ATM 網におけるバイナリ型スイッチを用いたレート制御方式の適用性の検討

大崎 博之、村田 正幸、宮原 秀夫

大阪大学 基礎工学部 情報工学科
〒 560 大阪府豊中市待兼山町 1-3

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

あらまし　現在、 ABR サービスクラスに適用される輻輳制御方式として、レート制御方式の標準化が進められている。レート制御方式の有効性は、その制御パラメータをいかに設定するかに大きく依存している。そこで本稿では、バイナリスイッチの制御パラメータが、レート制御方式の性能にどのような影響を与えるかを、数学的解析手法を用いて明らかにする。まず、異なる伝播遅延時間を持つ複数のコネクションが存在する場合に、伝播遅延時間や制御パラメータが、コネクション間の公平性与える影響を明らかにする。また、 ABR トラヒックよりも高い優先権を持つ CBR トラヒックが網に加わる場合に、スイッチにおけるセル廃棄を防ぐためのパラメータ設定条件を明らかにする。

和文キーワード　ABR サービスクラス、レート制御方式、制御パラメータ、バイナリ型スイッチ

Robustness of Rate-Based Congestion Control Algorithm with Binary-Mode Switch in ATM Networks

Hiroyuki Ohsaki, Masayuki Murata, 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. Even though the performance of the rate-based congestion control algorithm heavily depends on the selection of those 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/VBR traffic on ABR connections.

英文 key words　ABR Service Class, Rate-Based Congestion Control, Control Paramters, Binary-Mode Switch

1 Introduction

A rate-based congestion control algorithm has been standardized for ABR (Available Bit Rate) service class by the ATM forum [1, 2]. In the standard document [1], 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) which control envelopes of rate increase and decrease, respectively. During a connection establishment process, the source end system negotiates those control parameters with the network. Effectiveness of the rate-based congestion control is heavily dependent on a choice of control parameters as shown in [2]. If those 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 is not specified in the standard, and parameters should be determined intuitively unless a proper tool is provided.

In [3, 4], 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: no cell loss and full link utilization. Based on these results, we proposed a simple guideline for parameter tuning at the ATM Forum [5]. In addition to obtain high performance (in terms of cell loss and link utilization), fairness among connections is also an important issue. In reality, 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.

We further need to take into account existence of real-time applications such as motion video and voice in multimedia network environment. 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 queued in the buffer if a CBR/VBR cell exists in the switch buffer in the case that the switch has two independent buffers — one for CBR/VBR service class and the other for ABR service class. In other words, the bandwidth available to the ABR service class is limited by the existence of the CBR/VBR traffic. Hence, when a CBR or VBR connection is newly added into the network, the bandwidth available to the ABR service class is suddenly decreased, which must give a serious effect on the performance of the ABR connections. That is, the switch buffer for ABR cells may become overloaded for a while, which leads to a large queue buildup and eventually to cell losses due to the buffer overflow.

In this paper, we focus on the above two subjects. In Section 2, we first analyze the behavior of the rate-based congestion control for a single-hop network with a simple binary-mode switch but each group of connections is allowed to have the different propagation delay. In Section 3, we then analyze the maximum queue length at the switch after a new CBR connection is established in the network.

2 Multiple Groups of Connections

In this section, 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 short ramp-up time for an additional ABR connection.

2.1 Analysis

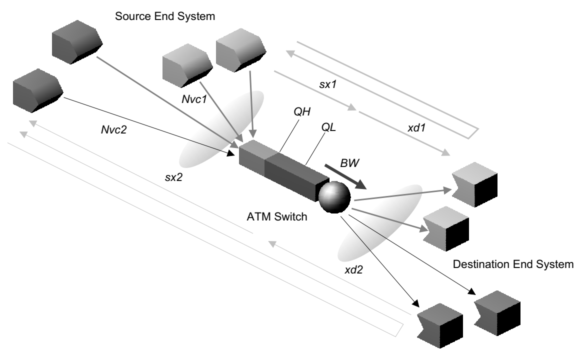

Fig. 1: Analytic Model for Multiple Groups for $N = 2$.

We divide ABR connections into N groups, each of which has different propagation delays. Within a group, connections have identical propagation delays. Figure 1 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 $\tau_{sx\,n}$ and $\tau_{xd\,n}$, respectively. For brevity, we introduce $\tau_n (= 2\tau_{sx\,n} + 2\tau_{xd\,n})$ and $\tau_{xds\,n} (= \tau_{sx\,n} + 2\tau_{xd\,n})$. The number of connections in group n is denoted by $N_{VC\,n}$. Thus, we have a relation.

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

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 $N_{RM\,n}$ as RIF, RDF and N_{RM} of group n, respectively.

We also assume $\tau_{sx\,i} \leq \tau_{sx\,j}$ and $\tau_{xd\,i} \leq \tau_{xd\,j}$ for any i and j $(i < j)$ without loss of generality.

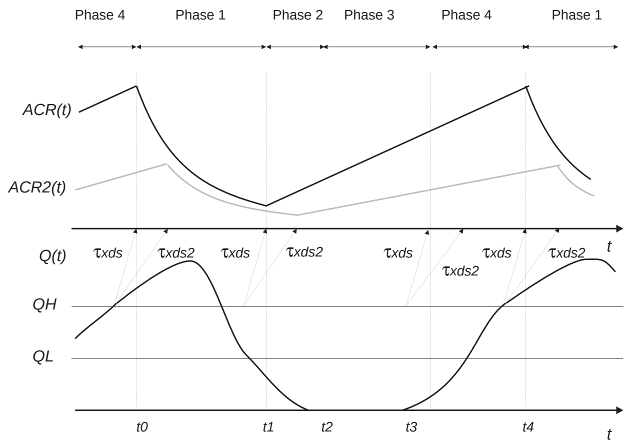

Fig. 2: Pictorial View of $ACR^n(t)$ and $Q(t)$.

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 shown in Fig. 2, $ACR^n(t)$ and $Q(t)$ have periodicity. We further introduce $ACR^n_i(t)$ and $Q_i(t)$ as the $ACR^n(t)$ and $Q(t)$ in Phase i, which are defined as

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

Because of the difference in propagation delays between the switch and the source via the destination ($\tau_{xds\,n}$), congestion information from the switch arrives at the sources of each group at different time. Hence, by defining Δ as $\tau_{xds\,n} - \tau_{xds1}$, $ACR^n_i(t)$ is obtained as follows (see [6] for the details of derivations).

$$\begin{aligned} ACR^n_1(t) &= ACR^n_1(\Delta) e^{-\frac{BW\,RDF_n}{N_{VC}\,N_{RM\,n}}(t-\Delta)} \\ ACR^n_2(t) &= ACR^n_2(\Delta) + \frac{BW\,RIF_n\,PCR}{N_{VC}\,N_{RM\,n}}(t-\Delta) \\ ACR^n_3(t) &\cong ACR^n_3(\Delta) e^{\frac{RIF_n\,PCR}{N_{RM\,n}}(t-\Delta)} \\ ACR^n_4(t) &= ACR^n_4(\Delta) + \frac{BW\,RIF_n\,PCR}{N_{VC}\,N_{RM\,n}}(t-\Delta) \end{aligned}$$

for

$$\Delta \leq \quad t \quad \leq \Delta + t_{i-1,i}.$$

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

$$\begin{aligned} Q_i(t) &= \max(Q_i(\tau_{sx1}) \\ &+ \int_{\tau_{sx1}}^{t} (\sum_{n=1}^{N} N_{VC\,n}\,ACR^n_i(x - \tau_{sx\,n}) - BW), 0), \\ &\qquad \tau_{sx1} \leq t < \tau_{sx1} + t_{i-1,i}. \end{aligned}$$

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

$$t_{i-1,i} = \begin{cases} Q_1^{-1}(Q_L) + \tau_{xds1} & i = 1 \\ \min(Q_2^{-1}(Q_H) + \tau_{xds1}, Q_2^{-1}(0) + \tau_{xds1}) & i = 2, 4 \\ ACR_3^{n-1}(BW/N_{VC}) + \tau & i = 3 \end{cases}$$

where $ACR^{n-1}_i(t)$ and $Q^{-1}_i(t)$ are defined as the inverse representations of $ACR^n_i(t)$ and $Q_i(t)$, respectively.

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 $ICR = PCR/20$ 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.

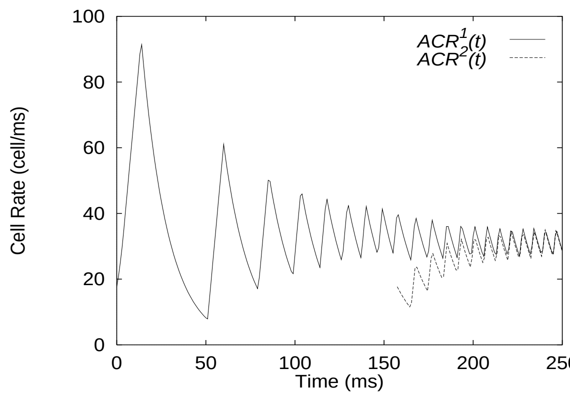

Fig. 3: Effect of Propagation Delay for $\tau_1 = 0.02$ms and $\tau_2 = 0.02$ms.

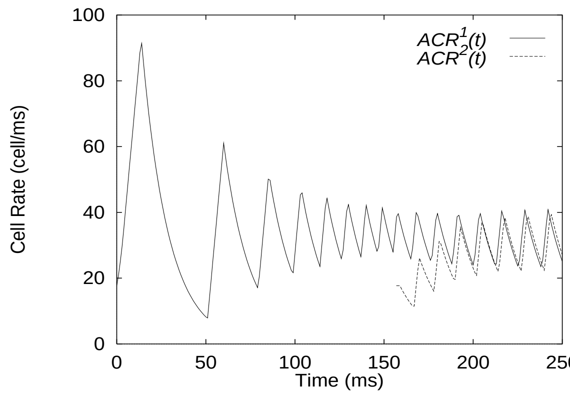

Fig. 4: Effect of Propagation Delay for $\tau_1 = 0.02$ms and $\tau_2 = 2.00$ms.

We first examine the effect of the propagation delay on the ramp-up time. In Figs. 3 and 4, we plot $ACR^n(t)$ for $RIF = 1/64$ and $RDF = 1/16$, which are chosen to satisfy two objectives — no cell loss and full link utilization — for group 1 [6]. Here, τ_{sxn} and τ_{xdn} are set to be 0.005 ms and 0.5 ms, respectively. To eliminate unfairness caused by the starting point of group 2, we add group 2 to the network when group 1 is at the beginning of the Phase 1. These figures indicate that the difference in propagation delays for group 2 has little effect on fairness and the rump-up time. That is, Fig. 4 still shows good fairness, and its ramp-up time is almost equivalent to Fig. 3 in spite of the large amplitude of $ACR^n(t)$.

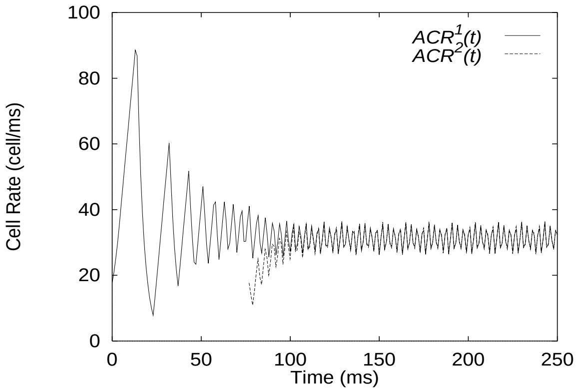

Fig. 5: Effect of Control Parameters for $RIF = 1/64$ and $RDF = 1/4$.

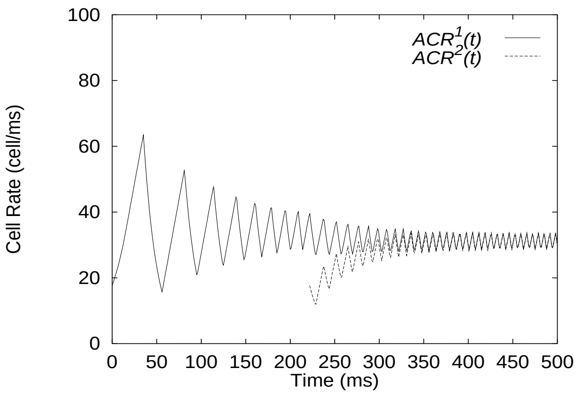

Fig. 6: Effect of Control Parameters for $RIF = 1/256$ and $RDF = 1/16$.

To demonstrate the effect of RIF and RDF on the additional ABR connection, we change these values. Figure 5 shows the case where a smaller value of RDF is used; that is, the rate decrease is faster than the case of Fig. 3. Here, $RDF = 1/4$ is changed from 1/16 while $RIF = 1/64$ is unchanged. On the other hand, slower rate increase is considered in Fig. 6 where we use $RIF = 1/256$ and $RDF = 1/16$. These parameter sets satisfy both no cell loss and 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. 6 (the time scale of this figure is different from others). Therefore, RIF should be set to a large value to fulfill good responsiveness.

3 Effect of CBR Traffic

In this section, by utilizing analytic results obtained in [6], we derive the maximum queue length at the switch when a CBR connection is newly established.

3.1 Analysis

We add a CBR connection to the model presented in Section 2 with $N = 1$ (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 $(1-p)BW$ 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

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

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. 2). 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

$$ACR \leftarrow \max(ACR - ACR \times CDF, MCR).$$

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. 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, that is, CBR traffic in this case.

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. $T_{RDF} \le T_{CDF}$; that is, 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. $T_{RDF} > T_{CDF}$; that is, 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.(2) for each forward RM cell. 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

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

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. (1), Q'_{max} depends on the initial values such as $Q(t' + \tau_{sx})$ and $ACR'(t')$ which is further depends on the time t'. In what follows, we derive the maximum of Q'_{max} for any t', which is defined as

$$\overline{Q'}_{max} = \max_{t'}(Q'_{max}). \qquad (2)$$

As shown in Fig. 2, ACR takes its maximum value at the end of Phase 4 (or at the beginning of Phase 1). In addition, $ACR(t')$ is maximized when the switch is not fully

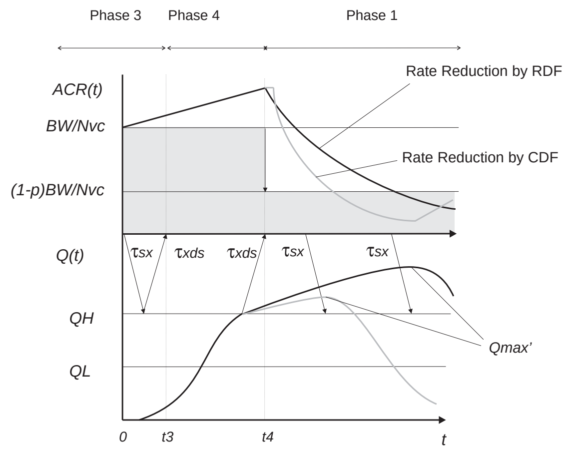

Fig. 7: Pictorial View of $ACR(t)$ and $Q(t)$ with CBR Traffic.

utilized since the large amplitude of $Q(t)$ means the large amplitude of $ACR(t)$. Therefore, $\overline{Q'}_{max}$ 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 $\overline{Q'}_{max}$ is given in a closed-form equation.

3.2 Numerical Examples

In the following numerical examples, propagation delays between source/destination end systems and the switch, 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, we use the same values employed in Section 2.

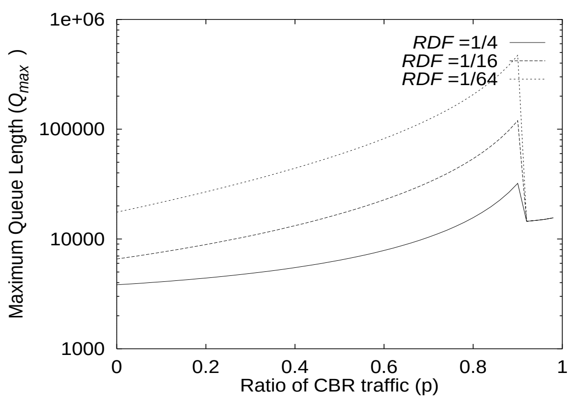

Fig. 8: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/64$ and $C_{RM} = 32$.

We first show the maximum queue length $\overline{Q'}_{max}$ obtained in Eq.(2) as a function of p in Fig. 8. In this figure, RIF is fixed at 1/64, and C_{RM} and CDF is at 32 and 1/2, respectively. RDF is varied to 1/4, 1/16 and 1/64. It can be found that $\overline{Q'}_{max}$ increases as p increases. For example, once a CBR connection that requires a half of the link bandwidth (75Mbit/s, in this case) is added, the switch would have 17,000 cells of buffer capacity to avoid cell loss for ABR connections with $RDF = 1/16$. It can also be found that $\overline{Q'}_{max}$ is suddenly reduced around $p = 0.9$ because of the rate reduction mechanism by CDF. Furthermore, one can find that the maximum queue length can be reduced by setting RDF to a large value.

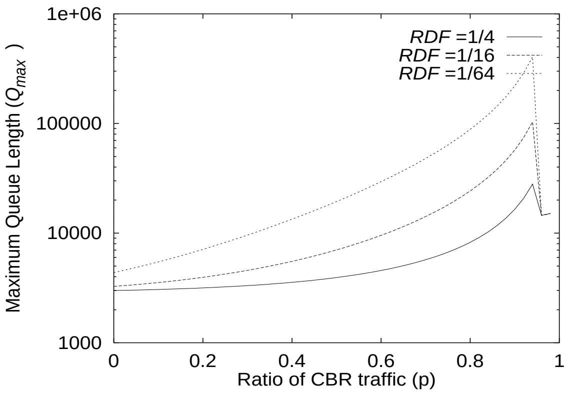

Fig. 9: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/1024$ and $C_{RM} = 32$.

In Fig. 9, RIF is changed from 1/64 to 1/1024. In this figure, the maximum queue length is decreased to some extent compared with Fig. 8. However, a large amount of buffer capacity is still required to avoid cell loss if p is large. No cell loss can be assured even when the CBR connection reserves the bandwidth close to the link capacity by setting C_{RM} properly. In Fig. 10, RIF is set to 1/64. 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 is limited even when p becomes large. For example, 12,000 cells of the buffer capacity is sufficient for achieving no cell loss with $RFD = 1/16$ even when the CBR connection requires the entire bandwidth.

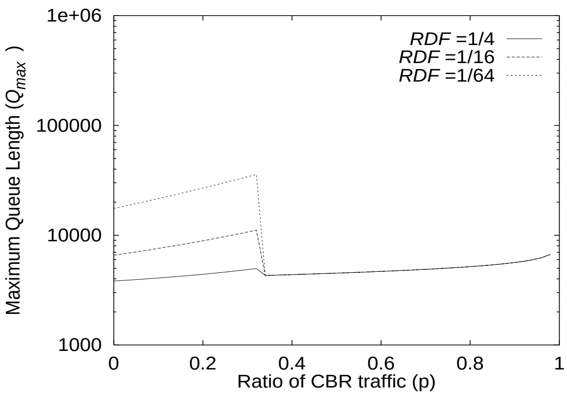

Fig. 10: The Maximum Queue Length vs. Ratio of CBR Traffic for $RIF = 1/64$ and $C_{RM} = 4$.

We plot $\overline{Q'}_{max}$ as the functions of C_{RM} and p in Fig. 11. In the figure, RIF and RDF is fixed at 1/64 and 16,

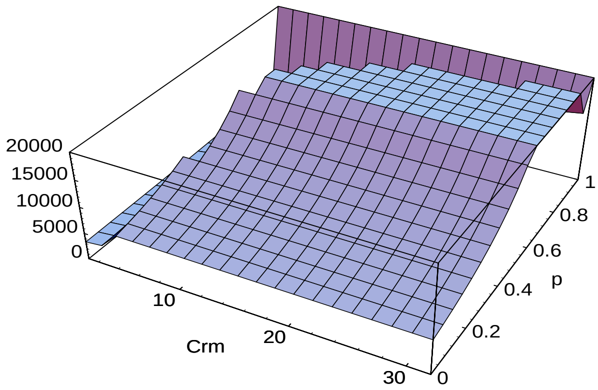

Fig. 11: The Maximum Queue Length for $RIF = 1/64$ and $RDF = 1/16$.

respectively. The z-axis is ranged from 0 to 20,000 cells. As can be found from this figure, C_{RM} can be given a smaller value to avoid cell losses completely for any traffic load of the CBR connection.

From the above observations, we can conclude that to limit the queue buildup, each of RIF and RDF should be small and large, respectively. Moreover, C_{RM} can be set to be a small value for avoiding cell loss resulting from CBR traffic.

References

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

[2] 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.

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

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

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

[6] H. Ohsaki, M. Murata, H. Suzuki, C. Ikeda, and H. Miyahara, "Parameter tuning of rate-based congestion control algorithms for ATM networks," *in preparation*, 1996.