Optimal Hop-by-Hop Flow Control in Computer Networks

ZVI ROSBERG AND INDER S. GOPAL

Abstract—**The problem of hop-by-hop flow control in a computer network is formulated as a Markov decision process with a cost function composed of the delay of the messages and the buffer constraints. The optimal control is shown to be a linear truncated function of the state and the explicit form is found when the arrival process of the messages is a Bernoulli process. For a renewal arrival process, the long-run average cost of any policy with a linear truncated structure is expressed by a set of linear equations.**

I. Introduction

FLOW control mechanisms [2], [6] are employed in computer networks in order to prevent the overloading of system resources and the consequent excessive time delays and buffer overflows. These mechanisms are classified in [2] into four levels: the session level, the network access level, the end-to-end level, and the hop-by-hop level. The purposes of each level are complementary and in most typical networks more then one level exists. The focus of this paper, however, is the hop-by-hop level, the purpose of which is to maintain a smooth flow of traffic between two adjacent nodes, and to avoid local buffer congestion. Note that, at this level, we do not have the ability of restricting traffic input into the network.

The specific hop-by-hop scheme studied in this paper is the following. Consider two adjacent nodes and, for the purposes of the explanation, assume that one node is the transmitting node and the other the receiving node. The transmitting node maintains at all times a counter which indicates the number of messages that the receiver is prepared to accept. We shall use the term "window" to refer to this counter. Whenever it transmits a message, it decrements the counter by one. When the counter reaches zero, it is not permitted to transmit any more messages. The receiver, periodically sends to the transmitter a message indicating the new number of messages that it is prepared to accept (a new "window"), and the transmitter upon receipt of this message updates its counter. If the receiver has indicated to the transmitter that it is prepared to accept a certain number of messages, it is required to set aside an amount of storage in its receive buffers sufficient to store that number of messages. Once the receiver has received the message and forwarded it to the next node on the message's path, it then releases the buffer storage used by that message.

The receiver keeps the history of the number of messages actually transmitted in each window. In making a decision on the current window size, the receiver uses this history as well as information on the number of messages currently waiting at the transmitter. The latter piece of information is sent by the transmitter to the receiver in the header of one of the messages that it transmits. However, because of transmission, propagation or queueing delays on the link, the information is typically outdated when it reaches the receiver and does not reflect the current number of messages present at the transmitter.

The scheme described above is analogous to the hop-by-hop mechanisms used in real networks such as GMDNET and TYMNET [2]. The mechanism may be applied once for all traffic that flows over a given link or may be applied individually for each virtual circuit that shares that link. We assume a single receive buffer pool that is completely shared by all links or virtual circuits that pass through the receiver.

There are clearly two conflicting forces that govern the receiver's actions. On the one hand, it is beneficial to commit to accept a large number of messages as this will reduce the possibility of the transmitter having to queue messages while awaiting permission to transmit from the receiver. On the other hand, if the receiver commits itself to receive too many messages it may find itself setting aside an excessive amount of storage, thereby preventing other links or virtual circuits from using that storage. The purpose of this paper is to understand this tradeoff and to find the best strategy that the receiver should adopt. We should note at this point that we have been able to derive the specific form of the optimal policy and not just the hysteretic properties as in many other studies. For a survey of previous and related studies, the reader is referred to [2] or [6].

The paper is structured in the following form. In Section II, we formalize the model and quantify the objective function. For the sake of analytic tractability, we are forced to place some simplifying restrictions on the model. In Section III, we formulate the optimal "window" selection problem as a discounted Markovian decision process and proceed to find the optimal discounted policy. In Section IV, we show how the results found for the discounted case apply to the long run average cost. In Section V, we show how the optimal long run policy can be computed for certain special cases. Finally, Section VI presents our conclusions.

II. The Model

Consider two adjacent nodes and, as before, assume one is a transmitter (XMTR) and the other a receiver (RCVR). Assume further that all messages in the system consist of packets of fixed length. In order to preserve analytical tractability, we shall assume that both XMTR and RCVR operate in synchronous, tightly regulated fashions. In particular, we shall assume that time is divided into slots, a slot being exactly long enough to transmit a single packet. A packet transmission may begin only on a slot boundary. We shall also assume that the window allocation process proceeds in "phases," a phase being a fixed predetermined number of slots (say, T slots). Prior to the start of every phase, RCVR informs XMTR of the window that it has chosen for that phase. This number represents the maximum number of packets that XMTR may transmit in that phase and also the number of buffer units of storage that RCVR must set aside for that phase. Clearly, the window size in any phase must be smaller than T.

The first packet that XMTR transmits in a phase includes the

Manuscript received December 13, 1984; revised September 15, 1985 and February 20, 1986. Paper recommended by Past Associate Editor, J. Walrand.

Z. Rosberg is with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598 on leave from the Department of Computer Science, Technion—Israel Institute of Technology, Haifa, Israel.

I. S. Gopal is with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598.

IEEE Log Number 8609615.

0018-9286/86/0900-0813$01.00

current number of packets that it has awaiting transmission. This information is used by the receiver in making the window size decision at the start of the next phase. Thus, the information that we use is always delayed by one phase. The receiver computes the window size and informs the transmitter of its decision through a control packet. This information must reach XMTR prior to the start of the next phase. Thus, T must be at least the sum of the propagation time from XMTR to RCVR, the processing time required for RCVR to understand the message and compute its decision, the propagation time from RCVR to XMTR, and the processing time at XMTR to understand the message.

The notion of a fixed duration phase is not entirely alien to real computer networks. For example, in TYMNET, each receiver sends a "backpressure vector" to each of its neighboring transmitters every half second. Based upon this "backpressure vector" the transmitter decides on the amount of traffic it can send to the receiver in the subsequent half second.

The choice of an appropriate value for T is determined by several factors. On one hand, since a decision of a window size is a commitment to reserve a number of buffers during the T slots, one would prefer a smaller T in order to have the possibility to change the decision when new information is revealed. However, as T gets smaller, decisions are made more often and more computational resources are required from the node and more control messages are sent through the channel. In any event, in order that the control mechanism be realizable, the T must satisfy the lower bound mentioned above. In this paper, we shall assume that some appropriate value of $T(>1)$ has already been chosen.

We assume that packet generation times forms a renewal process. That is, the number of packets generated during slot i, X_i, $i = 1, 2, \cdots$ are i.i.d r.v.'s, with a general distribution function F and a first moment λ.

As mentioned in the Introduction, we do not have the ability to discard packets at this flow control level. Thus, any new packet that arrives at XMTR is held in its buffer until it is transmitted irrespective of the flow control actions between XMTR and RCVR and the number of packets already existing in XMTR. If there was only a single link in the network, this would mean that the arrival process is independent of the system state and flow control actions. For a general network environment, there may be some dependence, but this is neglected in order to obtain a model which provides insight while preserving analytical tractability. Therefore, as an approximation, we assume that the arrival process is independent of the system state and flow control actions.

We shall also place a further restriction on the model that the messages that arrive in a particular phase may not be transmitted in that phase. This last restriction implies that packets may be forced to wait as much as one phase more than necessary. The average throughput and buffer requirement at the receiver will, however, remain unchanged. The restriction may be relaxed by changing the model so that the transmitter sends information to the receiver $T_1(\le T)$ slots before the start of the following phase and permits packets that arrive in the first $T - T_1$ slots of a phase to be transmitted in the same phase. The analytical tools presented in this paper can be used on this modified model.

For every k, $k = 1, 2, \cdots$, let Y_k be the number of packets generated at node XMTR during phase k. Clearly, $Y_k \sim F^{(T)}$, where $F^{(T)}$ is the T-convolution of F. Further, let N_k be the number of packets at node XMTR at the beginning of phase k, and w_k be the window size allocated during phase k. For definiteness, let $w_0 = N_0 = 0$. In order for RCVR to be able to inform XMTR of the window size it has selected for a phase, it is necessary for it to have completed any necessary computation prior to the start of the phase. Thus, w_k must be computed before the end of phase $k - 1$.

Define

$$S_k = (N_{k-1} - w_{k-1})^+, \qquad k \ge 1 \tag{2.1}$$

to be the state of the system at the start of phase k. S_k will be used to compute the value of w_k and can be trivially obtained by RCVR well before the end of phase $k - 1$. This is because XMTR sends N_{k-1} to RCVR in the first message of phase $k - 1$ and as T is sufficiently large, RCVR is guaranteed to receive this message before the start of phase k. In addition, w_{k-1} is computed by RCVR prior to the start of phase $k - 1$.

Since packets arriving during phase $k - 1$ are not allowed to use the window allocated during this phase we have

$$N_k = S_k + Y_{k-1}. \tag{2.2}$$

Let $p_i = \Pr(Y_k = i)$, $i = 0, 1, 2, \cdots$. The distribution of N_k given $S_k = s$ is given by

$$\Pr(N_k = n \mid S_k = s) = \begin{cases} \Pr(Y_{k-1} = n - s) = p_{n-s}, & n \ge s \\ 0, & \text{otherwise.} \end{cases} \tag{2.3}$$

To prevent an infinite expected delay on the link, we assume that

$$\lambda < 1. \tag{2.4}$$

It is clear that the lowest expected delay per packet is obtained when $w_k = T$ for every k. Under this policy we have a $GI/D/1$ system with batch arrivals and batch service time and it is known that (2.4) is a necessary and sufficient condition for the ergodicity of the system. (See, e.g., [1, pp. 369–414].) Thus, $\lambda \ge 1$ will provide an infinite expected delay per packet under any flow control policy.

We now define and motivate the objective function that we shall attempt to minimize in the next section. Recall that we have to balance the possible waste of buffers caused by over-allocation with the possible excessive time-delays caused by under-allocation. With this in mind, we define the cost incurred per phase, when there are n messages at node XMTR at the beginning of the phase and a window of size w is allocated to be $r(w, n)$, where

$$r(w, n) \triangleq b \cdot (w-n)^+ + h \cdot (n-w)^+ + c \cdot w \tag{2.5}$$

and b, h, and c are arbitrary nonnegative constants.

The first term adds to the cost function an additional cost of b units for every reserved buffer which is not utilized during the phase. This is due to the fact that the unused storage could be allocated to another link and is an attempt to capture the cost of over allocation. The second term adds to the cost function a cost of h units for every packet which is awaiting transmission at XMTR at the beginning of the phase but is not transmitted in that phase. This represents the waiting time of the packet and is an attempt to capture the cost of under allocation. Finally, the last term represents a fixed cost c, for every unit of buffer storage that RCVR reserves for XMTR in a phase. This last term attempts to capture the fact that allocating any window at all results in a decrease in the availability of the buffer pool. Clearly, the values of the constants can be adjusted to reflect the operating environment of the system. For example, if buffer storage is heavily constrained, it may be desirable to have a large b and c but a small h. Conversely, if time delays are paramount, h can be set large and b and c small. Note that the function $r(w, n)$ is reminiscent of the cost function in classical inventory problems. Nevertheless, we were unable to find any problem studied in the inventory literature which has the same structure as our problem.

At this point, the problem formulation may lead a reader to suspect (as we did!) that the optimal long-run average cost policy is a stationary one and has a linear truncated structure. In succeeding sections we formally prove that intuition. We use a direct approach in our proof as this enables us to derive explicitly the optimal policy and not just to establish hysteretic properties.

III. A Discounted Markov Decision Process

In this section we formulate a Markovian decision process and find the structure of the policy that minimizes the total discounted cost of the system.

A. Notation and Preliminary Results

Here, we derive the optimality equations for the Markovian decision process for both finite and infinite horizons.

Let S_k be the state for the decision problem at the beginning of phase k. The window size allocated by node RCVR to node XMTR at the beginning of phase k can be any integer w_k in $[0, T]$. The cost incurred per phase, when there are n messages at node XMTR at the beginning of the phase and a window of size w is allocated is as defined in (2.5).

Let $0 \leq \beta < 1$ be the interest rate used for discounting future cost, i.e., the present value of cost c incurred at phase k is $c \cdot \beta^k$. Also, let $V_k^\beta(s)$ be the minimum achievable total cost when the system is in state s, the discount factor is β, and there are k steps to go. Let

$$L(s, w) \triangleq E[r(w, s+Y)]$$

where Y is an integer valued random variable with distribution $F^{(T)}$. $L(s, w)$ is the expected cost per phase, when the state at the beginning of the phase is s, and a window of size w is allocated. From (2.2), (2.3), and (2.5),

$$\begin{aligned} L(s, w) &= b \sum_{n=0}^{w} (w-n)p_{n-s} + h \sum_{n=w+1}^{\infty} (n-w)p_{n-s} + c \cdot w \\ &= c \cdot w + E[d(w-s-Y)] \qquad (3.1) \end{aligned}$$

where $d(x) = b \cdot (x)^- + h \cdot (-x)^+$.

The optimality equations of dynamic programming yield

$$V_{k+1}^\beta(s) = \min_{0 \leq w \leq T} \{L(s, w) + \beta E[V_k^\beta([s-w+Y]^+)]\}.$$

From (3.1),

$$V_{k+1}^\beta(s) = \min_{0 \leq w \leq T} \left\{ L(s, w) + \beta V_k^\beta(0) \sum_{i=0}^{w-s} p_i + \beta \sum_{i=w-s+1}^{\infty} V_k^\beta(s-w+i)p_i \right\}. \qquad (3.2)$$

From the definition

$$V_k^\beta(s) = \sum_{i=1}^{k} \beta^{i-1} E_s[L(S_i, w_i^*)]$$

where the expectation is taken with respect to the probability measure of the stochastic process $(S_k, k \geq 1)$, given that $S_1 = s$, and the optimal window sizes, w_i^*, $1 \leq i \leq k$ are chosen.

Clearly, $L(s, w) \geq 0$ and, therefore, $V_k^\beta(s) \leq V_{k+1}^\beta(s)$. Thus, by the monotone convergence theorem, the following limit exists:

$$V_\infty^\beta(s) = \lim_{k \to \infty} V_k^\beta(s). \qquad (3.3)$$

In the next two lemmas, we show that optimality equations similar to (3.2) exist for $V_\infty^\beta(s)$ and that $V_\infty^\beta(s)$ is also the optimal cost in the infinite horizon case.

Let Π be any admissible control policy and $V_k^\beta(\Pi, s)$ the total discounted cost of using policy Π, and discount β, given that the system is at state s and there are k steps to go. We have

$$V_k^\beta(\Pi, s) = \sum_{i=1}^{k} \beta^{i-1} E_s[L(S_i, w_i)]$$

where w_i are the windows taken by policy Π.

Let

$$V^\beta(\Pi, s) = \lim_{k \to \infty} V_k^\beta(\Pi, s).$$

Lemma 3.1: For every $\beta < 1$, and every Π,

a) $V^\beta(\Pi, s) \leq \alpha + \gamma \cdot s$

b) $V_\infty^\beta(s) \leq \alpha + \gamma \cdot s$

where $\alpha, \gamma > 0$ are constants.

Proof: For every Π and horizon k

$$\begin{aligned} V_k^\beta(\Pi, s) &= \sum_{i=1}^{k} \beta^{i-1} E_s[L(S_i, w_i)] \\ &= \sum_{i=1}^{k} \beta^{i-1} E_s[b(w_i - N_i)^+ + h(N_i - w_i)^+ + cw_i] \\ &< \frac{T(c+b)}{1-\beta} + h \sum_{i=1}^{\infty} \beta^{i-1} E_s[N_i] \\ &\leq \frac{T(c+b)}{1-\beta} + h \sum_{i=1}^{\infty} \beta^{i-1}(s + i \cdot \lambda \cdot T) \\ &= \frac{T(c+b)}{1-\beta} + \frac{h \cdot \lambda \cdot T}{(1-\beta)^2} + \frac{h}{1-\beta} s. \end{aligned}$$

This completes the proof of a) and b). ■

Let $V^\beta(s) = \inf_\Pi V^\beta(\Pi, s)$, be the optimal cost when the system is at state s. In general, $V^\beta(s)$ is not necessarily equal to $V_\infty^\beta(s)$.

It is easy to verify that the conditions of Theorem 1 in [3] are satisfied (see also Section V there). Hence, $V^\beta(s)$ is the unique solution to the optimality equation and is achieved by a stationary policy

$$V^\beta(s) = \min_{0 \leq w \leq T} \{L(s, w) + \beta V^\beta(0) \sum_{i=0}^{w-s} p_i + \beta \sum_{i=w-s+1}^{\infty} V^\beta(s-w+i)p_i\}. \qquad (3.4)$$

From Lemma 3.1 b), (3.2), (3.4) and the uniqueness of $V^\beta(s)$,

$$V_\infty^\beta(s) = V^\beta(s). \qquad (3.5)$$

B. Properties of $L(s, w)$

In the next series of lemmas we will prove properties of $L(s, w)$ which will enable us to establish the convexity of $V^\beta(s)$. The following lemma is a straightforward result of (3.1).

Lemma 3.2: $L(s, s + w) = L(0, w) + c \cdot s$. ■

Lemma 3.3:

a) For every s, $L(s, w)$ is strictly convex in w.

b) $L(s, s)$, $L(s, 0)$ are linear (and consequently convex) in s.

c) $L(s, T)$ is linear (and consequently convex) is s for $s > T$.

Proof: a) $L(s, w)$ as a function of w is the classical cost function for the inventory problem. See, e.g., [5, p. 170] for its convexity. [It can also be easily verified from (3.1).]

b) From (3.1)

$$L(s, s) = h \cdot \lambda \cdot T + c \cdot s$$

and

$$L(s, 0) = h \cdot \lambda \cdot T + h \cdot s.$$

Linearity (and, hence, convexity) in s is apparent.

c) From (3.1)

$$L(s, T)=h \cdot \lambda \cdot T+(c-h)T+h \cdot s, \quad \text{for } s>T. \tag{3.6}$$

Linearity (and, hence, convexity) in s is again apparent. ■

Let $G_k(s, w)$ be the cost of a policy which allocates a window size of w at the first step when the state is s and then continues for the remaining $k-1$ steps according to the optimal policy. (We omit the factor β in order to simplify the notation.) We have

$$\begin{aligned} G_k(s, w) &= L(s, w)+\beta E[V_{k-1}^{\beta}([s-w+Y]^+)] \\ &= L(s, w)+\beta V_{k-1}^{\beta}(0)\sum_{i=0}^{w-s} p_i \\ &\quad +\beta \sum_{i=(w-s+1)^+}^{\infty} V_{k-1}^{\beta}(s-w+i)p_i. \end{aligned} \tag{3.7}$$

C. Convexity of $G_k(s, w)$ and $V_k^{\beta}(s)$

In this subsection, we prove that $G_k(s, w)$ is convex in w. This is a key result in the derivation of the optimal policy (in Section III-D). To show this we have to show that $V_k^{\beta}(s)$ is convex in s and make use of the properties of $L(s, w)$ established in subsection III-B. In addition, in our proof of convexity we use direct arguments and elaborate greatly on the structure of $G_k(s, w)$. This enables us to specify the optimal value of the window size $w_k^*(s)$ for which

$$\min_{0\le w\le T} G_k(s, w)=G_k(s, w_k^*(s))=V_k^{\beta}(s)$$

and to derive it in a closed form.

If we were interested only in showing the monotonicity property of $w_k^*(s)$ as a function of s, a more general approach such as proving submodularity (see, e.g., [7], [4]) of $G_k(s, w)$ would have sufficed.

From (3.7) and Lemma 3.2, we have the following lemma.

Lemma 3.4: $G_k(s, s+w) = G_k(0, w) + c \cdot s$. ■

Let w_1^* be the integer which satisfies

$$L(0, w_1^*)\le L(0, w) \quad \text{for every } 0\le w\le T. \tag{3.8}$$

To prove convexity of $G_k(s, w)$ and $V_\beta(s)$ we have first to show the convexity of the following function:

$$g(s)=\begin{cases} L(0, w_1^*)+c \cdot s & \text{for } 0\le s\le T-w_1^*, \\ L(0, T-s)+c \cdot s & \text{for } T-w_1^*<s\le T, \\ L(s, T) & \text{for } s>T. \end{cases}$$

Lemma 3.5: $g(s)$ is convex in s.

Proof: We have to show that for every s

$$g(s+2)-g(s+1)\ge g(s+1)-g(s). \tag{3.9}$$

From Lemma 3.3 a), $L(0, w)$ is convex in w. In the interval $[0, T]$, the first summand of $g(s)$ equals $L(0, w_1^*)$ until $s = T - w_1^*$ and then it varies from $L(0, w_1^*-1)$ to $L(0, 0)$. Therefore, $g(s)$ is convex in the interval $[0, T]$ as it is the sum of two convex functions.

From (3.6) $L(s, T)$ is also convex for $s > T$ and, thus, $g(s)$ is convex in the interval $[T, \infty]$.

Finally, it is left to show that (3.9) holds for $s = T-1, T$.

i) $s = T-1$. From the definition of $g(s)$, (3.1), and Lemma 3.2,

$$g(T+1)-g(T)=h$$

and

$$g(T)-g(T-1)=L(0, 0)-L(0, 1)+c=h-p_0(h+b).$$

Thus, (3.9) holds since $h, b > 0$.

ii) $s = T$. From the definition of $g(s)$,

$$g(T+2)-g(T+1)=h=g(T+1)-g(T).$$

Thus, the assertion is true. ■

In the next lemma we show that $V_k^{\beta}(s)$ is convex. The reader should note that general results from [7] can be used to prove convexity. However, we use a constructive proof which specifically derives the optimal window sizes. As mentioned previously, this approach is chosen because a more general approach (such as [7], [4]) can only provide monotonicity properties.

Lemma 3.6:

a) For every k and s, $G_k(s, w)$ is strictly convex in w.

b) For every k and $0 \le \beta \le 1$, $V_k^{\beta}(s)$ is convex in s.

Proof: The proof is by induction on k. First, we consider $k = 1$. Since $G_1(s, w) = L(s, w)$, from Lemma 3.3 a), part a) of this lemma is true for $k = 1$. To show part b), let w_1^* be as in (3.8). If $w_1^* = 0$, then from Lemma 3.2

$$L(s, s)\le L(s, s+w') \quad \text{for every } w'\ge 0. \tag{3.10}$$

Also, for $w \le s$,

$$L(s, w)=h \cdot \lambda \cdot T+h \cdot s+(c-h)w. \tag{3.11}$$

Thus, from (3.10) and (3.11)

$$V_1^{\beta}(s)=\min_{0\le w\le T} L(s, w)=\begin{cases} L(s, s) & \text{if } h\ge c, \\ L(s, 0) & \text{if } h<c. \end{cases}$$

From Lemma 3.3, in either of the above cases, $V_1^{\beta}(s)$ is convex.

Consider now the case where $w_1^* > 0$. From Lemmas 3.2 and 3.3 a)

$$L(s, s+w_1^*)\le L(s, s+w_1^*-1)$$

and

$$L(s, s+w_1^*)\le L(s, s+w_1^*+1).$$

Therefore, $L(s, s + w_1^*)$ is a global minimum and

$$V_1^{\beta}(s)=\min_{0\le w\le T} L(s, w) = \begin{cases} L(0, w_1^*)+c \cdot s & \text{for } s\le T-w_1^* \\ L(0, T-s)+c \cdot s & \text{for } T-w_1^*<s\le T \\ L(s, T) & \text{for } T<s. \end{cases}$$

From Lemma 3.5, it is clear that $V_1^{\beta}(s)$ is convex. Thus, we have proven assertion b) of the lemma for $k = 1$.

To prove the lemma for the general case, we will use an inductive argument. Suppose that the assertions a) and b) are true for k. We will now prove them for $k + 1$.

We shall first prove convexity of $G_{k+1}(s, w)$ and subsequently show that this implies strict convexity. From (3.7)

$$G_{k+1}(s, w)=L(s, w)+\sum_{i=0}^{\infty} \tilde{V}_k^{\beta}(i+s-w)p_i$$

where

$$\tilde{V}_k^{\beta}(j)=\begin{cases} V_k^{\beta}(0) & \text{for } j<0, \\ V_k^{\beta}(j) & \text{for } j\ge 0. \end{cases}$$

Since $\{p_i, i \ge 0\}$ is a probability measure and $V_k^{\beta}(j)$ is convex by induction, the second part of $G_{k+1}(s, w)$ is convex in w. Also, from Lemma 3.3 a), $L(s, w)$ is convex in w. Thus, $G_{k+1}(s, w)$ is convex in w as it is the sum of two convex functions, and assertion a) of the lemma is proven.

We will now prove that $V^{\beta}_{k+1}(s)$ is convex.

Let w^*_{k+1} be the integer which satisfies

$$G_{k+1}(0,\ w^*_{k+1}) \le G_{k+1}(0,\ w) \qquad \text{for every } 0 \le w \le T. \quad (3.12)$$

Proposition 1: If $w^*_{k+1} = 0$, then $V^{\beta}_{k+1}(s)$ is convex.

Proof: See the Appendix.

Proposition 2: If $w^*_{k+1} > 0$, then $V^{\beta}_{k+1}(s)$ is convex.

Proof: See the Appendix.

Propositions 1 and 2 prove the convexity of $V^{\beta}_k(s)$ and $G_k(s, w)$. Strict convexity of $G_k(s, w)$ follows from definition (3.7) and the fact that $L(s, w)$ is strictly convex. This completes the proof of Lemma 3.6. ■

D. The Optimal Policy

For every horizon $k \ge 0$ and discount factor $0 \le \beta \le 1$, let $w^*_k(\beta)$ be such that

$$\min_{0 \le w \le T} G_k(0,\ w) = G_k(0,\ w^*_k(\beta)). \quad (3.13)$$

Part a) of Lemma 3.6 guarantees that the minimum exists and is unique. Further, let $m^*_k(\beta)$ be the smallest integer m (if exists) such that

$$\beta \sum_{i=0}^{\infty} [\, V^{\beta}_{k-1}(m+1+i) - V^{\beta}_{k-1}(m+i)] p_i - (c-h) \ge 0.$$

If such an m does not exist, we define $m^*_k(\beta) = \infty$. Also, let $w^*_k(s, \beta)$ be the optimal window size when the system is at state s, the discount factor is β, and there are k steps to go.

From Lemma 3.6 and the proofs of Propositions 1 and 2 we have the following theorem which establishes the structure of the optimal policy for the finite horizon case.

Theorem 3.1: a) If $w^*_k(\beta) = 0$, then

$$w^*_k(s,\ \beta) = \begin{cases} 0, & \text{for } 0 \le s \le m^*_k(\beta), \\ s - m^*_k(\beta), & \text{for } m^*_k(\beta) \le s \le m^*_k(\beta) + T, \\ T, & \text{for } m^*_k(\beta) + T < s \end{cases}$$

and

$$V^{\beta}_k(s) = \begin{cases} G_k(s,\ 0), & \text{for } 0 \le s \le m^*_k(\beta), \\ G_k(s,\ s - m^*_k(\beta)), & \text{for } m^*_k(\beta) \le s \le m^*_k(\beta) + T, \\ G^k(s,\ T), & \text{for } m^*_k(\beta) + T < s. \end{cases}$$

b) If $w^*_k(\beta) > 0$, then

$$w^*_k(s,\ \beta) = \begin{cases} s + w^*_k(\beta), & \text{for } 0 \le s \le T - w^*_k(\beta), \\ T, & \text{for } T - w^*_k(\beta) < s \end{cases}$$

and

$$V^{\beta}_k(s) = \begin{cases} G_k(0,\ w^*_k(\beta)) + c \cdot s, & \text{for } 0 \le s \le T - w^*_k(\beta), \\ G_k(0,\ T - s) + c \cdot s, & \text{for } T - w^*_k(\beta) < s \le T, \\ G_k(s,\ T), & \text{for } T < s. \end{cases}$$

■

Next, we consider the infinite horizon for $\beta < 1$. Let

$$G(s,\ w) = \lim_{k \to \infty} G_k(s,\ w) = L(s,\ w) + \beta V^{\beta}(s) \sum_{i=0}^{w-s} p_i + \beta \sum_{i=w-s+1}^{\infty} V^{\beta}(s - w + i) p_i.$$

From Lemma 3.1 $G_k(s, w) < \infty$, and from (3.4) and (3.5) $G(s, w)$ is convex in w and $V^{\beta}(s)$ is convex in s as a limit of convex functions.

Let $w^*(\beta)$ be such that

$$\min_{0 \le w \le T} G(0,\ w) = G(0,\ w^*(\beta)).$$

Also, let $m^*(\beta)$ be the smallest integer m (if exists) such that

$$\beta \sum_{i=0}^{\infty} [\, V^{\beta}(m+1+i) - V^{\beta}(m+i) p_i - (c-h) \ge 0.$$

If such an m does not exist we define $m^*(\beta) = \infty$. Further, let $w^*(s, \beta)$ be the optimal window size when the system is at state s and the discount factor is β. (Recall from (3.4) that the optimal policy is stationary.)

As for the finite horizon we have the following theorem for the infinite horizon.

Theorem 3.2: a) If $w^*(\beta) = 0$, then

$$w^*(s,\ \beta) = \begin{cases} 0 & \text{for } 0 \le s \le m^*(\beta), \\ s - m^*(\beta) & \text{for } m^*(\beta) \le s \le m^*(\beta) + T, \\ T & \text{for } m^*(\beta) + T < s \end{cases}$$

and

$$V^{\beta}(s) = \begin{cases} G(s,\ 0) & \text{for } 0 \le s \le m^*(\beta), \\ G(s,\ s - m^*_k(\beta)) & \text{for } m^*(\beta) \le s \le m^*(\beta) + T, \\ G(s,\ T) & \text{for } m^*(\beta) + T < s. \end{cases}$$

b) If $w^*(\beta) > 0$, then

$$w^*(s,\ \beta) = \begin{cases} s + w^*(\beta) & \text{for } 0 \le s \le T - w^*(\beta), \\ T & \text{for } T - w^*(\beta) < s \end{cases}$$

and

$$v^{\beta}(s) = \begin{cases} G(0,\ w^*(\beta)) + c \cdot s & \text{for } 0 \le s \le T - w^*(\beta), \\ G(0,\ T - s) + c \cdot s & \text{for } T - w^*(\beta) < s \le T, \\ G(s,\ T) & \text{for } T < s. \end{cases}$$

■

We have established the structure of the optimal policy for the infinite horizon case. A graphical representation of the policy is given in Figs. 1 and 2. We have represented a discrete function by a continuous line and therefore for accuracy, one should only consider the integer points on the line. It is seen that the optimal window size is a linear function of the state, with a slope of unity, and truncated at 0 and at T. The value $m^*(\beta)$ represents the smallest state at which it is optimal to allocate a nonzero window. In Fig. 1, $w^*(\beta)$ is larger than zero and $m^*(\beta)$ is strictly zero. Thus, even if it is known that there are no waiting packets left in the previous phase it is optimal to allocate a nonzero window. In the next section, we shall show that for the nondiscounted average cost case, this is the situation that applies. The optimal window size in this case is simply the current system state s added to the optimal window size that would have been allocated if the state had been zero $w^*(\beta)$, truncated above by T. Thus, all that is necessary to compute the optimal control is to know the value of $w^*(\beta)$.

In Fig. 2, $w^*(\beta)$ is zero and $m^*(\beta)$ may be larger than zero. The structure of the optimal policy is similar, but for small values of the state variable s, it may be optimal to allocate a zero window.

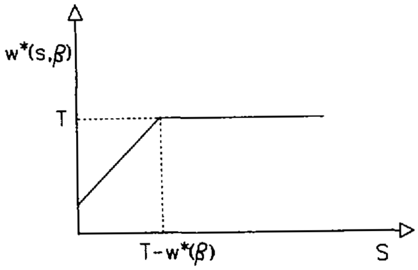


Fig. 1. The optimal policy for $w^*(\beta) > 0$.

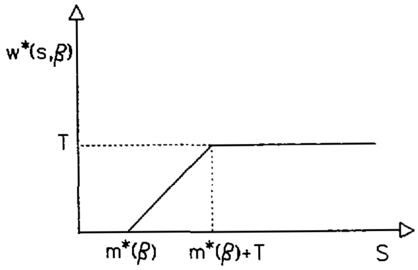


Fig. 2. The optimal policy for $w^*(\beta) = 0$.

IV. The Long Run Average Cost

In this section we extend the results of the previous section on the discounted cost criterion to the case which is usually of more interest in real computer networks, the long run average cost criterion. We show that the optimal flow control policy for this case is stationary and has a structure of the form given in Fig. 1.

Let P be the set of admissible policies for an infinite horizon. For $\Pi \in P$ and $0 \le \beta < 1$ we define $V_k^\beta(\Pi, s)$, $V^\beta(\Pi, s)$, and $V^\beta(s)$ as in Section III. Also, for $\beta = 1$, let

$$V_k(\Pi, s) = V_k^1(\Pi, s),$$

$$\bar{V}(\Pi, s) = \limsup_{k\to\infty} \frac{1}{k} V_k(\Pi, s)$$

and

$$\bar{V}(s) = \inf_P \bar{V}(\Pi, s).$$

We begin by showing the relation between the long run average cost function and some typical parameters of interest in computer networks such as average packet waiting time. For every Π under which the underlying Markov chain is ergodic, we have

$$\begin{aligned}\bar{V}(\Pi, s) &= \lim_{k\to\infty} \frac{1}{k} V_k(\Pi, s)\\ &= \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} E_s[b \cdot (w_i - N_i)^+ \\ &\quad + h \cdot (N_i - w_i)^+ + c \cdot w_i]\\ &= h \cdot \bar{N} + (c-h) \cdot \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} E_s[w_i]\\ &\quad + (b+h) \cdot \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} E_s[w_i - N_i]^+\end{aligned}$$

where w_i, $i \ge 1$ are the actions taken by policy Π, and $\bar{N} = \lim_{k\to\infty} 1/k \sum_{i=1}^{k} E_s[N_i]$, is the stationary expected number of packets in the system at the beginning of a phase. All expectations and probabilities are taken with respect to the probability measure of the Markov chain generated by the policy Π and the initial state s.

Note that if $h = 1/\lambda T$, we have from Little's theorem that the first term of $\bar{V}(\Pi, s)$ becomes the long run average waiting time of a packet in the system. The second term is the expected window size or, equivalently, the expected amount of buffer storage that the receiver must reserve for the transmitter. The last term represents the expected amount of unused storage.

Next, as in [3] we shall derive results for the long-run average cost case by letting $\beta \to 1$ in the discounted cost case.

First, we shall show that for β close to 1, only the cases presented in Fig. 1 or in Fig. 2 with $m^*(\beta) = 0$ are possible.

Lemma 4.1: If $\beta > 1 - h/c$ and $w^*(\beta) = 0$, then $m^*(\beta) < \infty$.

Proof: Suppose in contradiction that $m^*(\beta) = \infty$, then the optimal policy Π^*, always allocates a window size $w^*(s, \beta) \equiv 0$. From the definition

$$\begin{aligned}V^\beta(s) = V^\beta(\Pi^*, s) &= \sum_{k=1}^{\infty} \beta^{k-1} E_s[L(S_k, 0)]\\ &= \sum_{k=1}^{\infty} \beta^{k-1}[h \cdot s + h \cdot \lambda \cdot T + (k-1) h \cdot \lambda \cdot T]\\ &= \frac{h \cdot \lambda \cdot T}{(1-\beta)^2} + \frac{h}{1-\beta} \cdot s.\end{aligned} \tag{4.1}$$

Let Π_s^* be the policy which always allocates a window size of zero except at the first step when the state is s. In this first step it allocates a window size of s. Thus,

$$\begin{aligned}V^\beta(\Pi_s^*, s) &= L(s, s) + \beta \sum_{i=0}^{\infty} p_i \sum_{k=1}^{\infty} \beta^{k-1} E_i[L(S_k, 0)]\\ &= h \cdot \lambda \cdot T + c \cdot s + \beta \sum_{i=0}^{\infty} p_i \left(\frac{h \cdot i}{1-\beta} + \frac{h \cdot \lambda \cdot T}{(1-\beta)^2}\right)\\ &= h \cdot \lambda \cdot T + c \cdot s + \frac{\beta \cdot h \cdot \lambda \cdot T}{1-\beta} + \frac{\beta \cdot \lambda \cdot T}{(1-\beta)^2}\\ &= c \cdot s + \frac{h \cdot \lambda \cdot T}{(1-\beta)^2}.\end{aligned} \tag{4.2}$$

From (4.1) and (4.2) we have

$$V^\beta(\Pi_s^*, s) < V^\beta(\Pi^*, s)$$

which is a contradiction. Thus, $m^*(\beta) < \infty$. ∎

Lemma 4.2: If $\beta > 1 - h/c$ and $w^*(\beta) = 0$, then $m^*(\beta) = 0$.

Proof: Let τ be the first time when the total number of arrivals is at least $m^*(\beta)$. From Lemma 4.1 this τ is a finite stopping time.

From the definition of $V^\beta(s)$, its convexity and the structure of the optimal policy as given in Fig. 2,

$$\begin{aligned}&V^\beta(1) - V^\beta(0)\\ &= E_\tau\left\{\sum_{k=0}^{\tau-1} \beta^{k-1}[E_1[L(S_k, 0)] - E_0[L(S_k, 0)]]\right.\\ &\quad \left. + \beta^{\tau-1}[E_1 V^\beta(S_\tau) - E_0 V^\beta(S_\tau)]\right\}\\ &\ge E_\tau\left\{\sum_{k=0}^{\tau-1} \beta^{k-1}\left\{E_1[L(S_k, 0)] - E_0[L(S_k, 0)]]\right.\right.\\ &\quad \left.\left. + \beta^{\tau-1}[V^\beta(m^*(\beta)+1) - V^\beta(m^*(\beta))\right\}\right]\\ &= E_\tau\left\{h \sum_{k=0}^{\tau-1} \beta^{k-1} + \beta^{\tau-1}[L(m^*(\beta)+1, 1) - L(m^*(\beta), 0)]\right\}\\ &= E_\tau\left\{\frac{h(1-\beta^{\tau-1})}{1-\beta} + \beta_{\tau-1} c\right\} \ge \min\left\{\frac{h}{1-\beta}, c\right\} = c.\end{aligned} \tag{4.3}$$

Here, $E_\tau[f(\tau)]$ is the expectation of $f(\tau)$ with respect to the probability measure of τ.

From (4.3) and the convexity of $V^\beta(s)$,

$$\beta \sum_{i=0}^{\infty} [V^\beta(i+1) - V^\beta(i)]p_i \geq \beta \cdot c > \left(1 - \frac{h}{c}\right) c = c - h.$$

Hence, from the definition of $m^*(\beta)$ we have $m^*(\beta) = 0$. ■

From Lemma 4.2, the optimal flow control policy for $\beta > 1 - h/c$ has the form presented in Fig. 1 with $w^*(\beta) > 0$, or the form presented in Fig. 2 with $m^*(\beta) = 0$.

In the next theorem we shall show that the optimal policy for the long-run average cost has the same form.

Let $\beta_n > 1 - h/c$, $n \geq 1$, be a sequence of discount factors such $\beta_n \to 1$. For every β_n, $w^*(s, \beta_n)$ is the function defining the optimal flow control for the discount factor β_n. Since $w^*(s, \beta_n)$ is discrete and $w^*(s, \beta_n) = T$ for every $s \geq T$ and β_n, there exists a subsequence $\beta_{n'}$ such that

$$w^*(s, \beta_{n'}) \underset{n' \to \infty}{\longrightarrow} w^*(s).$$

Moreover, $w^*(s, \beta_{n'}) = w^*(s)$ for n' sufficiently large. Therefore, for $\beta_{n'}$ sufficiently close to 1, the optimal policies are all equal, and are as defined by the function $w^*(s)$, with the same form as given in Fig. 1, that is,

$$w^*(s, \beta) = \begin{cases} s + w^*(\beta) & \text{for } 0 \leq s \leq T - w^*(\beta), \\ T & \text{for } T - w^*(\beta) < s. \end{cases}$$

Theorem 4.1: The stationary flow control policy, which allocates a window $w^*(s)$, when the system is in state s, is the optimal policy for the long-run average cost. Moreover, the optimal cost is $\lim_{\beta \to 1} (1 - \beta) V^\beta(s) = \bar{V}(s) = \bar{V}(0)$.

Proof: All the assumptions and conditions of Theorem 2 in [3] are satisfied. Thus, the limit of the optimal discount policies is the optimal long-run average cost policy. Moreover, $\lim_{\beta \to 1} (1 - \beta) V^\beta(s) = \bar{V}(s) \equiv \bar{V}(0)$. ■

In the next section we shall evaluate the cost function of a policy that has the structure given in Fig. 1, and for a Bernoulli arrival process we shall explicitly specify the optimal policy.

V. The Evaluation of the Optimal Policy

In this section we consider only the long-run average cost case.

Let P^* be the set of all control policies defined by a function with a structure given in Fig. 1 (the optimal structure).

Clearly, a policy in P^* is defined by a single value w, $0 \leq w \leq T$, which is the allocated window when the system is in state 0. For every $0 \leq w \leq T$, let Π_w denote a policy in P^*.

Lemma 5.1: If $\Pi_w \in P^*$

$$V^\beta(\Pi_w, s) = V^\beta(\Pi_w, 0) + c \cdot s \qquad \text{for } 0 \leq s \leq T - w.$$

Proof: The assertion follows directly from the dynamic equation of the policy, Lemma 3.2, and (2.3). ■

From Theorem 4.1 the optimal policy, for the average cost criterion is the policy Π^*, which satisfies

$$(1-\beta) V^\beta(\Pi^*, 0) = \min_{\Pi_w \in P^*} (1-\beta) V^\beta(\Pi_w, 0) \tag{5.1}$$

for β close to 1. Clearly, the same policy also minimizes $V^\beta(\Pi, 0)$.

A. The Bernoulli Arrival Process

Suppose that X_i, $i \geq 1$, takes values 0, 1 with probability $1 - r$ and r, respectively, i.e., $F(0) = 1 - r$ and $F(1) = 1$.

Lemma 5.2: Under every policy $\Pi_w \in P^*$, if $S_1 \leq T - w$, then $S_k \leq T - w$ for every $k \geq 1$.

Proof: Suppose by induction that $S_k \leq T - w$. From (2.1) and (2.2)

$$S_{k+1} = (S_k + Y_{k-1} - w_k)^+ \tag{5.2}$$

where $Y_{k-1} \sim F^{(T)}$. Since X_i, $i \geq 1$ are Bernoulli variables, $Y_{k-1} \leq T$. From the definition of the policy Π_w, the inductive assumption, and (5.2), it follows that,

$$S_{k+1} = (Y_{k-1} - w)^+ \leq T - w.$$

■

From the dynamic equation of a policy Π_w and Lemma 5.1,

$$\begin{aligned} V^\beta(\Pi_w, 0) &= L(0, w) + \beta V^\beta(\Pi_w, 0) F^{(T)}(w) \\ &\quad + \beta \sum_{i=w+1}^{T} V^\beta(\Pi_w, i-w) p_i \\ &= L(0, w) + \beta V^\beta(\Pi_w, 0) + \beta \sum_{i=w+1}^{T} (i-w) \cdot c \cdot p_i, \end{aligned}$$

where $p_i = \binom{T}{i} r^i (1 - r)^{T-i}$ and $F^{(T)}(w) = \sum_{i=0}^{w} p_i$.

Hence,

$$(1-\beta) V^\beta(\Pi_w, 0) = c \cdot w + b \sum_{i=0}^{w} (w-i) p_i + (h + \beta \cdot c) \sum_{i=w+1}^{T} (i-w) p_i. \tag{5.3}$$

Let

$$w^* = \min \{ w | (b+c) F^{(T)}(w) > h(1 - F^T(w)) \}. \tag{5.4}$$

Roughly speaking, w^* is the value for which

$$(b+c) F^{(T)}(w) \sim h(1 - F^T(w)) \}.$$

This interprets the optimal policy in terms of the expected marginal costs.

Theorem 5.1: The policy Π_{w^*} is the optimal policy for the long-run average cost criterion and

$$\bar{V}(0) = c \cdot w^* + b \sum_{i=0}^{w^*} (w^* - i) p_i + (h + c) \sum_{i=w^*+1}^{T} (i - w^*) p_i.$$

Proof: From Theorem 4.1 it follows that the optimal policy is a policy, $\Pi_w \in P^*$, with w which minimizes the right-hand side of (5.3) when $\beta = 1$. It is easy to verify that w^* which is defined in (5.4) is the value of w that does it. The second part of the assertion follows from Theorem 4.1. ■

B. The General Renewal Arrival Process

Suppose that Y_k, $k \geq 1$, take integer values up to N (possibly infinite). Fix a β close to 1 and let

$$M(w, i) = V^\beta(\Pi_w, i).$$

Further, let

$$\gamma_n = L(n+1, T) - L(n, T), \qquad \text{for } n \geq 1,$$

$$\alpha_n = \sum_{i=T-w}^{\infty} [M(w, i+1) - M(w, i)] p_{i-n+w}, \qquad \text{for } n \geq 0.$$

$$a_n(\beta)=\sum_{i=T-w}^{\infty}\gamma_i p_{i-n+w}+\beta\cdot c\sum_{i=T-w}^{\infty}[F^{(T)}(2T-w-i-1)$$
$$-F^{(T)}(T-i-1)]p_{i-n+w}.$$

From the dynamic equations of Π_w, and Lemma 5.1, we have for $i \geq T-w$,

$$M(w,\ i+1)-M(w,\ i)=\gamma_i+\beta\sum_{j=T-i}^{\infty}[M(w,\ i+1-T+j)$$
$$-M(w,\ i-T+j)]p_j$$
$$=\gamma_i+\beta\cdot c[F^{(T)}(2T-w-i-1)$$
$$-F^{(T)}(T-i-1)]+\beta\alpha_{i-T+w}. \quad (5.5)$$

Multiplying both sides of (5.5) by p_{i-n+w} and summing up from $T-w$,

$$\alpha_n=a_n(\beta)+\beta\sum_{i=T-w}^{\infty}\alpha_{i-T+w}p_{i-n+w}$$
$$=a_n(\beta)+\beta\sum_{i=0}^{\infty}\alpha_i p_{i-n+w},\qquad \text{for } n\geq 0. \quad (5.6)$$

Let

$$A=\begin{bmatrix} p_T & p_{T-1} & p_{T+2} & \cdots & & & \\ p_{T-1} & p_T & p_{T+1} & \cdots & & & \\ p_{T-2} & p_{T-1} & p_T & \cdots & & & \\ \vdots & & & & & & \\ p_0 & p_1 & p_2 & \cdots & & & \\ 0 & p_0 & p_1 & p_2 & \cdots & & \\ 0 & 0 & p_0 & p_1 & \cdots & & \\ \vdots & & & & & & \\ 0 & 0 & \cdots & \cdots & p_i & p_{i+1} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$$a(\beta)=(a_0(\beta),\ a_1(\beta)a_2(\beta),\ \cdots)^T$$
$$\alpha=(\alpha_0,\ \alpha_1,\ \alpha_2,\ \cdots)^T$$

where V^T is the transpose of vector V. Then

$$\alpha=a(\beta)+\beta\cdot A\alpha. \quad (5.7)$$

Note that when $N<\infty$, (5.7) is a finite set of linear equations. Next, we shall represent $V^\beta(\Pi_w,\ 0)$ as a function of α. From (5.5) and Lemma 5.1, we have for $i\geq T-w$,

$$M(w,\ i+1)=M(w,\ 0)+c\cdot(T-w)+\beta\cdot c$$
$$\cdot\sum_{j=T-w}^{i}[F^{(T)}(2T-w-j-1)-F^{(T)}(T-j-1)]$$
$$+\beta\cdot\sum_{j=T-w}^{i}\alpha_{j-T+w}.$$

Hence,

$$V^\beta(\Pi_w,\ 0)=M(w,\ 0)=L(0,\ w)+\beta M(w,\ 0)F^{(T)}(w)$$
$$+\beta\sum_{i=w+1}^{\infty}M(w,\ i-w)p_i$$
$$=L(0,\ w)+\beta M(w,\ 0)F^{(T)}(w)$$
$$+\beta\sum_{i=w+1}^{T}(M(w,\ 0)+(i-w)c)p_i$$
$$+\beta\sum_{i=T+1}^{\infty}p_i\Big\{M(w,\ 0)+(T-w)c$$
$$+\beta\cdot c\sum_{j=T-w}^{i-w}(F^{(T)}(2T-w-j-1)$$
$$-F^{(T)}(T-j-1))+\beta\sum_{j+T-w}^{i-w}\alpha_{j-T+w}\Big\}$$
$$=L(0,\ w)+\beta V^\beta(\Pi_w,\ 0)$$
$$+\beta\cdot c\sum_{i=w+1}^{T}ip_i+\beta\cdot c(1-F^{(T)}(T))$$
$$-\beta\cdot c(1-F^{(T)}(w))$$
$$+\beta^2c\sum_{i=T+1}^{\infty}p_i\sum_{j=T-w}^{i-w}[F^{(T)}(2T-w-j+1)$$
$$-F^{(T)}(T-j-1)]$$
$$+\beta^2\sum_{i=T+1}^{\infty}p_i\sum_{j=T-w}^{i-w}\alpha_{j-T+w}.$$

Thus

$$(1-\beta)V^\beta(\Pi_w,\ 0)=L(0,\ w)+\eta(\beta,\ w)$$
$$+\beta^2\sum_{i=T+1}^{\infty}p_i\sum_{j=T-w}^{i-w}\alpha_{j-T+w} \quad (5.8)$$

where

$$\eta(\beta,\ w)=\beta\cdot c\sum_{i=w+1}^{T}i\cdot p_i+\beta\cdot c(1-F^{(T)}(T))$$
$$-\beta\cdot c(1-F^{(T)}(w))$$
$$+\beta^2\cdot c\sum_{i=T+1}^{\infty}p_i\sum_{j=T-w}^{i-w}[F^{(T)}(2T-w-j-1)$$
$$-F^{(T)}(T-j-1)].$$

Recall that

$$\lim_{\beta\to\infty}(1-\beta)V^\beta(\Pi_w,\ 0)=\bar{V}(\Pi_w,\ 0).$$

From Theorem 4.1, solving α in (5.7) for $\beta=1$, substituting the solution in (5.8) with $\beta=1$, we obtain the long-run expected cost for a policy $\Pi_w\in P^*$.

Note that all the expressions in (5.7) and (5.8) are well-defined and finite for $\beta=1$.

To find the optimal policy in the general case we have to evaluate numerically $\tilde{V}(\Pi_w, 0)$ for different w and find the best w by a binary search.

VI. Conclusion

In this paper we focused on a single hop and characterized the influence of the other hops that share the same storage resources by penalizing excessive storage usage in the cost function. This clearly simplifies the flow control problem and makes it mathematically tractable. It would obviously be preferable to explicitly include the other hops in the model and find an optimal policy that optimizes some global objective function while using information on the state of all hops to make its decisions. This general problem is extremly difficult and it is unlikely that clean solutions can be found. However, as understanding the single server queue enabled an understanding of networks of queues, it is hoped that our study of the single link problem will shed some light on the more general problem. Extending our model to include other links and queues is certainly a subject for further study.

Appendix

Proof of Proposition 1: From (3.12) and Lemma 3.4

$$G_{k+1}(s, s) \le G_{k+1}(s, s+w') \qquad \text{for } w' \ge 0. \tag{A.1}$$

Let

$$\alpha_m = \beta \sum_{i=0}^{\infty} [V_k^\beta(m+1+i) - V_k^\beta(m+i)]p_i$$

and let m^* be the smallest integer such that

$$\alpha_m \ge c-h.$$

From (3.1) and (3.7)

$$G_{k+1}(s, s-m-1) - G_{k+1}(s, s-m) = \alpha_m - (c-h), \qquad \text{for } m<s \tag{A.2}$$

and from the definition of m^*

$$\alpha_m < c-h, \qquad \text{for } m<m^*.$$

Therefore, as $G_{k+1}(s, w)$, is convex in w, (A.1) and (A.2)

$$\min_{w\ge 0} G_{k+1}(s, w) = \begin{cases} G_{k+1}(s, 0) & \text{for } s \le m^*, \\ G_{k+1}(s, s-m^*) & \text{for } s \ge m^*. \end{cases}$$

Thus,

$$V_{k+1}^\beta(s) = \min_{0\le w\le T} G_{k+1}(s, w)$$

$$= \begin{cases} G_{k+1}(s, 0) & \text{for } s\le m^* \\ G_{k+1}(s, s-m^*) & \text{for } m^* \le s \le m^*+T \\ G_{k+1}(s, T) & \text{for } m^*+T<s. \end{cases}$$

To prove convexity of $V_{k+1}^\beta(s)$ we have to show that the various functions from which it is composed are convex in their domain and that convexity is preserved on the boundaries. From (3.7)

$$G_{k+1}(s, 0) = L(s, 0) + \beta \sum_{i=0}^{\infty} V_k^\beta(s+i)p_i. \tag{A.3}$$

The second summand in (A.3) is convex in s by induction and the first summand is convex in s from Lemma 3.3 b). Thus, $V_{k+1}^\beta(s)$ is convex in $[0, m^*]$. Further,

$$G_{k+1}(s, s-m^*) = L(s, s-m^*) + \beta \sum_{i=0}^{\infty} V_k^\beta(m^*+i)p_i$$
$$= h\cdot\lambda\cdot T + (h-c)m^* + c\cdot s + \beta \sum_{i=0}^{\infty} V_k^\beta(m^*+i)p_i. \tag{A.4}$$

Thus, $V_{k+1}^\beta(s)$ is convex in $[m^*, m^*+T]$. Furthermore, for $s>T$

$$G_{k+1}(s, T) = L(s, T) + \beta \sum_{i=0}^{\infty} V_k^\beta(s-T+i)p_i, \tag{A.5}$$

which is convex in s by the inductive assumption and Lemma 3.3 c). Thus, $V_{k+1}^\beta(s)$ is convex in $[m^*+T, \infty]$.

To finish the proof of the convexity of $V_{k+1}^\beta(s)$, we still have to show that the following conditions hold on the boundaries (whenever meaningful):

i) $G_{k+1}(m^*+2, 2) - G_{k+1}(m^*+1, 1) \ge G_{k+1}(m^*+1, 1) - G_{k+1}(m^*, 0)$,

ii) $G_{k+1}(m^*+1, 1) - G_{k+1}(m^*, 0) \ge G_{k+1}(m^*, 0) - G_{k+1}(m^*-1, 0)$,

iii) $G_{k+1}(m^*+T+2, T) - G_{k+1}(m^*+T+1, T) \ge G_{k+1}(m^*+T+1, T) - G_{k+1}(m^*+T, T)$,

iv) $G_{k+1}(m^*+T+1, T) - G_{k+1}(m^*+T, T) \ge G_{k+1}(m^*+T, T) - G_{k+1}(m^*+T-1, T-1)$.

Case i): From (A.3) and (A.4)

$$G_{k+1}(m^*+2, 2) - G_{k+1}(m^*+1, 1) = G_{k+1}(m^*+1, 1) - G_{k+1}(m^*, 0) = c,$$

and Case i) holds.

Case ii): From (A.3) and (A.4)

$$G_{k+1}(m^*+1, 1) - G_{k+1}(m^*, 0) = c$$

and

$$G_{k+1}(m^*, 0) - G_{k+1}(m^*-1, 0) = h + \beta \sum_{i=0}^{\infty} [V_k^\beta(m^*+i) - V_k^\beta(m^*-1+i)]p_i.$$

Thus, Case ii) follows from the definition of m^*.

Case iii): From (3.6) and (A.5)

$$G_{k+1}(m^*+T+2, T) - G_{k+1}(m^*+T+1, T) = h + \beta \sum_{i=1}^{\infty} [V_k^\beta(m^*+2+i) - V_k^\beta(m^*+1+i)]p_i$$

and

$$G_{k+1}(m^*+T+1, T) - G_{k+1}(m^*+T, T) = h + \beta \sum_{i=0}^{\infty} [V_k^\beta(m^*+1+i) - V_k^\beta(m^*+i)]p_i.$$

Thus, Case iii) follows from the inductive assumption on the convexity of $V_k^\beta(s)$.

Case iv): From (A.4) and (A.5)

$$G_{k+1}(m^*+T+1, T) - G_{k+1}(m^*+T, T) = h + \beta \sum_{i=1}^{\infty} [V_k^\beta(m^*+1+i) - V_k^\beta(m^*+i)]p_i$$

and

$$G_{k+1}(m^*+T,\ T)-G_{k+1}(m^*+T-1,\ T-1)=c.$$

Thus, Case iv) follows from the definition of m^*.

The proof of Proposition 1 is complete.

Proof of Proposition 2: From (3.12) and Lemma 3.4

$$G_{k+1}(s,\ s+w_{k+1}^*)\le G_{k+1}(s,\ s+w_{k+1}^*+1)$$

and from the assumption in the proposition,

$$G_{k+1}(s,\ s+w_{k+1}^*)\le G_{k+1}(s,\ s+w_{k+1}^*-1).$$

Therefore, from the convexity of $G_{k+1}(s, w)$ in w, $G_{k+1}(s, s + w_{k+1}^*)$ is a global minimum. Moreover, from Lemma 3.4

$$V_{k+1}^{\beta}(s)=\min_{0\le w\le T} G_{k+1}(s,\ w)$$

$$=\begin{cases} G_{k+1}(0,\ w_{k+1}^*)+c\cdot s & \text{for } s\le T-w_{k+1}^*,\\ G_{k+1}(0,\ T-s)+c\cdot s & \text{for } T-w_{k+1}^*<s\le T,\\ G_{k+1}(s,\ T) & \text{for } s>T.\end{cases}$$

In the interval $[0, T]$, the first summand of $V_{k+1}^{\beta}(s)$ equals $G_{k+1}(0, w_{k+1}^*)$ up to $s = T - w_{k+1}^*$ and then it varies from $G_{k+1}(0, w_{k+1}^* - 1)$ to $G_{k+1}(0, 0)$. Since $G_{k+1}(s, w)$ is convex in w, $V_{k+1}^{\beta}(s)$ is the sum of two convex functions and is therefore convex in $[0, T]$.

From (3.6), (A.5) and the inductive assumption $G_{k+1}(s, T)$ is convex in s in (T, ∞) and so is $V_{k+1}^{\beta}(s)$.

To finish the proof we have to show the following boundary cases:

i) $G_{k+1}(T+2,\ T)-G_{k+1}(T+1,\ T)$

$\ge G_{k+1}(T+1,\ T)-G_{k+1}(0,\ 0)-c\cdot T,$

ii) $G_{k+1}(T+1,\ T)-G_{k+1}(0,\ 0)-c\cdot T$

$\ge G_{k+1}(0,\ 0)+c\cdot T-G_{k+1}(0,\ 1)-c\cdot(T-1).$

Case i): From (3.6) and (3.7)

$$G_{k+1}(T+2,\ T)-G_{k+1}(T+1,\ T)$$

$$=L(T+2,\ T)-L(T+1,\ T)+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(i+2)-V_k^{\beta}(i+1)]p_i$$

$$=h+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(i+2)-V_k^{\beta}(i+1)]p_i$$

and

$$G_{k+1}(T+1,\ T)-G_{k+1}(0,\ 0)-c\cdot T$$

$$=L(T+1,\ T)-L(0,\ 0)-c\cdot T+\beta\sum_{i=1}^{\infty}[V_k^{\beta}(i+1)-V_k^{\beta}(i)]p_i$$

$$=h+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(i+1)-V_k^{\beta}(i)]p_i.$$

Thus, Case i) follows from the inductive assumption on the convexity of $V_k^{\beta}(s)$.

Case ii): As in Case i),

$$G_{k+1}(T+1,\ T)-G_{k+1}(0,\ 0)-c\cdot T$$

$$=h+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(1+i)-V_k^{\beta}(i)]p_i$$

and

$$G_{k+1}(0,\ 0)+c\cdot T-G_{k+1}(0,\ 1)-c\cdot(T-1)$$

$$=c+L(0,\ 0)-L(0,\ 1)+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(i)-V_k^{\beta}(i-1)]p_i$$

$$-\beta V_k^{\beta}(0)p_0$$

$$=h+\beta\sum_{i=0}^{\infty}[V_k^{\beta}(i)-V_k^{\beta}(i-1)]p_i-\beta V_k^{\beta}(0)p_0-(h+b)p_0.$$

We know that h, b, $c > 0$, and that $V_k^{\beta}(s) > 0$ and by the inductive assumption $V_k^{\beta}(s)$ is also convex. Therefore, Case ii) is true.

The proof of Proposition 2 is complete.

Acknowledgment

We are indebted to B. Kadaba for presenting the problem to us and for many helpful discussions.

References

[1] J. W. Cohen, *The Single Server Queue.* Amsterdam, The Netherlands: North-Holland, 1969.

[2] M. Gerla and L. Kleinrock, "Flow control protocols," in *Computer Network Architectures and Protocols,* P. E. Green, Jr., Ed. New York: Plenum, 1982, pp. 361–412.

[3] S. A. Lippman, "On dynamic programming with unbounded rewards," *Management Sci.,* vol. 21, pp. 1225–1233, 1975.

[4] F. V. Lu and R. F. Serfozo, "M/M/1 queueing decision processes with monotone hysteretic optimal policies," *Operations Res.,* vol. 32, no. 5, pp. 1116–1132, Sept.-Oct. 1984.

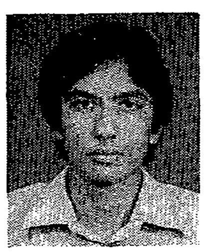
[5] S. M. Ross, *Applied Probability Models with Optimization Applications.* San Francisco, CA: Holden-Day, 1970.

[6] M. Schwartz, *Computer Communication Networks, Design and Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[7] D. M. Topkis, "Minimizing a submodular function on a lattice," *Operations Res.,* vol. 26, no. 2, pp. 305-321, Mar.-Apr. 1978.

Zvi Rosberg received the B.Sc., M.A., and Ph.D. degrees from the Hebrew University, Jerusalem, in 1971, 1974, and 1978, respectively.

From 1972 to 1978 he was a Senior Systems Analyst in the General Computer Bureau of the Israeli Government. From 1978 to 1979 he received a research fellowship at the Center of Operation Research and Econometrics (CORE), University of Louvain, Louvain, Belgium. From 1979 to 1980 he was a Visiting Assistant Professor at the University of Illinois, Urbana. Since 1980 he has been with the Department of Computer Science, Technion—Israel Institute of Technology, Haifa, Israel. At present he is on leave at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY.



Inder S. Gopal was born in Madras, India, in 1957. He received the B.A. degree in engineering science from Oxford University, Oxford, England, in 1977 and the M.S. and Ph.D. degrees in electrical engineering from Columbia University, New York, NY, in 1978 and 1982, respectively.

Since 1982, he has been a Research Staff Member in the Computer Sciences Department at the IBM T. J. Watson Research Center, Yorktown Heights, NY. At present, he is Manager of the Integrated Voice/Data Networking group. His current research interests are in communication networks and in distributed algorithms.