An Asymptotic Throughput Analysis for Input-buffered Switches

Duan-Shin Lee, Chao-Lin Yu, and Cheng-Shang Chang,
Institute of Communications Engineering
National Tsing Hua University
Hsinchu 300, Taiwan, R.O.C.
Email: lds@cs.nthu.edu.tw; clyu@gibbs.ee.nthu.edu.tw; cschang@ee.nthu.edu.tw

Abstract

The analysis of the maximum throughput of many input buffered switches requires to study discrete time queues with Poisson arrivals. The mean number of arrivals in a unit of time is ρF, where F is the maximum number of customers that can be served in one unit of time. For large F, we find an asymptotic formula for ρ such that the mean stationary queue length is equal to sF, where s is a positive integer.

Index Terms

discrete-time queues, asymptotic analysis, maximum throughput, input-buffered switches

I. Introduction

In this paper we present an asymptotic analysis of a discrete time queue with Poisson arrivals. Specifically, let $q(t)$ be the number of customers in the queue at time t and $q(t)$ satisfies the renowned Lindley equation:

$$q(t+1) = (q(t) - F)^+ + a(t), \tag{1}$$

where F is the maximum number of customers that can be served in a time slot and $a(t)$ is the number of new arrivals at time t. The symbol $(x)^+$ denotes $\max(x, 0)$. Assume that $a(t)$ is a sequence of independent Poisson random variables with mean ρF. From the classical queueing theory [1], this queue is stable and the stationary queue length distribution exists if $\rho < 1$. In this letter we present an asymptotic formula of ρ for large F such that

$$\mathsf{E}[q(\infty)] = sF, \tag{2}$$

where s is a positive integer.

In this paragraph we present a motivation for this study. Readers that are only interested in the queueing analysis can skip over this paragraph. Consider an input-buffered switch as shown in Figure 1. In a time slot at most one packet (*i.e.*, customer in queueing terms) can be switched from the head-of-line positions of input buffers to any output port. Clearly, if there are more than one packet destined to a particular output port, only one can reach the output port and the rest are blocked at the input buffers. This is called head-of-line (HOL) blocking. Many switch architectures with buffers placed at input ports suffer from the HOL blocking problem and thus can not achieve 100% throughput. To analyze the maximum throughput, assume that each input buffer already contains an infinite number of packets. We assume that a packet chooses its destination uniformly and independently. Then, it suffices to consider a particular output port due to symmetry. Let $q(t)$ be the number of packets in the HOL positions that destine to output port 1 at time t. As there are exactly N HOL packets at any time t, from symmetry we have $\mathsf{E}(q(t)) = 1$. Let $a(t)$ be the number of packets that become HOL packets and destine for output port 1. Then, $q(t)$

primary classification 90B22, secondary classification 90B18

This research is supported in part by the National Science Council, Taiwan, R.O.C., under Contract NSC-92-2213-E-007-076, and the program for promoting academic excellence of universities NSC 93-2752-E-007-002-PAE.

satisfies the Lindley equation (1) with $F = 1$. Finally, $a(t)$ is a random sum of Bernoulli random variables as each packet chooses its destination independently and uniformly. Thus, $a(t)$ converges to a Poisson random variable with mean ρ when $N \to \infty$. In this special case, we have the following famous formula for discrete time $M/GI/1$ queues (see e.g., [8])

$$\mathsf{E}[q(\infty)] = \frac{1}{2}\frac{2\rho - \rho^2}{1-\rho}. \tag{3}$$

This formula is the discrete time version of the well known Pollaczek-Khinchin formula. With this formula (2) can be easily solved in this special case ($s = 1$). It leads to the well-known $2 - \sqrt{2} = 58.6\%$ throughput in [5]. We mention that the throughput analysis of many switch designs based on input buffer architectures reduces to the solution of (2), where s depends on switches. We refer the reader to [2] for more examples.

The purpose of this letter is to present an asymptotic analysis of the root of (2) for large F. To this end, we consider the Lindley equation in (1). Let

$$p_j = \mathsf{P}(q(\infty) = j), \quad j = 0, 1, \ldots \tag{4}$$

and let $P(z)$ be the generating function of distribution $\{p_j, j = 0, 1, \ldots\}$. Assume that $a(1)$ is Poisson with mean ρF. By a routine algebraic manipulation, we can obtain an expression for $P(z)$ from (1), *i.e.*,

$$P(z) = e^{\rho F(z-1)} \frac{\sum_{j=0}^{F-1} p_j (z^F - z^j)}{z^F - e^{\rho F(z-1)}}. \tag{5}$$

The only unknowns in the above expression are $p_0, p_1, \ldots, p_{F-1}$. From Rouché's theorem the denominator in (5) has exactly F roots inside or on the unit circle on the complex plane. Specifically, let α_j denote the j^{th} and α_j is also the unique root of

$$z - e^{\mathbf{i}2\pi j/F} e^{\rho(z-1)} = 0. \tag{6}$$

Since $P(z)$ is analytic on or within the unit circle, the numerator in (5) must also vanish at α_j, $j = 0, 1, \ldots, F-1$. In addition, since the numerator in (5) is a polynomial, it must be of the form

$$P(z) = e^{\rho F(z-1)} \frac{c \prod_{j=0}^{F-1} (z - \alpha_j)}{z^F - e^{\rho F(z-1)}}, \tag{7}$$

where c is a constant. To determine c, one simply substitutes $z = 1$, applies L'Hospital's rule, and obtains

$$c = \frac{F(1-\rho)}{\prod_{j=1}^{F-1}(1-\alpha_j)}. \tag{8}$$

Differentiating (7), substituting $z = 1$ and applying L'Hospital's rule, we obtain the following formula for the mean queue length

$$\mathsf{E}[q(\infty)] = \frac{1 - F(1-\rho)^2}{2(1-\rho)} + \sum_{j=1}^{F-1} \frac{1}{1-\alpha_j}. \tag{9}$$

We call this formula the generalized Pollaczek-Khinchin mean queue length formula. We refer the reader to [7], [2] for more details in deriving (9). Reference [2] also contains numerical algorithms and bounds for the root of (2).

II. Asymptotic Analysis

In view of (9), the computation of ρ such that (2) holds involves a large number of roots and can be numerically difficult. In this section, we present an asymptotic analysis of the generalized Pollaczek-Khinchin mean queue length formula (9) for very large F. From this asymptotic formula we can compute the maximum throughput. To this end, we rewrite (9) as

$$\frac{\mathsf{E}[q(\infty)]}{F} = \frac{1-F(1-\rho)^2}{2F(1-\rho)} + \frac{1}{F}\sum_{j=1}^{F-1}\frac{1}{1-\alpha_j}. \tag{10}$$

For large F, we may approximate the sum in (10) by an integral. That is, we attempt to evaluate

$$\frac{1}{F}\sum_{j=1}^{F-1}\frac{1}{1-\alpha_j} \approx \int_0^1 \frac{d\theta}{1-\alpha(\theta)}, \tag{11}$$

where $\alpha(\theta)$ is the unique root of

$$z - e^{2\pi \mathrm{i}\theta} e^{\rho(z-1)} = 0 \tag{12}$$

on or within the unit circle. Differentiating (12), we obtain

$$\alpha'(\theta) = \frac{2\pi\mathrm{i}\alpha(\theta)}{1-\rho\alpha(\theta)}. \tag{13}$$

With the help of (13), we convert the integral in (11) into

$$\begin{aligned}\int_0^1 \frac{d\theta}{1-\alpha(\theta)} &= \int_0^1 \frac{1-\rho\alpha(\theta)}{(1-\alpha(\theta))(1-\rho\alpha(\theta))}\, d\theta \\ &= \int_0^1 \frac{1-\rho\alpha(\theta)}{1-\alpha(\theta)}\frac{\alpha'(\theta)}{2\pi\mathrm{i}\alpha(\theta)}\, d\theta \\ &= \frac{1-\rho}{2\pi\mathrm{i}}\int_0^1 \frac{\alpha'(\theta)}{1-\alpha(\theta)}\, d\theta + \frac{1}{2\pi\mathrm{i}}\int_0^1 \frac{\alpha'(\theta)}{\alpha(\theta)}\, d\theta. \end{aligned} \tag{14}$$

Note that the last two integrals in (14) are integrals of complex functions with respect to a real parameter θ. We can convert these integrals into two contour integrals. Specifically,

$$\int_0^1 \frac{\alpha'(\theta)}{\alpha(\theta)}\, d\theta = \int_D \frac{dz}{z} \tag{15}$$

$$\int_0^1 \frac{\alpha'(\theta)}{1-\alpha(\theta)}\, d\theta = \int_D \frac{dz}{1-z}, \tag{16}$$

where D is the curve $\{\alpha(\theta) | 0 \le \theta \le 1\}$. We summarize the properties of D in the following proposition.

Proposition 1 *D is a simple, smooth and closed contour that encircles the origin exactly once in the counter clockwise direction.*

This proposition is quite evident and we omit the proof. Since D is a simple contour, from the classical Cauchy integral formula [3][page 119], the value of the integral in (15) is

$$\frac{1}{2\pi\mathrm{i}}\int_D \frac{dz}{z} = 1. \tag{17}$$

The integral in (16) is called a singular integral because the integrand has a singularity on the contour D. This type of integrals was discussed in Gakhov [4]. Specifically, Gakhov showed that for a closed contour, if $f(z)$ satisfies the Hölder condition[1], the singular integral

$$\int_L \frac{f(z)}{z-z_0}\,dz \qquad (z_0 \in L)$$

exists in the sense of the Cauchy principal value. In addition, Gakhov [4][page 16, Eq. 3.11] showed that

$$\int_L \frac{f(z)}{z-z_0}\,dz = \int_L \frac{f(z)-f(z_0)}{z-z_0}\,dz + \mathrm{i}\pi f(z_0). \tag{18}$$

With the help of (18), the singular integral in (16) is

$$\frac{1-\rho}{2\pi\mathrm{i}} \int_D \frac{dz}{1-z} = -(1-\rho)/2. \tag{19}$$

Substituting (17), (19) and (11) into (10), we have

$$\frac{\mathsf{E}[q(\infty)]}{F} \approx \frac{1-F(1-\rho)^2}{2F(1-\rho)} + \frac{1+\rho}{2}. \tag{20}$$

To find the maximum throughput for large F, we solve

$$\frac{1-F(1-\rho)^2}{2F(1-\rho)} + \frac{1+\rho}{2} - s = 0. \tag{21}$$

If $s = 1$, (21) becomes

$$\left(\frac{1}{2F(1-\rho)}\right)[1-2F(1-\rho)^2] = 0. \tag{22}$$

The root of (22) is $F(1-\rho)^2 = 1/2$. This means that $\sqrt{F}(1-\rho) \approx \sqrt{2}/2$ for large F. If $s > 1$, (21) reduces to

$$1 - 2F(1-\rho)^2 - 2F(s-1)(1-\rho) = 0.$$

This quadratic equation with unknown $1-\rho$ can be solved:

$$1-\rho = \frac{-2F(s-1) \pm \sqrt{4F^2(s-1)^2 + 8F}}{4F}. \tag{23}$$

Using the following Taylor expansion

$$(1+t)^{1/2} = 1 + \frac{t}{2} - \frac{t^2}{8} + \frac{t^3}{16} + \ldots$$

one can simplify (23)

$$1-\rho = \frac{1}{2F(s-1)} - \frac{1}{4F^2(s-1)^3} + \ldots$$

This shows that $F(1-\rho) \approx 1/(2(s-1))$ for large F.

III. Conclusions

The throughput analysis of many input buffered switches requires to solve the root of an equation involving with the mean stationary queue length of discrete time queues. By the technique of complex contour integrals, we presented an asymptotic analysis to simplify the generalized Pollaczek-Khinchin mean queue length formula when F is large. Finally, we use this result to obtain an approximation for the maximum throughput of input buffered switches.

[1]Let L be a smooth contour and $\phi(z)$ be a function on it. The function $\phi(z)$ is said to satisfy on the curve the Hölder condition, if for two arbitrary points of this curve

$$|\phi(z_1) - \phi(z_2)| < A|z_1 - z_2|^{\gamma},$$

where A and γ are positive numbers. A is called the Hölder constant and γ is called the Hölder index.

REFERENCES

[1] F. Baccelli and P. Bremaud. *Elements of Queueing Theory*. New York: Springer-Verlag, 1994.
[2] C.-S. Chang, D.-S. Lee, and C.-L. Yu. Generalation of the Pollaczek-Khinchin formula for throughput analysis of input-buffered switches. In *Proc. IEEE Infocom*, 2005.
[3] R.V. Churchill and J.W. Brown, *Complex Variables and Applications*, 5-th edition, McGraw Hill, 1990.
[4] F.D. Gakhov, *Boundary Value Problems*, Dover, 1966.
[5] M. J. Karol, M. G. Hluchyj, and S. P. Morgan. Input versus output queueing on a space-division packet switch. *IEEE Trans. Commun.*, Vol. 35(12), Dec. 1987.
[6] C. Kolias and L. Kleinrock, Throughput analysis of multiple input queueing in ATM switches, *Proc. Broadband Communications, L.Mason and A. Casaca, Eds. London, U.K.:Chapman & Hall, 1996*, pp. 382-393.
[7] San-Qi Li, Performance of trunk grouping in packet switch design. *Performance Evaluation*, 12 (1991) 207-218, North-Holland.
[8] M. Schwartz. *Broadband Integrated Networks*. Prentice Hall, New Jersey, 1996.

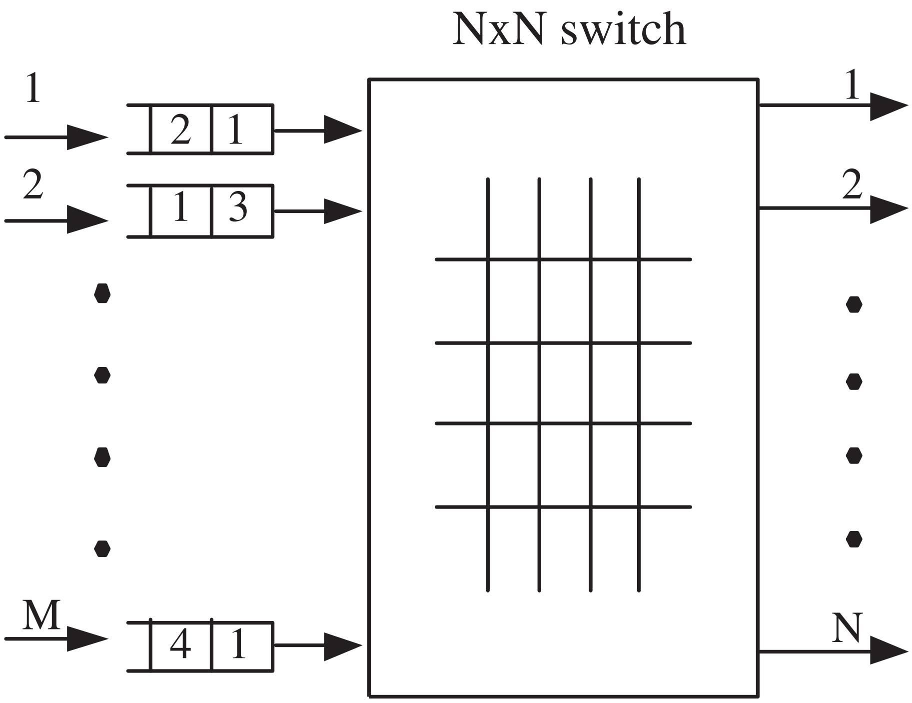

Fig. 1. An input buffered switch. The numbers in the input buffers denote the destination port numbers.