Adaptive Thresholds of Buffer to Solve the Beat-Down Problem of Rate Control in ATM Networks

Harry PRIHANTO†, Student Member and Kenji NAKAGAWA†, Member

SUMMARY ABR service is currently standardized to handle applications of data traffic in ATM network. As a flow control method, the rate-based flow control has been adopted and applied to manage the ABR service. Several control methods have been proposed, and the EPRCA is selected as one of the control methods by the ATM Forum. EPRCA is an excellent algorithm, but when the EPRCA is applied to the ATM network, several problems occur. One of the problems is the beat-down problem, which gives unfair allocation of transmission rate to connections. We propose a new control method which solves the beat-down problem. We will show that, by our proposed method, (i) the ACR is given fairly to every connection compared to the conventional method, and also (ii) the throughput is fair for both long-hop and short-hop connections, (iii) the ACR is proportional to the throughput, and finally (iv) the total throughput is larger than that of the conventional method. The fairness of the throughput in (ii) is measured by the fairness index. In (iii), being proportional means that the allocated ACR is close to the throughput and it is measured by the proportion index. The performance is evaluated by computer simulation.

key words: flow control, rate-based, ABR traffic, EPRCA, beat-down problem, ATM network

1. Introduction

To realize Broadband Integrated Services Digital Network (B-ISDN) which is intended to provide various services in telecommunications, Asynchronous Transfer Mode (ATM) is one of several data transfer technologies that are proposed and currently being developed [1]. In the ATM network, information is divided into a fixed length of data block, called cell. The ATM technology can provide high speed and high efficiency of the network because a simple data transfer protocol and wide bandwidth of optical fiber media are used in the network [2].

Several services are provided and defined in ATM networks. Five classes of service are defined as Constant Bit Rate (CBR), real-time Variable Bit Rate (rt-VBR), non-real-time Variable Bit Rate (nrt-VBR), Available Bit Rate (ABR), and Unspecified Bit Rate (UBR) [3]–[5]. The ABR service is designed to handle applications of data traffic in ATM network. As a flow control method, the ATM Forum has adopted and applied rate-based flow control to manage the ABR service. The rate-based flow control is a closed-loop control. Each source uses feedback information from a network to manage their transmission rate. A Resource Management (RM) cell conveys feedback information as explicit rate from switch to the source.

Several rate-based control methods have been proposed, e.g. EPRCA (Enhanced Proportional Rate Control Algorithm), CAPC (Congestion Avoidance using Proportional Control), ERICA (Explicit Rate Indication for Congestion Avoidance), etc. [4], [6]. EPRCA is a favorite control method in the ATM Forum at the present time [7]. EPRCA is an excellent algorithm, but when the EPRCA is applied in the ATM network, several problems occur [4], [6]. One of the problems is the beat-down problem, which gives unfair allocation of transmission rate to connections [10], [11]. To solve the beat-down problem, we propose a new control method based on EPRCA. We use variable thresholds of buffer to detect the congestion state of the switch.

We show through simulation results that our proposed method has better performance than conventional method. We will show that, by our proposed method, (i) the ACR is given fairly to every connection compared to the conventional method, and also (ii) the throughput is fair for both long-hop and short-hop connections, (iii) the ACR is proportional to the throughput, and finally (iv) the total throughput is larger than that of the conventional method. The fairness of the throughput in (ii) is measured by the fairness index. In (iii), being proportional means that the allocated ACR is close to the throughput and it is measured by the proportion index.

The rest of this paper is organized as follows. After describing ABR flow control in Sect. 2, we discuss an implementation of EPRCA and its problems in Sect. 3. We describe our proposed method in Sect. 4. We consider several network models and evaluate both performance of conventional and proposed method in Sect. 5. Finally, the conclusion will be described in Sect. 6.

2. ABR Flow Control

Flow control refers to the set of actions taken by the network to avoid congestion conditions. Congestion is defined as a state of network elements in which the network is not able to meet the negotiated performance objectives for already established connections [1]. The ABR flow control is operated between source end sys-
tem (SES) and destination end system (DES). The control is a closed-loop control because there are continuous feedback of control information between the network and the SES. Generally, the ABR flow control uses control information from the network to notify the SES to set an appropriate rate. Each switch in the network computes the appropriate rate that is called explicit rate (ER) for all connections which are sharing the same link. The control information such as ER is conveyed from the network to the SES by RM cells [3]–[5]. There are two directions of the RM cell flow, forward and backward, as shown in Fig. 1.

2.1 ABR Parameters

In the ABR service, each SES negotiates several operating parameters with the network at the time of connection setup. The important parameters are peak cell rate (PCR), minimum cell rate (MCR), allowed cell rate (ACR), and so on. PCR is the maximum rate for the SES to be allowed to transmit on the link. The SES can also request the MCR, which is the guaranteed minimum rate. During transmitting data, the SES is allowed to transmit at particular rate that is defined as ACR. The ACR value is dynamically changed between MCR and PCR depending on network condition. At the connection setup or after long idle interval, the ACR is set to the initial cell rate (ICR). The list of parameters used in the ABR service are shown in the Table 1 [3].

<table>
<thead>
<tr>
<th>Lable</th>
<th>Extension</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>Peak cell rate</td>
<td>-</td>
</tr>
<tr>
<td>MCR</td>
<td>Minimum cell rate</td>
<td>0</td>
</tr>
<tr>
<td>ACR</td>
<td>Allowed cell rate</td>
<td>-</td>
</tr>
<tr>
<td>ICR</td>
<td>Initial cell rate</td>
<td>PCR</td>
</tr>
<tr>
<td>CCR</td>
<td>Tagged cell rate</td>
<td>10 cell/s</td>
</tr>
<tr>
<td>Nrm</td>
<td>Number of cells between forward RM cells</td>
<td>32</td>
</tr>
<tr>
<td>RIF</td>
<td>Rate increase factor</td>
<td>1/16</td>
</tr>
<tr>
<td>RDF</td>
<td>Rate decrease factor</td>
<td>1/16</td>
</tr>
<tr>
<td>CDF</td>
<td>Cutoff decrease factor</td>
<td>1/16</td>
</tr>
<tr>
<td>ADTF</td>
<td>ACR decrease time factor</td>
<td>0.5 ms</td>
</tr>
<tr>
<td>TBE</td>
<td>Time buffer exposure</td>
<td>16,777,215</td>
</tr>
<tr>
<td>CRM</td>
<td>Missing RM-cell count</td>
<td>TBE/Nrm</td>
</tr>
<tr>
<td>FRTT</td>
<td>Fixed round-trip time</td>
<td>-</td>
</tr>
</tbody>
</table>

2.2 Control Mechanism

The basic mechanism of congestion control in ABR service requires the cooperation of SES and DES to the network to avoid congestion by controlling their rates. Then the SES generates RM cells in every $N_{rm} = 1$ (= 31) data cells to probe network condition. RM cells travel from the SES to the DES and back to the SES which carry feedback information of the network such as congestion state and rate allocation. The RM cell which travels from the SES to the DES are called forward RM (FRM) cell, and backward RM (BRM) cell from the DES to the SES. The RM cell contains a current cell rate (CCR), PCR, ER, and other parameters required by the SES. Details of the RM cell configuration can be found in [3].

The switch gives feedback information to the SES in every RM cell arrival. There are three types of switch to give the feedback information [3], [4], [8]:

- “binary” switch, using explicit forward congestion indication (EFCI) bit in the header of data cells to indicate the congestion state.
- “relative rate marking” switch, using congestion indication (CI) bit and no increase (NI) bit in the payload of RM cell to notify the congestion state.
- “explicit rate (ER)” switch, using other ER field in the payload of RM cell to give the allocation rate.

The ER switch is currently used to support ABR service. When a switch receives an RM cell, the switch may reduce the ER value in FRM and/or BRM cell to the calculated fair rate.

When the SES receives a BRM cell, the SES should adapt its rate to the information carried in the RM cell. If CI bit = 0, the SES can linearly increase its rate by rate increase factor (RIF), where RIF is determined at call setup. The increasing rate can reach the ER value in the RM cell but should never exceed the PCR. If CI bit = 1, the SES must exponentially decrease its rate by using rate decrease factor (RDF), where RDF is also determined at call setup. If the rate is still greater than the returned ER, the SES must further decrease its rate to the returned ER but should not be lower than the MCR.

3. Implementation and Problems of EPRCA

Several control methods have been proposed and the EPRCA is chosen as one of the control mechanisms by the ATM Forum [3]. The EPRCA is a closed-loop rate-based flow control. The RM cells are used to transfer information from the SES to the DES and vice versa. The EPRCA is implemented at ER switch in order to achieve fairness of rate allocation [6].
3.1 EPRCA

3.1.1 Switch Behavior

The basic features of EPRCA can be described as follows [3], [4], [6], [9]. In this algorithm all the connections are assumed to share a common queue with two congestion thresholds $QT$ and $DQT$. At every time an FRM cell arrives, the switch computes the mean allowed cell rate ($MACR$) using a running exponential weigth average as follows:

$$MACR = (1 - \alpha)MACR + \alpha CCR,$$

where $\alpha$ is averaging factor and generally is chosen to be 1/16. The FRM cells sent by the SES contain a desired ER in the ER field, the $CCR$ field set to the ACR, and the CI bit set to zero. The switch monitors its load by observing the queue length.

On arrival of BRM cell, when the queue length in the switch buffer exceeds the threshold value $QT$, congestion state is detected and only those connections whose $CCR$ values are above the fair share, computed as

$$fair\ share = DPF \times MACR,$$

are asked to reduce their rates. The other connections whose $CCR$ value are below the fair share, have no adjustments. $DPF$ is the down pressure factor of switch, and typically set to 7/8. Furthermore when the queue length in the switch buffer exceeds the second threshold value $DQT$, the heavy congestion is detected and all the connections are asked to reduce their rates.

3.1.2 SES Behavior

The SES adjusts its transmission rate when the SES receives BRM cells. If no congestion information is in the BRM cell, the SES can increase their rate by using the following equations to calculate their allowed cell rate ($ACR$):

$$ACR \leftarrow \max[MCR, \min(ACR + RIF \times PCR, ER, PCR)].$$

But if BRM cell carries some congestion information, the SES must decrease their rate with the following equation:

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

Default value of the ABR parameters are shown in the Table 1.

3.2 Problems of EPRCA

Several problems can be found when we apply the EPRCA in the network, such as unfairness problem [4], oscillation of ACR [6], beat-down problem [6], and etc. In this paper we focus on the beat-down problem. According to [6], the beat-down problem occurs in the binary switch when the switch manages several type of connections which traverse through different number of switches. The connections passing more switches (long-hop connection) have a higher probability of getting their RM cells marked than those passing fewer switches (short-hop connection). Therefore the long-hop connections are given the ACR which are smaller than that of short-hop connections.

We define the beat-down problem as the difference of average ACR between short-hop and long-hop connections is too large. According to our experiment, the ER switch also have the beat-down problem when EPRCA algorithm is applied [10], [11]. Then we try to solve the problem by proposing a new method which modifies the EPRCA as described in the next section.

4. Proposed Method

The beat-down problem occurs when there are various connections which pass different number of hops, if the switch handles those connections in the same manner without taking the number of hops into consideration. We propose a new method to solve the beat-down problem. We modify the EPRCA algorithm with making longer-hop connections more advantageous by applying adaptive thresholds.

The proposed method is described as follows:

• standard threshold

As the conventional EPRCA method, the proposed method sets 2 threshold values $QT$ and $DQT$ at the switch buffer to detect the congestion condition. The threshold values are set permanently in the conventional method, but the threshold values adaptively changed in the proposed method.

• adaptive threshold

Let $h$ and $d$ be integers. As shown in the Fig. 2, we consider other threshold values $QT_1, \ldots, QT_h, \ldots$ and $DQT_1, \ldots, DQT_h, \ldots$, as well as the original thresholds $QT$ and $DQT$. $QT_h$ and $DQT_h$ are defined as follows:

$$QT_h = QT + h \times d,$$

![Fig. 2 Proposed method.](image)
\[ DQT_h = DQT + h \times d. \]  
(5)

In the following algorithm, \( h \) means the number of hops that the FRM cell passed and \( d \) is the increment of threshold.

- **RM cell structure**
  
  An RM cell contains the information from the SES such as PCR, MCR, CCR, ER. We add the \( h \)-field for counting the number of passed switches of the RM cell in the up stream. Then the \( h \) value can be used to decide which threshold value should be used. The initial value of the \( h \)-field is 0.

- **switch behavior**
  
  When an FRM cell arrives, the switch computes the MACR and fair share using Eqs. (1) and (2). Then the switch reads the \( h \) value from the \( h \)-field of the FRM cell and computes \( QT_h \) and \( DQT_h \) by Eq. (5). The switch uses the \( QT_h \) and \( DQT_h \) as thresholds instead of \( QT, DQT \). If the queue length \( Q \) of the switch buffer is below \( QT_h \), the switch decides that there is no congestion and gives PCR to all connections. If \( QT_h < Q \leq DQT_h \), congestion state is detected and only those connections whose CCR values are above the fairshare are asked to reduce their rates. Furthermore, if \( Q > DQT_h \), the heavy congestion is detected and all the connections are asked to reduce their rate. Then, the switch increases the \( h \) value of the FRM cell by 1 and transmit it to the output link.

- **source behavior**

  We use Eqs. (3) and (4) for increasing and decreasing rate, respectively. Other parameters are used similarly as the conventional EPRCA method.

  In the conventional EPRCA, long-hop connections are given smaller ACR than short-hop connections. It is because the long-hop connections have a higher probability of decreasing rate than short-hop connections. Therefore, we use adaptive thresholds to make the rate of long-hop connections not strictly be choked at the switch. By making longer-hop connections more advantageous, we can solve the beatdown problem.

5. Simulation and Evaluation

We make simulation models to show that the proposed method provides better performance than the conventional method. We design several models to evaluate the fairness of ACR distribution and the throughput of each connection, and so on.

5.1 Simulation Model

We consider simulation models with 2, 3, and 4 switches. The 2 switch model is shown in Fig. 3(a). Each node contains 10 connections. The first switch (SW1) accommodates 10 SES’s. 5 SES’s are terminated to the first DES as short-connection (S1) and the other connections go to the next switch (SW2). New SES’s come to the SW2 as short-hop connection (S2), they are terminated to the last DES. Other 5 SES’s as a long-hop connection (L) travel from SW1 to SW2 and terminate to the final DES. The EPRCA is applied in the models and we use default EPRCA parameters as

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR</td>
<td>150 Mbps</td>
</tr>
<tr>
<td>MCR</td>
<td>1.5 Mbps</td>
</tr>
<tr>
<td>Nrm</td>
<td>32 cells</td>
</tr>
<tr>
<td>RIF, RDF</td>
<td>110, 100</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>110</td>
</tr>
<tr>
<td>DPF</td>
<td>5, 10 cells</td>
</tr>
<tr>
<td>QT, DQT</td>
<td>5, 10 cells</td>
</tr>
<tr>
<td>Switch buffer</td>
<td>35, 40, 45, 50 cells</td>
</tr>
<tr>
<td>Load per SES</td>
<td>0.06-0.12</td>
</tr>
<tr>
<td>Simulation time</td>
<td>1,000,000 cell times</td>
</tr>
</tbody>
</table>
Fig. 4 Average ACR in conventional method, 2 switch model.

Fig. 5 Average ACR in conventional method, 3 switch model.

Fig. 6 Average ACR in conventional method, 4 switch model.

Fig. 7 Average ACR in proposed method, 2 switch model.

Fig. 8 Average ACR in proposed method, 3 switch model.

Fig. 9 Average ACR in proposed method, 4 switch model.
shown in the Table 1. We assume that the EPRCA is applied in a local area network (LAN). Because the node distances are so short, then we can assume that the transfer delay can be ignored.

It is important to decide an appropriate value of QT and DQT for preventing a beat-down problem. If we set the QT and DQT too small, the network decides that the congestion occurs very frequently, then the connections are choked and the throughput will go down. On the other hand, if the QT is too large, the network misses the congestion, then the total ACR exceeds the link capacity. Furthermore, if the DQT is too large, the beat-down problem becomes worse. This is because the connections can update their fair shares many time before the queue length reaches DQT, so the difference of the fair shares between a short-hop and a long-hop connections enlarges. As our experiment results, we set QT and DQT value to 5, 10 cells respectively as the best condition of EPRCA method. Other parameters are shown in the Table 2.

Similarly to 2 switch model, 3 and 4 switch models contain long-hop and short-hop connections. 5 SES’s from the first node travel through all nodes to the final DES. New 5 SES’s enter to each node and leave immediately for the corresponding DES. We call them short-hop connections. The simulation models can be shown in Figs. 3(b), (c).

5.2 Comparison of ACR Allocation

The average ACR is defined as the distributed ACR to each connection (SES) averaged over the simulation time. The results are shown in Figs. 4–9. We set thresholds of switch buffer QT = 5 cells, DQT = 10 cells. The buffer size of each switch is variable from 35 to 50 cells. The average load of each SES is 0.1, then the total offered load of 10 SES’s is 1.0.

The beat-down problem appears as depicted in Fig. 4. The long-hop connection (L) is given smaller ACR value than the short-hop connections (S1, S2). This is because the long-hop connection passes many switches, so it has a high probability to meet the congestion at one of the switches. On the other hand, the problem are solved when we apply the proposed method as shown in the Fig. 7. Because of the advantage of L connections, they can get more ACR than the conventional method.

Similarly, for 3 switch model we have the same phenomenon. See Figs. 5 and Figs. 8. In the conventional method, the beat-down problem appears, especially at the short-hop connection S3. The S3 has the highest ACR among connections. We can solve the problem by applying the proposed method. Each connection has almost the same ACR values as seen in Fig. 8.

In the 4 switch model, we can also see that the beat-down problem still occurs in the conventional method, Fig. 6. The problem can be solved by applying the proposed method as shown in Fig. 9.

According to the results, the conventional method distributes higher ACR than the proposed method to short-hop connections, especially to the last short-hop connections, i.e., S2 in Fig. 4, S3 in Fig. 5, and S4 in Fig. 6. On the other hand, by the proposed method, the ACR are distributed fairly to every connection.

Now we evaluate the ACR which are given to each SES versus offered load. Figure 10, Fig. 11, and Fig. 12 show the average ACR of each connection at 2, 3, and 4 switch models, respectively. We find that the large ACR are distributed when the offered load is low. The ACR are choked at the high offered load because the heavy congestion is detected. The proposed method shows better performance by appropriate ACR allocation when offered load become high. For the high offered load ≥ 0.7, the proposed method allocates the ACR to long-hop and short-hop connections fairly compared to the conventional method. Because the long-hop connection is applied higher threshold QT, the ACR of the long-hop connection is not easy to be choked.

5.3 Comparison of Throughput

Next, let us investigate the comparison of throughput
performance between the conventional EPRCA and the proposed method. We define the throughput as the average transmitted information rate during the simulation. Figure 13, Fig. 14, and Fig. 15 show the throughput of each connection at 2, 3, and 4 switch models, respectively. We see that the difference of the throughputs between long-hop and short-hop connections by the proposed method is smaller than that by the conventional method. We will examine the difference numerically with the following fairness index.

5.3.1 Fairness Index

We here define the fairness index as the ratio of the maximum throughput among connections to the minimum throughput. If the fairness index is close to 1.0, we can say that the throughput of every connection is fair. We calculate the fairness indices for the conventional and proposed methods and show them in Tables 3, 4, and 5. We see from these tables that the fairness index of the proposed method is smaller than that of the conventional method. The fairness between the long-hop and short-hop connections is improved by the proposed method.

5.4 Allocation of ACR Proportional to Throughput

We define the proportion index as the ratio of the average ACR over all the connections to the average throughput. If the proportion index is close to 1.0, we can say that the ACR is allocated proportionally to the throughput for every connection. We show in Table 6 the proportion indices in the case of the offered load is 0.9. We see that the proportion index of the proposed method is smaller than that of the conventional method. The ACR is proportionally allocated by the
5.5 Total Throughput at a Switch

Total throughput at a switch is defined as the sum of the throughputs of all the connections which enter to the switch. We calculate the total throughputs at each switch and take the average over all switches and show the average values in Table 7 in the case of the offered load is 1.0. We see that the total throughputs of the proposed method are larger than the conventional method, either in the 2, 3, or 4 switch models. Although the throughputs of short-hop connection in the proposed method are smaller than that in the conventional method, the throughputs of long-hop connection are larger. As a result, the total throughput at a switch becomes larger than that of the conventional method.

6. Conclusion

When the original EPRCA is applied in the switch, several problems occur. The beat-down problem is one of the problems that appears not only in the binary switch but also in the ER switch. We show the problem by simulation models. Then we proposed a new algorithm based on the EPRCA to solve the beat-down problem. By the simulation, the average ACR is allocated fairly either to long-hop and short-hop connections. We find that the throughputs are fair for both long-hop and short-hop connections. Moreover, the proposed method allocates the ACR proportionally to the achieved throughput. We find also that the total throughput at each switch is larger than the conventional method.

References


Harry Prihanto was born in Jakarta, Indonesia, on 6 December, 1966. He received the B.Eng. degree in electronics from Yamanashi University in 1992 and the M.Eng. degree in electronics from Nagaoka University of Technology in 1996. In 1992, He joined the Association for Overseas Technical Scholarship (AOTS) training program at the Data Transmission Laboratories, NEC Co. Ltd. Japan. In 1993, he joined with the Directorate for Engineering Science, Agency for the Assessment and Application of Technology (BPPT), Jakarta, Indonesia. Since 1996 he has been a Doctorate student at the Department of Electrical Engineering, Nagaoka University of Technology. His research interests include network management and B-ISDN traffic management. He is a student member of IEEE. He is also a member of Indonesian Society on Electrical, Electronics, Communication and Information (IECI) and Institute for Science and Technology Studies (ISTECS).
Kenji Nakagawa was born in Yamashita, Japan, on 28 January, 1957. He received the B.S., M.S., and D.S. degrees in mathematics from Tokyo Institute of Technology in 1980, 1982, and 1986, respectively. In 1985, he joined the Basic Research Laboratories, NTT (Nippon Telegraph and Telephone Corp.). Since 1992, he has been an associate professor of Nagaoka University of Technology. His research interests include information theory, statistics, and queueing theory. Dr. Nakagawa is a member of the Society of Information Theory and its Application, the IEEE, and the Mathematical Society of Japan.