Analysis of Non-Uniform Cell Destination Distribution in Virtual Output Queueing Systems

Itamar Elhanany, *Student Member, IEEE*, and Dan Sadot, *Member, IEEE*

***Abstract*—In this letter we develop a discrete-time analytical queueing model for studying the performance of input-queued switches with nonuniform cell destination distribution. Virtual output queues are assumed at the ingress ports where cell arrivals are geometrically distributed and the service process is based on a work conserving random selection scheme. We consider the conditions for stability as basis for deriving closed-form expressions for the stationary queue size distributions from which the mean queue sizes and mean cell latencies are derived. We show that a very good agreement is obtained between simulation and analytical results.**

***Index Terms*—Input queued switches, nonuniform destination distribution, packet scheduling, virtual output queueing.**

I. Introduction

INPUT-QUEUED switching architectures, commonly realized using crossbar technologies, are widely deployed in high-performance switches and routers [1], [2]. In such systems, each input port may transmit to at most one output port at any given time. A scheduler, whether centralized or distributed, governs the switching process by determining the configuration of the crossbar matrix at any given time thus enabling data cells to traverse the switch fabric. A well-known phenomenon called head-of-line blocking [3] occurs in single queued ingress architectures where contention on a specific egress path prevents nonblocked queued cells from being transmitted, and in doing so limits the overall throughput of the switch. A common technique for overcoming the head-of-line blocking phenomena is virtual output queueing (VOQ). In VOQ a separate queue is maintained at the ingress port for each of the output destinations, as depicted in Fig. 1. Since the statistical nature of the traffic arriving at the virtual output queues has significant impact on the scheduler efficiency, it affects the switch performance particularly with respect to the latency, throughput and the amount of buffer memory required in the ingress modules.

In recent years, analysis of VOQ-based systems has received considerable attention, primarily in the context of scheduling algorithms [4]–[6]. Nonetheless, the scheduling performance is typically evaluated under the assumption that the destination distribution of arriving cells is uniform. This letter investigates the behavior of VOQ under traffic conditions characterized by cell arrivals which are nonuniformly distributed among the destinations. Based on the memoryless properties of the arrival and service processes, we devise an analytical model for obtaining closed-form expressions of the performance metrics. Validation of the results is demonstrated through very good matching between analytical and simulation results. The generic properties of the scheme allows for its application in the context of a wide range of scheduling algorithms.

Manuscript received May 6, 2002. The associate editor coordinating the review of this letter and approving it for publication was Dr. I. S. Venieris.

The authors are with the Electrical and Computers Engineering Department, Ben-Gurion University, Beer-Sheva 84120, Israel (e-mail: itamar@ieee.org; sadot@ee.bgu.ac.il).

Publisher Item Identifier 10.1109/LCOMM.2002.803490.

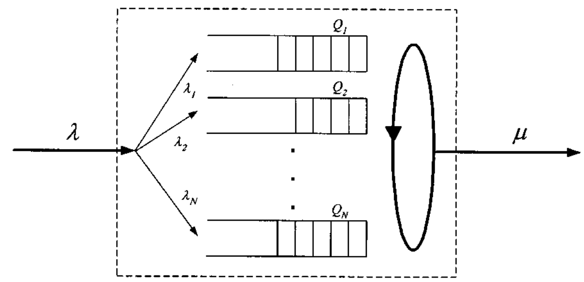

Fig. 1. Virtual output queueing based switch architecture with nonuniformly distributed cell arrivals.

II. Model Formulation and Analysis

We assume a discrete-time VOQ system with a single-server and infinite buffer capacity. The number of queues, N, corresponds to the number of distinct destinations in the system. All events occur at discrete time slot intervals in which at most a single arrival and a single service may occur. We employ an early-arrival model [7] whereby fixed-size cell arrivals obey a Bernoulli i.i.d. process. The global service discipline reflecting on the VOQ is also governed by an i.i.d. Bernoulli process, resulting in geometrically distributed interservice times. Let μ denote the homogeneous probability of service to the VOQ during each time slot. Once a service event occurs, an internal arbitration scheme determines which of the queues is to transmit a cell.

For any given queue, we let $Q(m)$ designate the queue occupancy at time slot m, such that

$$Q(m) = \max\left(Q(m-1) + A(m) - D(m), 0\right) \qquad (1)$$

where $A(m) \in \{0,1\}$ and $D(n) \in \{0,1\}$ are the number of arrivals and departures during time slot m, respectively. For stability, the arrival rate for each queue should converge to the departure rate, yielding the condition

$$\lim_{m\to\infty}\left(\frac{1}{m}\sum_{m=1}^{\infty} A(m)\right) = \lim_{m\to\infty}\left(\frac{1}{m}\sum_{m=1}^{\infty} D(m)\right). \qquad (2)$$

The statement of convergence between the arrival and departure rates, for geometric inter-arrival times and general state-dependent service discipline, can be expressed as

$$\lambda = \sum_{k=1}^{\infty} \mu_k \pi_k \tag{3}$$

where λ is the probability of arrival, π_k is the stationary probability of the queue size being k and μ_k is the probability of service given that the queue size is k.

Given a Geo/Geo/1 system in which $\mu_k = \mu$, an outcome of the steady-state balance equation for the early-arrival scenario can thus be written as

$$\lambda = \sum_{n=1}^{\infty} \mu\left(\pi_n\left(1-\lambda\right) + \pi_{n-1}\lambda\right) = \mu\left(1 - \pi_0(1-\lambda)\right) \tag{4}$$

where the term $\pi_0(1-\lambda)$ expresses the probability that the queue was empty prior to the arrival phase and no cell has arrived. When rearranged, (4) yields the well-known result for Geo/Geo/1 systems [8]

$$\pi_0 = \frac{\mu - \lambda}{\mu(1-\lambda)}. \tag{5}$$

An arriving cell may be destined to each of the N outputs. Let λ_k denote the probability that a cell is destined to destination k, such that $\sum_{k=1}^{N} \lambda_k = \lambda$. In the investigated arbitration scheme, during each time slot one of the nonempty queues in the VOQ is randomly selected for transmission, with probability μ. The underlying basic assumption is that the service discipline is memoryless, since for each time slot no information regarding previous service cycles is considered. It should be noted that the presented analysis, which assumes random selection between nonempty queues, can be broadened to apply for other VOQ arbitration algorithms that take into consideration additional queueing state information, such as weighted scheduling schemes [11].

We denote by $Q_k\{k = 1, 2, \ldots N\}$ the k^{th} queue within the VOQ, having probability of arrival λ_k. Due to the apparent memoryless properties of the arrival and service disciplines, these queues may be analyzed as Geo/Geo/1 systems. To that end, we focus our analysis on expressing the steady-state probability of Q_k being empty, i.e., $\pi_0^k = \Pr\{Q_k = 0\}$, which in turn will enable us to derive the stationary queue size distribution.

Three conditions must be met during a given time slot in order for Q_k to transmit a cell:

1) Q_k must be nonempty;
2) service must be granted to the VOQ system;
3) Q_k must prevail when contending with the other nonempty queues within the VOQ.

While the first two conditions are rather straightforward, the third condition requires some explanation. Assuming that Q_k is nonempty, it has an equal probability of being selected for transmission as any of the other nonempty queue. Let Γ_k be defined as the mean number of nonempty queues, excluding Q_k, such that

$$\Gamma_k = \sum_{j \neq k} (1 - \pi_0^j) = (N-1) - \sum_{j \neq k} \pi_0^j. \tag{6}$$

Given that Q_k is nonempty, the total number of contending queues is $\Gamma_k + 1$ and so we deduct that Q_k has a probability of service (μ_k)

$$\mu_k = \frac{\mu}{(\Gamma_k + 1)} = \frac{\mu}{N - \sum\limits_{j \neq k} \pi_0^j}. \tag{7}$$

Utilizing (4) and the three conditions for transmission stated above, we express the balance equation for Q_k as

$$\lambda_k = \mu \frac{\left(1 - \pi_o^k(1-\lambda_k)\right)}{N - \sum\limits_{j \neq k} \pi_0^j}. \tag{8}$$

Isolating π_0^k in (8), we have

$$\pi_0^k = \frac{1}{1-\lambda_k} - \frac{\lambda_k}{\mu(1-\lambda_k)} \left[N - \sum_{j \neq k} \pi_0^j \right]. \tag{9}$$

Letting $\Pi_N = \{\pi_0^1, \pi_0^2, \ldots, \pi_0^N\}^T$ and rearranging (9) for each of the queues, we attain a system solution in matrix form as

$$\begin{pmatrix} 1 & \alpha_1 & \alpha_1 & \alpha_1 & \ldots & \alpha_1 \\ \alpha_2 & 1 & \alpha_2 & \alpha_2 & \ldots & \alpha_2 \\ \alpha_3 & \alpha_3 & 1 & \alpha_3 & \ldots & \alpha_3 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \alpha_N & \alpha_N & \alpha_N & \ldots & \alpha_N & 1 \end{pmatrix} \Pi_N = \begin{pmatrix} \alpha_1 - \beta_1 N \\ \alpha_2 - \beta_2 N \\ \alpha_3 - \beta_3 N \\ \ldots \\ \alpha_N - \beta_N N \end{pmatrix}, \tag{10}$$

where

$$\alpha_k = -\frac{\lambda_k}{\mu(1-\lambda_k)}$$
$$\beta_k = -\frac{1}{1-\lambda_k}. \tag{11}$$

Solving this linear system, we find expressions for π_0^k which, by letting $\rho_k = 1 - \pi_0^k$ and utilizing the well known results for the Geo/Geo/1 model [8] which states that

$$\pi_n = \pi_0 \left(1 - \pi_0\right)^n = (1-\rho)\rho^n \tag{12}$$

we find the steady-state queue size distribution

$$\pi_n^k = (1 - \rho_k)\rho_k^n \tag{13}$$

the mean queue size,

$$E[Q_k] = \frac{\rho_k}{1-\rho_k}, \tag{14}$$

and, from Little's result, the mean cell delay

$$E[\tau_k] = \frac{E[Q_k]}{\lambda_k} = \frac{\rho_k}{\lambda_k(1-\rho_k)}. \tag{15}$$

III. Simulation Results

As means of validating the analytical framework with simulation results, we employ a nonlinear destination distribution model named Zipf's law [9]. The Zipf law states that the frequency of occurrence of some events, as a function of the rank (k) where the rank is determined by the above frequency of occurrence, is a power-law function: $P_k \sim 1/k^m$. A famous example of Zipf's law is the frequency of English words in a given

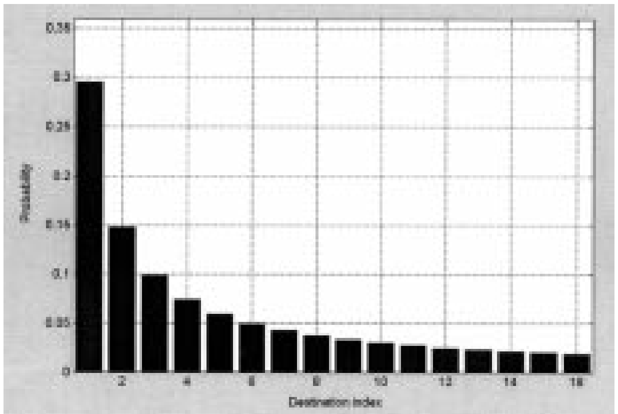

Fig. 2. The Zipf probability density function for $m = 1$ and $N = 16$.

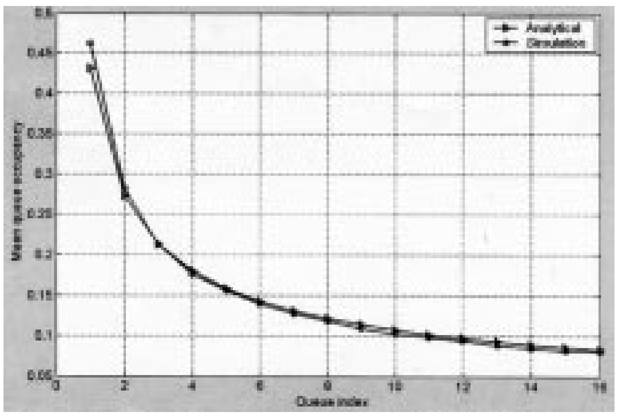

Fig. 3. Mean queue occupancy for each of the queues within the VOQ where $N = 16$, $\lambda = 0.8$, and $\mu = 0.85$.

text. Most common is the word "the", then "of", "to" etc. When the number of occurrence is plotted as the function of the rank ($k = 1$ most common, $k = 2$ second most common, etc.), the resulting form is a power-law function with exponential order typically close to 1. There has been evidence that the Zipf model accurately portrays web caching and access statistics [10], where the parameter m is close to unity. The probability that an arriving cell is heading to destination k is given by

$$\lambda_k = \frac{k^{-m}}{\sum_{j=1}^{N} j^{-m}}. \tag{16}$$

While $m = 0$ corresponds to uniform distribution, and as m increases the distribution becomes more biased toward a preferred destination. Fig. 2 depicts the Zipf distribution with $N = 16$ and $m = 1$. For the Zipf model we derive the following equation (10) parameters

$$\alpha_k = \frac{k^{-m}}{\mu\left(k^{-m} - \sum_{j=1}^{N} j^{-m}\right)} \quad \beta_k = \frac{\sum_{j=1}^{N} j^{-m}}{k^{-m} - \sum_{j=1}^{N} j^{-m}}. \tag{17}$$

Fig. 3 presents the mean queue occupancy for each queue in a given VOQ with $N = 16$. The generic probability of arrival to the VOQ, λ, is 0.8 while the probability of service, μ, is 0.9. Cell arrivals are distributed according to the Zipf model with parameter $m = 1$. Fig. 4 illustrates the mean queue occupancy for the first queue as a function of the probability of arrival where $\mu = 0.85$. We observe that there is good agreement between the analytical values and simulation results.

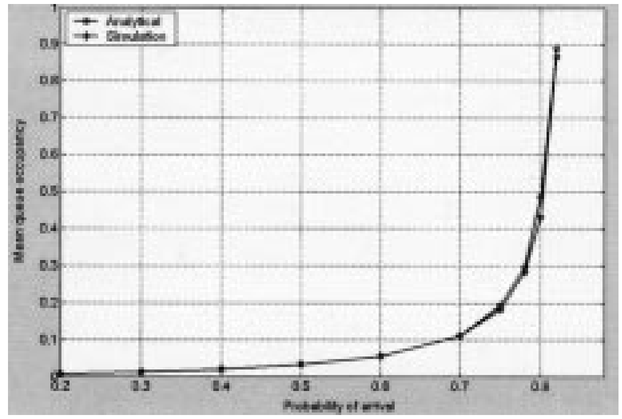

Fig. 4. Mean queue occupancy for first virtual output queue as a function of the probability of arrival, with $N = 16$ and $\mu = 0.85$.

IV. Conclusions

This letter investigates the behavior of discrete-time virtual output queueing systems with nonuniformly distributed Bernoulli i.i.d. cell arrivals. Random selection service discipline is assumed within the VOQ, yielding geometrical interservice times. Closed-form expressions for the queue size distributions, mean queue occupancy, and mean cell delay are obtained. Very good agreement is demonstrated between the analytical and simulations results.

References

[1] I. Elhanany and D. Sadot, "A contention-free Tbit/sec packet switching architecture for ATM over WDM networks," *IEICE Trans. Commun.*, vol. E83-B, no. 2, Feb. 2000.

[2] N. McKeown, "The iSLIP scheduling algorithm for input-queued switches," *IEEE Trans. Commun.*, vol. 7, pp. 188–201, Apr. 1999.

[3] M. J. Karol, M. G. Hluchyj, and S. P. Morgan, "Input versus output queueing in a space division switch," *IEEE Trans. Commun.*, vol. COM-35, pp. 1347–1356, Dec. 1987.

[4] C. Kolias and L. Kleinrock, "Throughput analysis of multiple input-queuing in ATM switches," *Proc. Int. IFIP—IEEE Conf. on Broadband Communications*, pp. 382–393, 1996.

[5] H. Kim, K. Kim, and Y. Lee, "Derivation of the mean cell delay and cell loss probability for multiple input-queued switches," *IEEE Commun. Lett.*, vol. 4, pp. 140–142, Apr. 2000.

[6] G. Nong, J. K. Muppala, and M. Hamdi, "Analysis of nonblocking ATM switches with multiple input queues," *IEEE/ACM Trans. Networking*, vol. 7, pp. 60–75, Feb. 1999.

[7] H. Takagi, *Queueing Analysis: A Foundation of Performance Evaluation, Volume 3: Discrete-Time Systems*. Amsterdam, The Netherlands: Elsevier Science, 1993.

[8] J. J. Hunter, *Mathematical Techniques of Applied Probability: Discrete Time Models: Techniques and Applications*. New York: Academic , 1983, vol. 2.

[9] G. K. Zipf, *Psycho-Biology of Languages*. Cambridge, MA: MIT Press, 1965.

[10] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "Web caching and Zipf-like distributions: Evidence and implications," in *Proc. 18th Annu. IEEE INFOCOM'99*, vol. 1, Mar. 1999, pp. 126–134.

[11] R. Schoenen and R. Hying, "Distributed cell scheduling algorithms for virtual-output-queued switches," in *Proc. IEEE GLOBECOM'99*, 1999, pp. 1211–1215.