Queueing Syst (2011) 68:375–384
DOI 10.1007/s11134-011-9234-1

Optimal scaling of average queue sizes in an input-queued switch: an open problem

Devavrat Shah · John N. Tsitsiklis · Yuan Zhong

Received: 9 May 2011 / Revised: 9 May 2011 / Published online: 30 June 2011
© Springer Science+Business Media, LLC 2011

Abstract We review some known results and state a few versions of an open problem related to the scaling of the total queue size (in steady state) in an $n \times n$ input-queued switch, as a function of the port number n and the load factor ρ. Loosely speaking, the question is whether the total number of packets in queue, under either the maximum weight policy or under an optimal policy, scales (ignoring any logarithmic factors) as $O(n/(1-\rho))$.

Keywords Input-queued switch · Average queue size · Maximum weight policy

Mathematics Subject Classification (2000) 60K25 · 90B36

1 Introduction

Stochastic processing networks, as formalized by Harrison [3], provide a general model that captures a variety of dynamic resource allocation scenarios. Generally speaking, in such a model there are several queues that need to be served, subject to certain constraints. The performance of such a queuing network is strongly dependent on the policy that determines which queues are to be served at each time slot.

D. Shah (✉) · J.N. Tsitsiklis · Y. Zhong
Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
e-mail: devavrat@mit.edu

J.N. Tsitsiklis
e-mail: jnt@mit.edu

Y. Zhong
e-mail: zhyu4118@mit.edu

The capacity region as well as throughput optimal[1] policies for such queuing networks are reasonably well understood, cf. [11]. However, the development of general performance analysis methods for estimating the distribution or the moments of the queue sizes induced by throughput optimal scheduling policies remains an important challenge.

In this note we put forth a particular performance analysis question. While the development of general analytical results may be too difficult, we focus on a special class of processing networks (input-queued switches) and on asymptotics. More concretely, we are interested in the way that the total queue size (in steady state) scales with the number of ports and with the load factor. Input-queued switches are, in our opinion, the simplest non-trivial example of a stochastic processing network. Over the years, it has served as a guiding example for designing as well as analyzing scheduling policies (cf. [9, 10]). Thus, we hope that making progress on the questions posed in this note will lead to further advances in the performance analysis of more general stochastic processing networks.

2 Input-queued switch model

An input-queued switch is a popular, and commercially available, architecture for switching packets in an Internet router. Abstractly, an $n \times n$ switch has n input ports and n output ports. At each time slot, each input port (respectively, output port) can be matched to at most one output port (respectively, input port) and packets are forwarded according to this matching. See Fig. 1 for an illustration of a 3×3 switch and some possible matchings.

The switch operates in discrete time, indexed by $\tau \in \{0, 1, \ldots\}$. At each time step, and for every port pair (i, j), unit-sized packets may arrive at input port i destined for output port j, according to an exogenous arrival process. Let $A_{i,j}(\tau)$ denote the cumulative number of such arriving packets until the beginning of time slot τ. We assume that the processes $A_{i,j}(\cdot)$ are independent for different pairs (i, j). Furthermore, for every input–output pair (i, j), $A_{i,j}(\cdot)$ is a Bernoulli process with parameter $\lambda_{i,j}$. In particular,

$$\lim_{\tau \to \infty} \frac{1}{\tau} A_{i,j}(\tau) = \lambda_{i,j}, \quad \text{with probability 1.}$$

Let $\boldsymbol{\lambda} = [\lambda_{i,j}] \in [0, 1]^{n \times n}$ denote the arrival rate vector. For every input–output pair (i, j), the associated arriving packets are stored in separate queues, so that we have a total of n^2 queues. Let $Q_{i,j}(\tau)$ be the number of packets waiting at input port i, destined for output j, at the beginning of time slot τ; let $\mathbf{Q}(\tau) = [Q_{i,j}(\tau)]$.

In each time slot, the switch can transmit a number of packets from input ports to output ports, subject to the following two constraints: (i) each input port can transmit at most one packet; and, (ii) each output port can receive at most one packet. In other

[1] For our purposes, a (Markovian) policy is called throughput optimal if the resulting Markov chain is positive recurrent whenever the network is underloaded.

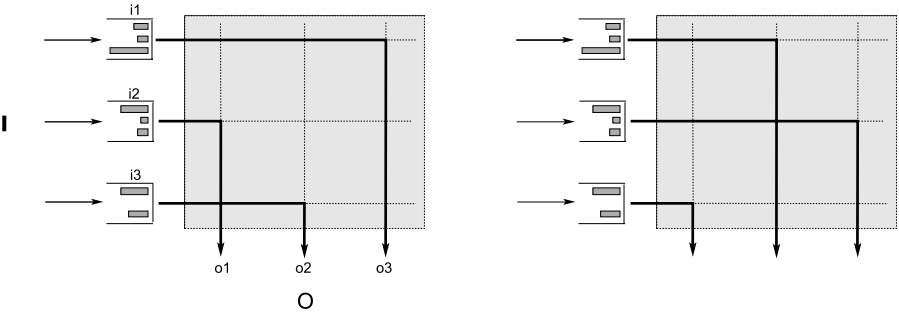

Fig. 1 An input-queued switch, and two example matchings of inputs to outputs

words, the actions of a switch at a particular time slot constitute a *matching* between input and output ports.

A matching, or *schedule*, can be described by a vector $\boldsymbol{\pi} \in \{0,1\}^{n\times n}$, where $\pi_{i,j} = 1$ if input port i is matched to output port j, and $\pi_{i,j} = 0$ otherwise. Thus, the set of all feasible schedules is

$$\mathcal{S} = \left\{ \boldsymbol{\pi} \in \{0,1\}^{n\times n} : \sum_k \pi_{i,k} \le 1,\ \sum_k \pi_{k,j} \le 1,\ \forall (i,j) \text{ with } 1 \le i,j \le n \right\}.$$

A scheduling policy (or simply *policy*) is a rule that, at any given time τ, chooses a schedule $\boldsymbol{\sigma}(\tau) = [\sigma_{i,j}(\tau)] \in \mathcal{S}$, based on the current queue vector $\mathbf{Q}(\tau)$. If $\sigma_{i,j}(\tau) = 1$ and $Q_{i,j}(\tau) > 0$, then one packet is removed from the queue associated with the pair (i,j). For simplicity we have restricted to so-called stationary Markovian policies. Under our restriction, for any given policy, $\mathbf{Q}(\cdot)$ is a Markov chain.

Regarding the details of the model, we adopt the following conventions. At the beginning of time slot τ, the queue vector $\mathbf{Q}(\tau)$ is observed by the policy. The schedule $\boldsymbol{\sigma}(\tau)$ is applied in the middle of the time slot. Finally, at the end of the time slot, the new arrivals occur. Mathematically, for all i, j, and $\tau \ge 0$, we have

$$Q_{i,j}(\tau+1) = Q_{i,j}(\tau) - \sigma_{i,j}(\tau)\mathbf{1}_{\{Q_{i,j}(\tau)>0\}} + A_{i,j}(\tau+1) - A_{i,j}(\tau). \qquad (1)$$

Without loss of generality, we can assume that $Q_{i,j}(0) = 0$, for all i, j.

2.1 Performance metrics

The overall performance goal of a scheduling policy is to keep the queue sizes small. The primary objective is usually to ensure the positive recurrence of the resulting Markov chain, for the largest possible set of arrival rates. This is because positive recurrence guarantees the existence of a unique stationary distribution and ergodicity (so that the queue sizes are prevented from drifting to ever increasing values).

To understand the nature of this primary objective, we note that since any scheduling policy must choose schedules or actions from $\mathcal{S}$, the resulting (time-average) service rate vector $\boldsymbol{\mu} = [\mu_{i,j}]$ must belong to the convex hull of $\mathcal{S}$. By the Birhoff–von

Neumann theorem [1, 12], this convex hull is the same as the set

$$\Lambda = \left\{ \boldsymbol{\pi} \in [0,1]^{n \times n} : \sum_k \pi_{i,k} \le 1, \ \sum_k \pi_{k,j} \le 1, \ \forall (i,j) \text{ with } 1 \le i, j \le n \right\}.$$

We define the *load factor*[2] associated with a given arrival rate vector $\boldsymbol{\lambda}$ to be

$$\rho(\boldsymbol{\lambda}) = \max_{1 \le i,j \le n} \left\{ \sum_k \lambda_{i,k}, \ \sum_k \lambda_{k,j} \right\}.$$

Clearly, if $\rho(\boldsymbol{\lambda}) > 1$, then the arrival rate vector $\boldsymbol{\lambda}$ does not belong to the set Λ of feasible service rate vectors. Thus, services cannot keep up with arrivals, and the system cannot be positive recurrent. On the other hand, if $\rho(\boldsymbol{\lambda}) < 1$, then the arrival rate vector $\boldsymbol{\lambda}$ can be accommodated by a suitable combination of matchings (with some extra margin to accommodate stochastic fluctuations). As a result, for every $\boldsymbol{\lambda}$ for which $\rho(\boldsymbol{\lambda}) < 1$, there exists a policy that results in a positively recurrent Markov chain. Interestingly, it turns out that one can find a *single* policy (independent of $\boldsymbol{\lambda}$) that guarantees this positive recurrence property [4, 11]. We call such policies *throughput optimal.*

Besides throughput optimality, an important secondary performance metric is the average queue size in steady state. Specifically, for any given $\boldsymbol{\lambda}$ with $\rho(\boldsymbol{\lambda}) < 1$, we are interested in the least possible value of $\overline{Q} = \mathbb{E}[\sum_{i,j} Q_{i,j}]$. Here, the expectation is with respect to the steady-state distribution of the queue-size vector $\mathbf{Q}$, which is well defined for policies that result in a positive recurrent Markov chain. (For a Markov chain which is not positive recurrent, we just let $\overline{Q} = \infty$.) We let $Q^*(n, \boldsymbol{\lambda})$ denote the optimal (over all policies) value of $\overline{Q}$, for given n and $\boldsymbol{\lambda}$.

Obtaining analytical expressions or somewhat detailed bounds on $Q^*(n, \boldsymbol{\lambda})$ seems to be very difficult. For this reason, we will focus on the asymptotics of $Q^*(n, \boldsymbol{\lambda})$, in the limit as $n \to \infty$ and $\rho(\boldsymbol{\lambda}) \to 1$.

2.2 The maximum weight scheduling policy

The maximum weight (MW, for short) scheduling policy was introduced in [11] and then studied in the context of input-queued switches in [4]. Under this policy, the schedule $\boldsymbol{\sigma}(\tau)$ chosen at time slot τ satisfies

$$\boldsymbol{\sigma}(\tau) \in \arg\max_{\boldsymbol{\pi} \in \mathcal{S}} \sum_{i,j} \pi_{i,j} Q_{i,j}(\tau),$$

breaking ties according to some prespecified rule. We note that the MW policy is stationary and Markovian, and does not require knowledge of the value of $\boldsymbol{\lambda}$. It is known to result in a positive recurrent Markov chain whenever $\rho(\boldsymbol{\lambda}) < 1$, and is therefore throughput optimal.

[2]This definition coincides with the natural definition of the load factor when the arrival streams are deterministic, as in the "static planning problem" in [3].

3 Problem statement

The basic problem of interest is to identify the best possible simultaneous dependence of $\overline{Q}$ on n and $\rho = \rho(\boldsymbol{\lambda})$. Loosely speaking, the issue is the following.[3] As discussed in the next section, there exist policies that attain $\overline{Q} = O(n^2/(1-\rho))$ and $\overline{Q} = O(n \log n/(1-\rho)^2)$. The question is whether there exist policies that combine the best features of the above two bounds, i.e., with $\overline{Q} = O(n^{1+\varepsilon}/(1-\rho))$ for arbitrarily small $\varepsilon > 0$, and whether this is achieved by the MW policy. A slightly different way of framing the question is to ask for the best possible scaling as a function of n, when we restrict to policies for which the dependence on ρ scales as $1/(1-\rho)$.

There are a variety of ways of formalizing the above questions. We state a few below.

1. Find β_0^*, the infimum over all positive numbers β for which there exists a constant $c > 0$ such that

$$Q^*(n, \boldsymbol{\lambda}) \leq c \cdot \frac{n^\beta}{1-\rho(\boldsymbol{\lambda})}, \tag{2}$$

   for all n and all $\boldsymbol{\lambda}$ with $\rho(\boldsymbol{\lambda}) \in (0, 1)$.
2. Find β_1^*, the infimum over all positive numbers β for which there exists a constant $c > 0$ such that

$$\limsup_{\rho(\boldsymbol{\lambda})\to 1} \big(1-\rho(\boldsymbol{\lambda})\big) \cdot Q^*(n, \boldsymbol{\lambda}) \leq c \cdot n^\beta, \tag{3}$$

   for all n. In the above, the limit superior is taken along sequences of $\boldsymbol{\lambda}$ that satisfy $\rho(\boldsymbol{\lambda}) < 1$.
3. Find β_2^*, the infimum over all positive numbers β for which there exists a constant $c > 0$ such that

$$\limsup_{n\to\infty} \frac{1}{n^\beta} \cdot Q^*(n, \boldsymbol{\lambda}) \leq \frac{c}{1-\rho(\boldsymbol{\lambda})}, \tag{4}$$

   for all $\boldsymbol{\lambda}$ that satisfy $\rho(\boldsymbol{\lambda}) < 1$.

It can be seen that (2) is a stronger requirement than (and thus implies) (3) and (4). For this reason, $\beta_1^* \leq \beta_0^*$ and $\beta_2^* \leq \beta_0^*$. As will be discussed shortly, all of these coefficients lie in the interval $[1, 2]$. We conjecture that $\beta_0^* = 1$, which would also imply that $\beta_1^* = 1$ and $\beta_2^* = 1$. The reason for introducing β_1^* and β_2^* is that an intermediate (weaker) conjecture, such as $\beta_1^* = 1$, may be easier to prove.

We also have the much stronger conjecture that β_0^* is equal to one even if we restrict to the MW policy (as opposed to considering optimal policies). In a further variation, the same questions can be posed for the case of uniform traffic, where $\lambda_{i,j} = \rho/n$, for all i and j.

[3] The reason why this discussion is loose is that the $O(\cdot)$ notation, for a function of two parameters, can admit different interpretations.

4 Known results

In this section we review the most relevant available results. We first discuss the reason why $1 \le \beta_i^* \le 2$. Then, in Sect. 4.3, we explain the reason why the exponents of interest are expected to be equal to one if the dependence on $\rho(\boldsymbol{\lambda})$ were to be ignored. To keep the notation simple, we will write ρ instead of $\rho(\boldsymbol{\lambda})$.

4.1 Lower bound: $\beta_i^* \ge 1$

Consider uniform loading $\boldsymbol{\lambda} = [\rho(\boldsymbol{\lambda})/n]$ with $\rho = \rho(\boldsymbol{\lambda}) \in (1/2, 1]$. Consider the aggregate queue size at input port i: $Q_i = \sum_k Q_{i,k}$. It follows from (1) that, for any $\tau \ge 1$,

$$Q_i(\tau) \ge A_i(\tau) - \tau, \tag{5}$$

where $A_i(\tau) = \sum_k A_{i,k}(\tau)$ is the aggregate arrival process at input port i. The random variable $A_i(\tau)$ is binomial with parameters $n\tau$ and ρ/n. It can be checked (either using Stirling's approximation or an argument along the lines of Lemma 2.1 in [8]) that there exists a positive constant $\beta > 0$ (independent of n and ρ) such that for any $\rho \ge 1/2$, any n, and any $\tau \ge 1$,

$$\mathbb{P}\big(A_i(\tau) \ge \rho\tau + \sqrt{\rho\tau}\big) \ge \beta. \tag{6}$$

From (5) and (6), by setting $\tau = \rho(1-\rho)^{-2}/4$, it follows that

$$\mathbb{P}\left(Q_i(\tau) \ge \frac{\rho}{4(1-\rho)}\right) \ge \beta. \tag{7}$$

Furthermore, for any $\tau' \ge \rho(1-\rho)^{-2}/4$, the exact same bound holds for $Q_i(\tau')$ (due to the stationarity of the Bernoulli process). Therefore, the steady-state expectation of Q_i (if well defined) must be at least $\beta\rho(1-\rho)^{-1}/4$. Due to the symmetry of the uniform traffic, it follows that the steady-state expectation of $\sum_{i,j} Q_{i,j}$ is lower bounded:

$$\mathbb{E}\left[\sum_{i,j} Q_{i,j}\right] \ge C_L \frac{n}{1-\rho}, \tag{8}$$

whenever $\rho \ge 1/2$ and for all n, where $C_L > 0$ is a universal constant.

4.2 Upper bound: $\beta_i^* \le 2$

In order to obtain an upper bound, it suffices to establish an upper bound under a particular policy. The following result is well known; cf. [4, 11]. We include a proof for completeness.

Theorem 1 *Under the maximum weight policy, we have*

$$\overline{Q} \le \frac{n^2}{1-\rho},$$

for all n and all $\boldsymbol{\lambda}$ with $\rho < 1$.

Proof The proof makes use of the Foster–Lyapunov moment bound (cf. [5]). We consider a quadratic Lyapunov function and define $F(\tau) = \frac{1}{2}\sum_{i,j} Q_{i,j}^2(\tau)$. A standard calculation [4, 11] shows that under the MW policy

$$\mathbb{E}\big[F(\tau+1) - F(\tau) \mid \mathbf{Q}(\tau)\big] \le -\frac{1-\rho}{n}\left(\sum_{i,j} Q_{ij}(\tau)\right) + n. \tag{9}$$

Therefore,

$$\mathbb{E}\big[F(\tau+1) - F(\tau)\big] \le -\frac{1-\rho}{n}\mathbb{E}\left[\sum_{i,j} Q_{i,j}(\tau)\right] + n. \tag{10}$$

By summing both sides of (10), for $\tau = 0, \ldots, T$, and with $\mathbf{Q}(0) = \mathbf{0}$, we obtain

$$\frac{1-\rho}{nT}\sum_{\tau=0}^{T-1}\mathbb{E}\left[\sum_{i,j} Q_{i,j}(\tau)\right] \le n. \tag{11}$$

Therefore,

$$\liminf_{T\to\infty}\frac{1}{T}\sum_{\tau=0}^{T-1}\mathbb{E}\left[\sum_{i,j} Q_{i,j}(\tau)\right] \le \frac{n^2}{1-\rho}. \tag{12}$$

Equation (9) and the Foster–Lyapunov criterion imply that $\mathbf{Q}(\cdot)$ is a positive recurrent Markov chain. It is also irreducible and aperiodic. Therefore, $\mathbf{Q}(\tau)$ converges in distribution to a random variable $\mathbf{Q}(\infty)$ that has the steady-state distribution. By Skorohod's representation theorem, $\mathbf{Q}(\tau)$ and $\mathbf{Q}(\infty)$ can be embedded in a common probability space on which $\mathbf{Q}(\tau) \to \mathbf{Q}(\infty)$ almost surely, as $\tau \to \infty$. Then, using Fatou's lemma,

$$\begin{aligned}\overline{Q} &= \mathbb{E}\left[\sum_{i,j}\mathbf{Q}_{ij}(\infty)\right] \\ &\le \liminf_{T\to\infty}\frac{1}{T}\sum_{\tau=0}^{T-1}\mathbb{E}\left[\sum_{i,j} Q_{i,j}(\tau)\right] \le \frac{n^2}{1-\rho}.\end{aligned} \tag{13}$$

□

Theorem 1 readily implies that $\beta_i^* \le 2$, for $i = 0, 1, 2$.

4.3 A batching policy

By comparing the results in the last two subsections, a natural question is whether the dependence of Q^* on n is of order n or n^2. In this subsection, we indicate that a $O(n \log n)$ scaling is possible, using a certain non-Markovian policy. On the other hand, the particular policy that leads to an $O(n \log n)$ bound has an undesirable dependence on ρ. In light of this, our open problem is essentially whether some other

policy can achieve an $O(n^{1+\varepsilon})$ scaling (for arbitrarily small $\varepsilon > 0$), without causing an undesirable dependence on ρ.

We will derive an upper bound by using a batching policy. Such a policy was first considered by Neely, Modiano, and Cheng [7], who established an $O(n \log n)$ upper bound for any fixed ρ, but without studying explicitly the detailed dependence of the upper bound on ρ. Here, we present a slight variant of the policy in [7], with a somewhat tighter analysis of the dependence on ρ. Without loss of generality, we assume that $1/2 \le \rho < 1$.

The batching policy operates as follows. Given $\boldsymbol{\lambda}$, choose a batching interval length equal to

$$T = \frac{25\rho \log n}{(1-\rho)^2}. \tag{14}$$

The policy serves all the packets that arrive during the interval $[kT, (k+1)T)$ (the "kth batch") separately for each $k \ge 0$. That is, the policy collects the packets in the kth batch; it starts serving them after the batching interval has elapsed (that is, after time $(k+1)T$), and after having served all packets in the $(k-1)$st batch. To keep the proof simple, we shall also require that each batch is served for at least Z time slots, where

$$Z = \rho T + 3\sqrt{T \log n}. \tag{15}$$

Let $L(k) = [L_{i,j}(k)]$ be a matrix whose typical entry, $L_{i,j}(k)$, equals the number of packets that arrived at input i, destined for output j, during the kth batch. For every i and j, let $R_i(k) = \sum_j L_{i,j}(k)$ and $C_j(k) = \sum_i L_{i,j}(k)$. Then, $R_i(k)$ (respectively, $C_j(k)$) is the sum of nT Bernoulli random variables, with mean $\mu_i \le \rho T$ (resp. $\mu_j \le \rho T$). Using a suitable variant of the Chernoff bound (see [6]), it follows that

$$\mathbb{P}\big(R_i(k) \ge \mu_i + \sqrt{\mu_i}(\sqrt{4\log n} + K)\big) \le \frac{1}{n^2} \exp\left(-\frac{K^2}{2}\right), \tag{16}$$

for any $K \ge 1$. Using (16) and the union bound, we obtain that

$$\mathbb{P}\Big(\max_{i,j}\big\{R_i(k), C_j(k)\big\} \ge \rho T + \sqrt{\rho T}(\sqrt{4\log n} + K)\Big) \le \frac{2}{n} \exp\left(-\frac{K^2}{2}\right), \tag{17}$$

for all $K \ge 1$.

A well known corollary of the Birkhoff–von Neumann theorem [1, 12] asserts that the total time required to serve all of the packets in the kth batch, is equal to $L^*(k) = \max_{i,j}\{R_i(k), C_j(k)\}$ (cf. [2]). Therefore, the service time $S(k)$ of the kth batch is $S(k) = \max\{L^*(k), Z\}$.

Using (17), an elementary calculation, and the assumption $\rho < 1$, it follows that for large enough n, there exist universal positive constants c_1, c_2 (independent of $\boldsymbol{\lambda}$, n, and T) such that

$$\rho T + 4\sqrt{T \log n} \le T - c_1(1-\rho)T, \tag{18}$$

$$\mathbb{E}\big[S(k)\big] \le \rho T + 4\sqrt{T \log n}, \tag{19}$$

$$\mathbb{E}\big[S^2(k)\big] \le \big(\rho T + 3\sqrt{T \log n}\big)^2 + c_2 T. \tag{20}$$

The inequality in (18), which is critical in the development that follows, made use of the definition of T, in (14). By definition $S(k) \geq Z = \rho T + 3\sqrt{T \log n}$. Using this inequality, together with (20), we obtain

$$\mathrm{var}\big(S(k)\big) \leq c_2 T. \tag{21}$$

Note that to obtain this particular variance bound of $S(k)$, we used the convenient requirement of service time being at least Z.

Under the batching policy, the resulting queue sizes are the same as if all of the arrivals in the kth batch were to arrive simultaneously at time $(k+1)T$. Thus, we can aggregate the arrivals in a batch and view them as an arrival of a single job, with a random processing time of $S(k)$. We are then faced with a D/G/1 queue, with interarrival times equal to T. We can now apply Kingman's upper bound on the waiting time of a batch, in steady state. (The waiting time is the time it takes between the arrival of the batch, until the beginning of the service of the batch.) Because the interarrival times have zero variance, Kingman's bound takes the form

$$\frac{\mathrm{var}(S(k))}{2(T - \mathbb{E}[S(k)])} \leq \frac{c_2 T}{2c_1(1-\rho)T} \leq c_4 T, \tag{22}$$

for some new absolute constant c_4.

Now, the waiting time of a packet is the sum of three contributions: (i) the time from the arrival of the packet until the end of the interval $[kT, (k+1)T)$ during which the packet arrived (and when the batch arrival gets recorded); (ii) the waiting time of the batch; (iii) the time from the beginning of the service of the batch until the packet gets served. The first contribution is bounded above by T. The second contribution is bounded above (in expectation) by $c_4 T$ (cf. (22)). The third contribution is somewhat more subtle, because a "typical" packet is more likely to belong to an uncharacteristically larger batch. Renewal theory (or the so-called random incidence formula) show that the expected service time of the batch that a typical arriving packet belongs to is equal to $\mathbb{E}[S(k)^2]/\mathbb{E}[S(k)]$. Using $S(k) \geq \rho T + 3\sqrt{T \log n}$ and (20), this term is also bounded above by a constant time T. We conclude that the waiting time of a typical packet is bounded above by cT, for some absolute constant c. Using Little's law, and the fact that the total arrival rate is bounded above by n, we obtain

$$\overline{Q} \leq cnT \leq c' \frac{n \log n}{(1-\rho)^2},$$

for some new absolute constant c'.

Note that the batching policy is not Markovian: its action at each time depends in a complicated manner on all of the past history, not just the current queue vector. On the other hand, it is often the case in dynamic programming theory that Markovian policies are no inferior to general policies. We are not aware of existing results of this kind that would apply directly to the problem at hand, but we conjecture this property to be true, so that Markovian policies can also deliver $O(n^{1+\epsilon})$ performance (for any $\epsilon > 0$) when ρ is held fixed.

Acknowledgements D. Shah would like to acknowledge numerous conversations on this topic with D. Wischik and B. Prabhakar over a period of several years. This research was partially supported by the NSF under grant CCF-0728554.

References

1. Birkhoff, G.: Tres observaciones sobre el algebra lineal. Univ. Nac. Tucumán. Rev, Ser. A **5**, 147–151 (1946)
2. Hajek, B., Sasaki, G.: Link scheduling in polynomial time. IEEE Trans. Inf. Theory **34**(5), 910–917 (1988)
3. Harrison, J. Michael: Brownian models of open processing networks: canonical representation of workload. Ann. Appl. Probab. **10**, 75–103 (2000)
4. McKeown, N., Anantharam, V., Walrand, J.: Achieving 100% throughput in an input-queued switch. In: Proceedings of IEEE Infocom, pp. 296–302 (1996)
5. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, New York (1993)
6. Motwani, R., Raghavan, P.: Randomized Algorithms. Cambridge University Press, Cambridge (1995)
7. Neely, M., Modiano, E., Cheng, Y.-S.: Logarithmic delay for $n \times n$ packet switches under the cross-bar constraint. IEEE/ACM Trans. Netw. **15**(3) (2007)
8. Shah, D., Tsitsiklis, J.N.: Bin packing with queues. J. Appl. Probab. **45**, 922–939 (2008)
9. Shah, D., Wischik, D.J.: Optimal scheduling algorithms for input-queued switches. In: Proceedings of IEEE Infocom (2006)
10. Shah, D., Wischik, D.J.: Switched networks with maximum weight policies: fluid approximation and multiplicative state space collapse. Ann. Appl. Probab. (2011, to appear)
11. Tassiulas, L., Ephremides, A.: Stability properties of constrained queuing systems and scheduling policies for maximum throughput in multihop radio networks. IEEE Trans. Autom. Control **37**, 1936–1948 (1992)
12. von Neumann, J.: A certain zero-sum two-person game equivalent to the optimal assignment problem. Contrib. Theory Games **2**, 5–12 (1953)