

APPROXIMATION OF LOSS PROBABILITY FOR MULTIPLEXING SYSTEM WITH LARGE NUMBER OF ON-OFF SOURCES REGULATED BY LEAKY BUCKET

Shoji Kasahara
Nara Institute of Science and Technology, TAO

Toshiharu Hasegawa
Kyoto University

(Received April 18, 1997; Revised November 10, 1997)

Abstract This paper considers the loss probability of the multiplexing system with the large size of the buffer and the large number of homogeneous ON-OFF sources. We model the multiplexing system as the fluid queueing system with the regulated input fluid traffic. Each traffic generated by an ON-OFF source (*original traffic*) is regulated by the leaky bucket and aggregated output traffic (*regulated traffic*) flows into the buffer of the multiplexer. Using the CDE approximation method, which consists of the loss probability for the bufferless system and that for the system with the large buffer, we explicitly derive the approximated probability of the buffer overflow for the multiplexing system. Our result includes the two important parameters of the leaky bucket; the size of a token pool and the token generation rate. Hence, our result provides the significant indication for designing the input source regulator with the leaky bucket.

1. Introduction

In the ATM-based Broadband ISDN (B-ISDN), it is expected that the different classes of traffic such as voice, video and data are multiplexed and carried to their destinations keeping their quality of service (QoS) satisfied the required level. There has been a number of works for the ATM-based B-ISDN from the many points of view such as the switching architecture, policing strategies, the performance of the ATM switch, the behavior of network system which consists of ATM switches, and so on. Among those researches, the theory of the effective bandwidth and its applications of the communication system have been recognized as powerful tools for the performance analysis of the ATM-based B-ISDN.

One of applications of the effective bandwidth is to approximate the probability of the buffer overflow for the multiplexing system. Concerning the approximation for the buffer overflow probability of the multiplexing system, Elwalid et al. [5] has proposed the Chernoff-dominant eigenvalue (CDE) method. The CDE method consists of two parts: one is obtained from the Chernoff's theorem and the other is the dominant eigenvalue of the multiplexing system, which is derived in the standard effective bandwidth argument. It has been suggested that the loss probability of the bufferless system is well approximated by the parameter obtained by Chernoff's theorem, while the loss probability of the system with the large buffer can be approximated by the dominant eigenvalue associated with the Markovian input traffic source. [5] also showed that the approximated probability of the buffer overflow is well matched with the simulation result, where the data of the video teleconferencing is modeled as an input traffic.

[9] considered the effective bandwidth of general Markovian traffic sources and studied the optimization problem of allocating token generation rates among the heterogeneous sources. They derived the explicit formula of the asymptotic log moment generating function (ALMGF) for the ON-OFF source, where ON and OFF periods are exponentially distributed

and the fluid traffic is generated constantly at ON time.

In this paper, we consider the loss probability of the multiplexing system, where the original input is an ON-OFF source and the generated traffic is regulated by leaky bucket. The leaky bucket has two types of buffers; the data buffer and the token pool. As the token pool is getting small, the delays in the data buffer becomes large while the output stream becomes smooth. From this property, the leaky bucket has been considered that it will be used as a policing device in the ATM network[6].

We model the considered multiplexing system as a fluid queueing one. We suppose that the regulator has the infinite buffer and finite token pool, and that the token is generated at the fixed rate. Note that our regulator model is the special case analyzed in [6]. We approximate the loss probability of the multiplexing system with the large number of sources and the large buffer using the CDE method. Supposing that ON and OFF periods are exponentially distributed, we can easily obtain the steady state probability for the output process from the regulator with leaky bucket. Using this probability, we derive the asymptotic loss probability of the bufferless multiplexing system by the Chernoff's theorem. On the other hand, from the standard argument of the effective bandwidth, we also approximate the loss probability of the multiplexing system with the large buffer using the explicit ALMGF derived in [9].

There are a number of studies for the theory of the effective bandwidth and its applications (see [1, 2, 7, 8]). In particular, the fluid model with the access regulator has been analyzed in [6]. In [7], the fluid queueing model with general Markovian traffic sources has been studied. The important element of the effective bandwidth is the ALMGF. [8] showed that the ALMGF can be computed if the considered process is a Markov regenerative process. Concerning the fluid queueing theory, [10] studied the first passage time in a fluid flow model where the input and output rates are governed by a finite state continuous time Markov chain. In particular, the Laplace transform of the distribution of the first passage time for an ON-OFF source has been derived.

The theory of the effective bandwidth is based on the large deviation theory. [4] has derived the asymptotic decay rate of the waiting time distribution under the general conditions where the stationarity of the input-output process and a Gärtner-Ellis condition are supposed. [3] has analyzed the queueing behavior with the large number of input sources. The authors derived the approximation of the loss probability using the scaling technique and applied the results to the queueing system with heterogeneous inputs. [12] has presented the fundamentals of the large deviation theory for applying to the computer and communication systems. For the more detailed theories and its applications, in particular communication systems, readers are referred to [11].

The paper is organized as follows. In Section 2, we describe our mathematical model in detail. We present the outline of the CDE method in Section 3. In Section 4, we analyze the loss probability for the bufferless system. In Section 5, we summarize the derivation of the ALMGF of the output from the regulator and analyze the loss probability for the system with the large buffer. Finally, we derive the expression of the loss probability for CDE method. We show some numerical examples in Section 6.

2. Fluid Queueing Model

We consider the fluid queueing system with multiple sources where each source is regulated by the leaky bucket scheme (Fig.1).

The input source is modeled as an ON-OFF source where ON and OFF periods are exponentially distributed with rates α and β , respectively. The source generates traffic at

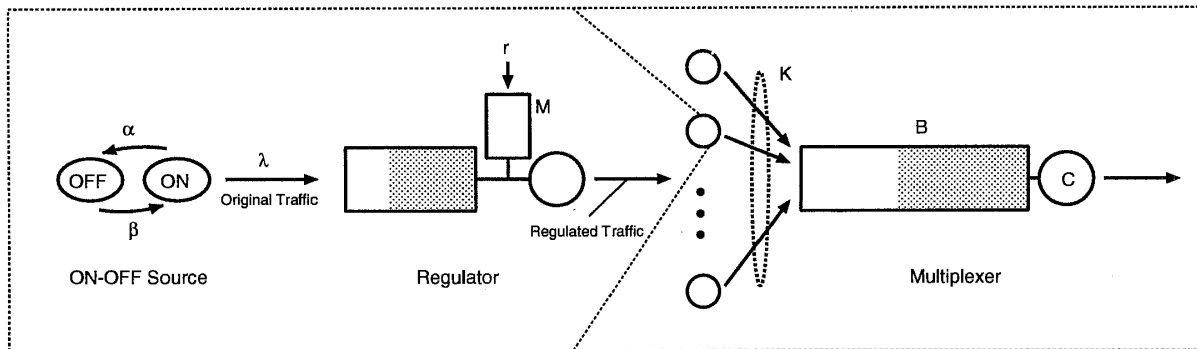


Figure 1: Fluid Model for Multiplexing System with Regulated Input Sources

rate λ when the state of the source is ON, and no traffic when OFF. We refer the fluid traffic generated by the ON-OFF source to the *original traffic*. We assume that $\lambda < \infty$.

The original traffic is regulated by the regulator called the leaky bucket. The regulator has an infinite data buffer and the token pool whose size is M . Tokens arrive at the token pool as fluid with the constant rate r . Throughout the paper, we suppose that $r < \lambda$. Arriving tokens which find the token pool full are lost. Arriving original traffic is sent to the multiplexer with rate λ until the token pool becomes empty. Note that the contents of the token pool decreases with rate $\lambda - r$. When the token pool becomes empty, the original traffic is sent to the multiplexer with the token generating rate r . We refer the output traffic from the regulator to the *regulated traffic*.

We assume that there are K independent and identical sources in our model. The regulated traffic from K sources is aggregated into the buffer of the multiplexer, whose size is B , and then transmitted with the constant rate C . Let b and c denote the buffer capacity and the transmission rate per source, respectively, i.e., $B = Kb$ and $C = Kc$.

Remark. [6] analyzed the regulator model with a finite data buffer and a finite token pool. Since we are interested in the behavior of the multiplexer's buffer with the regulated traffic, we don't need to consider the regulator model with a finite data buffer.

3. CDE Method

In this section, we restate the CDE method proposed in [5]. As for the precise proofs for the results of this section, readers are referred to section 2 and appendix in [5].

There are K identical homogeneous fluid sources in our network model. We also define $B(t)$ as the buffer content of the multiplexer at time t .

Then, it has been proved in [5] that if the process of the output rate of the regulator is Markovian and time-reversible, there exist easily calculated positive constants, C_1 and C_2 , such that for every $b > 0$,

$$(3.1) \quad \lim_{K \rightarrow \infty} \frac{1}{K} \log \Pr(B(\infty) \geq Kb) \leq -C_1 - C_2 b.$$

The lower bound of the left hand side of (3.1) is given by

$$(3.2) \quad \lim_{K \rightarrow \infty} \frac{1}{K} \log \Pr(B(\infty) \geq Kb) \geq -C_3 - C_2 b, \quad b \geq b_0,$$

where b_0 and C_3 are positive constants. Note that (3.1) and (3.2) don't depend on t , i.e., those are the equilibrium estimates.

Now we consider the following estimate of the stationary overflow probability of a buffer of size B derived from an infinite buffer analysis

$$(3.3) \quad G(B) = \Pr(B(\infty) \geq B).$$

For the output rate processes which is Markovian but not time-reversible, [5] has proposed the following approximation;

$$(3.4) \quad G(B) \approx Le^{zB},$$

where $L = e^{-KC_1}$, $z = -C_2$ and \approx means the simple approximation. The approximation in (3.4) is called the CDE method. It has been shown that L is the loss probability for bufferless multiplexing, which is estimated from the Chernoff's theorem, and that z is the dominant eigenvalue in the buffered multiplexer, which is known to determine the large buffer behavior in the logarithmic scale. In the same context of the logarithmic scale, [2] has given z in terms of the effective bandwidth. Let $h(\theta)$ denote the ALMGF of the input traffic and θ^* be the solution of $h(\theta)/\theta = C$, where the value $h(\theta)/\theta$ is called the effective bandwidth. In this context, z is given by $-\theta^*$.

Specifically, if $b \rightarrow 0$ in the upper bound (3.1), we have

$$(3.5) \quad \lim_{b \downarrow 0} \lim_{K \rightarrow \infty} \frac{1}{K} \log \Pr(B(\infty) \geq Kb) = -C_1.$$

The left hand side (LHS) of (3.5) can be obtained from the Chernoff's theorem. On the other hand, the standard arguments for Markovian traffic sources show that there is a positive constant C_2 such that, for each fixed K

$$(3.6) \quad \lim_{b \rightarrow \infty} \frac{1}{Kb} \log \Pr(B(\infty) \geq Kb) = -C_2.$$

The ALMGF plays an important role when we estimate the LHS of (3.6).

In the following two sections, L and z (i.e., θ^*) in (3.4) are separately considered.

Remark. In our model, the output rate process of the regulator is Markovian but not time-reversible. Hence (3.1) does not hold for our case. However, the numerical results in Section 6 show that the CDE method (3.4) gives the upper bound in our model.

4. Approximation of Loss Probability for Bufferless System

In this section, we consider the first part of the CDE method, the loss probability of the bufferless system. First, we analyze the stationary distribution of the virtual buffer process of the regulator. Then, we obtain the loss probability of the bufferless system using the Chernoff's theorem.

4.1. Stationary Distributions of Fluid Source with Leaky Bucket

In this subsection, we analyze the steady state analysis of the fluid queueing system with a leaky bucket whose input is an ON-OFF source [6, 7].

The original traffic is the single Markov modulated fluid source with state space $\{0, 1\}$ and infinitesimal generator Q , where 0 and 1 are corresponding to OFF and ON, respectively. Let Λ denote the 2×2 matrix whose diagonal elements represent the arrival rates. Then, we obtain

$$Q = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 0 & 0 \\ 0 & \lambda \end{pmatrix}.$$

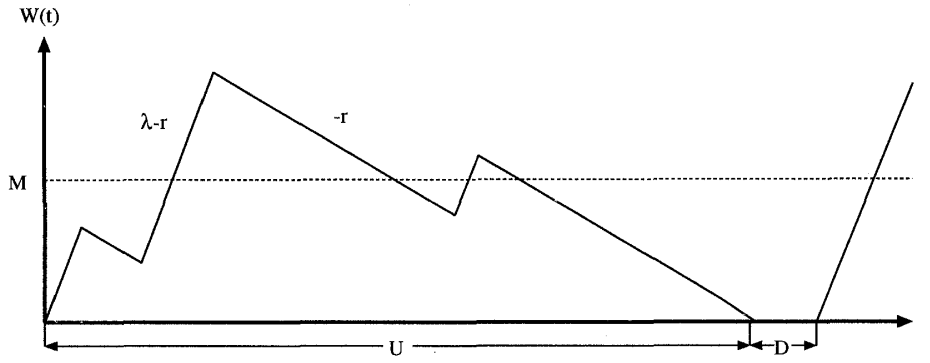


Figure 2: Sample Path of $W(t)$

Let $J(t)$ denote the state of the fluid source at time t . The source generates fluid traffic at the constant rate λ if $J(t) = 1$ and no traffic if $J(t) = 0$.

Let π denote the stationary state distribution of the ON-OFF source. Then, π satisfies the following equations

$$\pi Q = \mathbf{0}, \quad \pi \mathbf{e} = 1,$$

where $\mathbf{e} = (1, 1)^T$. We obtain

$$\pi = \left(\frac{\alpha}{\alpha + \beta}, \frac{\beta}{\alpha + \beta} \right).$$

Let $X(t)$ and $Y(t)$ denote the contents of the data buffer and the token pool at time t , respectively. We define the virtual buffer content process $W(t)$ as

$$(4.1) \quad W(t) = X(t) - Y(t) + M.$$

Note that $X(t)$ and $Y(t)$ are known from the value of $W(t)$. Fig.2 shows the typical sample path of $W(t)$. $W(t)$ increases at rate $\lambda - r$ when the source is ON. On the other hand, when the source is OFF, $W(t)$ decreases at rate r if $W(t) > 0$, or stays constant at 0 if $W(t) = 0$.

We assume that the limiting distribution of $(W(t), J(t))$ exists, and define

$$\pi_i(x) = \lim_{t \rightarrow \infty} \Pr(J(t) = i, W(t) \leq x), \quad i = 0, 1, x \geq 0.$$

Then $\pi(x) = (\pi_0(x), \pi_1(x))$ satisfies the following differential equation

$$(4.2) \quad \frac{d}{dx} \pi(x) \Gamma = \pi(x) Q.$$

where $\Gamma = \text{diag}(-r, \lambda - r)$. The i th diagonal element of Γ is the rate of the change of $W(t)$ when the source state is i . Γ is called the drift matrix [7].

It is known that the limiting distribution of $W(t)$ exists if $\pi \Gamma \mathbf{e} < 0$, that is,

$$(4.3) \quad \frac{\beta \lambda}{\alpha + \beta} < r.$$

If we define the traffic intensity ρ as

$$(4.4) \quad \rho = \frac{\beta \lambda}{(\alpha + \beta)r},$$

(4.3) means $\rho < 1$.

To solve the equation (4.2), we apply the method of the spectral expansion [6, 7, 10]. We try $\boldsymbol{\pi}(x) = \exp(\eta x)\boldsymbol{\phi}$ where η is a scalar and $\boldsymbol{\phi}$ is a two dimensional row vector. Substituting the value of $\boldsymbol{\pi}(x)$ into (4.2) yields

$$\eta e^{\eta x} \boldsymbol{\phi} \Gamma = e^{\eta x} \boldsymbol{\phi} Q,$$

and hence we obtain

$$(4.5) \quad \boldsymbol{\phi}(\eta \Gamma - Q) = \mathbf{0}.$$

A non-trivial vector $\boldsymbol{\phi}$ satisfying equation (4.5) exists if $\det(\eta \Gamma - Q) = 0$. Therefore we obtain

$$\eta = 0, \quad \frac{\beta\lambda - (\alpha + \beta)r}{r(\lambda - r)}.$$

Above values are known as the spectrum of the fluid flow model. Let η^* denote the non-zero value of η . The corresponding $\boldsymbol{\phi}$ with η^* is given by

$$\boldsymbol{\phi} = \left(\frac{\lambda - r}{r}, 1 \right).$$

It can be shown that the general solution to (4.2) is given by

$$(4.6) \quad \boldsymbol{\pi}(x) = \boldsymbol{\pi} + a e^{\eta^* x} \boldsymbol{\phi}.$$

Since $\lambda > r$, the token pool cannot be full when the source is ON and hence $\pi_1(0) = 0$. Using this condition, a in (4.6) is given by

$$a = -\frac{\beta}{\alpha + \beta}.$$

Thus, we obtain the following explicit expressions

$$\begin{aligned} \pi_0(x) &= \frac{1}{\alpha + \beta} \left\{ \alpha - \frac{\beta(\lambda - r)}{r} \exp\left(\frac{\beta\lambda - (\alpha + \beta)r}{r(\lambda - r)} \cdot x\right) \right\}, \\ \pi_1(x) &= \frac{\beta}{\alpha + \beta} \left\{ 1 - \exp\left(\frac{\beta\lambda - (\alpha + \beta)r}{r(\lambda - r)} \cdot x\right) \right\}. \end{aligned}$$

Using the above equations, the probability distribution of the virtual buffer contents in equilibrium is given by

$$(4.7) \quad \Pr(W \leq x) = \boldsymbol{\pi}(x)\mathbf{e} = 1 - \frac{\beta\lambda}{r(\alpha + \beta)} \cdot \exp\left(\frac{\beta\lambda - (\alpha + \beta)r}{r(\lambda - r)} \cdot x\right).$$

4.2. Loss Probability of Bufferless System

In this subsection, we consider the loss probability of the bufferless system using the Chernoff's theorem [11, 12]. More concretely, we consider the multiplexing system where the input fluid is the regulated traffic from K i.i.d. sources, the trunk capacity is C , and there is no buffer capacity in the system. We use c denote the trunk capacity per source, i.e., $C = Kc$.

First, we consider the rate process of the regulated traffic. Let $R(t)$ denote the rate of the regulated traffic at t . Then, $R(t)$ is given by

$$(4.8) \quad R(t) = \begin{cases} 0 & \text{if } J(t) = 0, W(t) \leq M, \\ \lambda & \text{if } J(t) = 1, W(t) \leq M, \\ r & \text{if } W(t) > M. \end{cases}$$

We assume that the following limiting distribution exists

$$\begin{aligned} \omega_0 &= \lim_{t \rightarrow \infty} \Pr(J(t) = 0, W(t) \leq M), \\ \omega_1 &= \lim_{t \rightarrow \infty} \Pr(J(t) = 1, W(t) \leq M), \\ \omega_2 &= \lim_{t \rightarrow \infty} \Pr(W(t) > M). \end{aligned}$$

Using $\pi_i(x)$, we obtain ω_i 's as

$$(4.9) \quad \omega_0 = \pi_0(M), \quad \omega_1 = \pi_1(M), \quad \omega_2 = 1 - \pi_0(M) - \pi_1(M).$$

Now we consider the i.i.d. random variables $x_1, \dots, x_K \in \{0, \lambda, r\}$, and assume that those have the common distribution $\{\omega_i; i = 0, 1, 2\}$, that is, x_i takes the value 0 w.p. ω_0 , λ w.p. ω_1 and r w.p. ω_2 . From the assumption that $r < \lambda < \infty$, the mean $E[x_1]$ exists. We define the following Legendre transform

$$(4.10) \quad \ell(c) = \sup_{\theta} \{ \theta c - \log E(e^{\theta x_1}) \}.$$

In our case, $\ell(c)$ is given by

$$(4.11) \quad \ell(c) = \sup_{\theta} \{ \theta c - \log(\omega_0 + \omega_1 e^{\lambda \theta} + \omega_2 e^{r \theta}) \}.$$

Note that $\ell(c)$ is nonnegative and convex. (For more details, see [2, 11, 12].) We consider the supremum of the right hand side (RHS) of (4.11). We define

$$(4.12) \quad g(\theta) = \theta c - \log(\omega_0 + \omega_1 e^{\lambda \theta} + \omega_2 e^{r \theta}).$$

Differentiating $g(\theta)$ with respect to θ yields

$$(4.13) \quad \frac{d}{d\theta} g(\theta) = \frac{c\omega_0 + (c - \lambda)\omega_1 e^{\lambda \theta} + (c - r)\omega_2 e^{r \theta}}{\omega_0 + \omega_1 e^{\lambda \theta} + \omega_2 e^{r \theta}}.$$

From the numerator of (4.13), if $c \leq \lambda$, there exists a point θ^* such that

$$c\omega_0 + (c - \lambda)\omega_1 e^{\lambda \theta^*} + (c - r)\omega_2 e^{r \theta^*} = 0,$$

and that $g(\theta^*)$ takes the maximum value at that point. If $c > \lambda$, then $\ell(c) = \infty$. Assuming $c \leq \lambda$, we obtain

$$(4.14) \quad \ell(c) = c\theta^* - \log(\omega_0 + \omega_1 e^{\lambda \theta^*} + \omega_2 e^{r \theta^*}).$$

Then, from the Chernoff's Theorem, we obtain the loss probability as

$$(4.15) \quad \Pr(x_1 + \dots + x_K \geq Kc) = e^{-K\ell(c) + o(K)},$$

where $o(K)$ satisfies

$$\lim_{K \rightarrow \infty} \frac{o(K)}{K} = 0.$$

Hence, we obtain the L part of the CDE method as

$$(4.16) \quad L = e^{-K\ell(c)}.$$

Remark. If $M \rightarrow \infty$ in (4.9), we obtain

$$(\omega_0, \omega_1) \rightarrow \boldsymbol{\pi}, \quad \omega_3 \rightarrow 0,$$

that is, the regulated traffic is becoming the same as the original traffic. In this case, x_i takes 0 w.p. $\alpha/(\alpha + \beta)$ and λ w.p. $\beta/(\alpha + \beta)$, and hence, x_i 's become i.i.d. Bernoulli random variables. This case has been treated in [12] as the elementary example of the large deviation theory.

5. Approximation of Loss Probability of Large Buffer System

In this section, we consider the second part of the CDE method, the loss probability of the system with the large buffer. First, we summarize the theory of the effective bandwidth and the probability of the buffer overflow. Second, we show the result of the ALMGF for the system where the original traffic process is a Markov regenerative process, which is developed by [8]. Then, we present the explicit formula of the ALMGF for the regulated traffic, which is derived in [9]. Finally, we derive the probability of the buffer overflow using those results.

Suppose that the trunk capacity is c . Let $A(t)$ denote the total amount of the regulated traffic during $[0, t]$. We define the ALMGF $h(\theta)$ as

$$(5.1) \quad h(\theta) = \lim_{t \rightarrow \infty} \frac{1}{t} \log E[e^{\theta A(t)}].$$

We are interested in the limiting tail distribution of the buffer process $B(t)$. It has been shown that this probability is such that

$$(5.2) \quad \Pr(B(\infty) \geq b) \sim e^{-\theta^* b}, \quad \text{as } b \rightarrow \infty,$$

where $f(x) \sim g(x)$ means that $f(x)/g(x) \rightarrow 1$ as $x \rightarrow \infty$. In (5.2) θ^* is given by the solution of $h(\theta)/\theta = c$. As for the rigorous proofs and the further discussions, readers are referred to [2, 3, 4, 11].

The function

$$(5.3) \quad a^*(\theta) = \frac{h(\theta)}{\theta}$$

is called the effective bandwidth function of the arrival process subject to the condition that the tail distribution of the buffer contents has the decay rate θ .

5.1. ALMGF for Output of Regulator

[8] showed that the ALMGF can be computed if the considered process is a Markov regenerative process. In [8], the ALMGF of ON-OFF source is explicitly derived where both ON and OFF periods are exponentially distributed. Let $h_{\text{org}}(\theta)$ denote the ALMGF of the original traffic, which is generated by an ON-OFF source. $h_{\text{org}}(\theta)$ is given as

$$(5.4) \quad h_{\text{org}}(\theta) = \frac{1}{2} \left(\lambda\theta - \alpha - \beta + \sqrt{(\lambda\theta - \alpha - \beta)^2 + 4\beta\lambda\theta} \right).$$

For further details, the readers are referred to [8].

In [9], the ALMGF of the output from the regulator for an ON-OFF source has been derived. The output process from the regulator with leaky bucket is modulated by the bivariate process $\{(W(t), J(t)); t \geq 0\}$ according to (4.8). Let $W(0) = 0$ and $J(0) = 1$ (see Fig.2). Then, $\{(W(t), J(t)); t \geq 0\}$ is a regenerative process that regenerates whenever it reaches the state $(0, 1)$. We define ON period U as

$$U = \inf\{t > 0 : W(t) = 0 \mid W(0) = 0, J(t) = 1\}.$$

Then the cycle time of the regenerative process is $U + D$ where D is the OFF period and exponentially distributed with rate β (see Fig. 2).

Since the token pool is full at time 0 and U , the total amount of tokens removed from the pool over $[0, U]$ is equal to the amount of tokens entering the pool during $[0, U]$. This means that the total output during U is given by rU . On the other hand, the total output during D is zero. Hence, the total output during the first regenerative cycle is rU . Thus, the output from the regulator with the leaky bucket can be considered as an ON-OFF source with U as ON time, D as OFF time and r as the arrival rate at ON time.

As for the U , which is identical with the busy period of this system, [10] analyzed the first passage times in fluid models and they derived the explicit expression of $U^*(s)$, the LST of U for the ON-OFF source. From the results of [10], we obtain

$$(5.5) \quad U^*(s) = \begin{cases} \frac{s + \beta + r\kappa(s)}{\beta} & \text{if } \xi^2 + 4s(s + \alpha + \beta)r(\lambda - r) \geq 0, \\ \infty & \text{otherwise,} \end{cases}$$

where

$$\kappa(s) = \frac{-\xi - \sqrt{\xi^2 + 4s(s + \alpha + \beta)r(\lambda - r)}}{2r(\lambda - r)},$$

and

$$\xi = (\lambda - 2r)s + (\lambda - r)\beta - r\alpha.$$

Let $h_{\text{reg}}(\theta)$ denote the ALMGF of the output from the regulator. Using the expression of $U^*(s)$ and $h_{\text{org}}(\theta)$, we obtain

$$(5.6) \quad h_{\text{reg}}(\theta) = \begin{cases} h_{\text{org}}(\theta) & \text{if } 0 < \theta \leq \zeta, \\ h_{\text{org}}(\zeta) - r\zeta + r\theta & \text{if } \theta > \zeta, \end{cases}$$

where

$$(5.7) \quad \zeta = \frac{\beta}{\lambda} \left(\sqrt{\frac{r\alpha}{\beta(\lambda - r)}} - 1 \right) + \frac{\alpha}{\lambda} \left(1 - \sqrt{\frac{\beta(\lambda - r)}{r\alpha}} \right).$$

Remarks. 1. As we stated, the ALMGF can be computed if the considered process is the Markov regenerative process. However, the output process of the regulator with Markov regenerative input is not regenerative in general. The ON-OFF process considered here is the exceptional case under which we can obtain the ALMGF of the regulated traffic explicitly.

2. The ALMGF of the output of the regulator with leaky bucket is identical to that of the output from the fluid queue with output rate r . From (5.6), we can see that the ALMGF of the output of the regulator is independent of M , the size of token pool. As [9] mentioned, this means that M does not play any role as far as reducing the effective bandwidth, but it acts strictly as a policing device that prevents arbitrarily large peak-rate bursts from entering the network.

5.2. Loss Probability of the System with K Input Sources

In this section, we consider the loss probability of the buffer with the K independent input sources regulated by leaky bucket.

Since we assume that K sources are independent, from [2], we can derive the ALMGF of the total amount of K sources as the sum of K ALMGFs, i.e., $h(\theta) = K \cdot h_{\text{reg}}(\theta)$.

Now we solve $h(\theta)/\theta = C$ with respect to θ . Since $C = Kc$, we solve $h_{\text{reg}}(\theta)/\theta = c$, i.e.,

$$\begin{aligned} h_{\text{org}}(\theta) - c\theta &= 0, \quad 0 < \theta \leq \zeta, \\ h_{\text{org}}(\zeta) - r\zeta + r\theta - c\theta &= 0, \quad \theta > \zeta. \end{aligned}$$

From (5.4), (5.7) and the assumptions in Section 4 that $\beta\lambda/(\alpha + \beta) < r < \lambda$ and $c < \lambda$, we obtain

$$(5.8) \quad \theta^* = \begin{cases} \frac{\beta\lambda - c(\alpha + \beta)}{c(c - \lambda)}, & \frac{\beta\lambda}{\alpha + \beta} < c \leq c^*, \\ \frac{(\alpha - \beta)r + \beta\lambda - 2\sqrt{\alpha\beta r(\lambda - r)}}{\lambda(r - c)}, & c^* < c < r, \end{cases}$$

where c^* is the value of the effective bandwidth at $\theta = \zeta$, i.e.,

$$(5.9) \quad c^* = a^*(\zeta) = \lambda \frac{\beta r - \sqrt{\alpha\beta r(\lambda - r)}}{(\alpha + \beta)r - \alpha\lambda}.$$

Using θ^* in (5.8), we can compute the second part of the CDE method, the loss probability of the system with the large buffer using the following

$$(5.10) \quad e^{zB} = \Pr(B(\infty) \geq Kb) \approx e^{-\theta^* Kb}.$$

Finally, we obtain the following expression of $G(B)$, the loss probability for the multiplexing system with the large number of sources and the large buffer as

$$(5.11) \quad G(B) \approx Le^{zB} = e^{-K\{\ell(c) + \theta^* b\}}.$$

Remarks. 1. As we noted in subsection 5.1, the ALMGF of the regulated traffic is independent of M and hence it is hard to investigate the loss probability with the leaky bucket parameter in the context of the standard effective bandwidth approach. Though (5.11) is the expression approximated by the CDE method, (5.11) includes the M and r , the parameters of the leaky bucket. Hence, this result can be considered as useful for designing the input source regulator with the leaky bucket.

2. In Fig. 3, we show the r - c region where θ^* exists. When (r, c) is in the region (1), θ^* is given by the first formula in (5.8). When (r, c) is in the region (2), θ^* is given by the second formula.

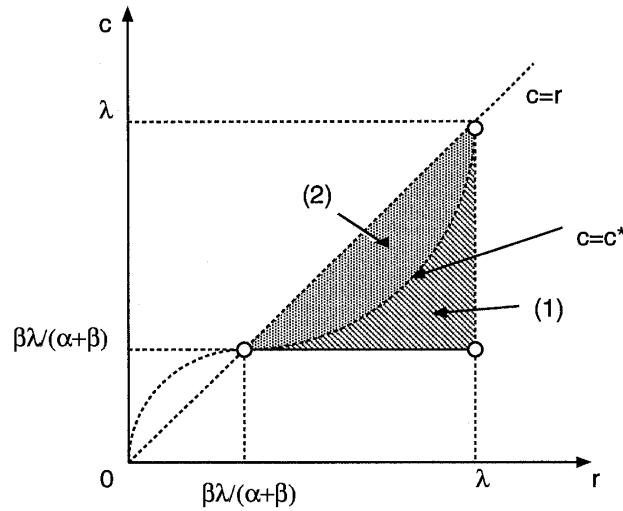
6. Numerical Examples

In this section, we show some numerical examples of the approximation by the CDE method. We set the parameters of the ON-OFF source as follows;

$$\alpha = 1, \quad \beta = 0.35/0.65.$$

We set the parameters of the multiplexer as follows;

$$K = 10, \quad b = 10.$$


 Figure 3: r - c Region Where θ^* Exists.

Changing the parameters of the regulator (M, r) where r satisfies the conditions of (5.8) and $\beta\lambda/(\alpha + \beta) < r < \lambda$ (see Fig. 3), we compute the logarithm of the loss probability of (5.11). We compare the loss probabilities calculated by the CDE method with e^{zB} , the approximation by the standard effective bandwidth (SEB) method.

Tables 1 and 2 show the logarithms of the loss probabilities under SEB and CDE methods changing the token generation rate r . In these tables, we set $\rho = 0.6$ and computed λ from (4.4) changing r . In Table 1, we set $c = 0.5$ and in Table 2, $c = 0.6$. As for the CDE method, we calculated the loss probabilities under $M = 0$ and 1.

 Table 1: Logarithms of Loss Probability: $c = 0.5$

r	SEB	CDE($M = 0.0$)	CDE($M = 1.0$)
0.76	-16.8630	-16.9342	-16.8941
0.77	-14.2590	-14.3106	-14.2818
0.78	-11.7618	-11.7973	-11.7776
0.79	-9.3647	-9.3875	-9.3749
0.80	-7.0619	-7.0750	-7.0679
0.81	-4.8480	-4.8543	-4.8509
0.82	-2.7179	-2.7199	-2.7188
0.83	-0.6669	-0.6671	-0.6670

In both tables, we observe that the loss probability becomes large as r getting large under both SEB and CDE methods. We can see that the value of r significantly affects the loss probability. This considerable variation is due to the number of sources, K . Even when the variation of r for each source is small, the aggregated traffic largely affects the buffer behavior of the multiplexer with the contribution of the K th power order (see (5.11)).

Table 2: Logarithms of Loss Probability: $c = 0.6$

r	SEB	CDE($M = 0.0$)	CDE($M = 1.0$)
0.92	-12.5957	-12.6534	-12.6226
0.93	-10.8312	-10.8743	-10.8515
0.94	-9.1266	-9.1575	-9.1412
0.95	-7.4787	-7.4997	-7.4887
0.96	-5.8849	-5.8981	-5.8912
0.97	-4.3425	-4.3498	-4.3460
0.98	-2.8491	-2.8522	-2.8506
0.99	-1.4023	-1.4031	-1.4027

We also observe that the CDE method gives lower values than SEB. In addition, we observe that the loss probability under $M = 0.0$ is smaller than that under $M = 1.0$. However, M does not improve the loss probability as well as r does.

Table 3: Loss Probability: Approximations and Simulation

r	SEB	M=0.0		M=1.0	
		CDE	Simulation	CDE	Simulation
0.80	8.571×10^{-4}	8.459×10^{-4}	$(5.716 \pm 0.410) \times 10^{-4}$	8.520×10^{-4}	$(6.085 \pm 0.350) \times 10^{-4}$
0.81	7.844×10^{-3}	7.794×10^{-3}	$(6.178 \pm 0.244) \times 10^{-3}$	7.821×10^{-3}	$(6.429 \pm 0.227) \times 10^{-3}$
0.82	6.601×10^{-2}	6.588×10^{-2}	$(5.826 \pm 0.133) \times 10^{-2}$	6.595×10^{-2}	$(5.962 \pm 0.113) \times 10^{-2}$

To validate the accuracy of the CDE method, the simulation was performed. In the simulation, we set $\rho = 0.6$ and $c = 0.5$. In Table 3, we compare the loss probabilities of SEB and CDE with the simulation results. From this table, the loss probability calculated by CDE is similar to that of SEB, while both probabilities are larger than the simulation result. This is because both b and K are not sufficiently large for applying SEB and CDE methods. However, we can observe that the loss probability of CDE becomes large when M is getting large, as is the case of the simulation. Hence the CDE method can capture the variation due to M more likely than SEB. Note that the CDE method gives the upper bounds of the loss probability in this table.

Fig. 4 shows the logarithms of the loss probabilities under SEB and CDE methods changing the token pool size M . In this figure, we set $\rho = 0.7$, $r = 0.68$ and $c = 0.5$. Since the SEB method does not include M , the loss probability under the SEB method takes the constant value. We observe that the loss probability becomes large as M getting large, and that the loss probability is increasing gradually when $M > 0.5$.

The token pool acts strictly as a policing device that prevents arbitrarily large peak-rate bursts from entering the network. Thus, setting $M \leq c$ prevents the system from instantaneous increasing of the buffer contents of the multiplexer. Since $c = 0.5$, M improves the loss probability when $M < 0.5$. On the other hand, M larger than c does not improve the loss probability.

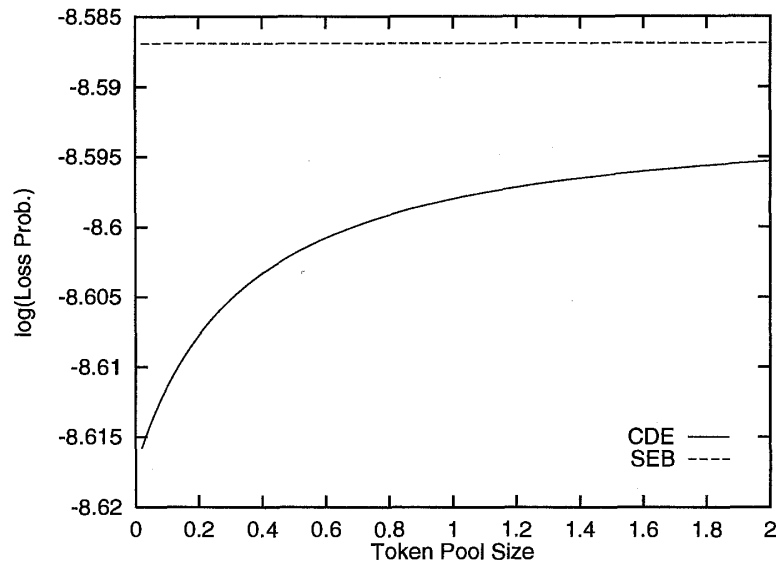


Figure 4: Logarithms of Loss Probability versus Token Pool Size

As we stated, r drastically improve the loss probability, while M does not improve the loss probability so much. From the view point of the improvement of the loss probability, adjusting the value of r is more effective than changing M .

7. Conclusion

In this paper, we considered the loss probability of the multiplexing system with the large buffer and the large number of homogeneous ON-OFF sources. Using the CDE method, we derived the approximation of the buffer overflow probability. From the numerical results, we observed that r drastically improve the loss probability, while M does not improve the loss probability so much.

In our analysis, we assumed that the input sources are homogeneous and that the process of the output rate of the regulator is Markovian for applying the CDE method. Although those assumptions should be relaxed, our results derived in this paper provides one of the significant indications for designing the input source regulator with the leaky bucket.

References

- [1] C. S. Chang, "Stability, Queue Length, and Delay of Deterministic and Stochastic Queueing Networks," *IEEE Trans. Automa. Contr.*, vol. 39, no. 5, pp. 913–931, 1994.
- [2] C. S. Chang and J. A. Thomas, "Effective Bandwidth in High-Speed Digital Networks," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 1091–1100, 1995.
- [3] D. D. Botvich and N. G. Duffield, "Large Deviations, the Shape of the Loss Curve, and Economies of Scale in Large Multiplexers," *Queueing Systems*, vol. 20, pp. 293–320, 1995.
- [4] P. W. Glynn and W. Whitt, "Logarithmic Asymptotics for Steady-state Tail Probabilities in a Single-server Queue," In *Studies in Applied Probability*, J. Gambos and J. Gani, eds., *J. Appl. Prob., Special vol. 31A*, 131-159, 1994.

- [5] A. I. Elwalid, D. Heyman, T. V. Lakshman, D. Mitra and A. Weiss, "Fundamental Bounds and Approximations for ATM Multiplexers with Applications to Video Teleconferencing," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 1004–1016, 1995.
- [6] A. I. Elwalid and D. Mitra, "Analysis and Design of Rate-based Congestion Control of High Speed Networks, I: Stochastic Fluid Models, Access Regulation," *Queueing Systems*, vol. 9, pp. 29–64, 1991.
- [7] A. I. Elwalid and D. Mitra, "Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks," *IEEE/ACM Trans. Networking*, vol. 1, no. 3, pp. 329–343, 1993.
- [8] V. G. Kulkarni, "Effective Bandwidths for Markov Regenerative Sources," *Queueing Systems*, vol. 24, pp. 137–153, 1996.
- [9] V. G. Kulkarni and N. Gautam, "Leaky Buckets: Sizing and Admission Control," Technical Report no. 95–10, Dept. of Operations Research, University of North Carolina, Chapel Hill, NC 27599, 1996.
- [10] A. Narayanan and V. G. Kulkarni, "First Passage Times in Fluid Models with an Application to Two Priority Fluid Systems," Technical Report no. 95–01, Dept. of Operations Research, University of North Carolina, Chapel Hill, NC 27599, 1995.
- [11] A. Shwartz and A. Weiss, *Large Deviations for Performance Analysis*, New York Chapman Hall, 1995.
- [12] A. Weiss, "An Introduction to Large Deviations for Communication Networks," *IEEE J. Select. Areas Commun.*, vol. 13, no. 6, pp. 938–952, 1995.

Shoji Kasahara †, ‡

†Graduate School of Information Science
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-01, Japan
kasahara@is.aist-nara.ac.jp

‡Telecommunications Advancement Organization of Japan (TAO)
Shiba 2-31-19, Minato-Ku, Tokyo, 105 Japan

Toshiharu Hasegawa

Department of Applied Systems Science
Graduate School of Engineering
Kyoto University
Kyoto 606-01, Japan
hasegawa@kuamp.kyoto-u.ac.jp