the "parking lot" configuration [20, 4]. This model consists of five interconnected switches and four connections with different numbers of hops. The connection VCn is established from SESn to DESn. Each VCn enters the network at SWn, and all exit from SW5. Since each connection has the different number of hops, unfairness among these connections may be caused, which is our main concern in this section. Note that the link between SW4 and SW5 possibly becomes bottleneck in this model. The operation algorithm of source and destination end systems follows the standard draft [2]. Each source end system is assumed to always have cells to transmit, by which we evaluate robustness of the rate-based control in the worst case condition.

Bandwidth of all links is fixed at 150Mbit/s (353.7 cell/ms), and propagation delays between the source and the switch, τ_{sx}, and between the destination and the switch, τ_{xd}, are also fixed at 0.00 1ms (about 0.2 km). On the other hand, propagation delays between two interconnected switches, τ_{xx}, are 0.01 ms or 1.00 ms (about 2 km and 200 km, respectively) as values for LAN and WAN environments. For intermediate switches, we model the binary mode switch with the FIFO scheduling, and provide 300 Kbyte (5,796 cells) of the buffer. Upper and lower threshold values in the buffer, Q_H and Q_L, are fixed at the half of the buffer size. Other control parameters used in our simulation are $PCR = 150$ Mbit/s, $MCR = PCR/1000$, $ICR = PCR/20$, $TCR = 0.01$ cell/ms, $Mrm = 2, Trm = 100$, $C_{RM} = 32$, $CDF = 1/2$ and $TOF = 2$. Refer to [2] for the description of these control parameters.

As we have shown in [11, 12], key parameters that determine the efficiency and stability of the rate-based congestion control are RIF and RDF. In these papers, we have analytically derived two boundary conditions for RIF and RDF to prevent cell loss and achieve full link utilization for the same model presented in Subsection 2.1. In [13], we have proposed a guideline for parameter tuning based on our analytic results and simulation experiments. Here, we summarize this guideline.

1. Estimate the round-trip delay, τ, and the number of active connections, N_{VC}, in the worst case condition.

2. Obtain two boundary conditions for preventing cell loss and achieving full link utilization for these parameters from our analysis [11, 12].

3. Set RDF to be a smaller value than 1/8, and determine RIF that satisfies the condition of preventing cell loss.

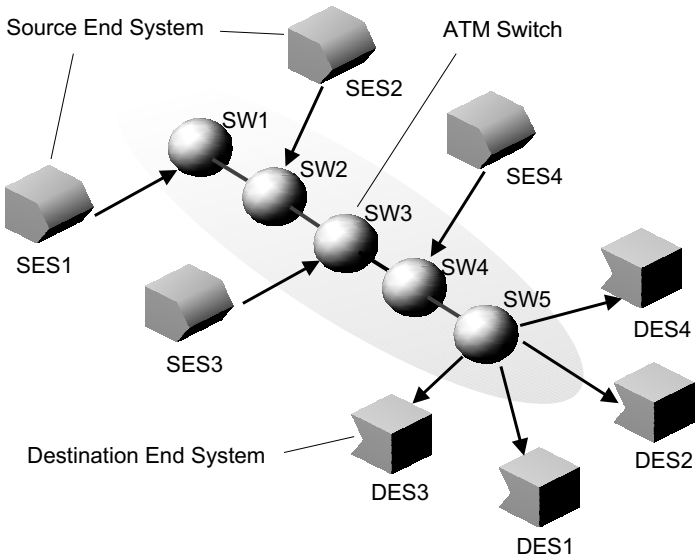

Figure 17: Parking Lot Configuration.

In our simulation, the number of active connections is set to be constant but the round-trip delay and the number of hops for each connection is varied. Thus, it is impossible to directly apply our analysis to the simulation model. In what follows, we investigate how our analytic methods, which is for a single-hop model and homogeneous sources, should be applied to a more generic model.

One problem is in determination of the round-trip delay, τ, that is used for obtaining two boundary conditions in our analysis. As we have shown in Subsection 2.2, the difference in propagation delays of connections has little effect on fairness. However, cell loss probability and link utilization are affected by the propagation delay since the larger round-trip delay implies the larger feedback delay [11, 12]. By letting τ_n be the round-trip delay for the nth connection, VCn, we consider three schemes for determining τ being applied to our analysis as follows.

Scheme 1: Adjust to the shortest connection

This scheme tunes parameters for the connection with the shortest round-trip delays. Thus, by assuming that VC1 has the shortest round-trip delay, we simply have

$$\tau = \tau_1$$

Scheme 2: Adjust equally to all connections

This scheme determines τ as an average of propagation delays of all connections. Thus, we have

$$\tau = \frac{1}{N_{VC}} \sum_{n=1}^{N_{VC}} \tau_n$$

Scheme 3: Adjust to the longest connection

This scheme is the opposite of Scheme 1; that is, parameters are tuned for the longest connection. Thus, by assuming that VCN has the longest round-trip delay, we have

$$\tau = \tau_N$$

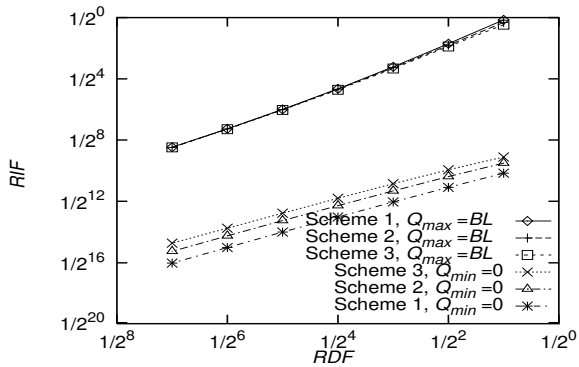

Figure 18: Analytic Results for Appropriate Parameters for $\tau_{xx} = 0.01$.

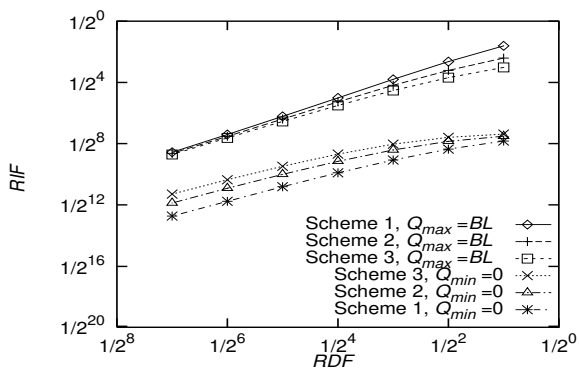

Figure 19: Analytic Results for Appropriate Parameters for $\tau_{xx} = 0.10$.

To compare these schemes, we plot two boundary lines for preventing cell loss and achieving full link utilization for $\tau_{xx} = 0.01$ ms, $\tau_{xx} = 0.10$ ms and $\tau_{xx} = 1.00$ ms in Figs. 18 through 20, respectively. N_{VC} is fixed at 4, and BL is 300 Kbyte to conform to the simulation parameters. The line labeled by "$Q_{max} = BL$" is the upper-bound of control parameters for preventing cell loss; that is, by selecting RIF and RDF from the lower region of this line, cell loss can be avoided. On the contrary, full link utilization can be fulfilled by selecting RIF and RDF from the upper region of the line labeled by "$Q_{min} = 0$". Hence, for preventing cell loss and achieving full link utilization, RIF and RDF should be chosen from the region between these two curves. One can find from these figures that when the round-trip delay is small, the boundary line for preventing cell loss (the "$Q_{max} = BL$" line) is nearly independent of schemes. However, the boundary line for full link utilization (the "$Q_{min} = 0$" line) is affected by schemes especially when the round-trip delay is large. Thus, for simulation of a WAN environment, we compare these three schemes although only Scheme 3 is used for simulation of a LAN environment.

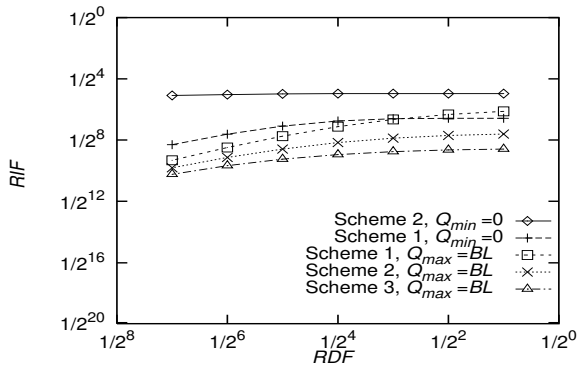

Figure 20: Analytic Results for Appropriate Parameters for $\tau_{xx} = 1.00$.

Table 2: Values of (RIF, RDF) for LAN Environment.

	fast down	moderate down	slow down
$Q_{max} = BL$	(1/4, 1/4)	(1/32, 1/16)	(1/256, 1/64)
$\vdots$	(1/32, 1/4)	(1/256, 1/16)	(1/1024, 1/64)
$Q_{min} = 0$	(1/512, 1/4)	(1/2048, 1/16)	—

3.2 Simulation Results

3.2.1 Case of LAN Environment

In this subsection, we show simulation results for a small propagation delay, $\tau_{xx} = 0.01$ ms, as a LAN environment. As described in the previous subsection, Scheme 3 is used for determining τ. However, in this subsection, we use three values of RIF for a given RDF to investigate how the values of RIF and RDF should be chosen form the region between two boundary lines. We first fix RDF to be 1/4 as a fast rate-down case. Then, three values of RIF (1/4, 1/32 and 1/512) are chosen from Fig. 18. That is, $RIF = 1/4$ is chosen from the "$Q_{max} = BL$" line for preventing cell loss, and $RIF = 1/512$ is from the "$Q_{min} = 0$" line for achieving full link utilization. $RIF = 1/32$ is chosen as the midst of these values. Note that $RIF = 1/4$ is slightly smaller than the "$Q_{max} = BL$" line, and $RIF = 1/512$ is larger than the "$Q_{min} = 0$" line because RIF and RDF is represented in a form of $1/2^n$ [2]. We also use $RDF = 1/16$ and 1/64 as moderate and slow rate-down cases, respectively. We summarize values of RIF and RDF used in this subsection in Table 2.

Figures 21 through 23 show the cell transmission rate of each connection and the queue length at the switch for $RIF = 1/4$, 1/32 and 1/512, respectively. It can be found from these figures that SW4 is fully utilized when $RIF = 1/4$ and 1/32, and that cell loss is prevented in all cases. However, fairness among connections is not fulfilled; longer connections (VC1 and VC2) transmit more cells than shorter connections (VC3 and VC4) (for

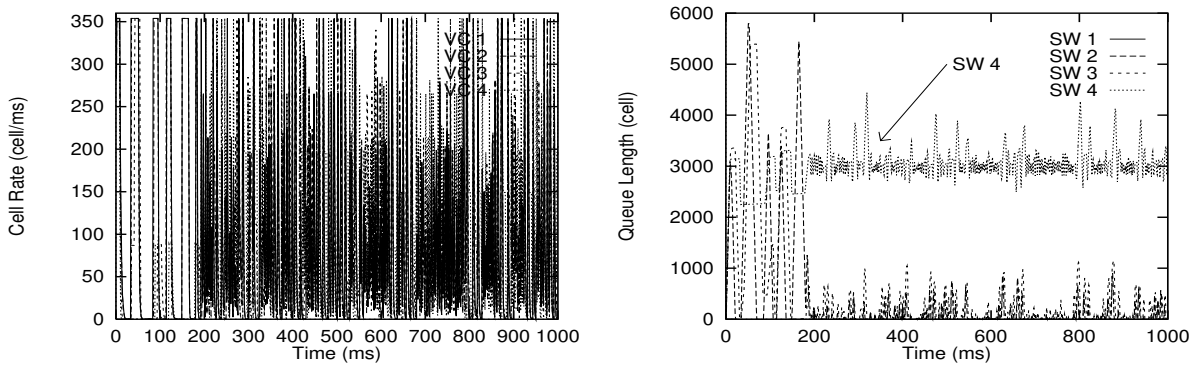

Figure 21: LAN Case for $RIF = 1/4$ and $RDF = 1/4$.

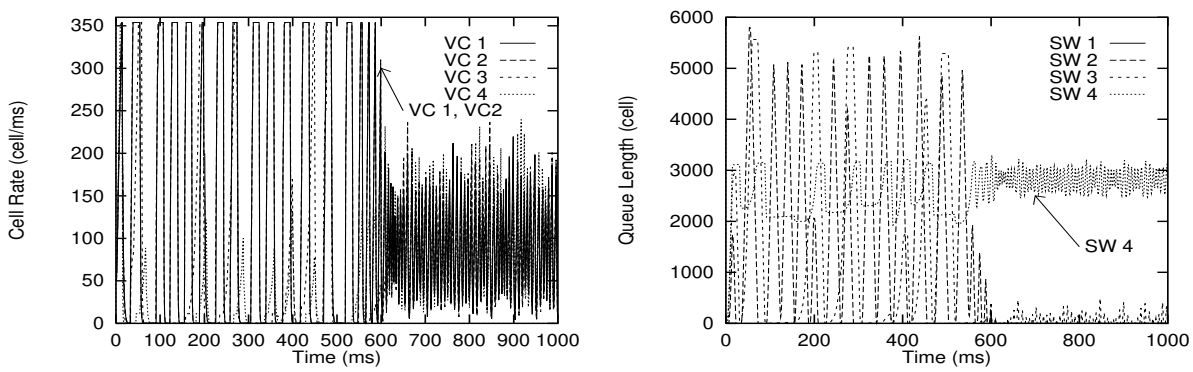

Figure 22: LAN Case for $RIF = 1/32$ and $RDF = 1/4$.

example, VC1 reaches PCR but VC4 does not in Fig. 21). it can be explained as follows. If SW4 becomes congested, each source decreases its rate by receiving backward RM cells of $CI = 1$. Because of different propagation delays, longer connections require more time to respond to congestion, and their ACR's remain high compared with shorter connections. Noting that an arrival rate of backward RM cells is proportional to its ACR, longer connections can receive much backward RM cells of $CI = 0$ after the congestion relief. Thus, longer connections can increase their ACR faster than the others, and it results in unfairness among connections.

We then change RDF to 1/16 for slower rate decrease (the third column of Table 2). Simulation results for $RIF = 1/32$, 1/256 and 1/2048 are plotted in Figs. 24 through 26, respectively. We choose the values of RIF similarly to the previous case. From the figures, we can observe that $RIF = 1/2048$ achieves good fairness as well as full utilization of the bottleneck link while $RIF = 1/256$ can do neither. Furthermore, some cells are lost when $RIF = 1/256$. Therefore, we conclude that too fast rate increase/decrease degrades fairness among connections and utilization of the bottleneck link.

Figures 27 and 28 show simulation results for $RIF = 1/256$ and 1/1024 when RDF is 1/64, which means much slower rate decrease. Both these figures indicate good fairness compared with the cases of $RDF = 1/4$

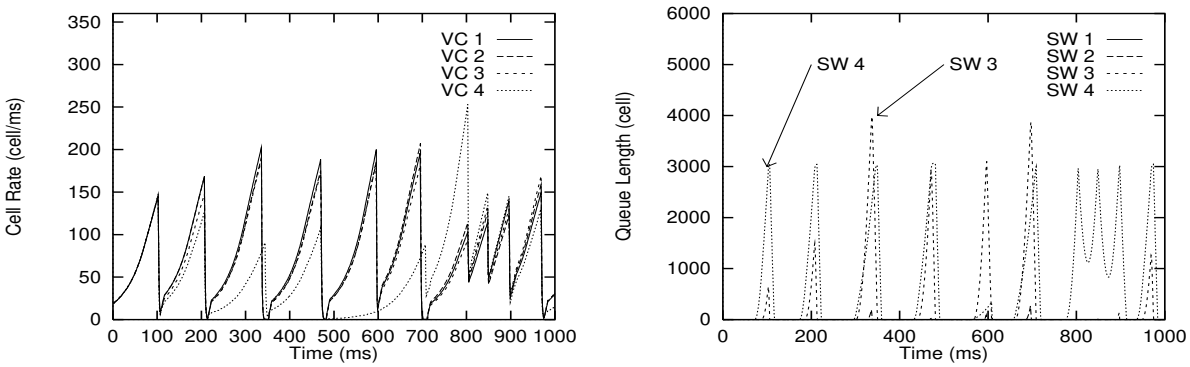

Figure 23: LAN Case for $RIF = 1/512$ and $RDF = 1/4$.

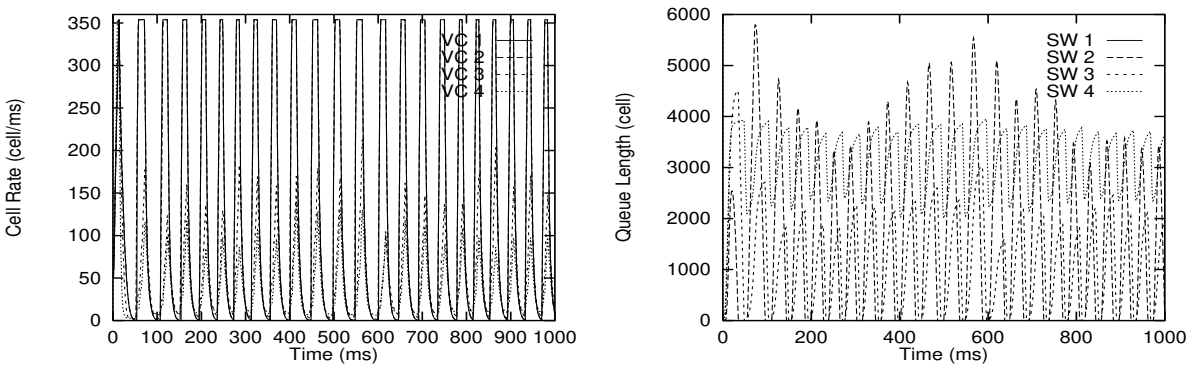

Figure 24: LAN Case for $RIF = 1/32$ and $RDF = 1/16$.

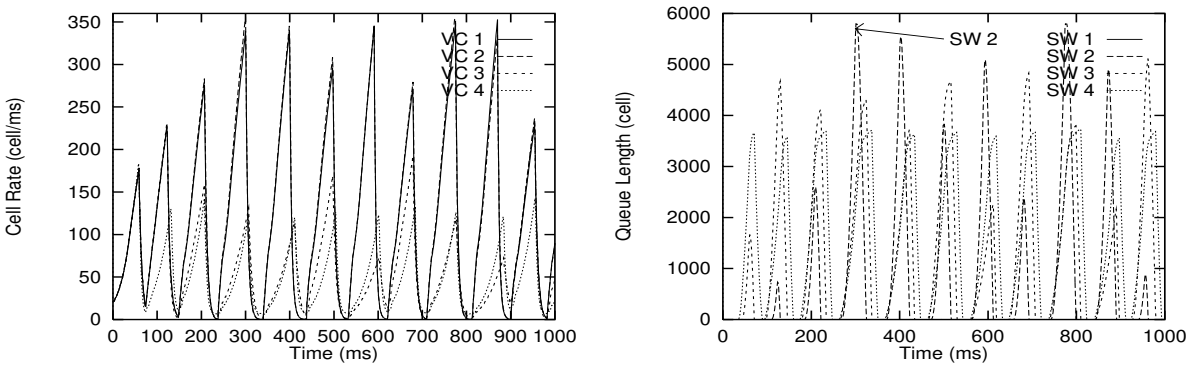

Figure 25: LAN Case for $RIF = 1/256$ and $RDF = 1/16$.

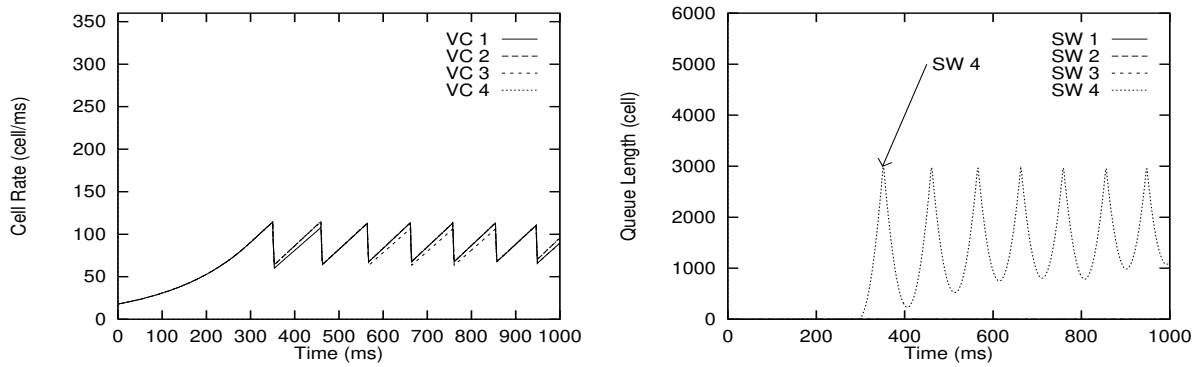

Figure 26: LAN Case for $RIF = 1/2048$ and $RDF = 1/16$.

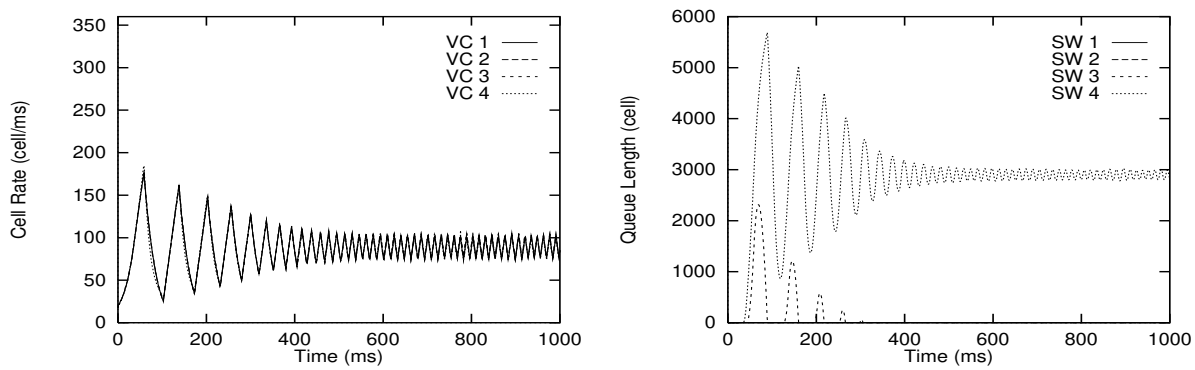

Figure 27: LAN Case for $RIF = 1/256$ and $RDF = 1/64$.

and 1/16. However, it should be noted that it takes longer time for each connection to be settled (around 500 ms in both cases). We conclude that a smaller value of RDF (i.e., slower rate decrease) is appropriate for achieving good fairness and stable operation, and that RIF should be chosen from the "$Q_{max} = BL$" line.

3.2.2 Case of WAN Environment

The objective in this subsection is to compare three schemes explained in Subsection 3.1, and to investigate an appropriate setting of RIF and RDF in the WAN environment. We set the propagation delay between switches, τ_{xx}, to 1.00 ms (about 200 km). As with the cases in Subsection 3.2.1, we use $RDF = 1/4$, 1/16 and 1/64. Then, for each of three schemes, we choose RIF from the "$Q_{max} = BL$" line in Fig. 20. The values of RIF and RDF are summarized in Table 3.

We first show simulation results for each scheme with $RDF = 1/4$ in Figs. 29 through 31. It is noted that in every scheme fairness among connections cannot be fulfilled, and that cell loss occurs regardless of RIF. In these figures, the case of $RIF = 1/128$ (Scheme 1) is at least better than others, but the use of these parameters should

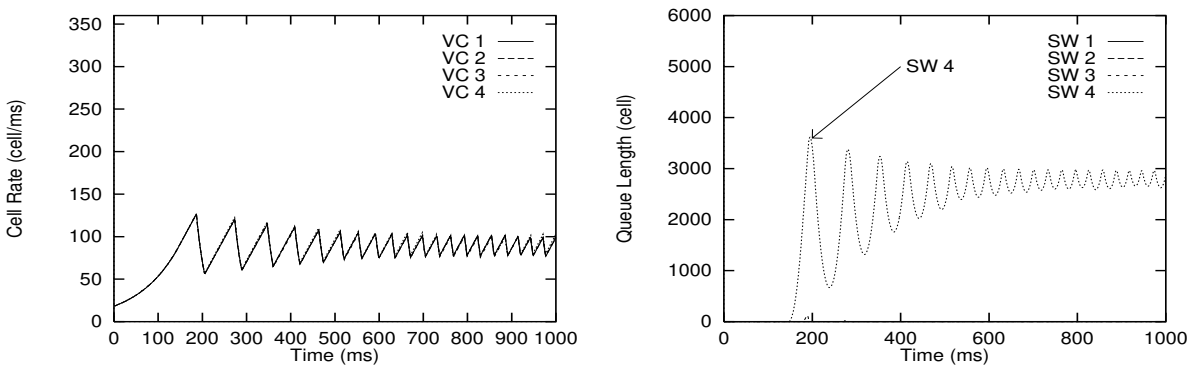

Figure 28: LAN Case for $RIF = 1/1024$ and $RDF = 1/64$.

Table 3: Values of (RIF, RDF) for WAN Environment.

	fast down	medium down	slow down
Scheme 1	(1/128, 1/4)	(1/256, 1/16)	(1/512, 1/64)
Scheme 2	(1/256, 1/4)	(1/512, 1/16)	(1/1024, 1/64)
Scheme 3	(1/512, 1/4)	(1/512, 1/16)	(1/1024, 1/64)

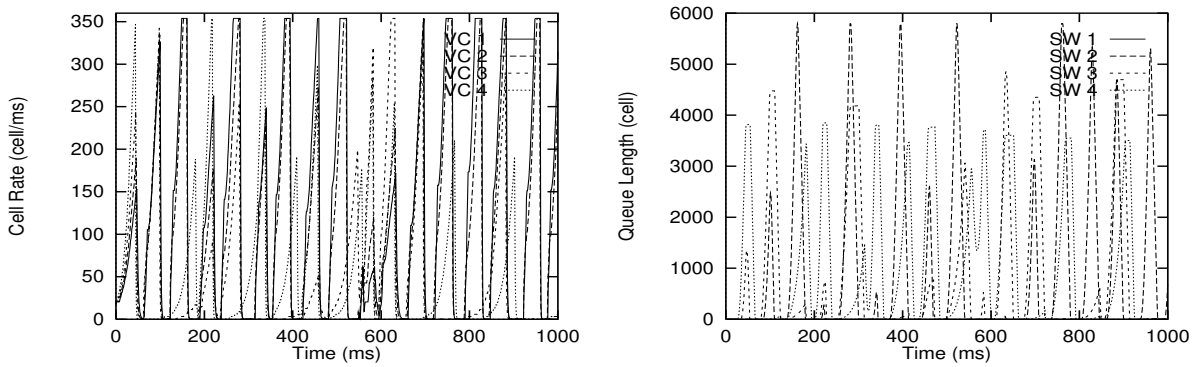

Figure 29: WAN Case for $RIF = 1/128$ and $RDF = 1/4$ (Scheme 1).

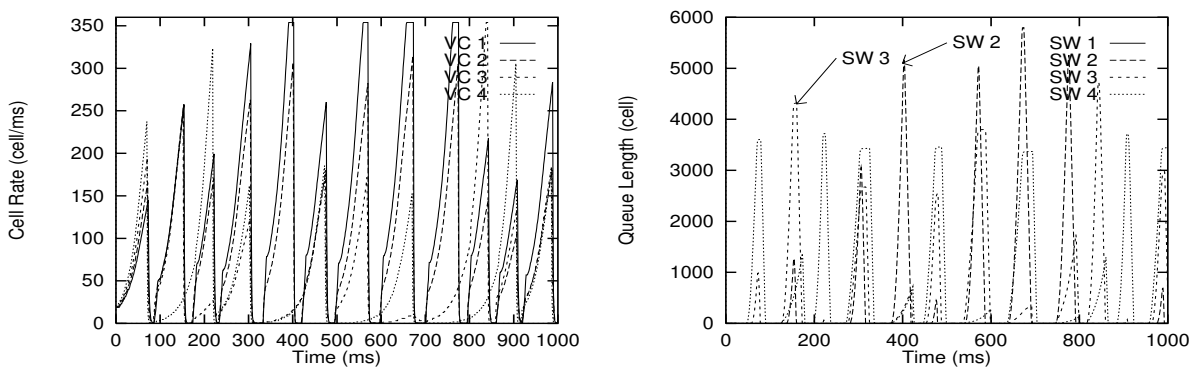

Figure 30: WAN Case for $RIF = 1/256$ and $RDF = 1/4$ (Scheme 2).

be avoided as explained later.

We next change RDF from 1/4 to 1/16, which means slower rate decrease, and plot simulation results for each scheme in Figs. 32 and 33. As can be found from these figures, there is little improvement over the cases of $RDF = 1/4$. Although cell loss can be avoided by setting $RIF = 1/512$ (Schemes 2 and 3), fairness among connections is not still accomplished.

Finally, we change the rate decrease to be much slower ($RDF = 1/64$). Results are shown in Figs. 34 (Scheme 1) and 35 (Schemes 2 and 3). It can be easily found that a fairness problem is dramatically improved compared with previous two cases with $RDF = 1/4$ and 1/16. Namely, when the rate decrease is slow as $RDF = 1/64$, every scheme shows good performance in terms of good fairness, no cell loss and full link utilization. However, it should be noted that the ramp-up time is quite large (i.e., around 100–200 ms), which indicates the binary-mode switch is unsuited for use in WAN environment. Finally, we conclude that rate-control parameters should be chosen from the "$Q_{max} = BL$" line with τ given by Scheme 2 and a smaller value of RDF (slow rate decrease) in multi-hop network configurations.

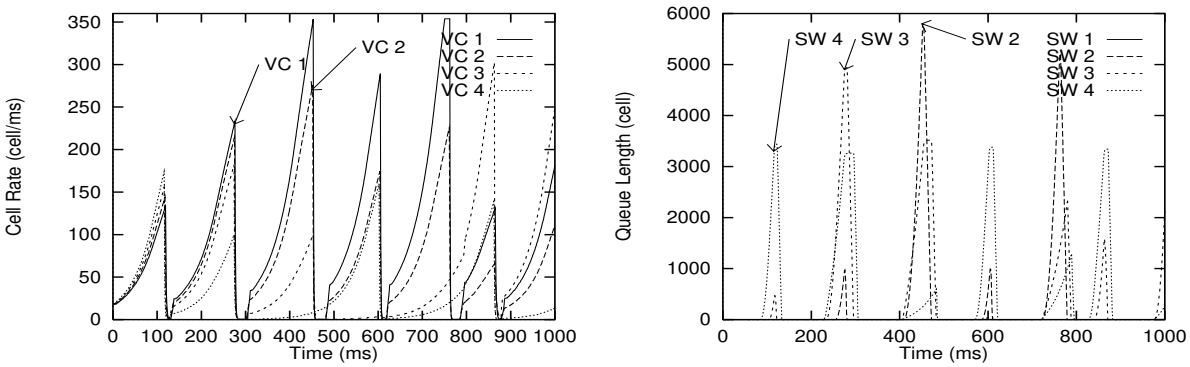

Figure 31: WAN Case for $RIF = 1/512$ and $RDF = 1/4$ (Scheme 3).

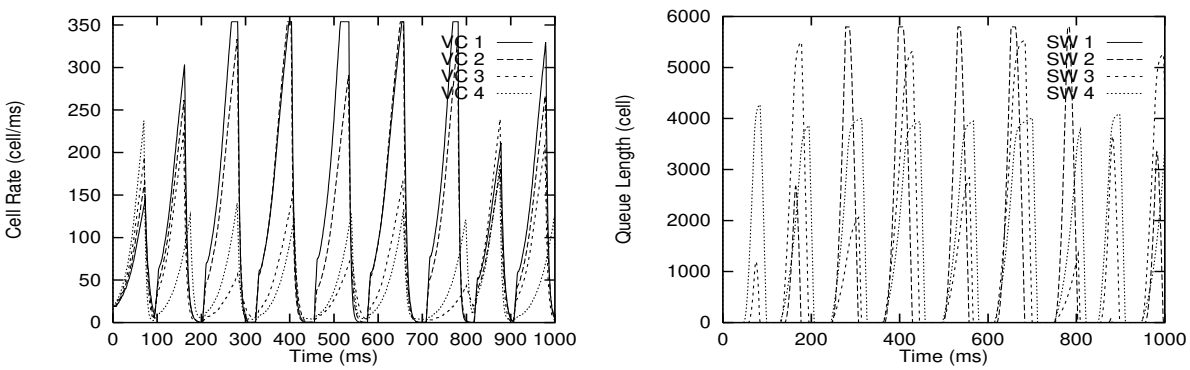

Figure 32: WAN Case for $RIF = 1/256$ and $RDF = 1/16$ (Scheme 1).

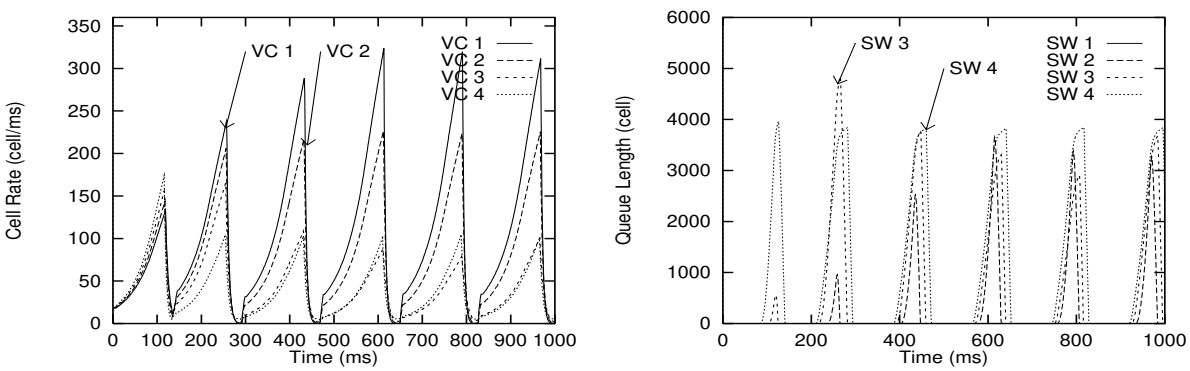

Figure 33: WAN Case for $RIF = 1/512$ and $RDF = 1/16$ (Schemes 2 and 3).

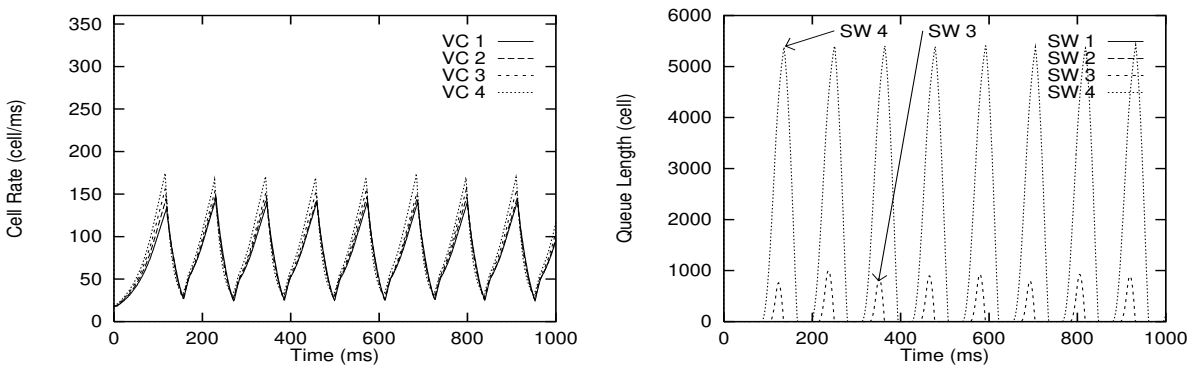

Figure 34: WAN Case for $RIF = 1/512$ and $RDF = 1/64$ (Schemes 1 and 2).

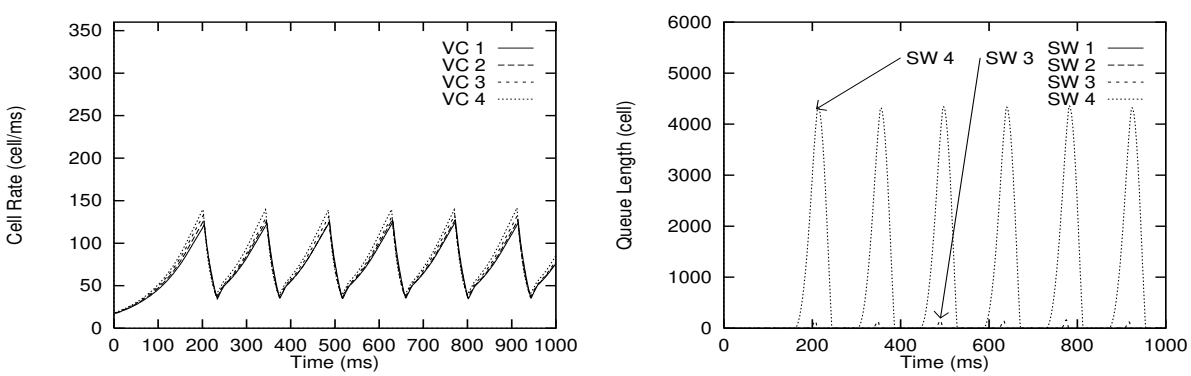

Figure 35: WAN Case for $RIF = 1/1024$ and $RDF = 1/64$ (Scheme 3).

4 Conclusion

In this paper, we have investigated an appropriate setting of control parameters for the rate-based congestion control with binary-mode switch. For this purpose, we have mainly focused on two rate-control parameters, RIF and RDF, which decides the envelope of rate increase/decrease.

First, we have presented two sorts of analyses. One was the analysis for the model with several groups of connections with different propagation delays in order to reveal the fairness problem among connections and the ramp-up time of an additional ABR connection. The other was the derivation of the maximum queue length at the switch buffer affected by an addition of background traffic such as CBR traffic. Through numerical examples, we have shown that a large value of RIF (i.e., fast rate increase) is helpful to shorten the ramp-up time, and that a small value of C_{RM} dramatically reduces the maximum queue length caused by CBR traffic.

Next, we have examined the appropriate setting of RIF and RDF by simulation experiments. As a simulation model, we have used the parking lot configuration having five interconnected switches and four connections with different numbers of hops. We have compared three schemes for applying our analysis to more generic network configurations. It has been shown that RDF should be set to a small value around 1/64 (i.e., slow rate decrease), and that RIF should be set to a large value as long as cell loss can be prevented — the maximum value of RIF is given by our analysis using Scheme 2 for parameter determination.

It should be noted that we have assumed PCR of each connection is identical to the link bandwidth, BW, in numerical examples. Since the amount of rate-increase is determined by $RIF \times PCR$ as in Eq. (1), RIF should be adjusted when PCR is not equal to BW as

$$RIF \leftarrow RIF \times \frac{PCR}{BW}. \tag{12}$$

At the end of this paper, we summarize the guideline for determining control parameters of the rate-based congestion control algorithm.

1. Estimate the number of active connections, N_{VC}, and their round-trip delays, τ_n.
2. Choose RDF around 1/64.
3. Calculate the average round-trip delay, τ, as
$$\tau = \sum_{n=1}^{N_{VC}} \frac{\tau_n}{N_{VC}}.$$
4. For these N_{VC}, τ, RDF and other given parameters, solve the equation $Q_{max} = BL$ in [10, 12] for RIF to obtain the maximum of RIF that can prevent cell loss.
5. Choose RIF smaller but closest to this solution.
6. Adjust RIF by Eq. (12) when PCR is not equal to BW.
7. C_{RM} can be set to a small value (for example, 2) for preventing buffer overflow caused by background traffic.

References

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

[2] The ATM Forum Technical Committee, "Traffic management specification version 4.0 (draft)," *ATM Forum Contribution 95-0013R10*, February 1996.

[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, pp. 60–72, April 1995.

[4] R. Jain, "Congestion control and traffic management in ATM networks: recent advances and a survey," *ATM Forum Contribution 95-0177*, 1995.

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

[6] K.-Y. Siu and H.-T. Tzeng, "Adaptive proportional rate control (APRC) with intelligent congestion indication," *ATM Forum Contribution 94-0888*, September 1994.

[7] R. Jain, S. Kalyanaraman, R. Viswanathan, and R. Goyal, "A sample switch algorithm," *ATM Forum Contribution 95-0178*, February 1995.

[8] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Analysis of rate-based congestion control methods in ATM networks, Part 1: steady state analysis," in *Proceedings of IEEE GLOBECOM '95*, pp. 296–303, November 1995.

[9] 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 —," in *Proceedings of IEEE GLOBECOM '95*, pp. 1095–1101, November 1995.

[10] H. Ohsaki, G. Hasegawa, M. Murata, and H. Miyahara, "Parameter tuning of rate-based congestion control algorithms and its application to TCP over ABR," in *Proceedings of First Workshop on ATM Traffic Management IFIP WG 6.2*, pp. 383–390, December 1995.

[11] H. Ohsaki, M. Murata, H. Miyahara, C. Ikeda, and H. Suzuki, "Parameter tuning for binary mode switch — analysis," *ATM Forum Contribution 95-1483*, 1995.

[12] H. Ohsaki, M. Murata, and H. Miyahara, "Parameter tuning analysis of a rate-based congestion control algorithm for atm networks," submitted to *IEEE/ACM Transactions on Networking*, February 1997.

[13] C. Ikeda, H. Suzuki, H. Ohsaki, and M. Murata, "Recommendation parameter set for binary switch," *ATM Forum Contribution 95-1482*, December 1995.

[14] H. Ohsaki, N. Wakamiya, M. Murata, and H. Miyahara, "Performance of an ATM LAN switch with backpressure function," in *Data Communications and their Performance: Proceedings of the 6th IFIP WG 6.3 Conference on Performance of Computer Networks* (S. Fdida and R. O. Onvural, eds.), pp. 99–113, Chapman & Hall, October 1995.

[15] 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.

[16] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Performance evaluation of rate-based congestion control algorithms in multimedia ATM networks," in *Proceedings of IEEE GLOBECOM '95*, pp. 1243–1248, November 1995.

[17] N. Yin, "Analysis of a rate-based traffic management for ABR service," in *Proceedings of IEEE GLOBECOM '95*, pp. 1076–1082, November 1995.

[18] A. Lin and C. Fang, "Simulation study of ABR robustness under binary switch modes," *ATM Forum Contribution 95-1019*, August 1995.

[19] C. Fang and A. Lin, "A simulation study of ABR robustness with binary-mode swiches: part II," *ATM Forum Contribution 95-1328R1*, October 1995.

[20] L. Wojnaroski, "Baseline text for traffic management sub-working group," *ATM Forum Contriubtion 94-0394r5*, October 1994.