Analysis of Output Queued Cell Switches with Random Arbitration and Generic Arrival Processes

Itamar Elhanany, *Member IEEE*
Department of Electrical and Computer Engineering
The University of Tennessee
Knoxville, TN 37996-2100
e-mail: itamar@ieee.org

***Abstract*— This paper presents an analysis of output queued cell switches which are introduced with generic non-uniformly distributed traffic. Random arbitration is employed whereby non-empty queues compete equally for service within each switching interval. In particular, we study the case of two-state Markov-modulated arrivals in which input ports generate bursty streams that are non-uniformly distributed. Under the assumption of a memoryless server, the probability generating function of the interarrival process is utilized to derive closed-form expressions for the queue size distribution. The methodology established in this paper forms a flexible tool in determining bounds on the behavior and expected performance of output queued switches under a range of traffic scenarios. The validity of the analytical inference is established through simulation results.**

Keywords- output queued switches; ON/OFF arrivals; packet scheduling; performance analysis

I. INTRODUCTION

Output queued (OQ) switches have been extensively studied in the literature. To a large extent, they represent the theoretical limit on the performance that can be achieved in any space-division switching fabric [1]. Consequently, analysis of input queued switching architectures is commonly carried out in comparison to that of an OQ switch [1][2]. Pragmatic output queueing schemes, such as shared memory architectures [3], have been deployed in switches and routers as well as studied at length in the literature. In such designs, a large shared memory space is utilized by all ports to read and write packets destined to the various output ports. The majority of the studies performed on OQ switches, however, consider traffic that obeys a Bernoulli (uncorrelated) process, and in most cases uniformly distributed such that all input ports offer the same load intensity to a given output port.

In a basic OQ switch, arriving cells traverse the switch directly to their designated output, without being queued or delayed in any way at the ingress ports. Such a scheme requires that a dedicated link, be it logical or physical, exist between each input port and each output port. In other words, N^2 links are needed for an $N \times N$ switch. A key advantage of OQ switches is that of minimal latency and controllable QoS provisioning [4]. Both are a result of the fact that arriving cells progress towards their destination without any impediment at the ingress stage.

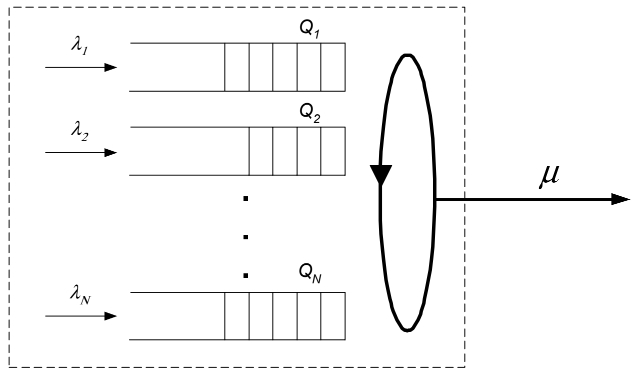

Figure 1. A basic model of the queueing architecture at an egress port of an output queued switch.

Figure 1 depicts a model of a typical output queueing configuration in an OQ switch. We let λ_k denote the mean offered load originating from input k, such that the aggregate load to the output is $\lambda = \sum_{k=1}^{N} \lambda_k$. In our context, traffic is said to be admissible if the aggregate load to each output port does not exceed 100%, i.e. $\lambda<1$. Although the terms scheduling and arbitration are commonly interchanged, we shall focus our attention on *arbitration,* pertaining to the process of determining which of N sources (queues) is granted service by a single (shared) server. Conversely, the task of *scheduling* addresses the matching of N sources to N destinations simultaneously, as is typically the case with input-queued switching architectures [2][5][6].

This paper presents an analysis of OQ switch introduced with a generic class of arrival processes that may be non-uniformly distributed as well as correlated. We apply random arbitration between the queues such that non-empty queues compete equally for service. Based on the probability generating functions of the interarrival times distribution of each queue, it is shown that precise characteristics of the queues' behavior can be obtained. In an aim to evaluate the performance of an output queued switch under realistic traffic conditions, we present analysis and simulation results for arrival patterns that obey a two-state Markov-modulated model. Moreover, the mean burst sizes and offered loads to the queues are heterogeneous.

The paper is organized as follows. Section II provides a description of the queueing model notation utilized throughout the paper along with some key assertions for the GI/Geo/1 queueing system. Section III describes the analysis of an OQ switch with random arbitration under a Markov modulated arrival process. Section IV presents and discusses simulation results while in section V the conclusions are drawn.

II. QUEUEING MODEL

A. Notation and Formulation

Packets may vary in size as they arrive at the switch ports. In typical switching platforms, a segmentation module partitions packets into fixed-size cells that are later reassembled at the egress modules prior to departing the switch. Processing fixed-size data units has proven both practical and easier to study. To that end, all data units traversing the switch fabric are assumed to be fixed in size.

We consider a discrete-time queueing system with N queues and a single-server of infinite buffer capacity, in which all events occur at fixed time slot intervals. Within each time slot, at most N arrivals may occur, originating from the N inputs. Since each output queueing system is associated with a unique switch port, at most a single departure may occur servicing one of the non-empty queues. We model the service discipline as a memoryless process for which there is a constant probability of service, μ, during each time slot. When $\mu = 1$, we have a basic output queueing model with no congestion. However, in the interest of expanding the analysis to address potential congestion on the output ports, we allow μ to be smaller than 1. Consequently, the interservice times are geometrically distributed with a mean of $1/\mu$.

Let $Q_k(n)$ denote the occupancy of queue k at time slot n, such that the evolution of the queue can be described as

$$Q_k(n+1)= Q_k(n)-A_k(n)-D_k(n), \qquad (1)$$

where $A_k(n)\in\{0,1\}$ and $D_k(n)\in\{0,1\}$ are the number of arrivals and departures to queue k during time slot n, respectively. In a stable system, the arrival rate must converge to the departure rate, such that

$$\lim_{T\to\infty}\left(\frac{1}{T}\sum_{n=1}^{\infty} A_k(n)\right)=\lim_{T\to\infty}\left(\frac{1}{T}\sum_{n=1}^{\infty} D_k(n)\right) \quad . \qquad (2)$$

If the latter does not hold, the queue occupancy either grows to infinity or, alternatively, converges to zero. Interpreting the above balance equation for a generic queueing system, we equate the mean probability of arrival to the mean probability of departure by writing

$$\begin{aligned}\Pr[arrival] &= \lambda = \Pr[departure] \\ &= \Pr[Service \cap (Q>0)] = (1-\gamma_o)\mu\end{aligned} \qquad (3)$$

from which we may isolate the expected stationary probability of the queue being empty,

$$\gamma_o = 1-\lambda/\mu \quad . \qquad (4)$$

B. The ON/OFF Arrival Process Revisited

It has been shown in the literature [7] that in a GI/Geo/1 discrete-time queueing system (general interarrival process and geometrically distributed service times), if f_n ($n\geq 1$) is the interarrival time distribution, with a p.g.f. $F(z)=\sum_{n=1}^{\infty} f_n z^n$,and the service times are geometrically distributed with parameter μ, then the stationary queue size distribution as viewed by an arriving cell, π_m, is in the form

$$\pi_m = (1-\rho)\rho^m \qquad m \geq 0 \qquad (5)$$

where ρ is a unique root of the equation

$$z = F\big(\mu z + (1-\mu)\big) \qquad (6)$$

that lies in the region (0,1). It has further been shown [8] that the queue size distribution, as viewed by an outside observer, is

$$\gamma_m = \begin{cases} 1-\xi & m=0 \\ \xi(1-\rho)\rho^{m-1} & m\geq 1 \end{cases} \quad . \qquad (7)$$

The latter is by definition independent of arrivals, hence its first moment

$$E[Q]=\sum_{m=1}^{\infty} m\gamma_m = \frac{\xi}{(1-\rho)} = \frac{\lambda}{\mu(1-\rho)} \qquad (8)$$

provides us with the mean queue occupancy from which, utilizing Little's results [9], we find the mean latency

$$E[W]=\frac{1}{\mu(1-\rho)} \quad . \qquad (9)$$

A late arrival model is considered, for reasons of convenience, such that within a time slot boundary a departure will always precede an arrival event, as shown in figure 2. We observe the queue size at the instances following the arrival phase such that the time slot boundaries are delimited by the observation instances.

Consider a discrete-time, two-state Markov chain generating arrivals modeled by an ON/OFF source which alternates between the ON and OFF states as shown in figure 3. Let the parameters p and q denote the probabilities that the Markov chain remains in states ON and OFF, respectively. An arrival is generated for each time slot that the Markov chain spends in the ON state. The result is a stream of correlated bursts of arrivals and silent periods, both of which are geometrically distributed in duration.

It can easily be shown that the parameters p and q are interchangeable with the mean arrival rate, $\lambda=(1-q)/(2-q-p)$, and mean burst size, $B=1/(1-p)$. Consequently, the offered load is identical to the steady-state portion of the time the chain spends in state ON.

Recalling the notation f_n for the interarrival times distribution, the probability of two consecutive arrivals occurring is identical to the probability that following an arrival the Markov chain remains in state ON, i.e. $f_1 = p$. Similarly, f_2 is the probability that following an arrival, the chain transitions

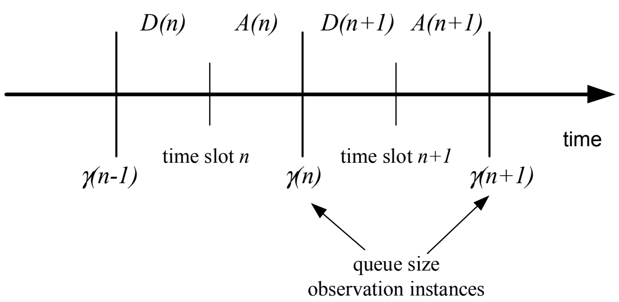

Figure 2. Late arrival model where within time slot boundaries a departure precedes an arrival event.

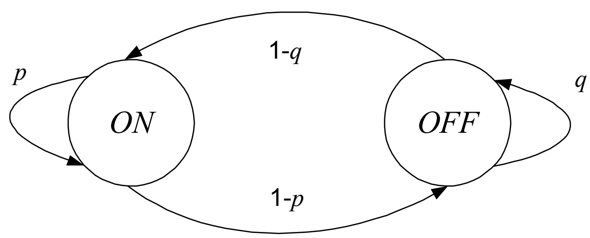

Figure 3. Basic two-state Markov modulated arrival process.

to the OFF state and then returns to the ON state. For $n>2$, it is apparent that following a transition from the ON state to the OFF state, there are $n-2$ time slots during which the chain remains in the OFF state before returning to the ON state. As a result, we obtain the following general expression for f_n:

$$f_n = \begin{cases} p & n=1 \\ (1-p)q^{n-2}(1-q) & n>1 \end{cases} . \quad (10)$$

The corresponding p.g.f. is

$$\begin{aligned} F(z) &= \sum_{n=0}^{\infty} f_n z^n = \sum_{n=1}^{\infty} f_n z^n \\ &= pz + \sum_{n=2}^{\infty} q^{n-2}(1-p)(1-q)z^n \\ &= pz + (1-p)(1-q)\frac{z^2}{1-qz} \end{aligned} . \quad (11)$$

The mean interarrival time can be found by differentiating $F(z)$ with respect to z and letting $z=1$,

$$\left.\frac{dF(z)}{dz}\right|_{z=1} = \sum_{n=1}^{\infty} n f_n = \frac{2-p-q}{1-q} , \quad (12)$$

which, as expected, equals to $1/\lambda$. Next we solve the equation $z = F(z\mu+(1-\mu))$ to find that the root in the region (0,1) is

$$\rho = \frac{(1-\mu)}{\mu}\left[\frac{1}{\mu(1-p-q)+q} - 1\right] . \quad (13)$$

Examining the condition $\rho<1$, which must be satisfied for stability yields the anticipated inequality $\mu > (1-q)/(2-p-q) = \lambda$.

III. OUTPUT QUEUEING ANALYSIS

In the investigated system, during each time slot in which a service event occurs, one of the non-empty queues is randomly selected for transmission in an unbiased manner. We shall refer to this arbitration scheme as *random arbitration*. Observably, the latter implies that the service discipline to each queue is also memoryless since during each time slot no information regarding previous service cycles is considered. As such, we will employ the results for the GI/Geo/1 queueing system to analyze the behavior of the individual output queues.

We focus our analysis on queue k observing that three conditions must be met during each time slot for the queue to be serviced: (1) service must be granted to the port (output link not congested), (2) the queue must be non-empty and, finally, (3) the queue must prevail when equally contending against the other non-empty queues. While the first two conditions are rather straightforward, the third condition requires some elaboration. Assuming that queue k is non-empty ($Q_k>0$), it has an equal probability of being selected for transmission as any other non-empty queue. The mean number of non-empty queues, excluding queue k, can be approximated by $\sum_{j\neq k}(1-\pi_{0,j})$ where $\pi_{0,j}$ denotes the stationary probability that queue j is empty. Given that queue k is non-empty, we add 1 to the mean number of non-empty queues to obtain the mean size of the contending set. By multiplying the expressions for the three conditions stated above, we find the probability of departure from queue k to be

$$\mu_k \cdot (1-\pi_{0,k}) = \frac{\mu(1-\pi_{0,k})}{\sum_{j\neq k}(1-\pi_{0,j})+1} = \frac{\mu(1-\pi_{0,k})}{N-\sum_{j\neq k}\pi_{0,j}} . \quad (14)$$

Since the arrival rate must converge to the departure rate we equate (14) to the rate of arrivals for each queue, yielding

$$\lambda_k = \frac{\mu(1-\pi_{0,k})}{N-\sum_{j\neq k}\pi_{0,j}} . \quad (15)$$

The latter offers N linear equations for the N variables $\pi_{0,j}$ ($j=1,2,\ldots,N$). Solving for $\pi_{0,j}$ we may directly obtain μ_k, the probability of service to each of the queues. We next turn our attention to the case where each input produces a stream of bursty arrivals modeled by a unique ON/OFF process, with respective parameters p_k and q_k. In view of the fact that queue sizes are geometrically distributed, based on (13) the respective parameters of these distributions are

$$\rho_k = \frac{1-\mu_k}{\mu_k}\left[\frac{1}{\mu_k(1-p_k-q_k)} - 1\right] . \quad (16)$$

We note that, as expected, the latter is a function of both the arrival model parameters and the rate at which the output queues are serviced.

IV. SIMULATION RESULTS AND DISCUSSION

Our simulations pertain to the scenario where traffic is both correlated and non-uniformly distributed between the inputs. Arrivals are generated by ON/OFF models which are independently operated for each input port. Thus two or more bursts may occur simultaneously originating from different inputs. As means of validating the analytical deductions with simulation results, we have selected a non-linear destination distribution model named Zipf's law [10]. The Zipf law states that the frequency of occurrence of some events, as a function of the rank (*m*) 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 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. The probability that an arriving cell is heading to destination k is given by

$$\lambda_k^{(m)} = \frac{k^{-m}}{\sum_{j=1}^{N} j^{-m}} \quad . \tag{17}$$

While m=0 corresponds to uniform distribution, as m increases the distribution becomes more biased towards preferred destinations. There has been recent evidence that the Zipf model accurately portrays web caching and access statistics [11], where the parameter m is close to unity. Previous work [1] has shown that the mean queue size in an output queued switch with First-in-First-Out (FIFO) arbitration and traffic arriving uncorrelated (Bernoulli) and uniformly distributed, is

$$E[Q]_{\substack{FIFO \\ Bernoulli}} = \frac{\lambda^2}{2(1-\lambda)} \quad . \tag{18}$$

An empirical study of uniformly distributed ON/OFF arrivals to an output queued switch employing FIFO arbitration has suggests that the mean queue size is

$$E[Q]_{\substack{FIFO \\ ON/OFF}} = \frac{\lambda B}{(1-\lambda)} \ , \tag{19}$$

where B is the mean burst size. By dividing each of the above mean queue sizes by the normalized offered loads we obtain their corresponding mean latencies. We shall refer to these assertions when comparing the latency of the FIFO discipline to that of random arbitration.

Figure 4 illustrates the mean queue occupancy for each of the output queues in a 16-port switch, where μ=0.8 and λ=0.7. Arriving traffic is distributed between the queues according to a $\text{Zipf}_{m=0.5}$ distribution with a mean burst size of 6 cells. As can be observed, the latency for each queue is well correlated with its share of the offered load. The simulation results clearly validate the analytical deductions.

In figure 5 the mean queue latency is presented as a function of the offered load for both random and FIFO arbitration. Traffic is assumed to be governed by the ON/OFF model with a mean burst size of 8 cells. Results for the case of

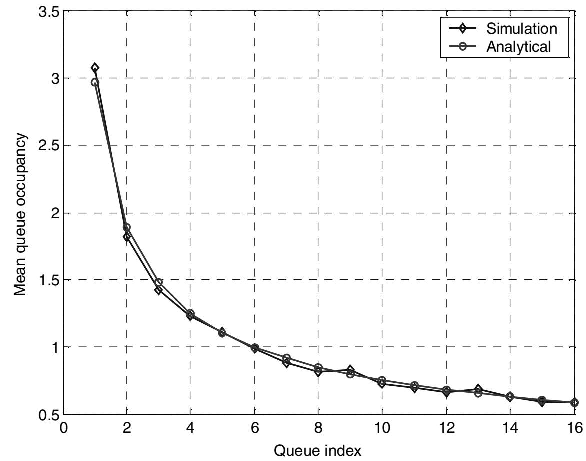

Figure 4. Mean queue occupancy for each of the output queues in a 16-port switch where λ = 0.8 and μ=1. Arriving traffic is distributed according to the $Zipf_{k=0.5}$ with mean burst sizes of 6 cells.

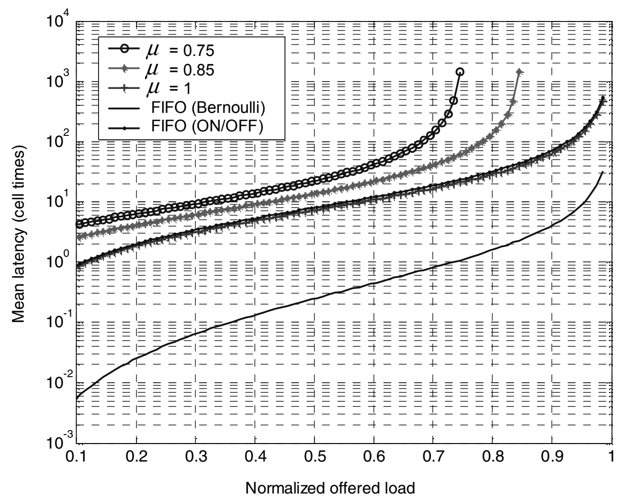

Figure 5. Mean queueing latency as a function of the offered load for random and FIFO arbitration schemes.

random arbitration are shown for different probability of service (μ) values, the impact of which is clearly observable in particular as the offered load approaches the service rate. Also shown are the mean latency metrics for FIFO arbitration, which in many cases constitute the theoretical lower limit on the delay through a space division switch. For the case of μ=1, FIFO and random arbitration yield almost identical results, suggesting that the presented methodology for obtaining the latency under generic traffic conditions provides valuable approximation to a pure FIFO system.

V. CONCLUSIONS

In this paper we present an analytical technique for evaluating the queueing behavior of an output queued switch that is introduced with non-uniformly distributed generic independent arrival processes. Under the assumption of a memoryless service discipline, the probability generating functions of the interarrival times distributions are utilized to derive per-queue closed-form expressions for the queue size distribution. To demonstrate the applicability of the proposed technique, we provide a detailed analysis of the case of non-uniformly distributed ON/OFF traffic. The results shown in this paper offer an insight on the performance envelope achievable by any egress arbitration scheme. To that end, any arbitration scheme which takes into consideration queue state is bound to yield better performance. Simulations are provided to validate the accuracy of the approach.

REFERENCES

[1] M. J. Karol, M. G. Hluchyj, and S. P. Morgan, "Input versus Output Queueing in a Space Division Switch," *IEEE Trans. Communications*, Vol. COM-35, pp. 1347-1356, Dec. 1987.

[2] I. Elhanany, D. Sadot, "DISA: A Robust Scheduling Algorithm for Scalable Crosspoint-Based Switch Fabrics," *IEEE Journal on Selected Areas in Communications,* Vol. 21, No. 4, pp. 535-545, May 2003.

[3] M. Arpaci and J. Copeland, "Buffer management for shared-memory ATM switches," *IEEE Comm. Surveys and Tutorials*, vol. 3, no. 1, First Quarter 2000.

[4] A. Parekh and R. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case," *IEEE/ACM Transactions on Networking*, 2(2):137-150, April 1994.

[5] N. McKeown, "The iSLIP Scheduling Algorithm for Input-Queued Switches," *IEEE/ACM Trans. on Networking*, Vol. 7, No. 2, pp. 188-201, April 1999.

[6] R. Rojas-Cessa, E. Oki, H. J. Chao, "CIXOB-k: Combined Input-Crosspoint-Output Buffered Packet Switch," *Proc. IEEE GLOBECOM 2001*, Vol. 4, pp. 2654-2660, Nov. 2001.

[7] J. J. Hunter, *Mathematical Techniques of Applied Probability: Discrete Time Models*: *Techniques and Applications*, Vol. 2, Academic Press, 1983.

[8] M. L. Chaudhry, U. C. Gupta, J. G. C. Templeton, "On the Relations Among the Distributions at Different Epochs for Discrete-Time GI/Geom/1 Queues," *Operations Research Letters*, Vol. 18, pp. 247-255, 1996.

[9] D. Gross, C. M. Harris, *Fundamentals of Queueing Theory*, John Wiley and Sons, 1985.

[10] G. K. Zipf, *Psycho-Biology of Languages*, Houghton-Miffin, MIT Press, 1965.

[11] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker, "On the Implications of Zipf's Law for Web Caching", *Proc. of IEEE INFOCOM '99*, New York, March 1999.