Asymptotic analysis for blocking probabilities of optical buffer with general packet-length distributions

Yasuji Murakami

The department of Telecommunications and Computer Networks, Osaka Electro-Communication University, Osaka, Japan

Email address:
mura@isc.osakac.ac.jp

Abstract: Asynchronous optical packet switching seems to be suitable as a transport technology for the next-generation Internet due to the variable lengths of IP packets. Optical buffers in the output port are an integral part for solving contention by exploiting the time domain. Fiber delay lines (FDLs) are a well-known technique for achieving optical buffers. This work aims to give a highly accurate approximation of the blocking probabilities of the optical buffers for a generally distributed packet length even when the offered load is extremely low. Such a tool is needed for investigating and designing realistic optical packet switches, which will be used for low-offered-load and low-packet-loss optical IP networks. We use the asymptotic expansion for the decay rate, resulting in a highly accurate approximation. By using the fourth order approximation of the decay rate, an accuracy within 10 % was obtained for both the exponential and uniform distribution cases of an offered load greater than 0.3. The approximations established in this work can be applied to investigate multiclass optical buffers for priority queueing.

Keywords: Asynchronous Optical Switching, Optical Buffers, Blocking Probabilities, General Packet-Length Distributions

1. Introduction

The ultimate capacity of the Internet may be constrained by energy density limitations and heat dissipation considerations rather than by the bandwidth of the physical components [1]. Recent studies show that optical packet switches do not appear to offer significant throughput improvements or energy savings compared to electronic packet switches [2]. However, the studies also show that optical switch fabrics generally become more energy efficient as the data rate increases [2]. We believe that development of the optical components could lead to a breakthrough in the optical packet switches, and use of all-optical packet switches, in which optical packets are buffered and routed in optical form, will solve the present Internet problems. Asynchronous optical packet switching appears to be suitable as a transport technology for the next-generation Internet due to the variable lengths of IP packets.

In optical packet switches, optical buffers in the output port are an integral part for solving contention by exploiting the time domain. Fiber delay lines (FDLs) are a well-known technique for achieving optical buffers because random access

memory (RAM) is not achievable with current optical technologies. However, optical buffers behave differently from electronic RAM. FDLs can only delay the packets for multiples of a discrete amount of time, called time granularity; the maximum delay is bounded and a packet will be discarded if the maximum delay is not sufficient to avoid contention.

Various analytical methods have been proposed to evaluate the performance of an asynchronous optical buffer [3-10]. One earlier work [3] investigates an optical buffer with Poisson arrivals and exponential packet lengths and presents an approximation using an iterative procedure. This exponential assumption has been extended to obtain an approximation of blocking probability [5]. Exact solutions in a Laplace-transformed expression for the blocking probability and delay have been given [6], where general packet-length distributions can be applied. However, these solutions are limited in the Laplace transformation form; numerical results, which are obtained by computer simulations, are not given by solving the inverse Laplace transformation. Exact solutions have also been obtained using the stochastic method of multi-regime Markov fluid queues [7], which provides a numerical algorithm and not a closed-form expression. In other research, closed-form expressions for certain optical

buffers, a numerical algorithm for general cases [8], and a heuristic approach, allowing for simple performance analysis [10], have been provided. However, numerical results only for exponentially distributed packet length were shown.

In our recent work [9], we established a close-form approximate expression for the blocking probabilities and the average delays of the optical buffer, where general packet-length distributions can be applied. For both the uniform and deterministic distributions, the accuracy was high enough in an offered load of greater than 0.7, but decreased when the offered load decreased.

In ITU-T Y.1541 [11], the objective of the end-to-end IP packet loss ratio (IPLR) is recommended to be less than 10^{-3}. Assuming that optical packets pass through a maximum of 100 switches for end-to-end network transmission, the IPLR required for one packet switch is less than 10^{-5}. Packet loss cannot be avoided for the finite FDL optical buffer, and it seems that an IPLR of less than 10^{-5} is quite difficult to achieve, except in a low offered load.

Our main purpose in this work is to give a highly accurate approximation in closed-form expressions for a generally distributed packet length, even when the offered load is extremely low. Such a tool is needed for investigating and designing realistic optical packet switches, which will be applied for ultra-high-capacity optical IP networks with low offered load and low packet loss. In other words, the present approximation aims at providing an engineering tool for real optical IP network design.

The probability $P(x > w) \equiv W(x)$ that an arriving customer has the waiting time w less than x can be approximated to be the exponential form $W_A(x) = 1 - Ce^{-\eta x}$, where η is a decay rate and C is a constant. In this work, we use the asymptotic expansion for the decay rate, and can then obtain a highly accurate approximation even when the offered load is extremely low.

Section 2 briefly reviews the theorems on which our analysis is based [6, 9]. In Section 3, we introduce an asymptotic analysis for the decay rate of the tail probabilities in an infinite optical buffer. The fourth order approximation of the blocking probabilities is given for the generally distributed packet length. In Section 4, we show numerical results and compare them with computer simulation results and exact calculations. Finally, Section 5 concludes this paper.
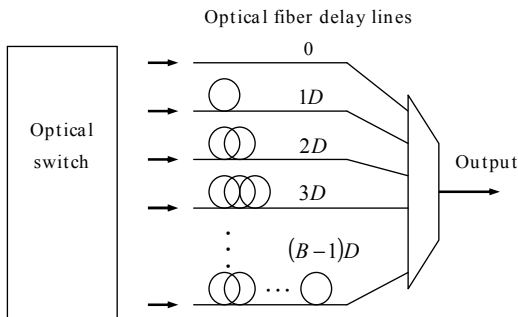

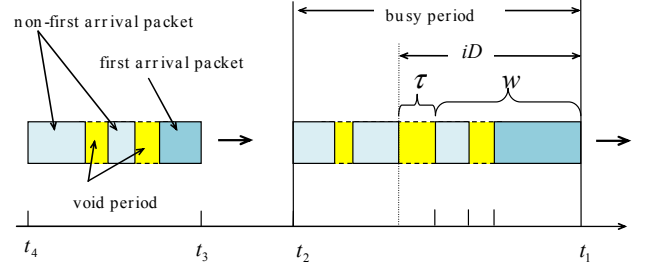**Fig. 1.** *Basic structure of optical buffer with fiber delay lines.*

Fig. 2. *An example of optical packet flow from optical buffer*

2. Optical Buffer Model and Theorems

Our approximation for blocking probabilities is based on the theorems derived by Liu et al. [6] and Murakami [9], so we first briefly review these theorems and notations.

2.1. Model

The basic structure of optical buffers, composed of FDLs, is shown in Fig. 1. Arrived packets can be delayed for a discrete time of $(i-1)D, 1 \le i \le B$, where D is the time granularity of FDLs and B is the buffer length, to avoid contention at the single output. The maximum delay is $T = (B-1)D$; packets with a delay time longer than T are discarded. All packets are served under the first-come first-served (FCFS) discipline in the optical buffer.

When a newly arrived packet has to wait for at least w units of time to be served, the packet is treated as follows.

1) If $(i-1)D \le w < iD$, the packet is switched to the $(i+1)$th FDL, and a void period τ is attached to the head of the packet, where

$$\tau = iD - w, \quad \left\lceil \frac{w}{D} \right\rceil = iD, \quad (1)$$

and $\lceil x \rceil$ means the smallest integer greater than x.

2) If $T = (B-1)D < w$, the packet is discarded.

In the case of an FCFS system with Poisson arrivals, the waiting time (that a fictitious packet, which arrived at an arbitrary point of time, has to wait) has the same statistics as the actual waiting time (that each packet spent in the system). Then, we deal with the distribution function of the "virtual" waiting time x instead of the actual waiting time. If a packet arrives, it would have to wait for a duration x to get service. The service times of buffered packets are prolonged by a void period τ, which introduces excess load compared to the real load. When the buffer is empty and the destination output is free, the packets are transmitted immediately and no void period is attached. The term "PACKET" is defined to indicate the effective service time. There are two kinds of PACKETs; *first-arrival packet*s, which have no void period and where service times are the real packet lengths, and *non-first-arrival packet*s, which arrive when the buffer is not empty and whose service times are prolonged by the void periods.

Figure 2 shows an example of the optical packet flow from the output of the optical buffer. In the figure, the packets flow from the left to the right. The head of the packet flow is at the

right. The output packet at t_1 belongs to *first-arrival packet*s and has no void period. The third packet from the head required waiting time w to be served when it arrived at the buffer, and the void period τ, given by (1), was attached to its head. The buffer becomes empty when the fourth packet flows from the output end at t_2. The busy period which consists of the four packets on the right side of the figure, therefore becomes the time from t_1 to t_2. At t_3, a *first-arrival packet* newly arrives and a new busy period starts.

2.2. Theorems in Infinite Optical Buffer Theorems in Infinite Optical Buffer

We first describe the case of the infinite buffer, i.e., $B \to \infty$, where no packet is lost. The following notations are used for the system.

λ Poisson arrival rate

s_0 real packet length

$g_0(x)$ probability density function (pdf) of packet length s_0

$G_0(x)$ cumulative distribution function (CDF) of packet length s_0

$\bar{s}_0$ mean value of s_0, so $\bar{s}_0 = \int_0^\infty x g_0(x) dx$

ρ offered load, i.e., $\rho = \lambda \bar{s}_0$

$l(x)$ pdf of void period τ

s_ν service time of *non-first-arrival packet*s, i.e., $s_\nu = s_0 + \tau$

$g(x)$ pdf of the service time s_ν, so

$$g(x) = g_0(x) \otimes l(x) \equiv \int_{-\infty}^\infty g_0(x-y) l(y) dy, \qquad (2)$$

where $\otimes$ shows convolution integral

$\bar{s}_\nu$ mean value of s_ν, so

$$\bar{s}_\nu = \int_0^\infty x g(x) dx = \bar{s}_0 + \bar{\tau} = \bar{s}_0 + \frac{D}{2} \qquad (3)$$

$v(x)$ steady state pdf of virtual waiting time x of PACKET

$V(x)$ steady state CDF of virtual waiting time x of PACKET

Q steady probability of system being empty

ρ_{eq} equivalent load of PACKET, i.e., $\rho_{eq} = 1 - Q$

The following equations were obtained for steady-state system conditions [6].

The equivalent load is given by

$$\rho_{eq} = \frac{\rho}{1 - \dfrac{D}{2\bar{s}_0} \rho}. \qquad (4)$$

When the system is in a stable condition, i.e., $\rho_{eq} < 1$, the Laplace transformation of $v(x)$ is given by

$$v^*(\theta) = \frac{\lambda Q[1 - g_0^*(\theta)]}{\theta - \lambda[1 - g^*(\theta)]}, \qquad (5)$$

where $g_0^*(\theta)$ and $g^*(\theta)$ are the Laplace transformations of $g_0(x)$ and $g(x)$, respectively. Equation (5) corresponds to the Pollaczek-Khinchin transform equation in the M/G/1 system [12]. If the denominator will be some rational function of θ, finding a particular factored form is desirable.

2.3. Theorems in Finite Optical Buffer

In the finite buffer, if the packet requires a virtual waiting time x longer than the maximum allowable delay $T = (B-1)D$, the packet is discarded. There are three types of PACKETs:

1) *first-arrival packet*s, $x = 0$ and are admitted past the buffer
2) *non-first-arrival packet*s and admitted packets, $0 < x \le T$
3) *non-first-arrival packet*s and discarded packets, $T < x$.

The notations used to describe the finite buffer system follow.

$\bar{w}_T$ mean packet delay of admitted packets without a void period, which is the average of delays of both type 1) and 2) PACKETs

P_B packet blocking probability

When $\rho_{eq} < 1$, the mean packet delay is given by

$$\bar{w}_T = T - \int_0^T \frac{V(x)}{V(T)} dx + \frac{D}{2}\left[1 - \frac{Q}{V(T)}\right], \qquad (6)$$

where the third term on the right-hand side is caused by the void periods of the admitted PACKETs, i.e., type 2) PACKETs, whose service times are prolonged.

When $\rho_{eq} < 1$, the packet blocking probability is given by

$$P_B = \frac{(1 - \rho')[1 - V(T)]}{1 - \rho'[1 - V(T)]}, \qquad (7)$$

where the *quasi-load* ρ' was introduced [9] as

$$\rho' \equiv \lambda \bar{s}_\nu = \rho\left(1 + \frac{D}{2\bar{s}_0}\right), \qquad (8)$$

and all the variables are associated with the infinite buffer system.

3. Our Asymptotic Analysis

The tail probability $1 - V(T)$ in (7) is obtained from the solution of the inverse Laplace transformation of (5). However,

solving the inverse Laplace transformations analytically is not easy except for specific functions. The probability $P(x > w) \equiv W(x)$, that an arriving customer has the virtual waiting time w less than x, can be approximated as the exponential function

$$W_A(x) = 1 - Ce^{-\eta x}, \tag{9}$$

where η is the decay rate and C the proportional constant [13-15]. Equation (9) holds for heavy traffic, i.e., traffic load ρ nears 1, and suitably large x in the M/G/1 system. If

$$W(x) \cong W_A(x) \tag{10}$$

and we use

$$W_A(0) = 1 - C = 1 - \rho, \tag{11}$$

we can obtain parameters such as

$$C = \rho \quad \text{and} \quad \eta = \rho/\overline{w}. \tag{12}$$

In our recent work [9], furthermore, we approximated the inverse Laplace transformation of $v^*(\theta)$ given by (5) as

$$V(x) = 1 - \rho_{eq} \exp\left[-\frac{2(1-\rho')}{1+C_g{}^2} \frac{x}{\overline{s}_v}\right], \tag{13}$$

where $C_g{}^2$ is the variance of the service time of the *non-first-arrival packet*s, which is normalized by the square of the mean service time $\overline{s}_v$; that is,

$$C_g{}^2 = \left(\left\langle s_v{}^2 \right\rangle - \overline{s}_v{}^2\right)\big/\overline{s}_v{}^2. \tag{14}$$

By the approximation of (13), we could obtain numerical solutions with high accuracy for the blocking probabilities and the average delays when the offered load was greater than 0.7.

3.1. Asymptotic Expansion for Decay Rate η

The method of heavy-traffic asymptotic analysis in solving the inverse Laplace transformation for the probability function, such as $v^*(\theta)$ of (5), has been already studied in [13, 14]. The decay rate η was expanded by a power series of $(1-\rho)$, and each coefficient of $(1-\rho)^n$ $(n = 0,1,2,\cdots)$ was calculated step by step. The first-order approximation in the expansion by $(1-\rho)$ is sufficient for heavy-traffic cases of $\alpha \to 1$ limit. When ρ is low, i.e., $(1-\rho)$ is large enough, however, the higher-order terms are needed to obtain highly-accurate calculations. If the denominator of (5) can be factorized by θ, the decay rate η is given by a negative root of denominator $= 0$. Then, we start to solve the negative root of $\theta - \lambda[1 - g^*(\theta)] = 0$ by asymptotically expanding $-\eta - \lambda[1 - g^*(-\eta)] = 0$ in powers of $(1-\rho)$.

The Laplace transformation $g^*(\theta)$ of the pdf for the service time of the *non-first-arrival packet*s can be expanded by

$$g^*(-\mu) = \int_0^\infty e^{\eta x} g(x)dx$$

$$= 1 + \frac{M_1}{1!}\eta + \frac{M_2}{2!}\eta^2 + \cdots + \frac{M_n}{n!}\eta^n + \cdots, \tag{15}$$

where M_n is the n th moment of $g(x)$, defined by

$$M_n \equiv \int_0^\infty x^n g(x)dx. \tag{16}$$

Equation (15) is called the moment-generating function of $g(x)$. For example, the first moment is the mean service time of the *non-first-arrival packet*s, such as

$$M_1 = \overline{s}_v = \overline{s}_0 + \frac{D}{2}, \tag{17}$$

and the second moment is given by

$$M_2 = \overline{s}_v{}^2\left(1 + C_g{}^2\right). \tag{18}$$

Inserting (15)~(17) into $-\eta - \lambda[1 - g^*(-\eta)] = 0$, we can obtain

$$1 - \rho' = \lambda\left[\frac{M_2}{2!}\eta + \frac{M_3}{3!}\eta^2 + \cdots + \frac{M_{n+1}}{(n+1)!}\eta^n + \cdots\right]$$

$$= \rho\left[\frac{M_2}{2!\overline{s}_v}\eta + \frac{M_3}{3!\overline{s}_v}\eta^2 + \cdots + \frac{M_{n+1}}{(n+1)!\overline{s}_v}\eta^n + \cdots\right], \tag{19}$$

where we use $\rho' = \lambda\overline{s}_v$. Equation (19) indicates that the asymptotic decay rate is expressed in powers of $(1-\rho')$ instead of $(1-\rho)$.

For expanding (19), we set

$$A \equiv \frac{M_2}{2!\overline{s}_v}\eta + \frac{M_3}{3!\overline{s}_v}\eta^2 + \cdots + \frac{M_{n+1}}{(n+1)!\overline{s}_v}\eta^n + \cdots$$

$$= b_1\eta + b_2\eta^2 + \cdots + b_n\eta^n +, \tag{20}$$

where

$$b_n = \frac{M_{n+1}}{(n+1)!\overline{s}_v}. \tag{21}$$

Assuming the conditions of $\eta, A < 1$, we can expand (19) in powers of A and obtain the equation in powers of η, i.e.,

$$1 - \rho' = A - A^2 + A^3 - A^4 + \cdots$$

$$= b_1\eta + \left(b_2 - b_1{}^2\right)\eta^2 + \left(b_3 - 2b_1b_2 + b_1{}^3\right)\eta^3$$

$$+ \left[b_4 - \left(2b_1b_3 + b_2{}^2\right) + 3b_1{}^2b_2 - b_1{}^4\right]\eta^4 + O\left(\eta^5\right), \tag{22}$$

where we show up to the fourth term of η.

In (22), $1-\rho'$ is given by the polynomial of η. According to the reverse transformation formula [16], η is also expressed by the polynomial of $1-\rho'$, i.e.,

$$\eta = a_1(1-\rho') + a_2(1-\rho')^2 + a_3(1-\rho')^3$$
$$+ a_4(1-\rho')^4 + O((1-\rho')^5), \qquad (23)$$

where each coefficient is given by

$$a_1 = \frac{1}{b_1} = \frac{2\bar{s}_v}{M_2} = \frac{2}{\left(1+C_g^2\right)}\frac{1}{\bar{s}_v}, \qquad (24)$$

$$a_2 = -\frac{b_2 - b_1^2}{b_1^3}, \qquad (25)$$

$$a_3 = \frac{2\left(b_2 - b_1^2\right)^2 - b_1\left(b_3 - 2b_1b_2 + b_1^3\right)}{b_1^5} \qquad (26)$$

$$a_4 = \frac{1}{b_1^7}\left[5b_1\left(b_2 - b_1^2\right)\left(b_3 - 2b_1b_2 + b_1^3\right)\right.$$
$$\left. - b_1^2\left(b_4 - b_2^2 - 2b_1b_3 + 3b_1^2b_2 - b_1^4\right) - 5\left(b_2 - b_1^2\right)^3\right] \quad (27)$$

Hereafter, we call the asymptotic series in which greater-than-n th terms are truncated, as in (23), the n th order approximation. Note that the first order approximation of (23) is used in (13)

3.2. Asymptotic Expansion for Proportional Constant C

Table 1. *Moments for service time distributions for non-first arrival packets*

moments	Exponential	Deterministic	Uniform
M_1	$1+\dfrac{D}{2}$	$1+\dfrac{D}{2}$	$1+\dfrac{D}{2}$
M_2	$2+D+\dfrac{D^2}{3}$	$1+D+\dfrac{D^2}{3}$	$\dfrac{4+3D+D^2}{3}$
M_3	$6+3D+D^2+\dfrac{D^3}{4}$	$\dfrac{(1+D)^4-1}{4D}$	$2+2D+D^2+\dfrac{D^3}{4}$
M_4	$24+12D+4D^2+D^3+\dfrac{D^4}{5}$	$\dfrac{(1+D)^5-1}{5D}$	$\dfrac{16}{5}+4D+\dfrac{8D^2}{3}+D^3+\dfrac{D^4}{5}$
M_5	$\begin{array}{c}120+60D+20D^2\\ +5D^3+D^4+\dfrac{D^4}{6}\end{array}$	$\dfrac{(1+D)^6-1}{6D}$	$\dfrac{(2+D)(4+2D+D^2)^2}{6}$

Two methods to obtain the approximation of C have been considered: the $x \to 0$ approximation and $x \to \infty$ approximation. In our previous paper, we used

$$C(x \to 0) = \rho_{eq}, \qquad (28)$$

as shown in (13).

However, the asymptotic constant was defined [15] as

$$C(x \to \infty) = \lim_{x \to \infty} e^{\eta x} P(w > x) \equiv \alpha \quad (0 < \alpha < 1). \qquad (29)$$

The asymptotic constant is a constant obtained by the inverse Laplace transformation for $v^*(\theta)$ of (5), and is a residue at the negative root of $\theta - \lambda[1 - g^*(\theta)] = 0$, i.e., $-\eta$.

If the Laplace transformation of $V^c(x) = 1 - V(x)$, which is the complementary distribution function of $V(x)$, is expressed by $V^{c*}(\theta)$, i.e.,

$$V^{c*}(\theta) = \frac{1 - v^*(\theta)}{\theta}, \qquad (30)$$

the asymptotic constant is given by the final-value theorem, i.e.,

$$\alpha = \lim_{\theta \to -\eta}(\theta + \eta)V^{c*}(\theta). \qquad (31)$$

With (5), we can obtain α as

$$\alpha = \frac{\lambda Q\left[1 - g_0^*(-\eta)\right]}{\eta\left[1 + \lambda g^{*'}(-\eta)\right]}, \qquad (32)$$

where the prime mark in $g^{*'}(\theta)$ means the differentiation with respect to θ, and then we can obtain $g^{*'}(-\eta)$ by using the moments

$$g^{*'}(-\eta) = -\left\{M_1 + \frac{M_2}{1!}\eta + \frac{M_3}{2!}\eta^2 + \cdots + \frac{M_{n+1}}{n!}\eta^n + \cdots\right\}. \quad (33)$$

In Eq. (32), $g^*(-\eta)$ is also given by

$$g_0^*(-\eta) = \frac{1}{1 - \bar{s}_0\eta} \quad \text{for the exponential distribution} \qquad (34)$$

$$= \frac{1}{2\bar{s}_0\eta}\left(e^{2\bar{s}_0\eta} - 1\right) \quad \text{for the uniform distribution,} \qquad (35)$$

and

$$= e^{\bar{s}_0 \eta} \quad \text{for the deterministic distribution.} \qquad (36)$$

3.3. Asymptotic Average Packet Delay

Using the asymptotic delay rate and the asymptotic constant, we can obtain the average packet delay of (6) as

$$\bar{w}_T = T\left[1 - \frac{1}{V(T)}\right] + \left(\frac{1}{\eta} + \frac{D}{2}\right)\left[1 - \frac{Q}{V(T)}\right], \qquad (37)$$

where we use (28) as C.

When the offered load is low, the blocking probability is in the range $P_B \ll 1$. Then, we can assume that the tail distribution of $1 - V(T) = \rho_{eq} e^{-\eta T}$ is small enough to use the asymptotic expansion of

$$\frac{1}{V(T)} \approx 1 + \rho_{eq} e^{-\eta T}, \qquad (38)$$

and we can obtain from (37)

$$\bar{w}_T \approx \rho_{eq}\left(\frac{1}{\eta} + \frac{D}{2}\right) - \rho_{eq} e^{-\eta T} \cdot T. \qquad (39)$$

The first term on the right-hand side of (39) indicates the average delay of optical packets in the infinite buffer length system ($B \to \infty$), where no packet is discarded and all packets are prolonged by $D/2$ on average. The second term indicates the product of the probability of packets discarded $\rho_{eq} e^{-\eta T}$ and the maximum delay T. Therefore, (39) gives the approximate average delay in which the discarded packets are excluded.

4. Numerical Examples

4.1. Accuracy of Asymptotic Analysis

Table 2. *Calculations of asymptotic decay rate when* $D = 0.3$

ρ	Exponential			Deterministic			Uniform		
	first	second	forth	first	second	forth	first	second	forth
0.1	0.8736	0.8828	0.8837	1.5305	1.9769	2.4098	1.2238	1.4974	1.7740
0.2	0.7601	0.7671	0.7676	1.3316	1.6695	1.9401	1.0647	1.2719	1.4447
0.3	0.6466	0.6517	0.6520	1.1327	1.3772	1.5348	0.9057	1.0556	1.1563
0.4	0.5330	0.5365	0.5366	0.9338	1.1001	1.1833	0.7467	0.8486	0.9018
0.5	0.4195	0.4217	0.4217	0.7350	0.8379	0.8761	0.5877	0.6508	0.6752
0.6	0.3060	0.3071	0.3071	0.5361	0.5909	0.6047	0.4287	0.4622	0.4711
0.7	0.1925	0.1929	0.1929	0.3372	0.3590	0.3621	0.2696	0.2829	0.2850
0.8	0.0790	0.0790	0.0790	0.1383	0.1420	0.1422	0.1106	0.1129	0.1130

Table 3. ρ_{eq} *and asymptotic constant* α *when* $D = 0.3$

ρ	ρ_{eq}	Exponential			Deterministic			Uniform		
		first	second	forth	first	second	forth	first	second	forth
0.1	0.102	-1.04	-1.89	2.85	-0.312	-0.884	0.624	-0.357	-0.960	0.699
0.2	0.206	-1.59	-784	0.826	-0.799	3.89	0.478	-0.889	5.04	0.511
0.3	0.314	-2.87	2.39	0.622	-1.88	1.32	0.499	-2.04	1.47	0.516
0.4	0.426	-11.4	1.37	0.587	-8.80	1.00	0.555	-9.34	1.07	0.563
0.5	0.541	6.21	1.05	0.616	5.32	0.891	0.632	5.55	0.927	0.633
0.6	0.659	2.50	0.929	0.686	2.30	0.861	0.721	2.36	0.879	0.719
0.7	0.782	1.58	0.896	0.788	1.52	0.878	0.820	1.55	0.885	0.818
0.8	0.909	1.18	0.931	0.909	1.16	0.935	0.925	1.17	0.936	0.923

We now calculate the blocking probabilities of (7) using the asymptotic decay rate and the asymptotic constants. In addition, we compare the calculations with simulation results, and confirm the accuracy of the asymptotic analysis. In our calculations, the time unit is set to be the mean packet length such that $\bar{s}_0 = 1$. Table 1 gives the moments of the service time distributions of the *non-first-arrival packet*s, whose packet lengths are prolonged by $D/2$ on average. The table also shows the moments of three service time distributions, exponential, uniform, and deterministic, where M_1 indicates the mean service time and is the same $1 + D/2$ for the three distributions.

Table 2 shows values of the asymptotic decay rate η against the offered load ρ, which were calculated using (23)~(27) and the moments in Table 1. For calculations, we set

$D = 0.3$, and then $\rho_{eq} = 1$ when ρ is 0.87. The table also shows results of the first, second, and fourth order approximations for the three service time distributions.

In the exponential distribution case, the decay rate shows little changes when the approximation order increases. For example, when $\rho = 0.3$, the difference between the values of the first order and fourth order approximations is only 0.8 %, and when $\rho = 0.8$, the difference is less than 0.1 %. Then, we see that the first order approximation of the exponential distribution already has good accuracy. We have already approximated the CDF of the virtual waiting time as the exponential function, which fits well with the exponential distribution of packet length.

In contrast, in the deterministic distribution case, the decay rate becomes larger when the approximation order increases. Moreover, the percentages of the decay-rate differences

between the approximation orders become smaller when the order increases or the offered load becomes heavier. For example, when $\rho = 0.3$, the difference between the values of the first order and second order approximations is 15 %, and that between the second order and fourth order is 10 %. When $\rho = 0.8$, the difference between the first order and the second order is 2.6 % and that between the second order and fourth order is 0.14 %. It is clear that the decay rate converges to a certain value and then the accuracy may increase when the approximation order increases or the offered load becomes heavier. When the deterministic distribution is expanded by the exponential function, the expansion coefficients remain large even for higher order terms, and that is why the high-order terms are needed for the decay rate to converge.

For the uniform distribution case, we can see that the dependencies of the decay rate on the approximation order and the offered load lies between those of the exponential and the deterministic distributions.

Table 3 shows values of proportional constants ρ_{eq} and α against ρ when $D = 0.3$. Equations (32)~(36) are used for calculations of the asymptotic constant α and results using the first, second, and fourth order approximations of asymptotically expanded $g^{*'}(-\eta)$, as given in (33), are shown for each service time distribution. α values approach those of ρ_{eq} when the approximation order increases or the offered load becomes heavier. When the order is low and the load is light, however, α values become negative or more than 1, which are over the definition of α, i.e., $0 < \alpha < 1$. This is because the low order approximation of asymptotically expanded $g^{*'}(-\eta)$ is low in accuracy. It is clear that even the fourth order approximation of (33) is not accurate enough. Then, we use ρ_{eq} values as C for calculations of the blocking probabilities.
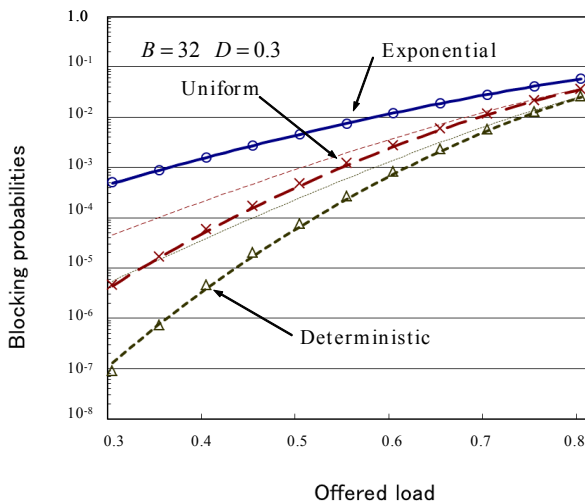

Fig. 3. *Blocking probabilities against the offered load with $B = 32$ and $D = 0.3$. Thick and thin lines respectively show the fourth order and first order approximation results. Marks are simulation results.*

Figure 3 shows the blocking probabilities against the offered load varied from the light-traffic case of 0.3 to the heavy-traffic case of 0.8. We set $B = 32$ and $D = 0.3$ and compared simulation results. We used $W_A(x)$ of (9) as $V(x)$, η of (23), and ρ_{eq} of (28) for the calculations. The solid, dashed, and dotted lines show the cases for the exponential, uniform, and deterministic packet-length distributions, respectively. The thick lines show the results of the fourth order approximation, and the thin lines show the results of the first order one, which have been already reported [9]. The circles, crosses, and triangles show the simulation results for the same respective cases. The simulations were carried out with 10^8 packets, and the countable packet loss limit became to be 10^{-8}.
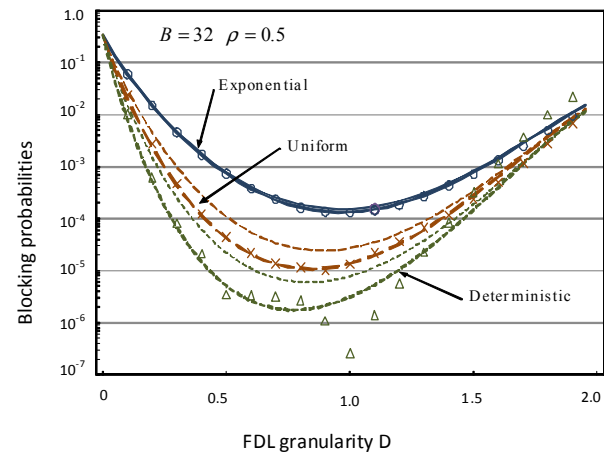

Fig. 4. *Blocking probabilities against D varied up to 2.0 with $B = 32$ and $\rho = 0.5$. Lines and marks are the same as those in Fig. 3.*

In both the uniform and deterministic distribution cases, the accuracy of the fourth order approximation was extensively improved compared with that of the first order approximation. The differences between the fourth order approximation and simulation results were within 10 % in the uniform distribution case for the entire offered load from 0.3 to 0.8. In the deterministic distribution case, the difference between them was also within 15% when the offered load was larger than 0.4. When the offered load was 0.3, however, the difference was 45%. The accuracy is not enough, but is extensively improved from the first order approximation.

Figure 4 shows the blocking probabilities against FDL granularity D varied up to 2.0 with $B = 32$ and $\rho = 0.5$. Due to the stable condition $\rho_{eq} < 1$, the value of D is limited by (4) as

$$D / \bar{s}_0 < 2(1/\rho - 1). \qquad (40)$$

When $\rho = 0.5$, the maximum D value is 2.0. The lines and marks have the same meanings as in Fig. 3.

Again, in both the exponential and uniform distribution cases, the accuracy of the fourth order approximation was excellent. The differences between the fourth order approximation and simulation results were within 10 %. In the deterministic distribution case, however, the differences between them were relatively large, especially around $D = 1$, where the FDL granularity is equal to the deterministic

packet length. Except around $D=1$, i.e. when the value of D is from 0.5 to 0.8 or from 1.3 to 1.5, the differences were within 60 %.
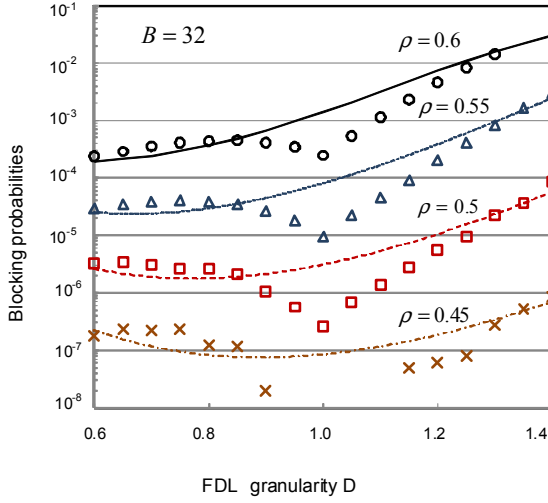

Fig. 5. *Blocking probabilities of deterministic packet-length distribution case around* $D=1.0$ *with* $B=32$ *and* $\rho = 0.45 \sim 0.6$ *. Lines show the fourth order approximation results and marks show simulation results.*

Figure 5 shows the blocking probabilities in detail around $D=1$ where ρ is set to be 0.45~0.6 as a parameter. At $D=1$, the simulation results have minimum values, which were about a one-tenth of the results of the fourth order approximation. It seems that, when the FDL granularity is equal to the deterministic packet length, delay periods of arrival packets may synchronize with the packet length and the blocking probabilities largely decrease even if arrivals are asynchronous. In this case, the assumption of an exponential function for the virtual waiting time, as shown in (9), can not to be applied. Except for $D=0.85 \sim 1.25$, however, the differences between approximations and simulations are within 60 %.
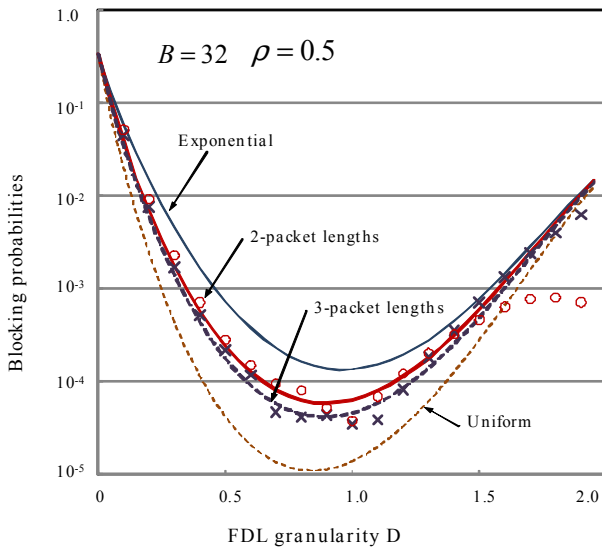

Fig. 6. *Blocking probabilities against* D *with* $B=32$ *and* $\rho = 0.5$ *. Thick lines show the forth order approximation results and marks show simulation results of 2-packet length and 3-packet length distributions.*

4.2. Applications to Real Traffic

In ITU-T Y.1541 (12/2011) Network performance objectives for IP-based services [11], the objective of the end-to-end IPLR is recommended to be less than 10-3 for all the QoS classes from 0 to 4. Assuming that packets pass through a maximum of 100 packet switches for end-to-end network transmission, the IPLR required for one packet switch is less than 10-5. Packet loss cannot be avoided for the finite FDL optical buffer, and, according to Fig. 3, it seems that an IPLR of less than 10-5 is quite difficult to achieve except in a low offered load.

In a realistic IP-packet-length distribution on the Internet, several sharp peaks occur at 40, 552, 576, and 1500 Bytes [17-19]. To design an optical buffer for a realistic network, we have to consider multiple fixed-packet-length distributions.

For n fixed packet lengths with $s_i (i=1,2,\cdots,n)$ and their arrival rates $\lambda_i (i=1,2,\cdots,n)$, the pdf of the packet length is

$$g_0(x) = \sum_{i=1}^{n} p_i \delta(x - s_i), \qquad (41)$$

where

$$p_i = \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} = \frac{\lambda_i}{\lambda} \text{ and } \lambda = \sum_{i=1}^{n} \lambda_i . \qquad (42)$$

The packet-length mean value $\bar{s}_0$ and each moment of the packet length distribution are given by

$$\bar{s}_0 = \sum_{i=1}^{n} p_i s_i = \sum_{i=1}^{n} \frac{\lambda_i s_i}{\lambda} = \frac{\rho}{\lambda}, \quad \rho \equiv \sum_{i=1}^{n} \lambda_i s_i \text{ and} \qquad (43)$$

$$M_2 = \sum_{i=1}^{n} p_i \left(s_i + \frac{D}{2} \right)^2 + \frac{1}{3} \left(\frac{D}{2} \right)^2$$

$$M_3 = \sum_{i=1}^{n} p_i \left(s_i + \frac{D}{2} \right)^3 + \left(\frac{D}{2} \right)^2 \left(\bar{s}_0 + \frac{D}{2} \right)$$

$$M_4 = \sum_{i=1}^{n} p_i \left(s_i + \frac{D}{2} \right)^2 \left[\left(s_i + \frac{D}{2} \right)^2 + 2 \left(\frac{D}{2} \right)^2 \right] + \frac{1}{5} \left(\frac{D}{2} \right)^4$$

$$M_5 = \sum_{i=1}^{n} p_i \left(s_i + \frac{D}{2} \right)^3 \left[\left(s_i + \frac{D}{2} \right)^2 + \frac{10}{3} \left(\frac{D}{2} \right)^2 \right] + \left(\frac{D}{2} \right)^4 \left(\bar{s}_0 + \frac{D}{2} \right). \quad (44)$$

Thus, we calculated the blocking probabilities in multiple fixed-packet-length distributions for two cases as follows.

Case 1. Two packet lengths; $s_1 = 64$ Bytes, $s_2 = 1518$ Bytes, and $p_1 = p_2 = 0.5$, so $\bar{s}_0 = 791$ Bytes.

Case 2. Three packet lengths; $s_1 = 64$ Bytes, $s_2 = 582$ Bytes, $s_3 = 1518$ Bytes, and $p_1 = p_2 = p_3 = 1/3$, so $\bar{s}_0 = 721$ Bytes.

Note that these packet lengths are the frame lengths in layer 2 of the network hierarchy model

Figure 6 shows the blocking probabilities with $B=32$ and

$\rho = 0.5$ against FDL granularity D. The offered load of $\rho = 0.5$ was selected as a low value, below which the present commercial IP networks are normally operated. The thick solid and dashed lines show the cases for the 2-packet-length and 3-packet-length distributions, and the circles and crosses represent the simulation results for the same respective cases. The thin solid and dashed lines show the results of the fourth order approximations for the exponential and uniform packet-length distributions, respectively.

The lines of the 2-packet-length and 3-packet-length distributions roughly agreed with those of the simulations when $D < 1.5$, and are situated between those of the exponential and uniform distributions. A realistic packet length distribution would be the shape of the 2-packet-length distribution, overlapping at the low levels with the uniform distribution [19]. Therefore, the blocking probabilities for the realistic distribution are situated between those for the 2-packet-length and the uniform distributions, and are greater than 10^{-5} for the $B = 32$ and $\rho = 0.5$ case.

Figure 7 shows the buffer length B, at which the blocking probabilities become less than 10^{-5}, against the offered load ρ. The FDL granularity D was selected to be 1.0 because, in Fig. 4, the blocking probabilities reach the minimum near $D = 1.0$. The lines have the same meanings as in Fig. 6. When B is small, the permitted offered load values strongly depend on the distributions. When B is large, however, the dependency on the distributions decreases. When $B = 40$, permitted offered load values lie in the range greater than 0.5 for the 2-packet-length and uniform distributions. Then, assuming that all-optical packet networks will be operated within the 0.5 offered load, the buffer length B is required to be greater than 40.
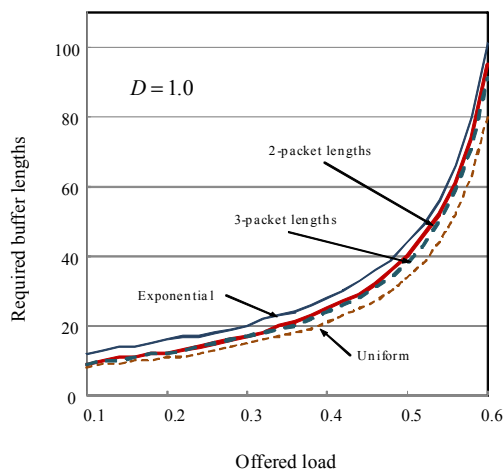

Fig. 7. *Buffer length* B, *at which blocking probabilities become less than* 10^{-5}, *against offered load* ρ. *Lines are the same as those in Fig. 6.*

5. Conclusions

This paper established a highly accurate approximation to calculate the blocking probabilities of optical-packet-switch buffers with generally distributed packet lengths. We focused on a simple exponential approximation for tail probabilities of the steady-state virtual waiting time in the infinite optical buffer. The decay rate of the tail probabilities was asymptotically expanded by a power series of $(1 - \rho')$, and approximations up to the fourth order were used for blocking-probability calculations. Calculation values were compared with simulation results.

The main results are as follows.

1) It was clarified that the higher-order terms in the asymptotic expansion reinforced the accuracy of the blocking-probability calculations for all packet-length distributions.

2) By using the fourth order approximation of the decay rate, an accuracy within 10 % was obtained for both the exponential and uniform distribution cases when an offered load is greater than 0.3. For the deterministic distribution case, however, the accuracy was within 60% except for $D = 0.85 \sim 1.25$.

3) The approximation was applied to design an optical buffer with realistic IP-network traffic distributions. It was clarified that, assuming that all-optical packet networks will be operated within a 0.5 offered load, the buffer length B is required to be greater than 40.

The approximations established in this paper can be applied to investigate multiclass optical buffers for priority queueing and to design wavelength-division multiplexing optical packet switches and networks with the maximum throughput.

References

[1] R. S Tucker et al, "Evolution of WDM optical IP networks: A cost and energy perspective," J. Lightwave Technol., Vol. 27, No. 3, pp. 243-252, 2009.

[2] R. S Tucker, "Scalability and energy consumption of optical and electronic packet switching," J. Lightwave Technol., Vol. 29, No. 16, pp. 2410-2421, 2011.

[3] F. Callegati, "Optical buffers for variable length packets," IEEE Commun. Lett., Vol. 4, No. 9, pp. 292-294, 2000.

[4] R. C. Almeida, J. U. Pelegrini, and H. Waldman, "A generic-traffic optical buffer modeling for asynchronous optical switching networks," IEEE Commun. Lett., Vol. 3, No. 2, pp. 175-177, 2005.

[5] X. Ma, "Modeling and design of WDM optical buffers in asynchronous and variable-length optical packets switches," Optical Commun., No. 269, pp. 53-63, 2007.

[6] J. Liu, T. T. Lee, X. Jiang, and S. Horiguchi, "Blocking and Delay Analysis of Single Wavelength Optical Buffer with General Packet Size Distribution," J. Lightwave Technol., Vol. 27, No. 8, pp. 955-966, 2009.

[7] H. E. Kankaya and N. Akar, "Exact analysis of single-wavelength optical buffers with feedback Markov fluid queues," J. Opt. Commun. Netw., Vol. 1, No. 6, pp. 530-542, 2009.

[8] W. Rogiest, and H. Bruneel, "Exact optimization method for an FDL buffer with variable packet length", Photon. Technol. Lett., Vol. 22, No. 4, pp. 242-244, 2010.

[9] Y. Murakami, "An approximation for blocking probabilities and delays of optical buffer with general packet-length distributions," J. Lightwave Technol., Vol. 30, No. 1, pp. 54-66, 2012.

[10] W. Rogiest et al., "Heuristic performance model of optical buffers for variable length packets," Photon Netw. Commun., Vol. 26, pp. 65-73, 2013.

[11] ITU-T Recommendation Y.1541 (12/2011), Network performance objectives for IP-based services.

[12] L. Kleinrock, "Queueing systems, Vol. 1: Theory", John Wiley & Sons, New York, 1975.

[13] A. A. Fredricks, "A class of approximations for the waiting time distribution in a GI/G/1 queueing system," Bell Syst. Tech. J. Vol. 61, pp. 295-325, 1982.

[14] G. L. Choudhury, and W. Whitt, "Heavy-traffic asymptotic expansions for the asymptotic decay rates in the BMAP/G/1 queue," Stochastic Models, Vol. 10, No. 2, pp. 453-498, 1994.

[15] J. Abate, G. L. Choudhury, and W. Whitt, "Exponential approximations for tail probabilities in queues, I: Waiting Times," Oper. Res., Vol. 43, No. 3, pp. 885-901, 1995.

[16] see 3.6.25 in M. Abramowitz and I. A. Stegun, "Handbook of Mathematical Functions, 10th printing," National Bureaus of Standards, U. S. Government Printing Office, Washington, D. C., 1972.

[17] F. Xue et al., "Design and experimental demonstration of a variable-length optical packet routing system with unified contention resolution," J. Lightwave Technol., vol. 22, no. 11, pp. 2570-2581, 2004.

[18] The cooperative association for the internet data analysis – Packet size distribution comparison between internet links in 1998 and 2008, www.caida.org/research/traffic-analysis/ , 2008.

[19] E. R. S. Castro, M. S. Alencar and I. E. Fonseca, "Probability density functions of the packet length for computer networks with bimodal traffic," International J. of Computer Networks & Communication, vol. 5, no. 3, pp.17-31, 2013.