Joint Conference of 1996 International Computer Symposium
December 19~21, Kaohsiung, Taiwan, R.O.C.

A Proportional Feedback Scheme for ATM Networks

Wei Kuang Lai[1] and Chuang-Chi Hwang
Institute of Computer and Information Engineering
National Sun Yat-Sen University

Abstract

Available Bit Rate (ABR) service class is proposed by ATM Forum to provide the "best-effort" service like traditional TCP/IP traffic. It is designed to make use of the remaining bandwidth not used by VBR and CBR. However, the bursty nature of ABR traffic makes it difficult to prevent buffers on ATM switches from overflowing. A number of schemes have been proposed for ABR traffic control. In this thesis, a proportional scheme which is rate-based hop-by-hop flow control is proposed. A mathematical model is built to analyze the hop-by-hop flow control scheme. It is shown the scheme can utilize full bandwidth without any cell loss. It is also shown the scheme has low queuing delay and fast transient response.

Key words: ABR service class, ATM, Hop-by-hop flow control, Adaptive buffer allocation.

I. Introduction

Flow control is essential for asynchronous transfer mode (ATM) [1] networks in providing "best-effort" services or ABR (Available Bit Rate) services in ATM Forum terminology. With proper flow control, computer users would be able to use an ATM network in the same way as they have been using conventional LANs. They would be able to acquire as many network resources as are available at any given moment, and compete equally for the available bandwidth.

In order to support different kinds of service models, users exchange network information with other users through the network. Since network resources are limited, users in return must response effectively to received network information. A number of mechanisms are considered to be crucial for an efficient flow control [2]. They are as follows:

1. Service scheduling mechanism.

Scheduling mechanism controls the order of serving individual cells from users. Two of the most popular scheduling algorithm are FIFO and fair queuing (FQ) [3] [4]. In FIFO, cells are served in the order of arrival. With FIFO, all users experience the same delay of queuing time although the overload condition is caused by a few of users. It hence provides no protection against uncooperative users. On the other hand, fair queuing provides good protection against greedy users. It tries to divide bandwidth evenly among all users through the round-robin mechanism [5] on a per VC basis. The round-robin scheduler has the effect of interleaving the coming cells. Thus FQ has the effect of reduces the degree of cells clumping. Hence, in the simulation of our proposed flow control scheme, fair queuing is assumed.

2. Buffer management mechanism.

Buffering is needed whenever the total cell arriving rates are faster than the available bandwidth on the output link. This can be written as $\sum$ input rate > available bandwidth. If the overloaded condition continues for a period of time, the buffer capacity may be exceeded and the cells have to be discarded. Thus, buffer management is important in congestion control. Two of the most popular schemes for buffer management are shared buffer scheme [6] and per-VC (virtual channel) allocation scheme. For shared buffer scheme, all flows share a common buffer pool in a first-come first-use (FCFU) order. In terms of buffer usage, this is more effective than per-VC scheme because all left buffer spaces can be used. The disadvantage is that it does not protect flows

[1] wklai@cie.nsysu.edu.tw

from each other. Once a flow occupies all the buffer spaces, others will be denied to access them. On the other hand, schemes with per-VC queuing reserve buffer spaces for each VC which can not be used by other VCs. It could be a problem about how to allocate proper buffer spaces to each flow efficiently and fairly. However, per-VC queuing schemes have the benefits of protecting flows from each other. Hence, some researchers have suggested partials of spaces are reserved for each VC and others are shared to take advantages of both. In this paper, we will instead concentrate on how to use buffer spaces efficiently in our scheme.

3. Feedback mechanism.

Feedback mechanism is another important mechanism in the flow control [7][8]. Without feedback, the upstream node would not be able to respond to the ever changing network states. For example, it would not be possible for upstream switch to decrease sending rates in response to a congestion condition in the downstream. Over the years there have been many different proposed schemes based on the feedback mechanism. They can be classified into two categories: end-to-end control [9][10][11] and hop-by-hop control [12][13][14][15]. End-to-end control takes the control action on the edge of the network. In general, end-to-end control is related to rate-based flow control. In rate-based flow control, feedback information is usually carried by resource management cells (RM cells). The feedback information is used to decide the maximum transmission rate that a source can send over each VC. At least two times end-to-end delays are needed to form a control loop. This is because control information must travel from one end of a network to the other end and back before the control information can take place. The time needed to form a control loop is also called round trip time (RTT). It takes a longer delay for control information to travel between two ends for end-to-end control than between two nodes for hop-by-hop control. Thus end-to-end control has a longer control loop. Due to the long propagation delay, it is not sensitive to the changing network states. This frequently leads to oscillations in throughputs and packet losses [16] [17].

As a result, it is not as suitable as hop-by-hop control for the bursty traffics such as ABR traffic or the LAN traffic. On the other end, hop-by-hop control reaches the lower bound for a control scope. Hence, it can react quickly to changing network status. Generally, hop-by-hop control implements a control scheme called credit-based flow control[12]. Each hop is composed of a sender and a receiver. Each of the VC maintains a separate queue or known as a per VC queuing. In order to forward any cells, it is required for the sender to receive credits from the receiver. The receiver sends credits to the sender based on the availability of the buffer spaces. On the other hand, each time the sender sends out a cell, its credit number is decreased by one. By maintaining a credit balance between sender and receiver, the scheme can prevent buffer overflow. Moreover, in this scheme, the sender can adapt itself to maximum available bandwidth. However, intermediate nodes and end nodes are involved in feedback control in hop-by-hop scheme. This is more complicated than end-to-end schemes in which intermediate nodes do not participate in feedback control. In this paper, a new scheme called a proportional feedback scheme is proposed for hop-by-hop flow control. This scheme is rate-based instead of credit-based.

In section II, we propose a proportional feedback scheme over the hop-by-hop control model. A mathematical model is built to analyze the scheme. It is shown the proposed scheme will guarantee no cell loss and full channel utility. Section II also describes how to implement the scheme. Section III concludes our work.

II. Proportional Feedback Scheme

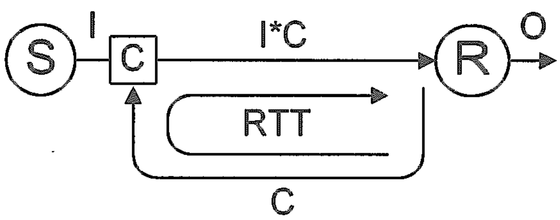

Figure 1. The functional diagram for a proportional feedback scheme.

In this section, we will develop a hop-by-hop scheme called a proportional feedback scheme.
Then we will build a mathematical model to analyze the scheme. By the mathematical mode, we will look at the scheme's cell loss, channel utility, and queuing delay. It is shown the proposed scheme would guarantee no cell loss and achieve full channel utility. Then we will look at the scheme's transient response.

1. A proportional feedback scheme

The functional diagram for proportional feedback control can be illustrated by Fig.1. In Figure 1, I is input, O is output, and B is the number of cells in the buffer. The reason why it's named so is because there is a proportional multiplier C, $0 \le C \le 1$, in this system. All the inputs to the multiplier will be multiplied by the factor C, which is updated by the feedback from the receiver dynamically, before forwarding it. Therefore, the allowed rate to flow onto the link is I*C after throttling control of the proportional multiplier. Figure 2 takes time into account. These shaded circles are the origins of I, C, and O at moment t. They can be written as I(t), C(t), and O(t), respectively. Besides, the number of cells in the buffer at moment t is represented as B(t). When the C(t) forms a loop by going to the sender and returning to the receiver (the blank circle), it should be c(t-RTT) in respect to the origin C(t). Here RTT is round trip time as described before. Similarly, I(t) should be I(t-RTT/2) when it arrives at receiver (blank circle) from the sender. The proportional control takes action when C(t) is feedback to the proportional multiplier (the © symbol in the Fig.2) on the sender. In this paper, without loss of generality, the maximum input rate of I(t) and output rate of O(t) is assumed to be K. This scheme requires a minimum receiver capacity of K*RTTT. The value of C(t) is decided by receiver capability and B(t). C(t) is set to be equal to (KRTT-B(t))/KRTT at time t. Since C(t) is a value between zero and one, the proportional multiplier can be regarded as the admission ratio for the input rate I(t). For example, if the proportional multiplier receives a feedback value of C(t)=1/2, then the output value from © is 1/2 I(t). Notice in real implementation, sending C(t) periodically is more feasible provided that we allocated a small amount of extra buffer spaces to tolerate the minor inaccuracy of C(t). In the following, a mathematical model is built to analyze the system.

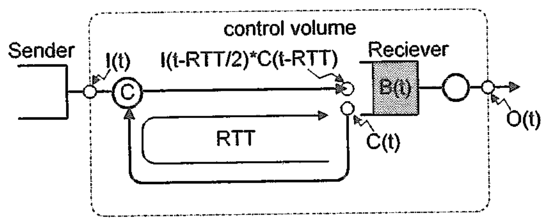

Figure 2. Block diagram for proportional feedback control.

1.1 Mathematical Modeling

The following equation is called buffer equation which is used to describe the change of cells in the buffer.

equation 1: $B'(t)=\frac{dB(t)}{dt}=$ *input rate - output rate* $= I(t\text{-}RTT/2)C(t\text{-}RTT)\text{-}O(t)$

The integral form of the eq.1 is:

equation 2: $\int B'(t)\ dt = \int [I(t\text{-}RTT/2)C(t\text{-}RTT)\text{-}O(t)]\ dt$

We assume the maximum value of I(t) and O(t) to be a constant value K. As a result, we have following relations.

$0\le I(t)\le K$ *and* $0\le O(t)\le K$

Besides, the value of proportional feedback C(t) lies in between zero and one.

$0\le C(t)\le 1$

Therefore, we combine them and have the condition 1.

Condition 1: $0\le I(t)\le K,\ 0\le O(t)\le K,\ 0\le C(t)\le 1$

No Cell Loss Guarantee

To ensure no cell loss, a flow control system must satisfy two conditions. First, under any network conditions, there should not be any lost cells. That is to say, for any I(t), O(t), no cells will be lost in any condition. Second, there should not be any lost cells before current control C(t) takes its effects. Since there is a RTT delay between a control is sent by a receiver and takes effects, it should be guaranteed that there will not be any lost cells during the "dead" time of control. Or the no cell loss control scheme can not be realized. It is equivalent to say in the following way:

$\int_t^{t+RTT} B'(t)\ dt \le$ free buffer capacity.

To sum up, the following two conditions are required to guarantee no cell loss.

Condition 2: ***For any I(t), O(t), no cell is lost in any condition.***

Condition 3: $\int_t^{t+RTT} B'(t)\ dt \le S - B(t)$ ***or*** $B(t+RTT) < S$; ***where S is the buffer capacity***

Combining eq.2 into condition 3, condition 3 becomes

equation 3: $\int_t^{t+RTT} B'(t)\ dt = \int_t^{t+RTT} [I(t\text{-}RTT/2)C(t\text{-}RTT)\text{-}O(t)]\ dt \le S - B(t)$

By substituting condition 1 into eq.3, we can find that

$$-KRTT\le \int_t^{t+RTT} B'(t)\ dt\le KRTT$$

To fit condition 2, S should not fall between -KRTT and KRTT. Thus, we can get

equation 4: $KRTT\le S$, *i.e.* $Min(S)=KRTT$

Therefore, to guarantee no cell loss, the minimum buffer requirements are KRTT. Now that we know the minimum buffer requirements are KRTT, in order to save the buffer capacity, the buffer capacity S is set to KRTT. After substituting S by KRTT, the eq.3 now becomes:

$$\int_t^{t+RTT} B'(t)\,dt = \int_t^{t+RTT} [I(t\text{-}RTT/2)C(t\text{-}RTT)\text{-}O(t)]\,dt \le KRTT - B(t)$$

We reorganize it into the following:

$$\int_t^{t+RTT} I(t\text{-}RTT/2)C(t\text{-}RTT)\,dt \le KRTT - B(t) + \int_t^{t+RTT} O(t)\,dt$$

And apply the worst case of condition 2 by letting I(t)=K and O(t)=0 then we get

$$K\int_t^{t+RTT} C(t\text{-}RTT)\,dt \le KRTT - B(t),$$

which can be written as

equation 5: $K\int_{t-RTT}^{t} C(t)\,dt \le KRTT - B(t).$

Since the physical meaning of C(t) is the admission ratio of the sender, therefore, in order to maximize the bandwidth utilization, the value of c(t) is set to be as large as possible. The eq.5 hence becomes:

equation 6: $K\int_{t-RTT}^{t} C(t)\,dt = KRTT - B(t),$

which is the no cell loss equation for proportional feedback scheme. This is the second time we do optimization. The first time is when we minimize the buffer requirements in eq. 4. Here we optimize the $\int_{t-RTT}^{t} C(t)\,dt$ to maximum. Equation 5 means that to satisfy the condition of no cell loss, the total admission cells in the past RTT can not be more than the free buffer spaces (KRTT-B(t)) in receiver. We want the total admission cells to be as many as possible. Thus, in eq.6 we let the total admission cells to be equal to the free buffer spaces. The physical meaning can be explained by picture in Fig. 3.

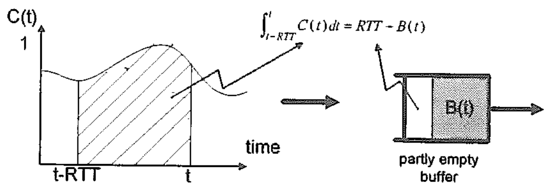

Figure 3. Interpretation of the proportional scheme.

1. 2 Analysis

In this section, we will analyze the performance of the proportional scheme. The analysis includes bandwidth utilization, queuing delay and transient response. For now, in our design of the flow control scheme, we have the following two equations to describe our system.

buffer equation: $B'(t)\,dt = I(t\text{-}RTT/2)C(t\text{-}RTT) - O(t)$
no cell loss equation:

$$K\int_{t-RTT}^{t} C(t)dt = KRTT - B(t)$$

Bandwidth Utilization Analysis

Bandwidth is said to be fully utilized if the allowed input rate can use all available bandwidth. Again, the basic equations for our system are the buffer equation and the no cell loss equation. By rearranging the buffer equation, we have

equation 7: $\dfrac{B'(t)+O(t)}{I(t-RTT/2)} = C(t-RTT)$

Now, we can rewrite the no cell loss equation because $\int_{t-RTT}^{t} c(t)dt = \int_t^{t+RTT} c(t\text{-}RTT)dt$ and we have

equation 8:

$$K\int_t^{t+RTT} C(t-RTT)dt = KRTT - B(t)$$

By combining eq.7 and eq.8, we get equation 9:

$$K\int_t^{t+RTT} \frac{B'(t)+O(t)}{I(t-RTT/2)}dt = KRTT - B(t)$$

In order to know the best bandwidth utility, I(t-RTT/2) is assumed to have maximum value K and Eq.9 becomes

$$\int_t^{t+RTT} (B'(t)+O(t))dt = KRTT - B(t)$$

$$B(t+RTT) - B(t) + \int_t^{t+RTT} O(t)dt = KRTT - B(t)$$

$$B(t+RTT) = k*RTT - \int_t^{t+RTT} O(t)dt$$

Let t = t'-RTT, we have equation 10:

$$B(t') = kRTT - \int_{t'-RTT}^{t'} O(t')dt'$$

By substituting eq.10 into no cell loss equation, we have $K\int_{t'-RTT}^{t'} C(t')dt' = \int_{t'-RTT}^{t'} O(t')dt'$, which implies that the maximum allowed cells are equal to available output bandwidth. Therefore, the output bandwidth is fully utilized. The result of eq.10 is shown in Fig. 4. As shown in the figure, the buffer is full while there is no output rate and the buffer is empty while output is in full bandwidth.

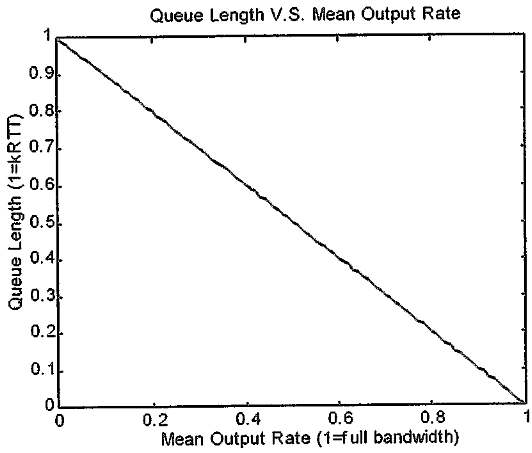

Figure 4. Mean output rate versus queue length.

Queuing Delay Analysis

Although in ATM network, ABR traffic does not have QoS guarantee, it's still desirable to reduce the cell transfer delay(CTD) as much as possible. The cell transfer delay in a connection is the sum of the propagation delay on the link, the queuing delay in the buffer and the cell service time in the switching unit. We are interested in the queuing delay.

At first, we have

equation 11: $\bar{O}$= mean O(t) over past RTT $= \frac{1}{RTT}\int_{t-RTT}^{t} O(t)\,dt$

For maximum queuing delay, constant ready input (eq.10) is applied. By combining eq. 10 and eq. 11, we have the following:

$\bar{B}$= mean B(t) over past RTT= $\frac{1}{RTT}\int_{t-RTT}^{t} B(t)dt$

$$= \frac{1}{RTT}\int_{t-RTT}^{t}(RTT - \int_{t-RTT}^{t} O(t)\,dt)\,dt$$

$$= \frac{(1-\bar{O})\,RTT^2}{RTT} = (1-\bar{O})\,RTT$$

Hence, the mean queuing delay $\bar{D}$ = $\frac{average\_buffer\_length}{average\_output\_rate} = \frac{\bar{B}}{\bar{O}} = \frac{(1-\bar{O})\,RTT}{\bar{O}}$

The result of above equation is shown in Fig. 5. In this figure, since the buffer space tends to be full when output rate goes down (as shown in Fig. 4), the mean queuing delay grows up exponentially when the mean output rate approaches zero. On the other hand, as the mean output rate goes up, the queuing delay goes to zero.

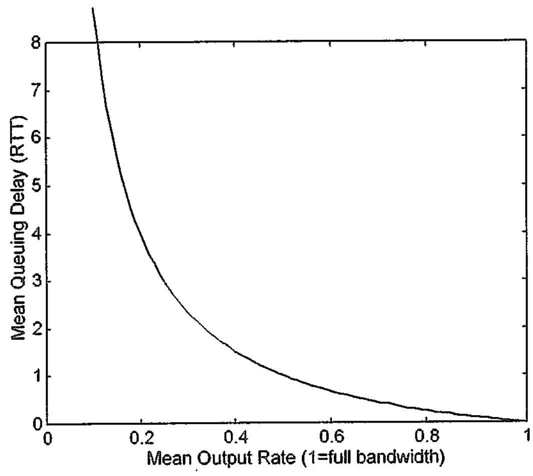

Figure 5. Mean queuing delay with different output bandwidth.

1.3 Transient Response Analysis

Transient response is regarding to how fast the system can react to the changes of network status. The network status may changes because available bandwidth varies or a newly joined VC wants to share the bandwidth. Let output bandwidth step up at time RTT and see how fast the system will react to the output changes. This is equal to the output is not blocked or the output is reconnected at time RTT. As shown in Fig. 6, when output bandwidth steps up to full bandwidth at time RTT, the cells in the buffer are still full and begin to decrease. According to Figure 3, the output notifies the input to begin transmission as soon as the buffer is not full. The input does not begin to transfer until it receives notice from the receiver at 1.5RTT. Cells in the buffer continue to decrease until its size reaches zero at 2RTT. This is because the output is in full rate between RTT and 2RTT and the cells from input at 1.5RTT do not arrive at the buffer before 2RTT. On the other hand, the input rate continues to increase from 1.5RTT to 2.5RTT.

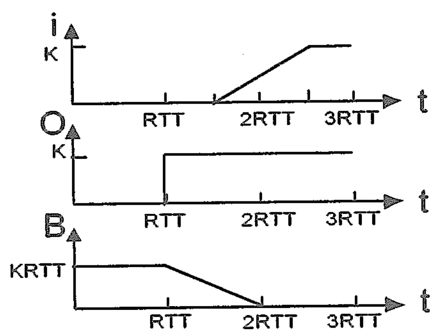

Figure 6. Transient response when the output bandwidth has a step-up change at time RTT.

III. Conclusion

In this paper, a proportional feedback scheme which is rate-based for hop-by-hop flow control is proposed. A mathematical model is built to analyze the proportional flow control scheme. It is shown this scheme can achieve 100% channel utility without any cell loss. The queuing delay of the scheme is shown to be low.

Reference:

[1] ATM Forum, "ATM User-Network Interface Specification," Version 3.0, (Prentice-Hall, Sep. 1993).

[2] C. Lefelhocz, "Congestion Control for Best-Effort Service: Why We Need A New Paradigm," IEEE Network, Jan./Feb. 1996, pp. 10-19.

[3] A. Demers, S. Keshav, and S. Shenker. "Analysis and Simulation of a Fair Queueing Algorithm," J. Internetworking: Res. and Exp., vol. 1, 1990, pp. 3-26; also in Proc. ACM SIGCOMM '89, pp. 3-12.

[4] M. Katevenis, "Fast Switching and Fair Control of Congested Flow in Broad-band Networks," IEEE JSAC, vol. 5, 1987, pp. 1315-26.

[5] E. Hahne, "Round-Robin Scheduling for Max-Min Fairness in Data Networks," IEEE JSAC, vol. 9, 1991.

[6] S. Morgon, "Queuing Disciplines and Passive Congestion Control in Byte-Stream Networks," Proc. Infocom '89, 1989, pp. 711-20.

[7] S. Shenker, "A Theoretical Analysis of Feedback Flow Control," Proc. Of the ACM Sigcomm Symposium, 1990.

[8] S. Keshav, "A Control-Theoretic Approach to Flow Control," Proc. SIGCOMM '91, Sept. 1991, pp. 3-15.

[9] K. K. Ramakrishnan and R. Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks," ACM Trans. On Computer Sys., May 1990.

[10] L. Roberts, "Enhanced Proportional Rate Control Algorithm," ATM Forum contribution 94-0735R1, Ottawa, Canada, Sept. 1994.

[11] Sui and Tzeng, 1995, "Intelligent Congestion Control for ABR Service in ATM Networks," U of C, Irvine.

[12] H. T. Kung, "The FCVC (Flow-Controlled Virtual Channels) Proposal for ATM Networks," Proc. Int. Conf. Network Protocols, San Francisco, Oct. 1993.

[13] H. T. Kung et al., "Use of Link-by-Link Flow Control in Maximizing ATM Network Performance: Simulation Results," Proc. IEEE Hot Interconnects Symp., Palo Alto, CA, Aug. 1993.

[14] H. T. Kung and K. Chang, "Receiver-Oriented Adaptive Buffer Allocation in Credit-Based Flow Control for ATM Networks," Proc. Infocom '95, April 1995, pp. 158-81.

[15] C. Ozveren, R. Simcoe, and G. Vorghese, "Reliable and Efficient Hop-by-Hop Flow Control," Proc. SIGCOMM '94, Aug. 1994, pp. 89-100.

[16] S. Shenker, L. Zhang, and D. Clark, "Some Observations on the Dynamics of a Congestion Control Algorithm," ACM Comp. Commun. Rev., vol. 20, no. 4, Oct. 1990, pp. 30-39.

[17] L. Zhang and D. Clark, "Oscillating Behavior of Network Traffic: A Case Study Simulation," J. Internetworking: Res. and Exp., vol. 1, 1990, pp. 101-12.