Dynamic Multithreshold Rate Control Mechanisms for Supporting ABR Traffic in ATM Networks

Pau C. Ting and Maria C. Yuang, *Member, IEEE*

***Abstract*— Existing feedback-based rate control schemes supporting the available bit rate (ABR) service in ATM networks mostly employ a single static buffer threshold at each switching node as the forewarning of congestion. In this paper, we first propose a continuous-based adaptive rate control mechanism, which employs, logically, an infinite number of thresholds. Each node periodically determines the precise permitted rate of immediate upstream nodes based on a simple fluid model aimed at satisfying both loss-free and starvation-free criteria. The scheme, as will be shown, achieves high utilization and low (zero) cell-loss probability under highly bursty (deterministic) traffic, but at the expense of a drastic increase in signalling overhead due to frequent adjustment of permitted rates. To reduce overhead, we further propose a so-called stepwise-based rate control mechanism adopting a limited number of movable thresholds, referred to as the threshold set. The threshold set shifts up (down) reflecting the increase (decrease) in departure rates. Compared to continuous-based control via simulation, stepwise-based control is shown to be efficient and accurate using a reasonably low number of thresholds. Moreover, we also display simulation results, which demonstrate that the stepwise-based mechanism outperforms existing single-static-threshold-based schemes in terms of cell-loss probability and link utilization.**

***Index Terms*— Available bit rate, binary rate control, cell-loss probability, feedback-based rate control, fluid model, quality of service.**

I. Introduction

THE available bit rate (ABR) [1]–[5] service in ATM networks [6], [7] has been deployed to allow efficient use of available bandwidth without degrading the quality of service (QOS) [6], [7] of admitted traffic. While the QOS of admitted traffic is guaranteed through admission control [8]–[10] and bandwidth allocation [11], [12], the ABR has been realized via the feedback-based rate control [2]–[4], [6], [13]–[28]. Feedback-based rate control deals with the dynamic adjustment of the granted rates of ABR sources as network loads fluctuate in an attempt to minimize the performance degradation of QOS-guaranteed services. Existing rate control mechanisms operate either on an end-to-end [2], [15]–[18], [20], [28]–[30] or hop-by-hop [2], [5], [19], [21]–[27] basis. While both classes of control mechanisms possess individual performance merit, hop-by-hop-based control has been considered to be more promising, due to its speedy reaction to the fluctuation of network loads [22], [23].

Manuscript received October 13, 1996; revised February 4, 1997.

The authors are with the Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, 30050 Taiwan, R.O.C.

Publisher Item Identifier S 0278-0046(98)00409-2.

Most hop-by-hop-based schemes adopt a static buffer threshold [2], [19], [22], [24]–[27] at each switching node as the forewarning of congestion. In these schemes, a switching node sends feedback messages to its immediate upstream nodes should the buffer occupancy exceed the predetermined threshold. The upstream nodes, in turn, adjust the cell departure rates on either a simple binary rate (i.e., start and stop) [2], [19], [26] or a specified rate [2], [22], [24], [25], [27] basis. Mishra *et al.* [22], [25] proposed a predictive rate control scheme for determining the permitted rate and illustrated that the buffer occupancy and throughput of a controlled connection converge to a desired operating point. Kawahara *et al.* [19] developed an analytical model based on binary rate control and showed significant performance improvement in terms of cell-loss probability (CLP) and resource utilization of the congested node, but at the expense of signalling overhead. Ko *et al.* [24] demonstrated the effectiveness of the proposed predictive control scheme in controlling congestion. Ideda and Suzuki [27] showed the effectiveness of a combination of the adaptive rate control with fast reservation protocol (FRP) in achieving high link utilization, while maintaining low head-of-line blocking probability, particularly on ATM local area networks. Exhibiting various performance credits, these schemes, however, result in improper rate determination due to the employment of one static threshold.

In this paper, we first propose a continuous-based adaptive rate control mechanism which employs, logically, an infinite number of thresholds. Each node periodically determines the precise permitted rate of immediate upstream nodes based on a simple fluid model aiming at satisfying both loss-free and starvation-free criteria. The scheme, as will be shown, achieves high utilization and zero CLP under deterministic traffic and low CLP under highly bursty traffic. The price paid is, however, a drastic increase in signalling overhead due to frequent adjustment of permitted rates. We further propose a so-called stepwise-based rate control mechanism adopting a limited number of movable thresholds, referred to as the threshold set. The threshold set shifts up (down) reflecting the increase (decrease) in departure rates. Compared to continuous-based control via simulation, stepwise-based control is shown to be efficient and accurate using only a reasonably low number of thresholds. Moreover, we also display simulation results which demonstrate that the stepwise-based mechanism outperforms existing single-static-threshold-based schemes in terms of CLP and link utilization.

The remainder of this paper is organized as follows. Section II presents the architecture of the feedback-based rate

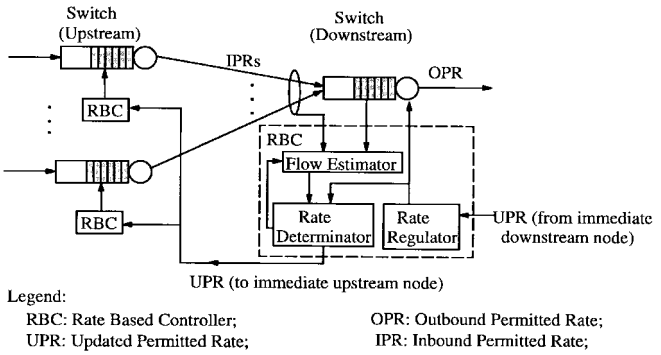


Fig. 1. System architecture.

controller and proposes the continuous-based rate control mechanism. Section III then introduces the stepwise-based rate control mechanism. Performance justification and comparisons through simulation results are provided in Section IV. Finally, concluding remarks are given in Section V.

II. System Architecture

Each switching node on which feedback-based rate control operates consists of a finite buffer and a rate-based controller (RBC), as shown in Fig. 1. In principle, in accordance with the buffer occupancy, the RBC of a switching node at each time unit (described later) determines the updated permitted transfer rate (UPR) for all immediate upstream nodes by sending feedback messages incorporating such rate. This rate then becomes the outbound permitted rate (OPR) of those immediate upstream nodes or the inbound permitted rate (IPR) of this current node throughout the next time unit.

The RBC is composed of a flow estimator, a rate determinator, and a rate regulator. At each time unit, the flow estimator predicts the aggregate flow of future incoming traffic based on the previous UPR and the current buffer occupancy. The rate determinator, in turn, determines the new UPR achieving two performance criteria (described later) based on a rate control law. Finally, the rate regulator ensures that the transfer rate never exceeds the granted OPR. In Sections II-A and B, we first introduce the two performance criteria, followed by the description of the design of the rate determinator and the flow estimator.

A. Performance Criteria

Basically, our rate-based control mechanism has been designed to achieve two performance criteria, i.e., loss free and starvation free. The loss-free requirement ensures that the buffer of each switching node never overflows. The starvation-free requirement then assures the achievement of a maximum of link utilization at all times. In the following, assumptions and notations used throughout the entire paper are first introduced, followed by the formulation of the two performance criteria.

In the system, the basic time unit for the determination of the UPR is the round-trip delay D (assumed to be a constant) between any two adjacent nodes. This time unit is referred to as a *superslot*. The maximum normalized transfer rate of node i is denoted as $R_i (R_i \leq 1)$. The state of each node is observed and modified upon receiving a feedback message at the beginning of each superslot. Consider a generic scenario [Fig. 2(a)] with sources having an infinite backlog of ABR traffic and a series of switching nodes, each of which (for example, node i) has a finite buffer of fixed size B and serves a set of M_{i-1} incoming sources. The data and control flows and variables used are depicted in Fig. 2(b). In the figure, the solid line and dotted line represent the data and feedback control flows, respectively. Let x_i^k denote the buffer occupancy of node i and μ_{i-1}^k the UPR for node $i-1$ computed by node i, at the beginning of superslot k. As shown in the figure, upon receiving the UPR μ_i^{k-1} sent from node $i+1$ at the beginning of superslot k, node i, in turn, determines the UPR μ_{i-1}^k for node $i-1$ according to the current buffer occupancy x_i^k. Now, we are at the stage of formally defining the two performance criteria.

1) Criterion 1—Loss Free: Significantly, the loss-free criterion comprises the upper bound of the UPR. Clearly, this criterion for node i throughout superslot $k+1$ is satisfied if and only if the buffer occupancy $x_i^n \leq B$, where $k+1 < n \leq k+2$. Based on the deterministic fluid model [29], [30], $x_i^n (k+1 < n \leq k+2)$ can be expressed as a function of x_i^{k+1}, the IPR (μ_{i-1}^k), and the OPR (μ_i^k). That is,

$$\begin{aligned} x_i^n &\simeq \begin{bmatrix} \text{Buffer occupancy} \\ \text{at the beginning of} \\ \text{superslot } (k+1) \end{bmatrix} \\ &\quad + \left[\begin{pmatrix} \text{Aggregate arrival rate} \\ \text{during} \\ \text{superslot } (k+1) \end{pmatrix} \right. \\ &\quad \left. - \begin{pmatrix} \text{Service rate} \\ \text{during} \\ \text{superslot } (k+1) \end{pmatrix} \right] \\ &\quad \times [\text{Observation interval}] \\ &\simeq x_i^{k+1} + (M_{i-1}\mu_{i-1}^k - \mu_i^k) \times (n-(k+1))D \\ &\leq B, \quad \text{for } k+1 < n \leq k+2. \end{aligned} \tag{1}$$

Rearranging (1), we get

$$\mu_{i-1}^k \leq \frac{\mu_i^k}{M_{i-1}} + \frac{B - x_i^{k+1}}{(n-(k+1))DM_{i-1}} \quad \text{for } k+1 < n \leq k+2. \tag{2}$$

Let $\overline{\mu_{i-1}^k}$ denote the upper bound of μ_{i-1}^k, occurring at $n = k+2$. Thus,

$$\overline{\mu_{i-1}^k} = \frac{\mu_i^k}{M_{i-1}} + \frac{B - x_i^{k+1}}{DM_{i-1}} = \frac{\mu_i^k + \mu_i^{k-1}}{M_{i-1}} + \frac{B - x_i^k}{DM_{i-1}} - \mu_{i-1}^{k-1}. \tag{3}$$

2) Criterion 2—Starvation Free: Essentially, the starvation-free criterion forms the lower bound of the UPR. In principle, the starvation-free criterion for a node is satisfied if the buffer of the immediate upstream node remains nonempty at all times. Thus, the criterion directly applies that, considering node i at superslot $k+1$, the buffer occupancy

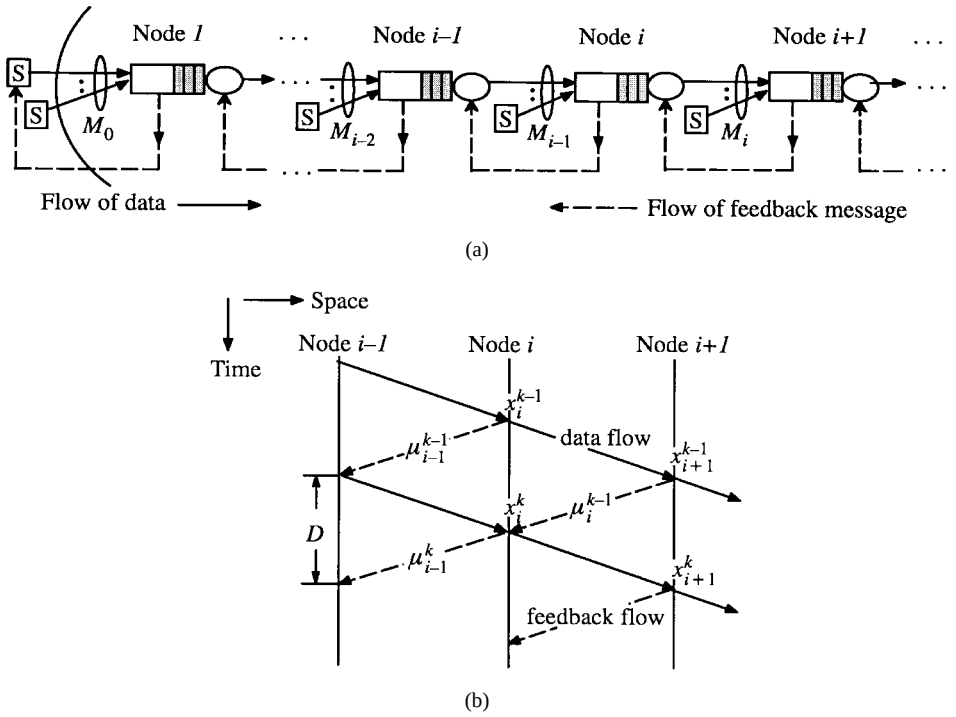


Fig. 2. A generic scenario. (a) Topology. (b) Data and control flows.

$x_i^n (k+1 < n \leq k+2)$ given in (1) should be greater than zero. That is,

$$x_i^n = x_i^{k+1} + (n-(k+1))D(\mu_{i-1}^k M_{i-1} - \mu_i^k) > 0, \quad \text{for } k+1 < n \leq k+2. \tag{4}$$

Rearranging (4), we get

$$\mu_{i-1}^k > \frac{\mu_i^k}{M_{i-1}} - \frac{x_i^{k+1}}{(n-(k+1))DM_{i-1}}, \quad \text{for } k+1 < n \leq k+2. \tag{5}$$

Let $\underline{\mu_{i-1}^k}$ denote the lower bound of μ_{i-1}^k, which occurs at $n = \overline{k+2}$. Obviously,

$$\underline{\mu_{i-1}^k} = \frac{\mu_i^k}{M_{i-1}} - \frac{x_i^{k+1}}{DM_{i-1}} = \frac{\mu_i^k + \mu_i^{k-1}}{M_{i-1}} - \frac{x_i^k}{DM_{i-1}} - \mu_{i-1}^{k-1}. \tag{6}$$

Based on (3) and (6), we obtain the following inequality:

$$\max\{0, \underline{\mu_{i-1}^k}\} < \mu_{i-1}^k \leq \min\{R_{i-1}, \overline{\mu_{i-1}^k}\}. \tag{7}$$

The set of UPR's satisfying (7) is depicted by the shaded area, referred to as the *safe region*, in Fig. 3.[1] Notice that UPR's violating the loss-free and starvation-free criteria fall into two other regions, namely, regions I and II, respectively. In Section II-B, we propose a continuous-based rate control mechanism for the determination of the UPR's within the safe region.

[1]Notice that function $\overline{\mu_{i-1}^k}$ $(\underline{\mu_{i-1}^k})$ can be expressed as a function of μ_i^{k-1}, μ_i^k, and x_i^k, as shown in (3), namely, by a three-dimensional irregular-surface diagram. By fixing μ_i^{k-1}, μ_i^k, and x_i^k as constants, we illustrate μ_{i-1}^k $(\overline{\mu_{i-1}^k}$ or $\underline{\mu_{i-1}^k})$ as a function of μ_{i-1}^{k-1} by a two-dimensional diagram, as shown in Fig. 3.

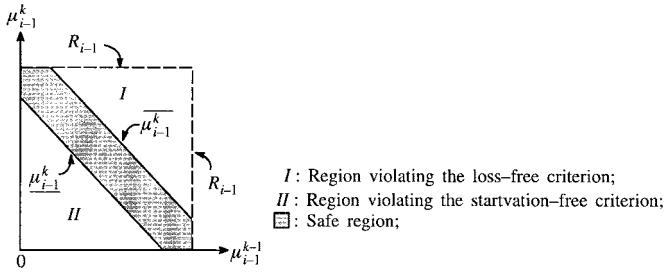


Fig. 3. The safe region of the UPR μ_{i-1}^k.

B. Continuous-Based Rate Control

Fundamentally, the UPR is dependent on the remaining buffer space and the OPR. Thus, at the beginning of superslot k at node i, the UPR is given as

$$\mu_{i-1}^k = \frac{\mu_i^k}{M_{i-1}} + \gamma \left[\frac{B - x_i^{k+1}}{D \cdot M_{i-1}} \right] \tag{8}$$

where M_{i-1} is the total number of multiplexed sources and parameter γ is the weight factor [22] with respect to the remaining buffer space. In particular, $\gamma = 1$ represents a fully linear relationship between the remaining buffer space and the UPR, whereas $\gamma = 0$ corresponds to an independent relationship instead. Basically, γ controls the rate of convergence to the buffer size in an attempt to achieve high link utilization. Notice that one can easily prove that the UPR at the beginning of any superslot would fall into the safe region if $0 < \gamma < 1$. Now, to compute μ_{i-1}^k in (8), we have to first obtain x_i^{k+1} and μ_i^k. Unfortunately, node i has no knowledge of x_i^{k+1} and μ_i^k at the beginning of superslot k. Therefore, the approximation

of μ_{i-1}^k, denoted as $\tilde{\mu}_{i-1}^k$, is reexpressed from (8) as

$$\tilde{\mu}_{i-1}^k = \frac{\hat{\mu}_i^k}{M_{i-1}} + \gamma\left[\frac{B - \hat{x}_i^{k+1}}{D \cdot M_{i-1}}\right] \tag{9}$$

where $\hat{x}_i^{k+1}$ and $\hat{\mu}_i^k$ are the predicted x_i^{k+1} and μ_i^k, respectively.

Based on the deterministic fluid model, $\hat{x}_i^{k+1}$ is simply evaluated as

$$\hat{x}_i^{k+1} = x_i^k + D \cdot (M_{i-1} \cdot \tilde{\mu}_{i-1}^{k-1} - \tilde{\mu}_i^{k-1}). \tag{10}$$

As for $\hat{\mu}_i^k$, we first make the following observation. Suppose node i located at m $(m \geq 0)$ hops upstream from the first node experiencing congestion, called the bottleneck. The bottleneck is supposed to send the throttle message at the beginning of superslot k'. It is worth noting that node i has transferred data by its maximum allowed transfer rate (R_i) until receiving this throttle message at superslot $k' + m$. Node i, in turn, reduces its transfer rate in accordance with the transfer rate of the bottleneck equally shared by $\Pi_{j=0}^{m-1} M_{i+j}$ upstream nodes. Thus, $\hat{\mu}_i^k$ can be estimated from its current OPR and the transfer rate of the bottleneck, excluding the consideration of factor γ. With γ taken into account, $\hat{\mu}_i^k$ is computed as

$$\hat{\mu}_i^k = \begin{cases} \tilde{\mu}_i^{k-1} + \gamma\left[\dfrac{R_{i+m}}{\prod_{j=0}^{m-1} M_{i+j}} - \tilde{\mu}_i^{k-1}\right], \\ \qquad \text{if } m \geq 1, k \geq k' + m + 1 \\ \tilde{\mu}_i^{k-1}(= R_i), \\ \qquad \text{otherwise.} \end{cases} \tag{11}$$

Replacing $\hat{x}_i^{k+1}$ and $\hat{\mu}_i^k$ by (10) and (11), respectively, $\tilde{\mu}_{i-1}^k$ in (9) becomes

$$\begin{aligned} \tilde{\mu}_{i-1}^k = {} & \tilde{\mu}_{i-1}^{k-1} + (1+\gamma)\left[\frac{\tilde{\mu}_i^{k-1}}{M_{i-1}} - \tilde{\mu}_{i-1}^{k-1}\right] \\ & + \gamma\left[\frac{B - x_i^k}{D \cdot M_{i-1}}\right] + \frac{\hat{\alpha}_i^{k-1}}{M_{i-1}} \end{aligned} \tag{12}$$

where $\hat{\alpha}_i^{k-1} = \hat{\mu}_i^k - \tilde{\mu}_i^{k-1}$.

In *Theorem 1*, we state and prove that the rate control mechanism governed by (12) truly guarantees the loss-free criterion. Throughout the proof, we assume that there exists one bottleneck in the network. This bottleneck (for example, node G) possesses a transfer rate of R_G and starts transmitting the throttle message at the beginning of superslot k'. We now observe node $G - m$, namely, m $(m \geq 0)$ hops upstream from the bottleneck. It is worth noting that node $G - m$ is loss free prior to superslot $k' + m$, i.e., before having received the throttle message.

Theorem 1: The buffer occupancy and UPR predicted based on (10) and (12) are conservative by nature. That is, considering node $G - m$ at the beginning of superslot k $(k \geq 0)$,

$$x_{G-m}^{k+1} \leq \hat{x}_{G-m}^{k+1}, \text{ and } \mu_{G-m-1}^k \geq \tilde{\mu}_{G-m-1}^k. \tag{13}$$

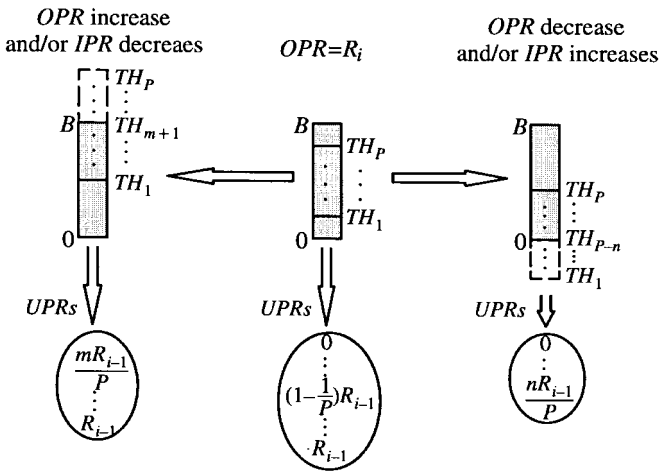


Fig. 4. Sets of movable thresholds and UPR's at node i.

That is to say, the rate control mechanism governed by (12) satisfies the loss-free criterion.

Proof: See the Appendix.

III. Stepwise-Based Rate Control

The continuous-based rate control mechanism proposed above performs rapid readjustment of UPR's on a superslot basis. Consequently, frequent computation and adjustment of UPR's results in a drastic increase in signaling overhead. To reduce the overhead, we introduce the notion of multiple movable thresholds and propose a stepwise-based rate control mechanism.

A. Basic Concept

In stepwise-based rate control, each node, for example, node i, assigns a set of evenly distributed $(P+1)$ UPR's, namely, $\{0, R_{i-1}/P, 2R_{i-1}/P, \cdots, R_{i-1}\}$, to respective $P+1$ partitions of the buffer space, delimited by a set of P movable thresholds. Basically, the threshold set is altered should the granted OPR or IPR be modified. Upon receiving the signalling of the reduction (increase) of the OPR, the threshold set shifts down (up) reflecting more (less) stringent constraint on the departure rate. On the other hand, as the granted IPR is reduced (increased), the threshold set shifts up (down) reflecting more (less) stringent constraints on the arrival rate. Fig. 4 depicts the threshold sets and legitimate UPR's with respect to the alteration of the granted OPR and/or IPR at node i.

As shown at the right side of Fig. 4, the legitimate UPR set shifts up as the OPR declines and/or the IPR rises due to the reduction of prositive thresholds. On the other hand, the legitimate UPR set shifts down as the OPR increases and/or the IPR declines, as shown at the left side of Fig. 4. Accordingly, at the beginning of each superslot, a new UPR is reassigned and notified to immediate upstream nodes only when the current buffer occupancy is altered to a different partition (this may occur when either the buffer occupancy or the threshold set is changed). Thus, if the buffer occupancy settles in the same partition, the granted UPR remains the same, resulting in the elimination of transferring feedback messages and, thus, the reduction in signalling overhead.

Two problems have been considered in the design of the stepwise-based rate control. The first problem is the determination of the most advantageous number of partitions P subject to the round-trip delay D and traffic characteristics. Apparently, the step-wise based mechanism using an infinitely large P is logically identical to the continuous-based mechanism, whereas the mechanism using $P = 1$ becomes a simple start–stop rate control mechanism. The impact of various P's on the system performance will be evaluated in Section IV. The second problem is the computation of dynamic threshold sets, which is discussed next in great detail.

B. Determination of Threshold Sets

Let $\tilde{v}_{i-1}^k$ denote the UPR computed by node i at the beginning of superslot k. Assume that $\tilde{v}_{i-1}^k$ falls into the range of $[(1-n/P)R_{i-1}, (1-(n-1)/P)R_{i-1}]$. That is,

$$\left(1-\frac{n}{P}\right)\cdot R_{i-1} \le \tilde{v}_{i-1}^k \le \left(1-\frac{n-1}{P}\right)\cdot R_{i-1} \quad \text{where } 1 \le n \le P. \tag{14}$$

Replacing $\tilde{v}_{i-1}^k$ by the right-hand side of (12) and then rearranging the inequality, we obtain

$$x_i^k \ge B - \left(1+\frac{1}{\gamma}\right)(M_{i-1}R_{i-1} - \tilde{v}_i^{k-1})D + M_{i-1}(R_{i-1} - \tilde{v}_{i-1}^{k-1})D + \frac{\tilde{\alpha}_i^{k-1}D}{\gamma} + (n-1)\frac{M_{i-1}R_{i-1}D}{\gamma\cdot P} \tag{15}$$

$$x_i^i \le B - \left(1+\frac{1}{\gamma}\right)(M_{i-1}R_{i-1} - \tilde{v}_i^{k-1})D + M_{i-1}(R_{i-1} - \tilde{v}_{i-1}^{k-1})D + \frac{\tilde{\alpha}^{k-1}D}{\gamma} + n\cdot\frac{M_{i-1}R_{i-1}D}{\gamma\cdot P}. \tag{16}$$

The expression at the right-hand side of inequality (15) is designated as $TH_{i,n}^k$, defined as the nth threshold of node i during superslot k. Similarly, the expression at the right-hand side of inequality (16) is denoted as $TH_{i,n+1}^k$. That is,

$$\begin{aligned} TH_{i,n+1}^K &= TH_{i,n}^k + \Delta_i \quad \text{for } 1 \le n \le P-1 \\ TH_{i,1}^k &= B - \left(1+\frac{1}{\gamma}\right)(M_{i-1}R_{i-1} - \tilde{v}_i^{k-1})D \\ &\quad + M_{i-1}(R_{i-1} - \tilde{v}_{i-1}^{k-1})D + \frac{\tilde{\alpha}_i^{k-1}D}{\gamma} \end{aligned} \tag{17}$$

where $\Delta_i = (M_{i-1}R_{i-1}D)/(\gamma P)$ is the fixed increment between two adjacent thresholds, referred to as the *interthreshold width*. Equation (17) is hereinafter referred to as the *threshold-setting* function. Notice that node i initially sets $TH_{i,1}^1$ as $B-(1+1/\gamma)(M_{i-1}R_{i-1}-R_i)D$. Clearly, since each threshold value is a function of the previous UPR $(\tilde{v}_{i-1}^{k-1})$ and the current OPR $(\tilde{v}_i^{k-1})$, the number of possible threshold sets generated by the threshold-setting function is bounded by $(P+1)^2$. In addition, the threshold-setting function is monotonically nondecreasing with respect to the OPR and monotonically nonincreasing with respect to the IPR.

Input : buffer size– B (cells);
number of partitions– P;
maximum transfer rate– R (cell/slot);
Variable : x : buffer occupancy (cells);
TH_n : nth threshold (cells) where $1 \le n \le P$;
UPR_{old} : previous UPR (cell/slot);
UPR_{new} : new UPR (cell/slot);
OPR : outbound permitted rate (cell/slot);
Initiation : Δ : the inter–threshold width (cells);
$TH_0 = 0$;
$TH_{P+1} = B$;
For super–slot k
Do
{ (i). **If** UPR_{new} is received from immediate downstream node **then**
{ Set $OPR = UPR_{new}$;
Calculate TH_1 based on Equation (16);
For $n = 2$ to P **do** $TH_n = TH_{n-1} + \Delta$; }
(ii). **For** $n = 0$ to P **do**
{ **If** $(TH_n \le x < TH_{n+1})$ **then**
{ $UPR_{new} = (1-\frac{n}{P})R$; } }
(iii). **If** $(UPR_{new} \ne UPR_{old})$ **then**
Notify UPR_{new} to immediate upstream nodes; }

Fig. 5. Stepwise-based rate control algorithm.

Upon determining the threshold set, the UPR $(\tilde{v}_{i-1}^k)$ can afterward be assigned according to the current buffer occupancy. As a result,

$$\tilde{v}_{i-1}^k = \begin{cases} R_{i-1}, & \text{if } 0 \le x_i^k < TH_{i,1}^k. \\ \left(1-\frac{n}{P}\right)R_{i-1}, & \text{if } TH_{i,n}^k \le x_i^k < TH_{i,n+1}^k \\ & \quad \text{where } n = 1, 2, \cdots, P-1 \\ 0, & \text{if } TH_{i,P}^k \le x_i^k \le B \end{cases} \tag{18}$$

as shown by the initial case $(\mathrm{OPR} = R_i)$ in Fig. 4. As the OPR increases and/or IPR decreases until $TH_{i,n'}^k > B$ is satisfied $(1 \le n' \le P)$,

$$\tilde{v}_{i-1}^k = \begin{cases} R_{i-1}, & \text{if } 0 \le x_i^k < TH_{i,1}^k. \\ \left(1-\frac{n}{P}\right)\cdot R_{i-1}, & \text{if } TH_{i,n}^k \le x_i^k < TH_{i,n+1}^k \\ & \quad \text{and } 1 \le n \le n'-1 \end{cases} \tag{19}$$

as shown at the left part of Fig. 4. On the other hand, as the OPR decreases and/or IPR increases until $TH_{i,n''}^k < 0$ is satisfied $(1 \le n'' \le P)$,

$$\tilde{v}_{i-1}^k = \begin{cases} \left(1-\frac{n}{P}\right)\cdot R_{i-1}, & \text{if } TH_{i,n}^k \le x_i^k < TH_{i,n+1}^k \\ & \quad \text{and } n'' \le n \le P-1 \\ 0, & \text{if } TH_{i,P}^k \le x_i^k < B \end{cases} \tag{20}$$

as shown at the right side of Fig. 4. The stepwise-based rate control algorithm based on the above dynamic thresholds is formally presented in Fig. 5. Finally, in the following theorem, we prove that the stepwise-based rate control mechanism satisfies the loss-free criterion.

Theorem 2: Given the threshold-setting function, the UPR determined from (19) and (20) is more conservative than the UPR computed from (12). That is, considering node i at the beginning of superslot $k(k>1)$,

$$\tilde{\mu}_{i-1}^k \ge \tilde{v}_{i-1}^k. \tag{21}$$

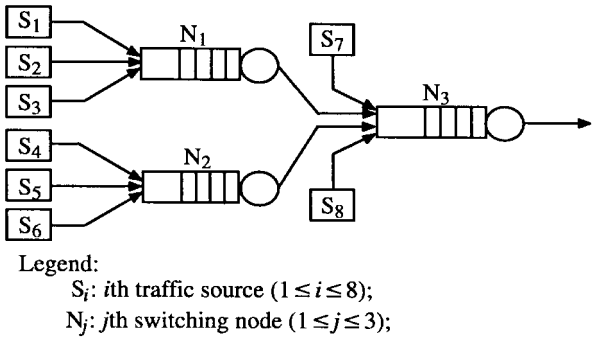
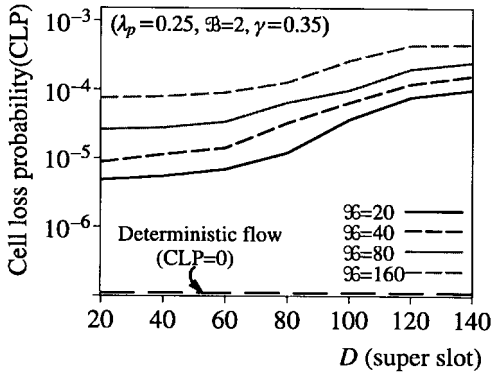
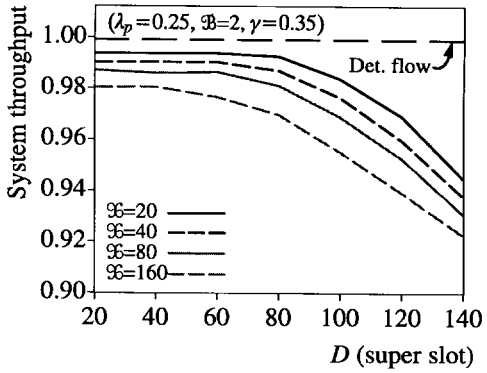


Fig. 6. Configuration for simulation.



(a)



(b)

Fig. 7. Performance of continuous-based rate control. (a) CLP. (b) System throughput.

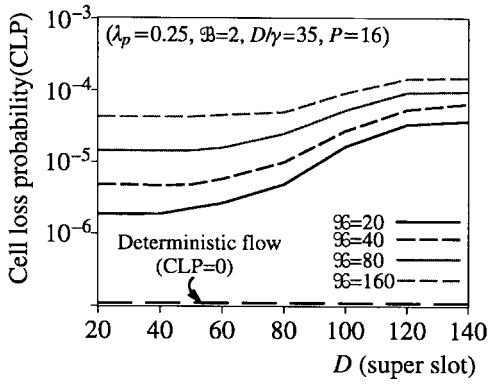
That is to say, the stepwise-based rate control mechanism satisfies the loss-free criterion.

Proof: Notice that rate $\tilde{\mu}_{i-1}^k$ determined based on continuous-based rate control must fall between the two bounds in the inequality in (14). Moreover, from (18), $\tilde{v}_{i-1}^k$ is taken as a low bound of legitimate UPR's. Thus, $\tilde{\mu}_{i-1}^k \geq \tilde{v}_{i-1}^k$ holds for $k > 1$. Hence, the more conservative stepwise-based rate control indeed satisfies the loss-free criterion. ■

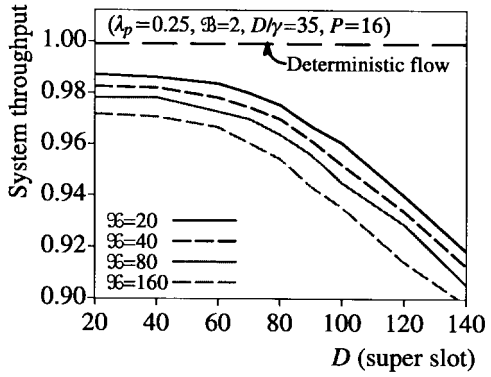
IV. Experimental Results

We first examine the effectiveness of our continuous and stepwise-based rate control under both deterministic fluid flow and realistic bursty traffic via simulation. Next, for the stepwise-based mechanism, we further investigate the impact of the number of partitions, weight factor, and traffic burstiness, on three performance metrics, namely, CLP, system throughput, and signalling overhead. System throughput is defined as the ratio of the total number of successfully transmitted cells to the total number of generated cells. Signalling overhead is quantized as the ratio of the total number of feedback messages generated using stepwise-based control to those using continuous-based control. Finally, we draw performance comparisons with respect to the three aforementioned metrics between the stepwise-based mechanism using one movable dynamic threshold and the existing binary rate control scheme [19], referred to as BRC, using one static threshold.

In simulation, any bursty traffic was modeled by an interrupted Bernoulli process (IBP) [31]. Such arrival process can be characterized in terms of mean burst length $(\mathcal{X})$, mean silence length $(\mathcal{L})$, peak arrival rate (λ_p), and the burstiness $(\mathcal{B})$. Accordingly, for an IBP arrival, the mean arrival rate and burstiness are given by $\lambda_p \times (\mathcal{X}/\mathcal{X}+\mathcal{L})$ and $\mathcal{X}+\mathcal{L}/\mathcal{X}$, respectively. The network configuration used in simulation is depicted in Fig. 6. In the experimented network, each switching node was assumed to possess a finite buffer of 200 cells and a maximum transfer rate of 1 cell/slot. Traffic from different sources was assumed to be homogeneous, and each has an infinite backlog of traffic. In addition, the total offered load, defined as the sum of the mean arrival rate from each source, is set as unity, designating the heavy traffic condition of the network.



(a)



(b)

Fig. 8. Performance of stepwise-based rate control. (a) CLP. (b) System throughput.

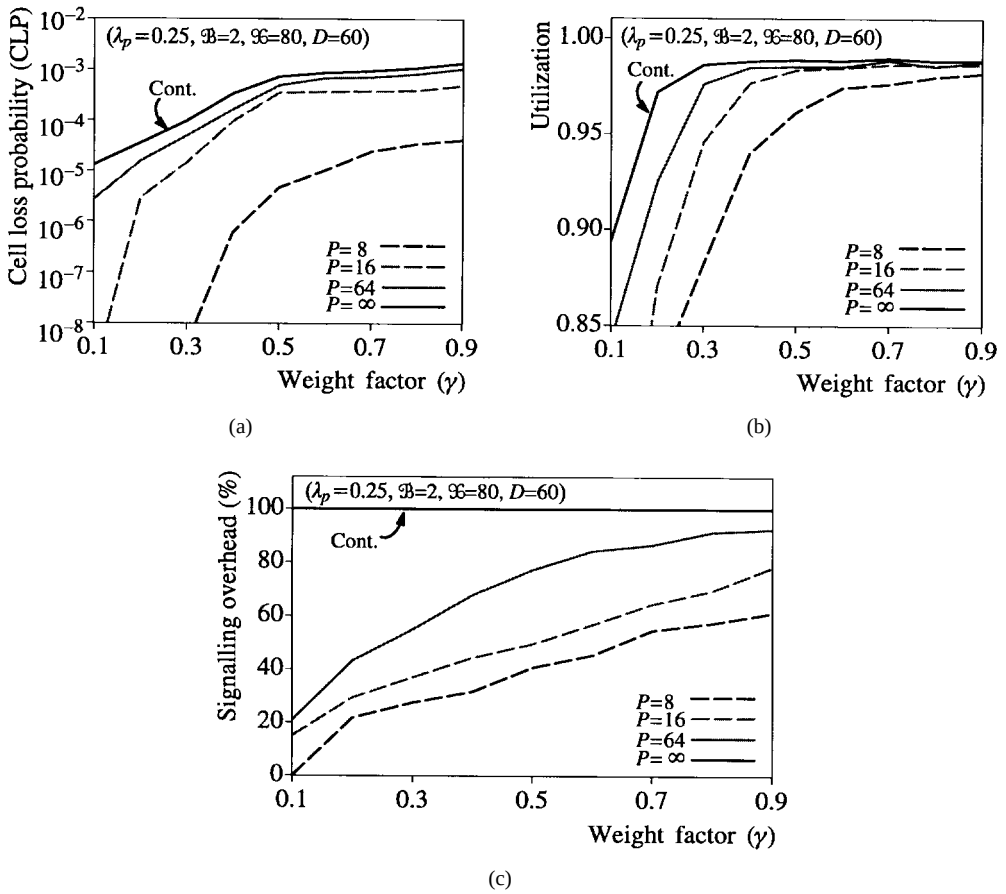


Fig. 9. Performance comparisons between continuous-based and stepwise-based control. (a) CLP. (b) Utilization. (c) Signalling overhead.

Fig. 7 illustrates CLP and system throughput under various burst lengths, $\mathcal{X}$ based on continuous-based rate control adopting a weight factor γ of 0.35. The figure shows that, under deterministic traffic, loss-free $(\mathrm{CLP} = 0)$ transmissions and a system throughput of unity can be achieved, as was justified by the analysis. Under bursty traffic, on the other hand, the mechanism yields nonzero CLP and system throughput of less than one. In particular, the higher the burstiness, the greater the CLP and the smaller the system throughput. This is because an increase in traffic burstiness results in a decrease in statistical multiplexing gain [7]. Moreover, smaller D's yield better CLP and system throughput due to faster adjustment of UPR's.

Fig. 8 demonstrates the impact of round-trip delay D on the performance of stepwise-based control under various mean burst lengths. In simulation, γ was altered in correspondence with the round-trip delay D yielding the ratio of D to γ, a constant. The rationale behind this is that the conservativeness or aggressiveness of the control mechanism is dependent on the dual consideration of D and γ, namely, D/γ in this case. The figure exhibits that simulation results for the deterministic traffic justify the correctness of the above analysis. Moreover, traffic with greater mean burst length $\mathcal{X}$ yields deteriorate performance. More significantly, compared to the performance of continuous-based control, stepwise-based rate control, having been proved to be more conservative, results in better CLP, but poorer system throughput.

Fig. 9 draws performance comparisons between continuous-based and stepwise-based control for a variety of partitions under various weight factors. The results show that the increase of the weight factor γ results in profound improvement in link utilization, however, at the expense of degradation in CLP and signalling overhead. This result can be justified as follows. Notice that the weight factor γ designates the significance of the buffer occupancy to the determination of the UPR. Consequently, higher γ yields more aggressive UPR and, in turn, results in higher link utilization and poorer CLP due to frequent fluctuation of the buffer occupancy. It is worth noticing that stepwise-based control performs compatibly to continuous-based control as P increases, resulting in better link utilization, but poorer CLP and signalling overhead.

Finally, we draw performance comparisons, via simulation, between our stepwise-based control mechanism using single dynamic threshold and the existing static-threshold BRC. In simulation, we employed a simpler network, with sources S_3, S_6, S_7, and S_8 removed from the network shown in Fig. 6. In the simulation of BRC, the static threshold at each buffer was set as $[B + 1 - (D + 1)\times$ (the number of multiplexed sources $-1)]$. In the simulation of our mechanism, we set the interthreshold width as 40, namely, $D/\gamma = 20$. Simulation results are depicted in Fig. 10. As shown in Fig. 10(a) and (b), BRC imposes higher CLP due to its deficiency in the adaptability to the throttled departure rate. As a result, the node which is located upstream from the congested node suffers from high

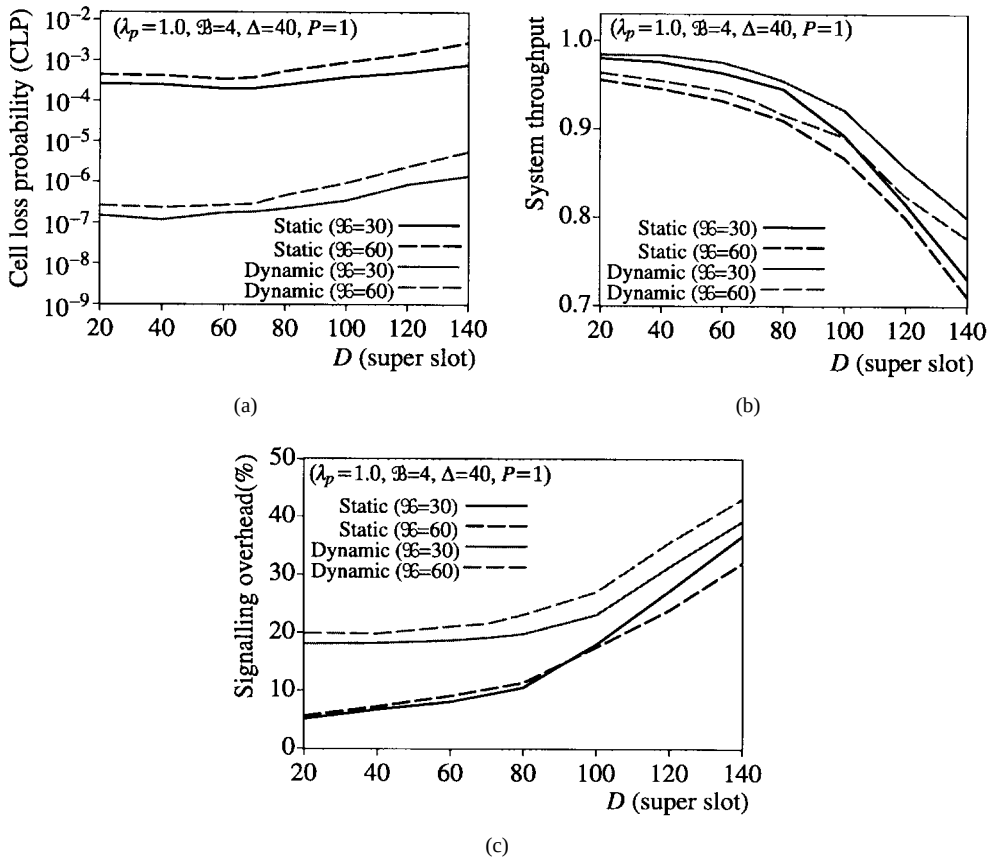


Fig. 10. Performance comparisons between static-based and dynamic-based control. (a) CLP. (B) System throughput. (c) Signalling overhead.

cell loss. By contrast, our dynamic-threshold-based control achieves much reduced CLP, however, with the price of greater signalling overhead being paid, as shown in Fig. 10(c). Moreover, Fig. 10(b) exhibits that our mechanism achieves higher system throughput than BRC.

V. Conclusions

In this paper, we first proposed a continuous-based adaptive rate control mechanism logically employing an infinite number of thresholds. Each node periodically determines the precise UPR of immediate upstream nodes based on a simple fluid model aiming to satisfy both loss-free and starvation-free criteria. The scheme, as has been shown, achieves high utilization and low CLP under even highly bursty traffic. We further proposed the stepwise-based rate control mechanism, adopting a limited number of movable thresholds. Simulation results justified the performance compatability of stepwise-based control to continuous-based control with signalling overhead much reduced. Moreover, simulation results also demonstrated the superiority of the stepwise-based mechanism over the existing single-static-threshold-based scheme in terms of cell loss probability and link utilization at the expense of tolerable signalling overhead.

Appendix
Proof of Theorem 1

It is worth reminding that node $G-m$ transfers data by R_{G-m} until receiving the throttle message at the beginning of the $(k'+m)$th superslot. Let the actual aggregate arriving rate of node $G-m$ at time t be denoted as $\lambda_{G-m}(t)$. Thus, for $0 \le k < k'+m$ and $m \ge 0$,

$$\begin{aligned} x_{G-m}^{k+1} &= x_{G-m}^{k} + \int_{(k-1)D}^{kD} (\lambda_{G-m}(t) - R_{G-m})\, dt \\ &\le x_{G-m}^{k} + D \cdot (M_{G-m-1} \cdot R_{G-m-1} - R_{G-m}) \\ &= \hat{x}_{G-m}^{k+1}. \end{aligned} \tag{22}$$

After the $(k'+m)$th superslot, the node employs a transfer rate of $\tilde{\mu}_{G-m}^{n}$ $(n \ge k'+m-1)$. That is, for $k \ge k+m$ and $m \ge 0$,

$$\begin{aligned} x_{G-m}^{k+1} &\lesssim x_{G-m}^{k} + D \cdot (M_{G-m-1} \cdot \tilde{\mu}_{G-m-1}^{k-1} - \tilde{\mu}_{G-m}^{k-1}) \\ &= \hat{x}_{G-m}^{k+1}. \end{aligned} \tag{23}$$

Equations (22) and (23) yield

$$x_{G-m}^{k+1} \le \hat{x}_{G-m}^{k+1}, \quad \text{for } k \ge 0, m \ge 0. \tag{24}$$

Consider the bottleneck first, i.e., $m = 0$. By directly applying the result of (24) to (8) and expanding $\hat{x}_{G}^{k+1}$ based on (10), we get

$$\begin{aligned} \mu_{G-1}^{k} &\ge \tilde{\mu}_{G-1}^{k-1} + (1+\gamma)\left[\frac{\tilde{\mu}_{G}^{k-1}}{M_{G-1}} - \mu_{G-1}^{k-1}\right] \\ &\quad + \gamma\left[\frac{B - x_{G}^{k}}{D \cdot M_{G-1}}\right] = \tilde{\mu}_{G-1}^{k}, \quad \text{for } k \ge 1. \end{aligned} \tag{25}$$

Now, consider one node upstream from the bottleneck, i.e., node $G-1$, at the beginning of superslot $(k'+1)$. Applying

(25), we obtain the following result:

$$\begin{aligned}\mu_{G-1}^{k} - \tilde{\mu}_{G-1}^{k-1} &\gtrsim \mu_{G-1}^{k-1} - \tilde{\mu}_{G-1}^{k-1} + \gamma\left[\frac{R_G}{M_{G-1}} - \tilde{\mu}_{G-1}^{k-1}\right] \\ &\gtrsim \gamma\left[\frac{R_G}{M_{G-1}} - \tilde{\mu}_{G-1}^{k-1}\right] = \hat{\alpha}_{G-1}^{k-1}, \quad \text{for } k \geq k'+1. \end{aligned} \tag{26}$$

Therefore, we get

$$\mu_{G-1}^{k} \gtrsim \tilde{\mu}_{G-1}^{k-1} + \hat{\alpha}_{G-1}^{k-1} = \hat{\mu}_{G-1}^{k}, \quad \text{for } k \geq k'+1. \tag{27}$$

As for $k < k'+1$ at node $G-1$, we have $\hat{\mu}_{G-1}^{k} = \tilde{\mu}_{G-1}^{k-1} = \tilde{\mu}_{G-1}^{k-1} = \mu_{G-1}^{k-1} = R_{G-1}$ as given in (11). Hence, (27) holds for $k \geq 1$. Subsequently, we directly apply this result and (24) to (8) and replace $\hat{x}_{G-1}^{k+1}$ by (10). That is,

$$\begin{aligned}\mu_{G-2}^{k} &\geq \frac{\hat{\mu}_{G-1}^{k}}{M_{G-2}} + \gamma\left[\frac{B - \hat{x}_{G-1}^{k+1}}{D \cdot M_{G-2}}\right] \\ &= \tilde{\mu}_{G-2}^{k-1} + (1+\gamma)\left[\frac{\tilde{\mu}_{G-1}^{k-1}}{M_{G-2}} - \tilde{\mu}_{G-2}^{k-1}\right] \\ &\quad + \gamma\left[\frac{B - x_{G-1}^{k}}{D \cdot M_{G-2}}\right] + \frac{\hat{\alpha}_{G-1}^{k-1}}{M_{G-2}} \\ &= \tilde{\mu}_{G-2}^{k}, \quad \text{for all } k \geq 1. \end{aligned} \tag{28}$$

By the same token, we have

$$\mu_{G-m}^{k} \gtrsim \tilde{\mu}_{G-m}^{k-1} + \hat{\alpha}_{G-m}^{k-1} = \hat{\mu}_{G-m}^{k}$$

and

$$\mu_{G-m-1}^{k} \gtrsim \tilde{\mu}_{G-m-1}^{k}, \quad \text{for } k \geq 1. \tag{29}$$

Combining (25), (28), and (29), the permitted rate inequality in (13) is directly derived. ■

REFERENCES

[1] T. M. Chen, S. S. Liu, and V. K. Samalam, "The available bit rate service for data in ATM networks," *IEEE Commun. Mag.*, vol. 34, pp. 56–71, May 1996.

[2] F. Bonomi and K. W. Fendick, "The rate-based flow control framework for the available bit rate ATM service," *IEEE Networking Mag.*, vol. 9, pp. 25–39, Mar./Apr. 1995.

[3] H. T. Kung, "The credit-based flow control framework for the available bit rate ATM service," *IEEE Networking Mag.*, vol. 9, pp. 40–55, Mar./Apr. 1995.

[4] K. Y. Siu and H. Y. Tzeng, "Intelligent congestion control for ABR service in ATM networks," *Comp. Commun. Rev.*, vol. 24, pp. 81–105, Oct. 1995

[5] J. Calvignac, J. Cherbonnier, I. Iliadis, J.-Y. Le Boudec, and D. Orsatti, "ATM best-effort service and its management in the LAN," in *Proc. EFOC&N'94*, June 1994, pp. 155–160.

[6] M. D. Prycker, *Asynchronous Transfer Mode Solution for Broadband ISDN*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[7] H. Saito, *Teletraffic Technologies in ATM Networks*. Norwood, MA: Artech House, 1994.

[8] H. Saito and K. Shiomoto, "Dynamic call admission control in ATM networks," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 982–989, Sept. 1991.

[9] T. Murase, H. Suzuki, S. Sato, and T. Takeuchi, "A call admission control algorithm for ATM network using a simple quality estimate," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1461–1470, Dec. 1991.

[10] T. H. Lee, K. C. Lai, and S.-T. Duann, "Real time call admission control for ATM networks with heterogeneous bursty traffic," in *Proc. IEEE ICC'94*, 1994, pp. 80–85.

[11] J.-L. C. Wu, L.-D. Chou, and S.-J. Tzeng, "Two-level dynamic step sizes of virtual paths on bandwidth allocation in ATM networks," *Proc. Inst. Elect. Eng.*, vol. 142, pp. 158–164, June 1995.

[12] Y. H. Kim and C. K. Un, "Analysis of bandwidth allocation strategies with access restrictions in broadband ISDN," *IEEE Trans. Commun.*, vol. 41, pp. 771–781, May 1993.

[13] P. Newman, "Backward explicit congestion notification for ATM local area networks," in *Proc. IEEE GLOBECOM'93*, 1993, pp. 719–723.

[14] K. K. Ramakrishnan and R. Jain, "A binary feedback scheme for congestion avoidance in computer networks," *ACM Trans. Comput. Syst.*, vol. 8, no. 2, pp. 303–313, 1990.

[15] N. Yin and M. G. Hluchyi, "On closed-loop rate control for ATM networks," in *Proc. INFOCOM'94*, 1994, pp. 99–108.

[16] K. W. Fendick and M. A. Rodrigues, "Asynptotic analysis of adaptive rate control for diverse soruces with delayed feedback," *IEEE Trans. Inform. Theory*, vol. 40, pp. 2008–2025, Nov. 1994.

[17] K. W. Fendick and M. A. Rodrigues, "An adaptive framework for dynamic access to bandwidth at high speed," in *Proc. ACM SIGCOMM'93*, Sept. 1994.

[18] F. Bonomi, D. Mitra, and J. B. Seery, "Adaptive algorithms for feedback-based flow control in high-speed, wide-area networks," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1267–1283, Sept. 1995.

[19] K. Kawahara, Y. Oie, M. Murata, and H. Miyahara, "Performance analysis of reactive congestion control for ATM networks," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 651–661, May 1995.

[20] L. Benmohamed and S. M. Meerkov, "Feedback control of congestion in packet switching networks: The case of a single congested node," *IEEE/ACM Trans. Networking*, vol. 1, pp. 693–708, Dec. 1993.

[21] I. Iliadis, "A new feedback congestion control policy for long propagation delay," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1284–1295, Sept. 1995.

[22] P. P. Mishra, H. Kanakia, and S. K. Tripathi, "On hop-by-hop rate-based congestion control," *IEEE/ACM Trans. Networking*, vol. 4, pp. 224–239, Apr. 1996.

[23] G. Ramamurthy and B. Sengupta, "A predictive hop-by-hop congestion control policy for high speed networks," in *Proc. IEEE INFOCOM'93*, 1993, pp. 1033–1041.

[24] K. Ko, P. P. Mishra, and S. K. Tripathi, "Interaction among virtual circuits using predictive congestion control," *Comput. Networks ISDN Syst.*, vol. 25, pp. 681–699, Jan. 1993.

[25] K. Ko, P. P. Mishra, and S. K. Tripathi, "Predictive congestion control in high speed wide area networks," in *Proc. 2nd IFIP WG6.1/W664 Int. Workshop Protocols for High Speed Networks*, Nov. 1990.

[26] M. D. Schroeder *et al.*, "Autonet: A high-speed, self-configuring local area network using point-to-point links," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 1318–1335, Oct. 1991.

[27] C. Ikeda and H. Suzuki, "Adaptive congestion control schemes for ATM LAN's," in *Proc. IEEE INFOCOM'94*, 1994, pp. 829–838.

[28] R. S. Pazhyannur and R. Agrawal, "Feedback-based flow control of B-ISDN/ATM networks," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1252–1266, Sept. 1995.

[29] J-C. Bolot and A. U. Shankar, "Analysis of a fluid approximation to flow control dynamics," in *Proc. IEEE INFOCOM'92*, 1992, vol. 3, pp. 2398–2407.

[30] Y. Gong and I. F. Akyildiz, "Dynamic traffic control using feedback and traffic prediction in ATM networks," in *Proc. IEEE INFOCOM'94*, 1994, pp. 91–98.

[31] O. Hashida and S. Shimogawa, "Switched batch Bernoulli process (SBBP) and the discrete-time SBBP/G/1 queue with application to statistical multiplexer," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 394–401, Apr. 1991.

Pau C. Ting was born in Tai-Tong, Taiwan, R.O.C., in 1967. He received the B.S. and M.S. degrees in computer science and information engineering in 1993 and 1994, respectively, from the National Chiao Tung University, Hsinchu, Taiwan, R.O.C., where he is currently working toward the Ph.D. degree.

His current research interests include congestion control and bandwidth management in high-speed networks and performance modeling and analysis.

Maria C. Yuang (M'91), for a photograph and biography, see this issue, p. 3.