Virtual Input Queued Packet Switches with Non-Uniform Arrivals and Bursty Service

Itamar Elhanany*, Ortal Arazi†
Dept. of Electrical & Computer Engineering
University of Tennessee
Knoxville, TN 37919
email: {*itamar, †ortal}@ieee.org

Michael Kahane
Dept. of Electrical & Computer Engineering
Ben-Gurion University,
Beer-Sheva, Israel
email: michaelk@bgumail.bgu.ac.il

ABSTRACT

This paper presents a performance analysis of output queued packet switch architectures employing virtual input queueing (VIQ), whereby arrival rates are non-uniformly distributed between the sources and the service intervals are bursty. In particular, we study the case of a two-state Markov-modulated service discipline, reflecting on several pragmatic scenarios such as noisy packet radio networks. We show that by exploiting an extended Markov-modulated service process and the Geo/GI/1 queueing model, closed-form expressions for the mean queueing latencies can be obtained. The methodology established in this paper can be extended to derive additional performance metrics and expected behavior of more complex packet switching architectures.

1. INTRODUCTION

Output queued (OQ) switches have been extensively studied in the literature over the last two decades. To a large extent, they constitute the theoretical benchmark on the performance that can be achieved in any space-division switching fabric, in particular with regard to work conservation and minimization of the average queueing delay [1]. Consequently, analysis of input queued switching architectures is commonly carried out in comparison to that of an OQ switch [1][2][3]. Pragmatic output queueing schemes, such as shared memory architectures [4], have been deployed in switches and routers as well as studied at length. In such realizations, a large shared memory space is utilized by all input and output ports to read and write packets traversing the switching fabric.

In order to minimize the computational complexity of the output-link scheduling that is required at each output port, virtual input queueing (VIQ) is commonly employed. In VIQ, a unique logical queue is maintained at every output for packets arriving from each of the inputs, as shown in figure 1. This allows for FIFO scheduling, and other monotonic scheduling policies, to be deployed with great

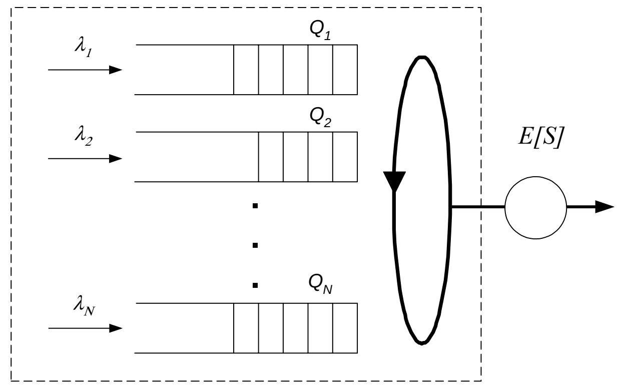

Figure 1. A basic model of the virtual input queueing architecture at an egress port of an output queued switch. Each logical queue corresponds to a different input port.

ease. It should also be noted that each input can generate a different rate of arrivals to the studied output port, denoted here by λ_1, λ_2, … λ_N.

The majority of the studies conducted on OQ switches, however, consider a deterministic server, that in each time slot serves one of the non-empty logical queues with probability 1. The latter assumption is somewhat incoherent with realistic networking scenarios, such as radio packet networks and optical burst networks, in which, due to channel conditions, the server is not necessarily available at any time to serve a packet from the output buffers.

In some wireless networks, a shared channel may be busy such that a switch is required to wait a given amount of time before reattempting to transmit packets. Hence, it is widely acknowledged that bursty conditions characterize many pragmatic wireless channels [5]. Consequently, the service interval durations are interpreted as bursty in nature, whereby such burstiness can be modeled as a Markovian process.

This paper presents an analysis of OQ switches introduced with non-uniformly distributed Bernoulli arrivals and a Markov-modulated service process. We apply random arbitration between the queues such that non-empty queues compete equally for service. Utilizing the Geo/GI/1 queueing model, it is shown that accurate characterization of the queues' behavior is attained.

The paper is organized as follows. Section 2 provides a description of the VIQ formulation used throughout the paper along with some key assertions pertaining to the Geo/GI/1 model. Section 3 describes the analysis of the output queued switch with Markov-modulated service intervals, and in section 4 the conclusions are drawn.

2. VIRTUAL INPUT QUEUEING MODEL

2.1 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 of fixed size.

We consider a discrete-time virtual input queueing system with N queues of infinite buffer capacity and a single-server, 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 VIQ system is associated with a unique switch port, at most a single departure may occur servicing one of the non-empty queues at each time slot.

Arrivals obey a Bernoulli process and are non-uniformly distributed between the inputs. In order for the system to be stable, we further require that the service rate, μ, exceeds the aggregate rate of arrivals such that

$$\sum_{j=1}^{N} \lambda_j < \mu \qquad (1)$$

Given that the system is stable, we can deduct that for each queue the arrival rate must converge to the departure rate, such that

$$\lambda_i = \mu_i (1 - \pi_0^{(i)}) \qquad (2)$$

where μ_i denotes the probability that queue i is serviced given that the queue is non-empty, and $\pi_0^{(i)}$ is the probability that the queue is empty.

The service discipline, S, is modeled as a Markov-modulated process, with a mean service time $E[S]$, which is shared between the queues. Note that when $E[S]$=1, we have a canonical output queueing model with no congestion. However, in the interest of expanding the analysis to address potential congestion on the output ports, we allow $E[S]$ to be smaller than 1.

We employ the Geo/GI/1 queueing model with late arrivals to analyze the queueing behavior of the system. The interarrival times are geometrically distributed by definition (due to the Bernoulli process which generates them), and it will be shown that the bursty service discipline results in generally distributed independent service times. Letting S denote the service time random variable under stationary conditions, it has been shown that for a Geo/GI/1 queueing system, the mean queue size, $E[Q]$, is given by [6]

$$E[Q] = \rho + \frac{\lambda^2}{2(1-\rho)}\left\{E[S^2] - E[S]\right\} \qquad (3)$$

where $\rho = \lambda E[S]$, $E[S] = 1/\mu$ (mean service time) and $E[S^2]$ is the second moment of the service time random variable. Thus, for each VIQ, we are required to obtain the first and second moments of its service times, in order to establish its mean queue size. Using Little's theorem [7], we can then find the average waiting time for queue i as $E[T_i]=E[Q_i]/\lambda_i$.

2.2 Single-Queue Model with Markov-Modulated Service Events

We first investigate the case whereby there is only one queue in the system. Consider a discrete-time, two-state Markov chain generating service events modeled by an ON/OFF source which alternates between the ON and OFF states. Let the parameters p and q denote the probabilities that the Markov chain remains in states ON and OFF, respectively. A service event occurs in each time slot that the Markov chain spends in the ON state. The result is a stream of correlated bursts of service events and silent periods, both of which are geometrically distributed.

It can easily be shown that the parameters p and q are interchangeable with the mean service rate, $\mu=(1-q)/(2-q-p)$, and mean service interval duration, $B=1/(1-p)$. Consequently, the mean service time is identical to the steady-state portion of the time the chain spends in state ON. In order to find the first and second moment of the service time, we first find the probability generating function of the service times. Letting f_n denote the probability that n time slots separate two consecutive service events, it can be shown that the service times are distributed as

$$f_n = \begin{cases} p & n = 1 \\ (1-p)q^{n-2}(1-q) & n > 1 \end{cases} \qquad (4)$$

with the corresponding probability generating function (p.g.f),

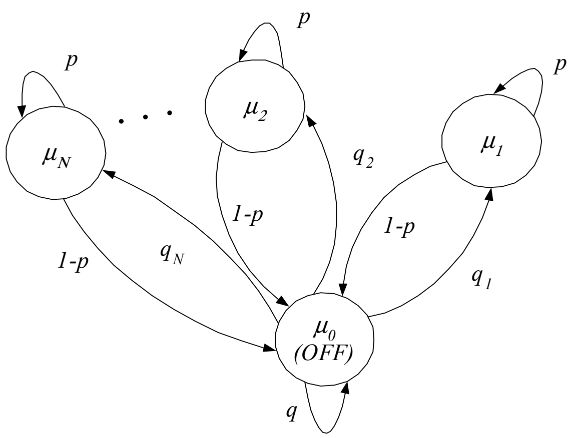

Figure 2. The Markov chain governing the service discipline. Each queue receives a service portion of μ_i.

$$F(z) = pz + (1-p)(1-q)\frac{z^2}{1-qz} \cdot \quad (5)$$

In order to find the first and second moment of the service times, we use the following well-known identities

$$\frac{dF(z)}{dz}\Big|_{Z=1} = E[S] = \frac{2-p-q}{1-q} = 1/\mu \quad (6)$$
$$\frac{dF^2(z)}{dz^2}\Big|_{Z=1} = E[S^2] - E[S] = \frac{2(1-p)(1+q(2-q))}{(1-q)^2}$$

2.3 VIQ Model with Markov-Modulated Service Events

Next, we extend the foundations presented in section 2.2 to address the case of multi-queue architecture. Letting N denote the number of queues, a burst is defined as a sequence of consecutive service events granted to the same queue. We further define the service for each queue by the portion of the overall service rate it receives, μ_k ($k=1,2,..N$), and a mean service interval duration size, B. We construct a Markov chain corresponding to the behavior of the service process, as illustrated in figure 2. The chain consists of N+1 states, N of which represent service events granted to the N queues, while the last state is the OFF (idle) state. We label the ON states as $\kappa_1,\kappa_2,\ldots\kappa_N$, and the OFF state as κ_0. The probability of remaining in the OFF state is q while the probability of remaining in each of the ON states is p.

To complement the latter, the probability of returning from any ON state to the OFF state is (1-p) while a transition from the OFF state to any of the ON occurs with probability q_i. Thus, we can represent the Markov chain via a $(N+1)\times(N+1)$ transition probability matrix, P, where each element, p_{ij}, denotes the probability of transitioning from state i to state j.

The first row of P, with the exception of its first element, consists of the probabilities of transitioning from the OFF state to each of the ON states, signifying a beginning of a service burst. The first column, with the exception of its first element, contains the probability of returning from each of the ON states to the OFF state (i.e. terminating of a burst). The first element on the diagonal is the probability of remaining in the OFF state while the rest of the diagonal elements are the probabilities of remaining in the ON states. Accordingly, the examined transition probability matrix is

$$P = \begin{pmatrix} q & q_1 & q_2 & \cdots & q_N \\ 1-p & p & 0 & \cdots & 0 \\ 1-p & 0 & p & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1-p & 0 & 0 & \cdots & p \end{pmatrix} \quad (7)$$

In equilibrium, we observe that any pair (κ_0, κ_i) must satisfy $\mu_i(1-p) = (1-\mu)q_i$. Consequently, by finding μ_i and q_i, we may fully construct P. Each element q_i is a product of three conditions: (1) the server is activated (transition from the OFF state to any of the ON states), (2) queue i is non-empty and (3) queue i is randomly selected from the set non-empty queues (given that it is non-empty). This product can be written as

$$q_i = (1-q)\times\left(1-\pi_0^{(i)}\right)\times\left(N - \sum_{j\neq i}\pi_0^{(j)}\right)^{-1} \cdot \quad (8)$$

From the state diagram and the equilibrium that must exist between any pair of states, we obtain N-1 independent linear equations for the N variables $\pi_0^{(i)}$. The missing equation is $\sum_{i=1}^{N}\pi_0^{(i)} = 1$. Substituting the latter in (2) directly yields μ_i. We would like to find, for each queue, the p.g.f. of the service times distribution. The latter is done by utilizing the k-step transition matrix, $P^{(k)}$, in which each element, $p_{ij}^{(k)}$, represents the probability of transitioning from the i^{th} state to the j^{th} state in precisely k-steps, with no restrictions made on passing through state j in any of the intermediate steps. In accordance with the Chapman-Kolmogorov equation [6] we have $P^{(k)} = P^k$ ($k\geq 1$), for which the matrix-form p.g.f. is

$$P(z) = \sum_{n=0}^{\infty}(zP)^n = [I - zP]^{-1} \quad (9)$$

where $|z|<1$. We next define the k-step *first passage time* probability matrix [6], $F^{(k)}$, the elements of which, $f_{ij}^{(k)}$, are the probabilities of transitioning from state i to state j in *precisely* k-steps with the constraint that prior to the k^{th}-step the process has not visited state j. In other words, $f_{ij}^{(k)}$ denotes the probability of the first transition from state i to state j occurs in precisely k steps. It can be shown that $f_{ij}^{(k)}$ is

$$f_{ij}^{(k)} = \sum_{s_1 \neq j} \sum_{s_2 \neq j} \cdots \sum_{s_{k-1} \neq j} p_{is_1} p_{s_1 s_2} \cdots p_{s_{k-2} s_{k-1}} p_{s_{k-1} j} \qquad (10)$$

with $f_{ij}^{(1)} = p_{ij}$ and the s terms are the intermediate states between i and j. Since each diagonal element, $f_{ii}^{(k)}\big|_{i>1}$, is by definition the probability of k steps separating two consecutive service events to queue i, it is identical to the definition of the service time distribution of the i^{th} queue. It has been shown that the following relationship exists between $p_{ii}(z)$ and $f_{ii}(z)$ [6]:

$$f_{ii}(z) = 1 - \frac{1}{p_{ii}(z)} \qquad (11)$$

Accordingly, as a first step in finding $F_{ii}(z)$, we need to find $P(z) = [I - zP]^{-1}$. Algebraic exploration of the latter yields the following generic result for $p(z)_{ii}$,

$$p(z)_{ii}\big|_{i>1} = \frac{1 - zp_{11} - \sum_{j=2, j\neq i}^{N+1} \varphi_j(z)}{\left[1 - zp_{11} - \sum_{j=2}^{N+1} \varphi_j(z)\right](1 - zp_{ii})}, \qquad (12)$$

where

$$\varphi_j(z) = \frac{z^2 p_{j1} p_{1j}}{1 - zp_{jj}}$$

from which we find $F_{ii}(z)$ using (11). The latter offers the required inter-service time distribution p.g.f., for each of the N queues. Applying this result to (3) for each queue yields the mean queue size, from which we directly obtain the mean waiting time experienced by the respective arriving packets.

3. SUMMARY

In this paper we presented an analytical framework for evaluating the behavior of output queued switches which employ virtual input queueing. Arrivals are independent and non-uniformly distributed between the ports, while a bursty service discipline is applied using a Markov-modulated process. We utilized the probability generating functions of the inter-service times distributions, in the context of Geo/GI/1 queueing models, to derive per-queue expressions for the mean queue size and waiting time. The methodology presented in this paper may be broadened to address additional traffic scenarios and switch architectures.

4. 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] S. T. Chuang, A. Goel, N. McKeown, and B. Prabhakar, “Matching output queueing with a combined input output queued switch,” *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1030–1039, Dec. 1999.

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

[5] Xiaoqiao Meng, Zhenghua Fu and Songwu Lu, "Robust Packet Scheduling in Wireless Cellular Networks", *ACM Mobile Networks and Applications (MONET),* Vol. 9, No. 2, pp. 113-123, April 2004.

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

[7] D. V. Lindley, “The Theory of Queues with a Single Server,” *Proc. Cambridge Philosophy Society*, Vol. 48, pp. 277-289, 1952.G. K. Zipf, *Psycho-Biology of Languages*, Houghton-Miffin, MIT Press, 1965.