

Goodput Analysis of a Fluid Queue with Selective Discarding and a Responsive Bursty Source

Parijat Dube* and Eitan Altman†

Abstract— In this paper we analyse a feedback system consisting of a finite buffer fluid queue and a responsive source. The source alternates between silence periods and active periods. At random epochs of times the source becomes ready to send a burst of fluid. The length of the bursts (length of the active periods) are independent and identically distributed with some general distribution. The queue employs a threshold discarding policy in the sense that only those bursts at whose commencement epoch (the instant at which the source is ready to send), the workload (i.e., the amount of fluid in the buffer) is less than some preset threshold are accepted. If the burst is rejected then the source backs off from sending. Using techniques from Volterra Integral Equations we obtain an explicit characterization of the queue length distribution at commencement epochs of bursts from which we obtain an explicit characterization of the goodput ratio associated with such a feedback system. For the particular case of exponential distribution of on-periods we are able to obtain explicit closed form expression for the goodput ratio. Our explicit characterizations shall be quite helpful in studying the sensitivity of goodput ratio to different parameters, in selecting optimal discarding threshold etc. which will further provide useful “engineering” guidelines for better network designing.

Index Terms— on-off fluid, finite buffer, level-crossing arguments, Volterra integral equations, workload distribution, goodput ratio, numerical examples.

I. INTRODUCTION

In this paper we analyse a feedback system consisting of a finite buffer fluid queue and a responsive source. The source alternates between silence periods and active periods. At random epochs of times the source becomes ready to send a burst of fluid. The length of the bursts (length of the active periods) are independent and identically distributed with some general distribution. The queue employs a threshold discarding policy in the sense that only those bursts at whose commencement epoch¹, the workload (i.e., the amount of fluid in the buffer) is less than some preset threshold are accepted. If the burst is rejected then the source backs off instantaneously from sending and goes into a silence (back-off) period. The input to the queue can be modeled as an on-off source which generalized distribution of the on time. The off period distribution is characterized as follows. After an exponentially distributed silence time, a new batch arrives; if it is accepted (i.e. the amount of fluid is below the threshold) then the off

period ends. Otherwise a new exponentially distributed silence time is taken and so on. The off period is then the sum of the consecutive silence periods.

One particular motivation for our model is the performance analysis of selective message discarding policies with responsive sources. Selective message discarding policies have been proposed [2] and implemented in routers (e.g., in Cisco BP 8600 series, Fujitsu ATM switches) to prevent network congestion. This is particularly the case with the router supporting UBR (Unspecified bit rate) service class of ATM where message (i.e., a frame) discarding is employed to achieve the twin goals of reduced network congestion and increased goodput [7]. In ATM context, message discarding is based on the idea that loss of a single packet results in the corruption of the entire message (to which it belongs) and hence it is advantageous to discard the entire remaining message.

The explicit closed form expressions which we seek in this paper helps in studying the sensitivity of the goodput² to different source and network parameters, in buffer dimensioning, in selecting optimal discarding threshold (say for the optimality criterion being the maximization of goodput) etc. The solutions can then serve in providing useful “engineering” guidelines for designing networks.

Two discarding mechanisms have frequently been used: the *partial discarding*, in which packets that belongs to an already corrupted message is discarded, and the *early discarding*, in which in addition to partial discarding, an admission control is applied to reject an entire message if upon arrival of its first packet, the queue exceeds some threshold value K (*threshold discarding*) [7]. We have focused both on the discrete as well as on the fluid analysis (back-to-back message arrival with exponentially distributed message lengths) of the first mechanism in [12] and of the second in [10], [9] with the goal of obtaining explicit closed form expressions of the performance metrics in particular the goodput ratio. In all the models analysed in [12], [10], [9] the system was open loop in the sense that the source was nonresponsive. The discarding policies worked independently without any cooperation from the source (the source continues sending even if its data is being rejected by the network node). Also in the fluid models in [12], [9] we have back-to-back messages with exponential distribution of length and thus the fluid arrival rate was deterministic. In this scenario goodput ratio for packet (fluid) models associated with a message discarding policy is defined as *the ratio of good packets (fluid corresponding to good messages*³, i.e.,

*IBM T. J. Watson Research Center, Yorktown Heights, NY, 10598, USA. E-mail: pdube@us.ibm.com.

†INRIA, B.P. 93, 06902, Sophia Antipolis Cedex, France, and C.E.S.I.M.O., Universidad de Los Andes, Facultad de Ingenieria, Merida, Venezuela. E-mail:Eitan.Altman@sophia.inria.fr.

¹By commencement epoch of a burst we mean the time instant when the modulating process goes from off to on state.

²In the paper we shall also be using the term goodput and goodput ratio interchangeably

³In the paper we shall be using the term message and burst interchangeably

messages which do not suffer any packet (fluid) loss due to threshold discarding or buffer overflow) exiting the node to the total packets (fluid) arriving at its input.

In view of the rich recent research on improving the performance by some sort of congestion notification [4], [14] to sources, it becomes interesting to study the performance of message discarding policies with responsive sources. Thus we need to model the feedback system consisting of a network node with selective burst (the burst can be seen as a message) discarding and a source which responds to congestion signals (in our case the congestion signal being positive if the queue length at the network node is higher than some preset threshold and is negative if the queue length is less than some preset threshold). This will help us in understanding the improvement in performance achievable with combining selective burst discarding with congestion feedback to sources and responsiveness on the part of sources to backoff from sending.

A first step in this direction was made in [11] where we analysed a responsive bursty source and (i) a variant of the early discarding in which there is only threshold discarding (ii) a combination of both threshold discarding and partial discarding with an on-off source (an on period corresponds to a message). In the model in [11] the burst lengths were exponentially distributed and the analysis was based on working with Poisson Counter Driven Stochastic Differential Equations [13] describing the workload dynamics. We obtained closed form expressions for the distribution of the workload process by first finding the Laplace-Stieltjes Transform (LST) of the stationary workload process and then inverting it to get the density of the workload process.

In this paper we go way beyond the results of our previous work in [11] not only in extending the model to generally distributed burst sizes, but also in obtaining closed form expressions for the goodput ratio. We note that the approach used in [11] turned out to be quite useful for computing the stationary workload distribution (for the case of exponentially distributed on periods) but it did not provide an expression for the goodput ratio. In this paper we propose two notions of the goodput ratio: one from the source point of view and one from the overall system point of view. Using a completely different analytical approach that employs *Volterra Integral Equations* methodology we are able to obtain explicit characterization of the goodput ratios for general distribution of burst sizes. In particular, we express the distribution of a Markov chain (associated with workload) embedded at arrival epochs of messages by Volterra Integral Equations from which we characterize the goodput ratio. By the arrival epoch of a message we mean the epoch at which the source changes/wants to change from silence period to active period (the source gets ready to send a burst, whether it is accepted or not).

As a special case of our model (with no discarding) we can obtain the probability density function of the workload process at arrival epochs of bursts/messages in an finite buffer fluid queue with on-off fluid source with general distribution of on periods and exponential distribution of off periods. Modeling of PCM (pulse code modulation) coded voice sources as a two state Markov Modulated Poisson Process (MMPP) is a

standard acceptable practice and also there have been works on modeling the aggregate arrival due to the superposition of ATM traffic as a two state MMPP (see [3],[5]). In such scenarios our model will be handy in fluid analysis of corresponding queues.

The paper is organised as follows. In Sec. II we formally define our model and define two notions of goodput ratios and obtain expressions for them. The model is analysed in Sec. III and a characterization of the workload process at arrival epochs of bursts is obtained. Explicit calculations for the distribution of the workload process and closed form expressions for the goodput ratios are also provided for exponential distribution of active periods. Numerical examples are provided in Sec. IV. In Sec. V we analyse a policy which has partial discarding of bursts in addition to threshold burst discarding (like in our model)⁴. We finally conclude in Sec. VI.

II. MODEL: FORMAL DEFINITION

The fluid arrival rate is h when the source (actually) sends a burst and 0 when the source is silent (which models either the thinking time of the source or forced backoff by the source due to positive congestion feedback); the server has a constant capacity c . Let the buffer size be B (maximum amount of fluid) and the threshold be K , $K < B$. Let the message lengths have distribution $F_1(\cdot)$ and the silence-period (off-period) between messages have distribution as characterized in Sec. I. The discarding policy is such that if at the commencement epoch of a message the workload process $v(t)$ is less than K , the message is admitted, otherwise not. We assume that $c < h$ ⁵. Fig 1 explains the model (the source behavior and the workload process in the queue). We shall first define and characterize two notions of the goodput ratio.

A. Goodput Ratio: From Source and System Point of View

We call a message as *good* if during its period of arrival to the queue the process $v(t)$ always remain below B . The fluid associated with good message is called good fluid. We next define:

- *Goodput ratio perceived by the source \mathcal{G}_S* : The ratio of “good fluid” to the “total fluid” that could have been sent by the source if there was no backoff by the source (due to congestion feedback) and no fluid loss at the queue. Thus total fluid is the fluid sent by the source in a system with infinite buffer and no burst discarding.
- *Goodput ratio of the feedback system \mathcal{G}_{FS}* : The ratio of “good fluid” to the “actual total fluid” sent by the source. Thus actual total fluid is the fluid actually sent by the source (and hence it does not include the fluid corresponding to bursts that were not sent due to source backoff).

Thus \mathcal{G}_S can be viewed as the Quality of Service (QoS) perceived by the source and \mathcal{G}_{FS} reflects the efficiency of the overall feedback system. We next evaluate expressions for \mathcal{G}_S and \mathcal{G}_{FS} .

⁴Recall that under partial discarding once the buffer starts overflowing the source backs off instantaneously and goes into silence period.

⁵For the case $c \geq h$ the workload will always be 0 w.p. 1

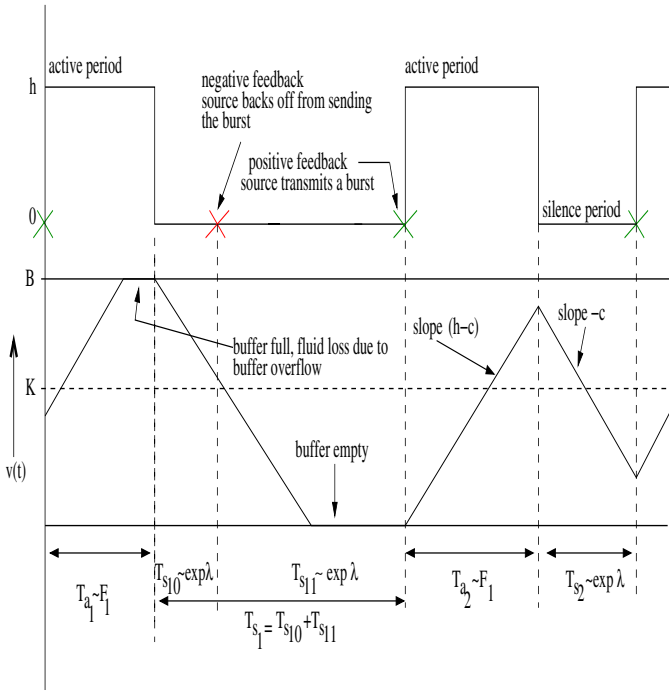


Fig. 1. The dynamics of the arrival process and the workload in our model. T_{a_n} and T_{s_n} denote the n th active and silence periods respectively.

Let V_f be a random variable representing the success of a message, $V_f = 1$ for a good message, and $V_f = 0$ for a message which has lost some fluid. Let W be the random variable denoting the length of a message. We next define the sub-distribution function $H(w, 1)$ as the probability that a message is of length $\leq w$ and is good, i.e.,

$$H(w, 1) = P(W \leq w, V_f = 1).$$

Then we can write the goodput ratio \mathcal{G}_S as

$$\mathcal{G}_S = \frac{\int_0^\infty w dH(w, 1)}{\int_0^\infty w dF_1(w)}. \quad (1)$$

And since a source backoff if the queue length is greater than K we have:

$$\mathcal{G}_{FS} = \frac{\int_0^\infty w dH(w, 1)}{F(K) \int_0^\infty w dF_1(w)}. \quad (2)$$

where $F(\cdot)$ is the distribution of the queue length at arrival epochs of messages. Observe that:

$$H(w, 1) = \quad (3)$$

$$\begin{aligned} & P(V_f = 1 | W \leq w) P(W \leq w) \\ &= \int_0^B P(V_f = 1 | W \leq w, \mathcal{V} = v) dF(v) \\ & \int_0^w f_1(u) du \end{aligned} \quad (4)$$

where f_1 is the message length density and \mathcal{V} is the queue length at the arrival epoch of a message with ρ being the queue length density at the arrival epoch of the message. Let M be the mean message length.

Lemma 1: For $K > 0$, the goodput ratios can be expressed as:

$$\begin{aligned} \mathcal{G}_S &= M^{-1} \int_0^{\frac{B}{h-c}} w dH(w, 1) \\ \mathcal{G}_{FS} &= \frac{\mathcal{G}_S}{F(K)} \end{aligned}$$

and for $K = 0$, $\mathcal{G}_S = \mathcal{G}_{FS} = 0$.

Proof: For $K = 0$, no burst will be accepted at any time and $\rho(0) = 1$. Thus $H(w, 1) = 0$ for all w and hence $\mathcal{G}_S = \mathcal{G}_{FS} = 0$. We will next look at $K > 0$ case. Observe that in (3), the integral can be taken up-to K and not B as,

$$P(V_f = 1 | W \leq w, \mathcal{V} = v) = 0 \quad \text{for } v \geq K$$

Observe that, for $v < K$,

$$P(V_f = 1 | W \leq w, \mathcal{V} = v) = \begin{cases} 1 & \text{if } w \in [0, \frac{B-v}{h-c}] \\ P(W < \frac{B-v}{h-c} | W < w) & \text{if } w > \frac{B-v}{h-c}. \end{cases}$$

Or in other words, for $w \in [0, \frac{B-K}{h-c}]$, $P(V_f = 1 | W \leq w, \mathcal{V} = v)$ is

$$= \begin{cases} 1 & \text{if } v < K \\ 0 & \text{if } v \geq K \end{cases}$$

for $w \in (\frac{B-K}{h-c}, \frac{B}{h-c}]$, $P(V_f = 1 | W \leq w, \mathcal{V} = v)$ is

$$= \begin{cases} 1 & \text{if } v \in [0, B - w(h-c)] \\ \frac{P(W < \frac{B-v}{h-c})}{P(W < w)} & \text{if } v \in (B - w(h-c), K) \\ 0 & \text{if } v \geq K. \end{cases}$$

And for $w > \frac{B}{h-c}$, $P(V_f = 1 | W \leq w, \mathcal{V} = v)$ is

$$= \begin{cases} \frac{P(W < \frac{B-v}{h-c})}{P(W < w)} & \text{if } v \in [0, K) \\ 0 & \text{if } v \geq K \end{cases}$$

Thus we write, for $w \in [0, \frac{B-K}{h-c}]$,

$$H(w, 1) = F_1(w) \int_0^K \rho(v) dv, \quad (5)$$

for $w \in (\frac{B-K}{h-c}, \frac{B}{h-c}]$, $H(w, 1) = H_a(w) + H_b(w)$, where,

$$H_a(w) = F_1(w) \int_0^{B-w(h-c)} \rho(v) dv \quad (6)$$

And,

$$H_b(w) = \int_{B-w(h-c)+}^K F_1\left(\frac{B-v}{h-c}\right) \rho(v) dv \quad (7)$$

and for $w > \frac{B}{h-c}$

$$H(w) = \int_0^K F_1\left(\frac{B-v}{h-c}\right) \rho(v) dv \quad (8)$$

Observe that for $w > \frac{B}{h-c}$, from (8)

$$dH(w, 1) = 0$$

Thus \mathcal{G}_S from (1) can be written as in Lemma 1 and also \mathcal{G}_{FS} from (2) .

Observe that from Lemma 1 the expression for \mathcal{G}_S and \mathcal{G}_{FS} requires an expression for $H(w, 1)$ and $F(K)$. From (3) an expression for $H(w, 1)$ requires the distribution F of \mathcal{V} , the workload at arrival epochs of messages. In the next section we shall provide an approach for finding the distribution function F .

III. MODEL ANALYSIS

Let \mathcal{V}_n be the random variable denoting the workload at the instant when the source is ready to send the burst at the n th time. Let T_a be the random variable for the length of an active period. Observe that $F(\cdot)$ is the stationary distribution of \mathcal{V}_n . We shall next find the distribution function F . We first establish an interesting equivalence between the stationary distribution F of \mathcal{V}_n and the stationary distribution of the workload at arrival epochs, V_a , in a “pseudo” $M/G/1$ type queue with constrained workload (the maximum amount of unfinished work is bounded by B) defined below.

Consider a (pseudo) $M/G/1$ queue with:

- The arrival process of packets is Poisson with parameter λ (same as the parameter for the exponential distribution of a backoff period. A silence period can have one or more backoff periods.)
- The packet sizes (X) have general distribution G which is related to distribution F_1 of the active periods in the (original) fluid model as:

$$G(x) = P(X \leq x) = P(T_a \leq \frac{x}{h-c}) = F_1\left(\frac{x}{h-c}\right)$$

Recall that F_1 is distribution function for the active periods in the on-off fluid arrival process whereas G is the distribution of the jumps (packet sizes) for the $M/G/1$ queue.

- Only those packets which sees the workload in the queue less or equal to K are accepted. Also if the packet size is greater than the vacant place in the queue (as the maximum workload is bounded by B) the excess data of the packet is lost. Fig. (2) makes this view point clear.

We thus have the following Lemma.

Lemma 2: The stationary distributions of \mathcal{V}_n is same as the distribution of V_a .

By Poisson Arrivals See Time Averages (PASTA) we have that the stationary distribution of V_a is same as the (time) stationary distribution of the workload process in the pseudo $M/G/1$ system. Thus we have the following Proposition

Proposition 1: The stationary distribution of \mathcal{V}_n is same as the (time) stationary distribution of the workload process V in the pseudo $M/G/1$ queue described above.

A. Characterisation of the Distribution Function of V

From Proposition 1 the stationary distribution of \mathcal{V}_n is same as the distribution of V in the pseudo $M/G/1$ queue. We next obtain the distribution of V . Since the distribution is same we

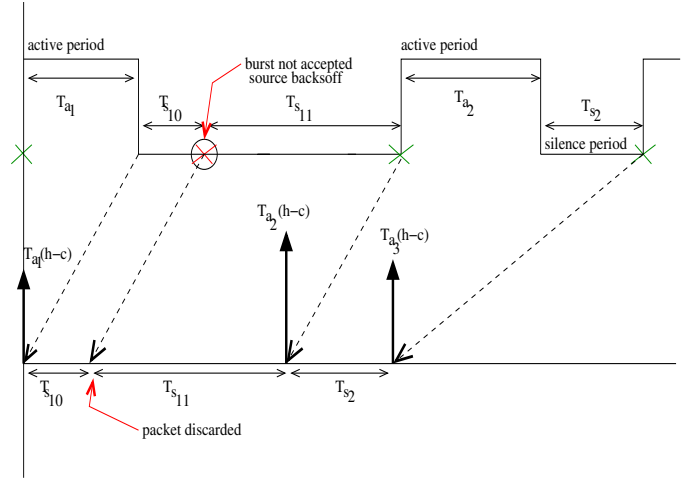


Fig. 2. Our view point of looking at the on-off fluid queue as a $M/G/1$ queue. T_{s10}, T_{s11} are exponential backoff periods, $T_{s1} = T_{s10} + T_{s11}$ is a silence period, T_{s2} is a silence period with a single backoff and hence its length is exponentially distributed. T_{a1}, T_{a2}, T_{a3} are active periods with distribution $F_1(\cdot)$.

use the same notation F for the distribution function of V and stationary distribution of \mathcal{V}_n .

We derive differential equations for $F(v)$ for $v < K$ and $v \geq K$. We use the classical up and down crossing argument. Fix a time interval of length $\Delta > 0$ (assuming the process is in equilibrium). Whenever the workload increases from a level less than or equal to v to a level more than v we say that an up-crossing of the level v has occurred. Similarly, if the workload size decreases from more than v to less than or equal to v we say that a down crossing of the level v has occurred. Thus for $v \in [0, K)$ we have

$$\lambda \Delta \int_0^\infty P\{v - (h-c)g < V \leq v\} dF_1(g) = (1 - \lambda \Delta) P(v < V \leq v + \Delta c) + o(\Delta) \quad (9)$$

and for $v \in [K, B)$, we have $P(v < V \leq v + \Delta c) + o(\Delta) =$

$$\begin{aligned} & \lambda \Delta \int_0^\infty P\{v - (h-c)g < V < K\} dF_1(g) \\ &= \lambda \Delta \int_{\left(\frac{v-K}{h-c}, \infty\right)} P\{v - (h-c)g < V < K\} \\ & \quad \times dF_1(g) \end{aligned} \quad (10)$$

(9) can be interpreted as follows: the left hand side is the “up-crossing” probability of the level v within a short period of length Δ : the probability of having an arrival in the equivalent $M/G/1$ queue within a time interval $[t, t + \Delta]$ is approximately $\lambda \Delta$ for very small Δ . The distribution of the amount of workload brought by the arrival is $F_1(x/(h-c))$. If (at stationary regime) the amount of workload at the queue is within the interval $(v - (h-c)g, v]$ and an arrival containing an amount g of workload arrives within time $[t, t + \Delta]$ (with Δ very small) then the workload at the queue will exceed the level v during this time interval; this is called an up-crossing. The right hand side of the equation has a similar interpretation of a “downcrossing” of level v ; the equality represents the fact that at steady state, the upcrossing and downcrossing

probabilities should be the same (these crossing arguments are quite standard in queueing theory, see e.g. [16]). (10) has a similar interpretation.

Dividing (9) and (10) by Δ , we get for $v < K$:

$$\lambda \int_0^\infty P\{v - (h-c)g < V \leq v\} dF_1(g) = \lim_{\Delta \rightarrow 0} \left(\frac{P(v < V \leq v + \Delta c)}{\Delta} - \lambda P(v < V \leq v + \Delta c) \right)$$

and for $v \in [K, B)$,

$$\lim_{\Delta \rightarrow 0} \frac{P(v < V \leq v + \Delta c)}{\Delta} = \lambda \int_{((\frac{v-K}{h-c}), \infty)} P\{v - (h-c)g < V < K\} dF_1(g)$$

Assuming the existence of derivative of $F(v)$ for $v \in (0, K)$ and $[K, B)$ we have $\frac{dF(v)}{dv} =$

$$\begin{cases} \frac{\lambda}{c} \int_0^\infty P\{v - (h-c)g < V \leq v\} \times dF_1(g) & v \in (0, K) \\ \frac{\lambda}{c} \int_{((\frac{v-K}{h-c}), \infty)} P\{v - (h-c)g < V < K\} \times dF_1(g) & v \in [K, B) \end{cases}$$

which gives $\frac{dF(v)}{dv} =$

$$\begin{cases} \frac{\lambda}{c} (F(v) - \int_0^\infty P(V \leq v - (h-c)g) dF_1(g)) & v \in (0, K) \\ \frac{\lambda}{c} \left(F(K^-) (1 - F_1\left(\frac{v-K}{h-c}\right)) - \int_{((\frac{v-K}{h-c}), \infty)} P(V \leq v - (h-c)g) dF_1(g) \right) & v \in [K, B) \end{cases}$$

We shall first solve for $v \in [K, B)$. We have $\frac{dF(v)}{dv} =$

$$\begin{aligned} & \frac{\lambda}{c} F(K^-) \bar{F}_1\left(\frac{v-K}{h-c}\right) \\ & - \frac{\lambda}{c} \int_{((\frac{v-K}{h-c}), (v/(h-c))]} P(V \leq v - (h-c)g) dF_1(g) \\ & = \frac{\lambda}{c} F(K^-) \bar{F}_1\left(\frac{v-K}{h-c}\right) \\ & - \frac{\lambda}{c} \int_{[0, K)} F_1\left(\frac{v-z}{h-c}\right) dF(z) \\ & = \frac{\lambda}{c} (F(K^-) - F_1\left(\frac{v}{h-c}\right) \rho(0)) \\ & - \frac{\lambda}{c} \int_{[0, K)} F_1\left(\frac{v-z}{h-c}\right) dF(z) \end{aligned} \quad (11)$$

Observe that in (11) we have expressed the density for $v \in [K, B)$ in terms of the density for $v \in (0, K)$. We next solve

for density for $v \in (0, K)$. With $\rho(0) \equiv F(0)$, we have $\frac{dF(v)}{dv}$

$$\begin{aligned} & = \frac{\lambda}{c} (F(v) - F_1\left(\frac{v}{h-c}\right) \rho(0)) \\ & - \frac{\lambda}{c} \int_0^v F_1\left(\frac{v-z}{h-c}\right) dF(z) \\ & = \frac{\lambda}{c} \rho(0) \left(1 - F_1\left(\frac{v}{h-c}\right) \right) \\ & + \frac{\lambda}{c} \int_0^v \left(1 - F_1\left(\frac{v-z}{h-c}\right) \right) dF(z) \\ & = \frac{\lambda}{c} \rho(0) \bar{F}_1\left(\frac{v}{h-c}\right) \\ & + \frac{\lambda}{c} \int_0^v \bar{F}_1\left(\frac{v-z}{h-c}\right) dF(z) \end{aligned} \quad (12)$$

From (12) and (11) we see that F is continuous at K and thus there is no probability mass at $v = K$.

Remark 1: If there was no threshold discarding, i.e., $K = \infty$ then (13) is the standard *Takacs Integro-differential equation* (see for e.g., [1], [6]).

This is a standard Volterra Integral Equation [15] whose solution is given as ⁶:

$$\begin{aligned} \frac{dF(v)}{dv} = & \left[\sum_{i=1}^{\infty} \left(\frac{\lambda}{c}\right)^{i+1} \int_0^v M_i(v, u) \rho(0) \bar{F}_1\left(\frac{u}{h-c}\right) du \right] \\ & + \frac{\lambda}{c} \rho(0) \bar{F}_1\left(\frac{v}{h-c}\right) \end{aligned} \quad (14)$$

where the iterated kernels M_i are given as

$$\begin{aligned} M_1(v, g) & = \bar{F}_1\left(\frac{v-g}{h-c}\right) \\ M_2(v, g) & = \int_0^v \bar{F}_1\left(\frac{v-t}{h-c}\right) M_1(t, g) dt \\ & \vdots \\ M_n(v, g) & = \int_0^v \bar{F}_1\left(\frac{v-t}{h-c}\right) M_{n-1}(t, g) dt \\ & \vdots \end{aligned}$$

Substituting the expression for $\frac{dF(v)}{dv}$ from (14) in (11) we get $\frac{dF(v)}{dv}$:

$$\begin{aligned} & = \frac{\lambda}{c} (F(K) - F_1\left(\frac{v}{h-c}\right) \rho(0)) \\ & - \frac{\lambda}{c} \int_0^K F_1\left(\frac{v-z}{h-c}\right) \times \\ & \left(\left[\sum_{i=1}^{\infty} \left(\frac{\lambda}{c}\right)^{i+1} \int_0^z M_i(z, u) \rho(0) \bar{F}_1\left(\frac{u}{h-c}\right) du \right] \right. \\ & \left. + \frac{\lambda}{c} \rho(0) \bar{F}_1\left(\frac{z}{h-c}\right) \right) dz \end{aligned} \quad (15)$$

Observe that there is no probability mass at $v = B$, the workload hits B at isolated points of time (at jumps) and never stays there. Recall that we are working with the stationary

⁶See Appendix A for related theory

distribution of \mathcal{V}_n . Thus integrating (15) from K to B we shall have

$$\begin{aligned}
1 - F(K) = & \\
& + \frac{\lambda}{c} \left(F(K) \int_K^B dv - \rho(0) \int_K^B F_1 \left(\frac{v}{h-c} \right) dv \right) \\
& - \frac{\lambda}{c} \int_K^B \int_0^K F_1 \left(\frac{v-z}{h-c} \right) \left(\left[\sum_{i=1}^{\infty} \left(\frac{\lambda}{c} \right)^{i+1} \right. \right. \\
& \left. \left. \int_0^z M_i(z, u) \rho(0) \bar{F}_1 \left(\frac{u}{h-c} \right) du \right] \right. \\
& \left. + \frac{\lambda}{c} \rho(0) \bar{F}_1 \left(\frac{z}{h-c} \right) \right) dz dv \quad (16)
\end{aligned}$$

Secondly, integrating (14) from 0 to K we get:

$$\begin{aligned}
F(K) - \rho(0) = & \\
& \frac{\lambda}{c} \rho(0) \int_0^K \bar{F}_1 \left(\frac{v}{h-c} \right) dv + \rho(0) \left[\sum_{i=1}^{\infty} \left(\frac{\lambda}{c} \right)^{i+1} \right. \\
& \left. \times \int_0^K \int_0^v M_i(v, u) \bar{F}_1 \left(\frac{u}{h-c} \right) dudv \right] \quad (17)
\end{aligned}$$

Solving (17) and (16) we get expression for $\rho(0)$:

$$\begin{aligned}
& \left[\left(1 + \frac{\lambda}{c} \int_0^K \bar{F}_1 \left(\frac{v}{h-c} \right) dv + \sum_{i=1}^{\infty} \left(\frac{\lambda}{c} \right)^{i+1} \times \right. \right. \\
& \left. \int_0^K \int_0^v M_i(v, u) \bar{F}_1 \left(\frac{u}{h-c} \right) dudv \right) \times \\
& \left(1 + \frac{\lambda}{c} \int_K^B \bar{F}_1 \left(\frac{v-K}{h-c} \right) dv \right) \\
& - \frac{\lambda}{c} \left(\int_K^B F_1 \left(\frac{v}{h-c} \right) dv \right. \\
& \left. + \frac{\lambda}{c} \int_K^B \int_0^K F_1 \left(\frac{v-z}{h-c} \right) \left(\left[\sum_{i=1}^{\infty} \left(\frac{\lambda}{c} \right)^i \int_0^z M_i(z, u) \right. \right. \right. \\
& \left. \left. \left. \bar{F}_1 \left(\frac{u}{h-c} \right) du \right] + \bar{F}_1 \left(\frac{z}{h-c} \right) \right) dz dv \right) \right]^{-1} \quad (18)
\end{aligned}$$

which when substituted in (18) will give an expression for $F(K)$. Having obtained this we can get explicit expressions for density for $v \in (0, K)$ and $v \in [K, B)$ from (14) and (15) respectively.

Remark 2: For some distributions we get closed form expression for $\sum_{i=1}^{\infty} \left(\frac{\lambda}{c} \right)^{i+1} M_i(v, g)$. We shall call this sum as the resolvent and denote it by $\Gamma(v, g; \frac{\lambda}{c})$.

Having obtained a characterization of the density function of the workload process at the instant when a source is ready to send the burst, we can obtain an explicit expression for \mathcal{G}_S and \mathcal{G}_{FS} from Lemma 1. In the next section we provide explicit closed form expressions for distribution F of \mathcal{V} (or V) and hence of \mathcal{G}_S for exponential distribution of burst sizes.

B. Explicit Calculations for Exponential Distribution of On Periods

We have the following:

Proposition 2: For $F_1(x) = 1 - e^{-\mu x}$ we have $\frac{dF(v)}{dv} =$

$$\begin{cases} \frac{\lambda}{c} \rho(0) e^{-\frac{\mu v}{h-c}} \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}} \right) & v \in (0, K) \\ \frac{\lambda}{c} (F(K) - (1 - e^{-\mu(\frac{v}{h-c})}) \rho(0)) \\ - \rho(0) \left(\frac{\lambda}{c} \right)^2 \left[-e^{-\mu K/(h-c)} \left(\frac{h-c}{\mu} \right) \right. \\ + e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \left(K - \frac{c}{\lambda} \right) \\ - e^{-\mu v/(h-c)} \left(K + \left(K - \frac{c}{\lambda} \right) e^{\lambda K/c} \right) \\ \left. - \frac{\frac{\mu}{h-c} - \frac{\lambda}{c}}{\frac{\mu}{h-c} - \frac{\lambda}{c}} e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \right. \\ \left. \left(K + \frac{\frac{2\lambda}{c} - \frac{\mu}{h-c}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right) \frac{\lambda}{c}} \right) \right. \\ \left. + \left(\frac{h-c}{\mu} \right) + \frac{c}{\lambda} (1 - e^{-\mu v/h-c}) \right. \\ \left. + \frac{\left(\frac{\mu}{h-c} \right) \left(\frac{2\lambda}{c} - \frac{\mu}{h-c} \right)}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)^2 \frac{\lambda}{c}} \right] & v \in [K, B) \end{cases} \quad (19)$$

where $\rho(0) = \frac{1}{A_1 A_2 + A_3}$ and $F(K) = \frac{A_3}{A_1 A_2 + A_3}$, with:

$$A_1 = 1 + (B - K) \frac{\lambda}{c} \quad (20)$$

$$\begin{aligned}
A_2 = & \left(\frac{e^{-\frac{\mu K}{h-c}} - e^{-\frac{\mu B}{h-c}}}{\frac{\mu}{h-c}} \right) \left(\frac{2\lambda}{c} + \left(\frac{\lambda}{c} \right)^2 \right. \\
& \left. \left(K + \left(K - \frac{c}{\lambda} \right) e^{\frac{\lambda K}{c}} \right) \right. \\
& - (B - K) \frac{\lambda}{c} \left(1 + \frac{\lambda}{c} \left(-e^{-\frac{\mu K}{h-c}} \left(\frac{h-c}{\mu} \right) \right. \right. \\
& \left. \left. + e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \left(K - \frac{c}{\lambda} \right) \right) \right. \\
& \left. - \frac{\frac{\mu}{h-c}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)} e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \right. \\
& \left. \left(K + \frac{\frac{2\lambda}{c} - \frac{\mu}{h-c}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right) \frac{\lambda}{c}} \right) + \left(\frac{h-c}{\mu} \right) + \frac{c}{\lambda} \right. \\
& \left. + \frac{\left(\frac{\mu}{h-c} \right) \left(\frac{2\lambda}{c} - \frac{\mu}{h-c} \right)}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)^2 \frac{\lambda}{c}} \right) \right), \quad (21)
\end{aligned}$$

$$\begin{aligned}
A_3 = & 1 + \frac{\lambda}{c} \left(\left(\frac{h-c}{\mu} \right) (1 - e^{-\mu K/h-c}) \right. \\
& \left. + \frac{\frac{\lambda}{c}}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \left(\frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} - e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \right) \right. \\
& \left. \left(K + \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) \right) \quad (22)
\end{aligned}$$

Proof: Refer to Appendix B.

And further:

Proposition 3: For $F_1(x) = 1 - e^{-\mu x}$, \mathcal{G}_S is given by:

$$\begin{aligned}
= & \mu^2 \rho(0) \left[I_0 \left(1 + \frac{\lambda(h-c)}{c\mu} (1 - e^{-\frac{\mu K}{h-c}}) \right) \right. \\
& - \frac{\left(\frac{\lambda}{c} \right)^2}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)} \times \\
& \left. \left(e^{-K(\frac{\mu}{h-c} - \frac{\lambda}{c})} \left(K + \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) - \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) \right) \quad (23)
\end{aligned}$$

$$\begin{aligned}
& +I_3 - \frac{\left(\frac{\lambda}{c}\right)^2}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \left(e^{-B\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \times \right. \\
& \left. \left(\left(B + \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) I_1 - (h-c)I_2 \right) - \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} I_3 \right) \\
& \left. + \frac{\lambda(h-c)}{\mu c} (I_3 - I_4) \right] \quad (23)
\end{aligned}$$

where I_0, I_1, I_2, I_3 are defined as:

$$\begin{aligned}
I_0 &= \int_0^{\frac{B-K}{h-c}} w e^{-\mu w} dw & I_1 &= \int_{\frac{B-K}{h-c}}^{\frac{B}{h-c}} w e^{-w\left(\frac{h-c}{c}\right)\lambda} dw \\
I_2 &= \int_{\frac{B-K}{h-c}}^{\frac{B}{h-c}} w^2 e^{-w\left(\frac{h-c}{c}\right)\lambda} dw & I_3 &= \int_{\frac{B-K}{h-c}}^{\frac{B}{h-c}} w e^{-\mu w} dw \\
I_4 &= \int_{\frac{B-K}{h-c}}^{\frac{B}{h-c}} w dw
\end{aligned}$$

and $\rho(0)$ is obtained by solving (33) and (34).

Proof: Refer to Appendix C.

Remark 3: From (23) we observe that the \mathcal{G}_S is not defined when the load $\frac{\lambda(h-c)}{\mu c}$ at the queue is equal to 1.

C. A Bound for General Distribution of On Periods

Observe that from (14) the expression for $\frac{dF(v)}{dv}$ requires the calculation of M_i for all $i \geq 1$. Thus a natural interest is in bounding the error if the summation in (14) is truncated to some finite number, say n of terms. Let us denote by $S_i = \left(\frac{\lambda}{c}\right)^{i+1} \int_0^v M_i(v, u) \rho(0) \bar{F}_1\left(\frac{u}{h-c}\right) du$. We have the following result:

Lemma 3: The absolute value of the general term S_i in (14) is less than:

$$\left(\frac{\lambda}{c}\right)^i \frac{K^{i-1}}{(i-1)!} \int_0^K \left(\frac{\lambda}{c}\right) \rho(0) \bar{F}_1\left(\frac{s}{h-c}\right) ds$$

And thus if in the series in (14) only those terms are included which contain powers of $\left(\frac{\lambda}{c}\right)$ up-to the $n+1$ th-term then the magnitude of error will not exceed

$$\left(\frac{\lambda}{c}\right)^2 \int_0^K \rho(0) \bar{F}_1\left(\frac{s}{h-c}\right) ds \sum_{i=1}^n \left(\frac{\lambda}{c}\right)^{i-1} \frac{K^{i-1}}{(i-1)!}. \quad (24)$$

Also as $\bar{F}_1(\cdot) \leq 1$ the error magnitude from (24) will be less than:

$$\left(\frac{\lambda}{c}\right)^2 K \rho(0) \sum_{i=1}^n \left(\frac{\lambda}{c}\right)^{i-1} \frac{K^{i-1}}{(i-1)!}. \quad (25)$$

Proof: Follows from the theory of Volterra Integral Equations in [15].

IV. NUMERICAL EXAMPLES

We shall next plot the stationary density of the workload process \mathcal{V} at arrival epochs of bursts and the goodput ratios $\mathcal{G}_S, \mathcal{G}_{FS}$ using our analytical expressions for an example with exponential distribution of burst sizes. We take $h = 30, c = 10, \lambda = 20, \mu = 30, B = 20, K = 10$. We observe in Fig. 3 that density has a peak at the discarding threshold K . We next plot \mathcal{G}_S as a function of K in Fig. 4. We observe that with appropriate choice of K \mathcal{G}_S as high as 0.73 can be obtained.

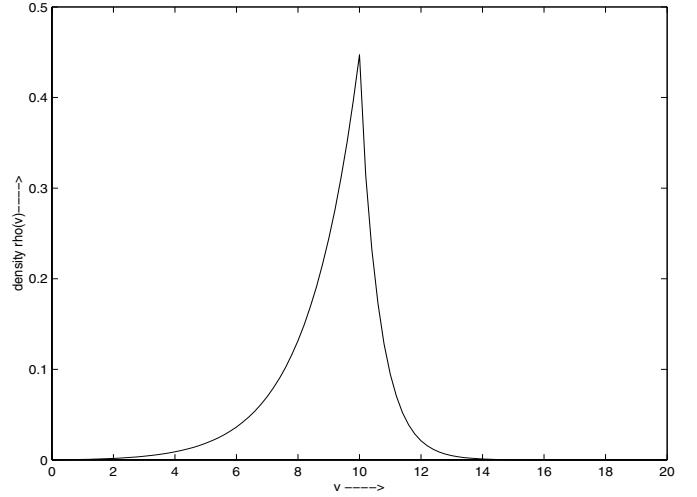


Fig. 3. The density function of the stationary workload process for the example with $\rho(0) = 7.5321 \times 10^{-5}$. Observe that there is a peak at discarding threshold $K = 10$.

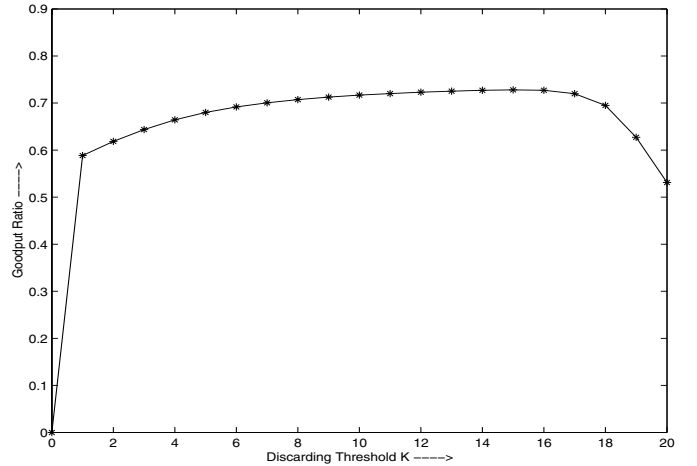


Fig. 4. The goodput ratio \mathcal{G}_S as a function of K with $h = 30, c = 10$. For $K = 15$, the goodput is maximum at 0.7281.

We next plot \mathcal{G}_{FS} as a function of K in Fig. 5. We observe that the optimum value of K is anything from 1 to 15 as \mathcal{G}_{FS} is almost 1 for these values of K . Further \mathcal{G}_{FS} decreases with K for $K > 15$. For this example we find that $K = 15$ is the optimum as it maximizes both \mathcal{G}_S and \mathcal{G}_{FS} . Also, in the absence of any threshold discarding, $K = B$, both \mathcal{G}_S and \mathcal{G}_{FS} are low and hence having a selective discarding policy improves both the performance as perceived by the source and the overall performance. For this particular example in Fig. 4 it seems that the choice of the threshold K scarcely influence the performance. However we will demonstrate that it may not be true for other values of parameters. Let us now take $h = 100$ while keeping all other parameters the same and plot \mathcal{G}_S as a function of K in Fig. 6. We observe that for this example \mathcal{G}_S is very sensitive to the choice of K calling for the need of optimal discarding threshold for such a system.

Remark 4: From the explicit closed form expressions for \mathcal{G}_S (and hence \mathcal{G}_{FS}) from Proposition 3 for the exponential

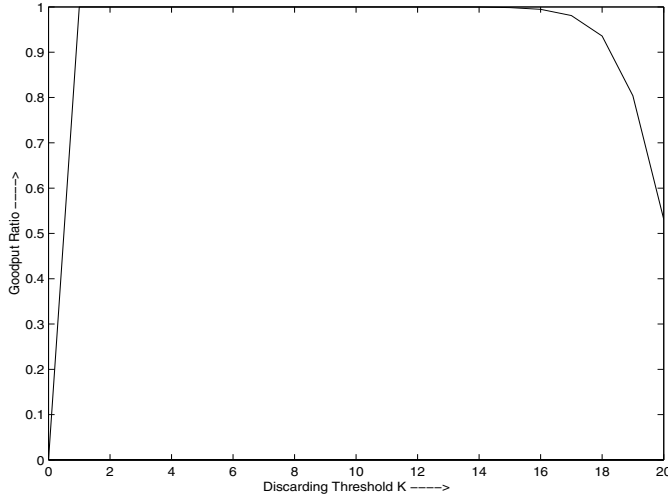


Fig. 5. The goodput ratio \mathcal{G}_{FS} as a function of K . For $K = 1$ to 15 the goodput is very close to 1.

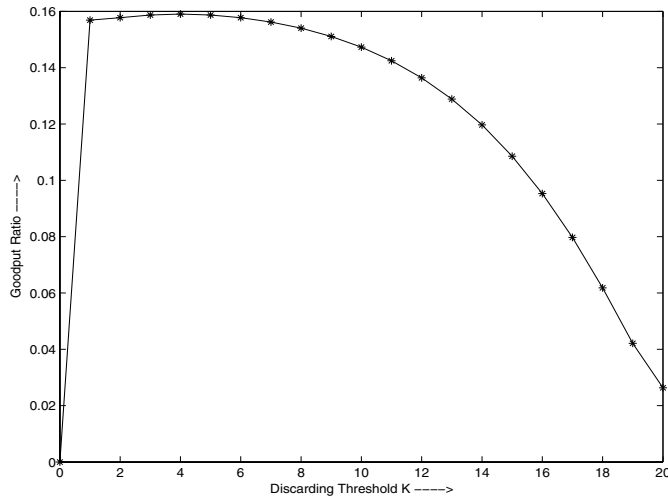


Fig. 6. The goodput ratio \mathcal{G}_S as a function of K with $h = 100$, $c = 10$. For $K = 4$, the goodput is maximum at 0.1590.

case, the optimal discarding threshold can be obtained by solving $\frac{d\mathcal{G}_S}{dK} = 0$.

V. AN IMPROVED FEEDBACK SYSTEM: A COMBINATION OF THRESHOLD AND PARTIAL DISCARDING WITH RESPONSIVE SOURCES

In our model we have considered the case where a feedback to the source to stop sending data is sent only at the arrival epochs of bursts. However observe from Fig. 1 that there is a positive probability that some (accepted) messages will become bad if during their arrival period the buffer overflows and then the fluid queue stays at B . Next consider a feedback system where apart from threshold discarding (like in our model) there is also partial discarding of messages. Thus whenever the buffer level becomes B a feedback signal is sent to the source to stop the current burst transfer and the source backs off (responsive source) and enters a silence period. Thus the probability mass that we have in Fig. 1 disappears. This policy will certainly improve the overall system performance

or in other words \mathcal{G}_{FS} for this policy will be higher than that for our model. We shall next show how one can obtain the performance metrics for this policy from our analysis. We shall denote the previous policy that we analysed by *TMDF* (Threshold Message Discarding with Feedback) and this improved policy by *PTMDF* (Partial and Threshold Message Discarding with Feedback). A little thinking clears that the workload process at arrival epochs for both *TMDF* and *PTMDF* has the same distribution. Thus we have:

Lemma 4: The stationary distribution of the workload process at arrival epochs of messages in *PTMDF* is given by (14) and (11) for $v \in (0, K)$ and $[K, B)$ respectively.

Having obtained this we now proceed to obtain the two goodput ratios for this system. Observe that \mathcal{G}_S will be the same for both *TMDF* and *PTMDF*. However \mathcal{G}_{FS} will be different as here the source also backs off once the queue hits B . We have:

Proposition 4: For *PTMDF* we have:

$$\begin{aligned}\mathcal{G}_S &= M^{-1} \int_0^{\frac{B}{h-c}} wdH(w, 1) \\ \mathcal{G}_{FS} &= \frac{\int_0^{\frac{B}{h-c}} wdH(w, 1)}{E[Z]}\end{aligned}$$

where M is the mean length of the message and Z is a random variable defined as:

$$Z = \begin{cases} W & \text{if } \mathcal{V} < K, W < \frac{B-\mathcal{V}}{h-c} \\ \frac{B-\mathcal{V}}{h-c} & \text{if } \mathcal{V} < K, W \geq \frac{B-\mathcal{V}}{h-c} \\ 0 & \text{o.w.,} \end{cases} \quad (26)$$

with E as the expectation.

Proof: Under *PTMDF* the actual amount of fluid for a message sent by the source, depends on both \mathcal{V} and the (potential) length of the message. Thus if $\mathcal{V} < K$ and the potential length of the message W can be completely accepted then $Z = W$, if $\mathcal{V} < K$ and the message cannot be completely accepted then $Z = \frac{B-\mathcal{V}}{h-c}$ and if $\mathcal{V} \geq K$, $Z = 0$.

Corollary 1: For *PTMDF* with $F_1(x) = 1 - e^{-\mu x}$, we have \mathcal{G}_S given by (23) and $\mathcal{G}_{FS} = \frac{\mathcal{G}_S}{E[Z]}$ where

$$\begin{aligned}E[Z] &= \rho(0) \left[1 - e^{-\frac{\mu B}{h-c}} + \frac{\lambda}{c} \left(\frac{h-c}{\mu} (1 - e^{-\frac{\mu K}{h-c}}) \right. \right. \\ &\quad \left. \left. - e^{-\frac{\mu B}{h-c}} \left(\frac{c}{\lambda} + e^{\frac{K\lambda}{c}} \left(K - \frac{c}{\lambda} \right) \right) \right) - K e^{-\frac{\mu B}{h-c}} \right. \\ &\quad \left. + \frac{\lambda}{c} \left(\frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)^2} - \frac{e^{-K \left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)} \right. \right. \\ &\quad \left. \left. \left(K + \frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)} \right) \right) \right]. \quad (27)\end{aligned}$$

Proof: We have from (26),

$$\begin{aligned}Z &= WI \left(\mathcal{V} < K, W < \frac{B-\mathcal{V}}{h-c} \right) \\ &\quad + \left(\frac{B-\mathcal{V}}{h-c} \right) I \left(\mathcal{V} < K, W \geq \frac{B-\mathcal{V}}{h-c} \right). \quad (28)\end{aligned}$$

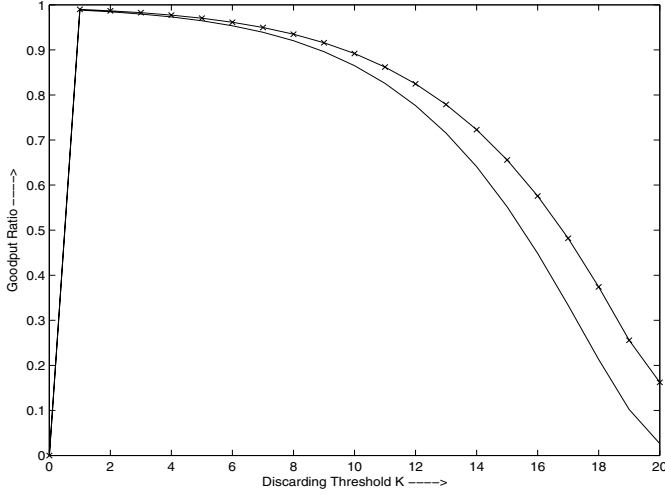


Fig. 7. The goodput ratio \mathcal{G}_{FS} for *TMDF* and *PTMDF* as a function of K with $h = 100$, $c = 20$. Observe the improvement in \mathcal{G}_{FS} achievable by combining partial with threshold discarding.

We have $W \sim \mu e^{-\mu w}$. Thus

$$\begin{aligned}
 E[Z] &= \int_0^K \left[\int_0^{\frac{B-v}{h-c}} \mu e^{-\mu w} w dw \right. \\
 &\quad \left. + \int_{\frac{B-v}{h-c}}^{\infty} \left(\frac{B-v}{h-c} \right) \mu e^{-\mu w} dw \right] dF(v) \\
 &= \int_0^K \left[\frac{1}{\mu} - e^{-\mu \left(\frac{B-v}{h-c} \right)} \left(\frac{B-v}{h-c} + \frac{1}{\mu} \right) \right. \\
 &\quad \left. + \left(\frac{B-v}{h-c} \right) e^{-\mu \left(\frac{B-v}{h-c} \right)} \right] dF(v) \\
 &= \frac{1}{\mu} \int_0^K \left(1 - e^{-\mu \left(\frac{B-v}{h-c} \right)} \right) dF(v).
 \end{aligned}$$

From Prop. 4, distribution of \mathcal{V} is given by (14) and (11). Thus substituting the expression for $dF(v)$ we get

$$\begin{aligned}
 E[Z] &= \frac{\rho(0)}{\mu} \left[\left(1 - e^{-\frac{\mu B}{h-c}} \right) + \frac{\lambda}{c} \int_0^K \left(1 - e^{-\mu \left(\frac{B-v}{h-c} \right)} \right) \times \right. \\
 &\quad \left. e^{-\frac{\mu v}{h-c}} \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}} \right) dv \right] \\
 &= \frac{\rho(0)}{\mu} \left[\left(1 - e^{-\frac{\mu B}{h-c}} \right) + \frac{\lambda}{c} \left(e^{-\frac{\mu v}{h-c}} \right. \right. \\
 &\quad \left. \left. + \frac{\lambda}{c} v e^{-v \left(\frac{\mu}{h-c} - \frac{\lambda}{c} \right)} - e^{-\frac{\mu B}{h-c}} \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}} \right) \right) \right] dv,
 \end{aligned}$$

from which we have (27).

Fig. 7 shows the improvement in \mathcal{G}_{FS} for *TMDF* and *PTMDF* for the same example source with $h = 100$.

VI. CONCLUSION AND SCOPE FOR FURTHER RESEARCH

We analyse a feedback system with a fluid queue having a selective burst discarding policy and a responsive source that backs off when the burst is not accepted by the queue. Using techniques from Volterra Integral Equations we characterize the distribution of the queue length at arrival epochs of bursts

(at times when the source gets ready to transmit). For general distribution of active periods we propose an approach for calculating the queue length distribution and for the special case of exponential distribution we obtain explicit closed form expressions for the distribution of the workload at arrival epochs of bursts and the goodput ratios. Our analytical expressions are useful in studying the sensitivity of the performance to different parameters of the source and network (if we model the bottleneck as the fluid queue), in selecting optimal discarding threshold and also in buffer dimensioning. As we saw in our numerical investigations, there are parameters for which almost optimal goodput is obtained for a large range of choices of the discarding threshold (Fig. 5), whereas for other parameters, the performance is quite sensitive to the choice of threshold (Fig. 6 and Fig. 7).

In our analysis we have assumed that there is a zero feedback delay from the queue to the source and the source backs off instantaneously. This can be a good approximation for controlling access networks [8], since distances between users/clients and network access points are relatively short, feedback delay due to propagation is negligible. A further extension of our work can be to analyse active queue management policies with probabilistic discarding of bursts. For e.g., we can have two thresholds K_1 and K_2 such that if the workload is below K_1 , a burst is not discarded, if its above K_2 , the burst is always discarded and between K_1 and K_2 , the discarding can be probabilistic.

APPENDIX A

We state the following theorem for the existence and uniqueness of the solution to a general integral equation of Volterra type from [15].

Theorem 1: Consider an integral equation

$$\phi(x) - \beta \int_a^v D(x, s) \phi(s) ds = f(x), \quad (29)$$

whose solution by the method of successive approximations be given by (the n th approximate)

$$\phi_n(x) = f(x) + \sum_{m=1}^n \beta^m \int_a^v D_m(x, s) f(s) ds \quad (30)$$

where $D_m(x, s)$ is determined by the recurrence relationship

$$D_1(x, s) = D(x, s); \quad D_m(x, s) = \int_a^v D(x, t) D_{m-1}(t, s) dt.$$

Then if $D(x, s)$ is bounded and $f(x)$ is absolutely integrable, i.e.,

$$\int_a^v |f(s)| ds < \infty$$

then successive approximations of the form $\phi_n(x)$ (30) converge uniformly for all values of β and limit of successive approximations is the solution of (29), and this solution is unique.

APPENDIX B

We will next calculate the kernels $M_i(v, g)$ and provide a general expression for $i = n$. Observe that $M_1(v, g) = e^{-\mu(v-g)/(h-c)}$ and

$$\begin{aligned} M_2(v, g) &= \int_0^v e^{-\mu(v-t)/(h-c)} e^{-\mu(t-g)/(h-c)} dt \\ &= v e^{-\mu(v-g)/(h-c)} \\ M_3(v, g) &= \int_0^v e^{-\mu(v-t)/(h-c)} t e^{-\mu(t-g)/(h-c)} dt \\ &= \frac{v^2}{2!} e^{-\mu(v-g)/(h-c)} \\ &\vdots \end{aligned}$$

and, in general

$$M_n(v, g) = \frac{v^{n-1}}{(n-1)!} e^{-\mu(v-g)/(h-c)}$$

Hence the resolvent

$$\begin{aligned} \Gamma\left(v, g; \frac{\lambda}{c}\right) &= e^{-\mu(v-g)/(h-c)} \sum_{m=1}^{\infty} \left(\frac{\lambda}{c}\right)^{m+1} \frac{v^{m-1}}{(m-1)!} \\ &= e^{-\mu(v-g)/(h-c)} e^{\lambda v/c} \left(\frac{\lambda}{c}\right)^2 \\ &= \left(\frac{\lambda}{c}\right)^2 e^{v\left(\frac{\lambda}{c} - \frac{\mu}{h-c}\right)} e^{\frac{\mu g}{h-c}}. \end{aligned}$$

Thus we get for $v \in (0, K)$:

$$\frac{dF(v)}{dv} = \frac{\lambda}{c} \rho(0) e^{\frac{-\mu v}{h-c}} \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}}\right) \quad (31)$$

And from (15) for $v \in [K, B)$:

$$\begin{aligned} \frac{dF(v)}{dv} &= \quad (32) \\ &\frac{\lambda}{c} (F(K) - F_1\left(\frac{v}{h-c}\right) \rho(0)) \\ &- \frac{\lambda}{c} \int_0^K F_1\left(\frac{v-z}{h-c}\right) \rho(0) \frac{\lambda}{c} \times \\ &\left[e^{-\mu z/(h-c)} + e^{z\left(\frac{\lambda}{c} - \frac{\mu}{h-c}\right)} \frac{\lambda z}{c} \right] dz \\ &= \frac{\lambda}{c} (F(K) - (1 - e^{-\mu\left(\frac{v}{h-c}\right)}) \rho(0)) - \rho(0) \left(\frac{\lambda}{c}\right)^2 \\ &\left[(e^{-\mu z/(h-c)} - e^{-\mu v/(h-c)}) \left(z + \left(z - \frac{c}{\lambda}\right) e^{\lambda z/c}\right) \right. \\ &\left. + \frac{\mu}{h-c} \int e^{-z\mu/(h-c)} \left(z + \left(z - \frac{c}{\lambda}\right) e^{\frac{\lambda z}{c}}\right) dz \right]_0^K \\ &= \frac{\lambda}{c} (F(K) - (1 - e^{-\mu\left(\frac{v}{h-c}\right)}) \rho(0)) - \rho(0) \left(\frac{\lambda}{c}\right)^2 \\ &\left[-e^{-\mu z/(h-c)} \left(\frac{h-c}{\mu}\right) + e^{-z\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \left(z - \frac{c}{\lambda}\right) \right. \\ &\left. - e^{-\mu v/(h-c)} \left(z + \left(z - \frac{c}{\lambda}\right) e^{\lambda z/c}\right) \right. \\ &\left. - \frac{\frac{\mu}{(h-c)}}{\frac{\mu}{h-c} - \frac{\lambda}{c}} e^{-z(\mu/h-c-\lambda/c)} \left(z + \frac{\frac{2\lambda}{c} - \frac{\mu}{h-c}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right) \frac{\lambda}{c}}\right) \right]_0^K \end{aligned}$$

From which we get (19). Integrating both sides of the last equation from K to B we get:

$$A_1 F(K) + A_2 \rho(0) = 1 \quad (33)$$

with A_1 and A_2 as defined in Prop. 2. Further integrating (31) from 0 to K we have :

$$F(K) = \rho(0) A_3 \quad (34)$$

with A_3 as defined in Prop. 2. Solving (33) and (34) we get $F(K) = \frac{A_3}{A_1 A_3 + A_2}$ and $\rho(0) = \frac{1}{A_1 A_3 + A_2}$ which when substituted in (31) and (32) gives explicit expressions for density for $(0, K)$ and $[K, B]$ respectively.

APPENDIX C

Proof of Corollary 3 We have $F_1(x) = 1 - e^{-\mu x}$. From (5) and (31) we have for $w \in [0, \frac{B-K}{h-c}]$

$$\begin{aligned} H(w, 1) &= \\ &(1 - e^{-\mu w}) \rho(0) \left[1 + \frac{\lambda}{c} \int_0^K e^{\frac{-\mu v}{h-c}} \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}}\right) dv \right], \end{aligned}$$

from which we get $\frac{dH(w,1)}{dw}$

$$\begin{aligned} &= \mu e^{-\mu w} \rho(0) \left[1 + \frac{\lambda(h-c)}{c\mu} (1 - e^{\frac{-\mu K}{h-c}}) \right. \\ &\left. + \left(\frac{\lambda}{c}\right)^2 \int_0^K v e^{-v\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} dv \right] \\ &= \mu e^{-\mu w} \rho(0) \left[1 + \frac{\lambda(h-c)}{c\mu} (1 - e^{\frac{-\mu K}{h-c}}) - \frac{\left(\frac{\lambda}{c}\right)^2}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \right. \\ &\left. \left(e^{-K\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \left(K + \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) - \frac{1}{\frac{\mu}{h-c} - \frac{\lambda}{c}} \right) \right] \quad (35) \end{aligned}$$

from (6) we have $H_a(w) =$

$$\begin{aligned} &(1 - e^{-\mu w}) \rho(0) \left[1 + \frac{\lambda}{c} \int_0^{B-w(h-c)} e^{\frac{-\mu v}{h-c}} \right. \\ &\left. \left(1 + \frac{\lambda}{c} v e^{\frac{v\lambda}{c}} \right) dv \right], \end{aligned}$$

from which we get $\frac{dH_a(w,1)}{dw}$

$$\begin{aligned} &= \rho(0) \left(\mu e^{-\mu w} \left(1 + \frac{\lambda}{c} \left(\frac{1 - e^{\frac{-\mu}{h-c}(B-w(h-c))}}{\frac{\mu}{h-c}} \right) \right. \right. \\ &\left. \left. + \frac{\lambda}{c} \left(\frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)^2} - \frac{e^{-(B-w(h-c))\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \right) \right) \right. \\ &\left. \times \left(B - w(h-c) + \frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)} \right) \right) \\ &- (1 - e^{-\mu w}) \frac{\lambda}{c} (h-c) e^{\frac{-\mu}{h-c}(B-w(h-c))} \\ &\left(1 + \frac{\lambda}{c} (B - w(h-c)) e^{(B-w(h-c))\frac{\lambda}{c}} \right) \quad (36) \end{aligned}$$

and from (7) we have $H_b(w) =$

$$\frac{\lambda}{c}\rho(0)\int_{(B-w(h-c))^+}^K (1 - e^{-\mu(\frac{B-v}{h-c})}) e^{-\mu v/h-c} \left(1 + \frac{\lambda}{c}ve^{v\lambda/c}\right) dv,$$

which gives $\frac{dH_b(w)}{dw}$:

$$= \frac{\lambda}{c}\rho(0)(h-c)(1 - e^{-\mu w})e^{-\frac{\mu}{h-c}(B-w(h-c))} \times \left(1 + \frac{\lambda}{c}(B-w(h-c))e^{(B-w(h-c))\lambda/c}\right) \quad (37)$$

Thus from Eqs. (36) and (37) we have for $w \in (\frac{B-K}{h-c}, \frac{B}{h-c}]$, $\frac{dH(w,1)}{dw} = \frac{dH_a(w)}{dw} + \frac{dH_b(w)}{dw}$ and thus

$$\begin{aligned} \frac{dH(w,1)}{dw} = & \quad (38) \\ & \rho(0)\mu e^{-\mu w} \left(1 + \frac{\lambda}{c} \left(\frac{1 - e^{-\frac{\mu}{h-c}(B-w(h-c))}}{\frac{\mu}{h-c}}\right) \right. \\ & + \frac{\lambda}{c} \left(\frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)^2} - \frac{e^{-(B-w(h-c))\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)}}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)}\right) \\ & \left. \times \left(B - w(h-c) + \frac{1}{\left(\frac{\mu}{h-c} - \frac{\lambda}{c}\right)}\right)\right) \end{aligned}$$

Now from Lemma 1 we can write

$$\mathcal{G} = \mu \left(\int_0^{\frac{B-K}{h-c}} wdH(w,1) + \int_{\frac{B-K}{h-c}}^{\frac{B}{h-c}} wdH(w,1) \right) \quad (39)$$

Thus we get (23) from (35), (38) and (39).

REFERENCES

- [1] J. W. Cohen. *The Single Server Queue*. North-Holland Publishing Company, 1982.
- [2] S. Floyd and A. Romanow. Dynamics of TCP traffic over ATM networks. *ACM SIGCOMM'94*, pages 79–88, September 1994.
- [3] H. Heffes, D. M. Lucantoni. A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance. *IEEE Journal on Selected Areas in Communications*, pages 856–867, 1986.
- [4] K. K. Ramakrishnan, S. Floyd, D. Black. The Addition of Explicit congestion Notification (ECN) to IP. *RFC 3168, Proposed Standard*, available at <ftp://ftp.isi.edu/in-notes/rfc3168.txt>, September 2001.
- [5] K. S. Lee, M. Kim. Matching Technique for MMPP Modeling of Heterogeneous ON-OFF Sources. *IEEE GLOBECOM*, 1994.
- [6] L. Kleinrock. *Queueing Systems, Vol. I: Theory*. John Wiley & Sons, 1975.
- [7] Y. Lapid, R. Rom, and M. Sidi. Analysis of Discarding Policies in High-Speed Networks. *IEEE JSAC*, 16:764–777, 1998.
- [8] W. Scheinhardt M. Mandjes, D. Mitra. Simple models of network access, with applications to design of joint rate and admission control. *proc. of Infocom 2002*, June 2002.
- [9] P. Dube, E. Altman. Fluid Analysis of Early Message Discarding Policy Under Heavy Traffic. *proc. of IEEE INFOCOM 2002, June 2002*.
- [10] P. Dube, E. Altman. Queueing Analysis of Early Message Discard Policy. *proc. of IEEE ICC 2002, April/May 2002*.
- [11] P. Dube, E. Altman. On the Workload Process in a Fluid Queue with Bursty Input and Selective Discarding. In *Proc. of Seventeenth International Teletraffic Congress ITC 17, Salvador de Bahia, Brazil, Dec. 2001*. A longer version submitted to *Probability in the Engineering and Informational Sciences*, Jan 2002.
- [12] P. Dube, E. Altman. Queueing and Fluid Analysis of Partial Message Discard Policy. *proc. of 9th IFIP Working Conference on Performance Modeling and Evaluation of ATM and IP Networks, longer version submitted to Queueing Systems*, June 2001.
- [13] R. W. Brockett, W. B. Gong, Y. Guo. New analytical methods for queueing systems. *IEEE Conference on Decision and Control*, 1999.
- [14] S. Floyd. TCP and Explicit Congestion Notification. *ACM Computer Communication Review*, 24(5):10–23, October 1994.
- [15] S.G.Mikhlin. *Integral Equations, International Series of Monographs on Pure and Applied Mathematics, Vol. 4*. Pergamon Press, 1964.
- [16] M. Shalmon. The GI/GI/1 queue and its variations via the levs preemptive resume discipline. *Prob. Engng. Inform. Sc.* 2, pages 215–230, 1988.