Maximum Size Matching is Unstable for Any Packet Switch

Isaac Keslassy, *Member, IEEE*, Rui Zhang-Shen, *Member, IEEE*, and Nick McKeown, *Senior Member, IEEE*

***Abstract*—Input-queued packet switches use a matching algorithm to configure a nonblocking switch fabric (e.g., a crossbar). Ideally, the matching algorithm will guarantee 100% throughput for a broad class of traffic, so long as the switch is not oversubscribed. An intuitive choice is the maximum size matching (MSM) algorithm, which maximizes the instantaneous throughput. It was shown, by McKeown *et al.* in 1999, that with MSM the throughput can be less than 100% when $N \geq 3$, even with benign Bernoulli i.i.d. arrivals. In this letter, we extend this result to $N \geq 2$, and hence show it to be true for switches of any size.**

***Index Terms*—Instability, maximum size matching (MSM), switching algorithms.**

I. Introduction

HIGH-SPEED Internet routers commonly use virtual output queueing (VOQ), a crossbar switch, and (internally) fixed-size cells. Time is slotted with one cell transmission per time-slot. At each time-slot a matching algorithm finds a match between N inputs and N outputs ($N \geq 2$, since there is no need for a matching algorithm when $N = 1$), and cells are transferred according to this match.

This letter is about switches that are unstable even though no input or output is over-subscribed. It is known that for a broad class of traffic, a switch is stable (for $N \geq 2$) if the maximum weight matching (MWM) algorithm is used [1], [3]. On the other hand, it is known that with the maximum size matching (MSM) algorithm, a switch can be unstable for $N \geq 3$ [1] (if ties are broken randomly).[1] This is surprising because MSM maximizes the instantaneous throughput by transferring the maximum number of cells during each time-slot.

The instability result in [1] is based on a counterexample that holds for $N \geq 3$. In this letter we extend the proof to $N \geq 2$, and hence prove that MSM is unstable for any switch. We also derive the exact throughput formula for the $N = 2$ case. Our results are mainly of theoretical interest (it is unusual to build 2×2 switches); the letter completes existing results, extending them to switches of any size.

Manuscript received March 9, 2003. The associate editor coordinating the review of this paper and approving it for publication was Prof. C. Douligeris.

The authors are with the Computer Systems Laboratory, Stanford University, Stanford, CA 94305-9030 USA (e-mail: keslassy@stanford.edu; rzhang@stanford.edu; nickm@stanford.edu).

Digital Object Identifier 10.1109/LCOMM.2003.817330

[1]We assume here that MSM breaks ties randomly. In [2] it is shown that otherwise, MSM could be stable for $N \geq 2$.

II. Problem Statement

We will consider a packet switch with two inputs and two outputs, i.e., $N = 2$.

Notation: Time-slot t represents the interval $[t-1, t)$. Let Q_{ij} denote the VOQ at input i destined to output j. Q_{ij} contains $L_{ij}(t)$ packets at the end of time-slot t, with $L_{ij}(0) = 0$ for all i, j by convention. $A_{ij}(t)$ packets arrive at Q_{ij} at the beginning of time-slot t and $D_{ij}(t)$ packets depart from it at the end of the time-slot, with $0 \leq A_{ij}(t)$, $D_{ij}(t) \leq 1$. The service indicator $S_{ij}(t)$ is 1 if Q_{ij} is serviced at time t, and 0 otherwise. There is a departure from Q_{ij} if it both receives a service and is nonempty. As a consequence, for $t \geq 1$, $L_{ij}(t)$ satisfies the following equation:

$$\begin{aligned} L_{ij}(t) =& L_{ij}(t-1) + A_{ij}(t) - D_{ij}(t) \\ =& [L_{ij}(t-1) + A_{ij}(t) - S_{ij}(t)]^+ \end{aligned} \tag{1}$$

where the notation $[x]^+$ is equivalent to $\max\{x, 0\}$.

Arrivals: For our counterexample, it is sufficient to assume that the arriving traffic follows a Bernoulli i.i.d. distribution with mean rate λ_{ij} arriving to Q_{ij}. We will consider the following type of traffic:

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} a & b \\ b & 0 \end{bmatrix} \tag{2}$$

where a and b are positive constants. It is assumed that no input or output is oversubscribed, i.e., $a + b < 1$.

Services: VOQs are serviced according to a MSM algorithm with ties broken randomly.

Stability: A queue is said to be *unstable* if after a finite time, its occupancy never returns to zero with probability one. Note that with Bernoulli traffic, this is implied by the queue having a positive drift, which happens if the service rate is less than the incoming traffic rate. A switch is said to be *unstable* if any of its queues is unstable.

III. Instability of MSM When $N = 2$

Our approach is to find values of a and b such that the service rate of Q_{11} is less than its arrival rate.

Lemma 1: At the end of a time-slot t, at least one of the two queues Q_{12} and Q_{21} is empty

$$L_{12}(t) \cdot L_{21}(t) = 0. \tag{3}$$

Proof: By induction. The case when $t = 0$ is clear. Assume that this property holds until the end of some time-slot $t \geq 0$. Consider the following two cases.

Case 1) At least one of the two queues is empty after the arrivals at time-slot $t+1$. It will still be empty after departures, hence the property holds for time-slot $t+1$.

Case 2) Both queues are occupied after the arrivals in time-slot $t+1$. MSM will choose the configuration with size two that serves Q_{12} and Q_{21} (because Q_{22} is always empty). By assumption, at least one of the two queues was empty at the end of time-slot t, therefore this queue will be empty after the service, and the property holds for $t+1$. ■

Let $x(t)$ be the service probability of Q_{11} at time-slot t, i.e., $x(t)=\Pr\{S_{11}(t)=1\}$, and $p(t)$ the probability that both Q_{12} and Q_{21} are empty at time-slot t, i.e., $p(t) = \Pr\{L_{21}(t) = L_{12}(t)=0\}$. The following two lemmas provide upper bounds on $x(t)$ and $p(t)$ that will be useful for showing the instability of MSM.

Lemma 2: (Bound on $x(t)$) *For any* $t \geq 1$,

$$x(t) \leq \frac{(1-b)(1+p(t-1))}{2}. \quad (4)$$

Proof: Let $t \geq 1$, and consider whether Q_{11} is served at time-slot t. There are two cases.

Case 1) With probability $p(t-1)$, $L_{12}(t-1) = L_{21}(t-1) = 0$. There are three possibilities. If both Q_{12} and Q_{21} have arrivals, which happens with probability b^2, Q_{11} will not be served. With probability $2b(1-b)$, only one of the two queues $\{Q_{12}, Q_{21}\}$ has an arrival, then Q_{11} is served with probability 1/2, if it is nonempty. Finally, with probability $(1-b)^2$, neither of $\{Q_{12}, Q_{21}\}$ has an arrival, in which case the probability that Q_{11} is served cannot exceed 1. Hence, an upper bound on the probability that Q_{11} receives service in this case is: $\Pr(S_{11}(t)=1 \mid L_{12}(t-1) = L_{21}(t-1)=0) \leq 2b(1-b)\cdot 1/2 + (1-b)^2 = 1-b$.

Case 2) With probability $1-p(t-1)$, $L_{12}(t-1)+L_{21}(t-1) > 0$. By Lemma 1 , only one of the two queues $\{Q_{12}, Q_{21}\}$ is nonempty. There are two possibilities. If the empty one has an arrival (which occurs with probability b), then Q_{11} will not be served. However, if the empty one does not have an arrival and remains empty (with probability $1-b$), then Q_{11} is served only if it is nonempty, and then only with probability 1/2. Thus, we get the upper bound of $x(t)$ in this case, $\Pr(S_{11}(t)=1 \mid L_{12}(t-1)+L_{21}(t-1) > 0) \leq (1-b)/2$.

Combining the two cases yields $x(t) \leq p(t-1)\cdot(1-b) + (1-p(t-1))\cdot(1-b)/2 = (1-b)(1+p(t-1))/2$. ■

Lemma 3: (Bound on $p(t)$) For any $t \geq 1$,

$$p(t) \leq 1 - \frac{a \cdot b}{2}. \quad (5)$$

Proof: Let $t \geq 1$. We will show that $\Pr(L_{12}(t)+L_{21}(t) > 0) \geq (a\cdot b)/2$, by considering two cases.

Case 1) $L_{12}(t-1) = 0$. Consider the following possible succession of events: $A_{11}(t) = A_{21}(t) = 1$ (which implies $A_{12}(t)=0$), and $S_{11}(t)=1$. This succession of events happens with probability $a\cdot b\cdot 1/2$, and after this succession of events it is clear that $L_{12}(t)+L_{21}(t) > 0$. We did not consider other possible events, therefore, $\Pr(L_{12}(t)+L_{21}(t) > 0 \mid L_{12}(t-1)=0) \geq ab/2$.

Case 2) $L_{12}(t-1) > 0$. In this case, Q_{12} will remain nonempty as long as $A_{12}(t) = 1$, which happens with probability b. Therefore, $\Pr(L_{12}(t)+L_{21}(t) > 0 \mid L_{12}(t-1) > 0) \geq b > ab/2$.

Hence $\Pr(L_{12}(t)+L_{21}(t) > 0) \geq ab/2$, proving the lemma.

Theorem 4: MSM is unstable for $N=2$ whenever

$$\frac{1-b}{1+\frac{b(1-b)}{4}} < a < 1-b. \quad (6)$$

Proof: From Lemmas 2 and 3, we get $x(t) \leq \hat{x}$, where $\hat{x} \stackrel{\text{def}}{=} (1-b)(1-ab/4)$, for all $t \geq 1$. If we can find a tuple (a,b) such that $0 < \hat{x} < a < 1-b$, then Q_{11} would have more arrivals than services, and MSM would be unstable. Solving $\hat{x} < a < 1-b$ yields (6), which is true over a nonempty set for any $b \in (0,1)$. ■

For example, when $b=0.5$, the switch is unstable for $a > 8/17 \simeq 0.471$, i.e., for a load $\rho = a+b$ such that $0.971 < \rho < 1$. As we will see shortly, this bound is not tight.

Corollary 5: MSM is unstable for any switch of size $N \geq 2$.

Proof: The case $N \geq 3$ was proved in [1], and $N=2$ in Theorem 4. ■

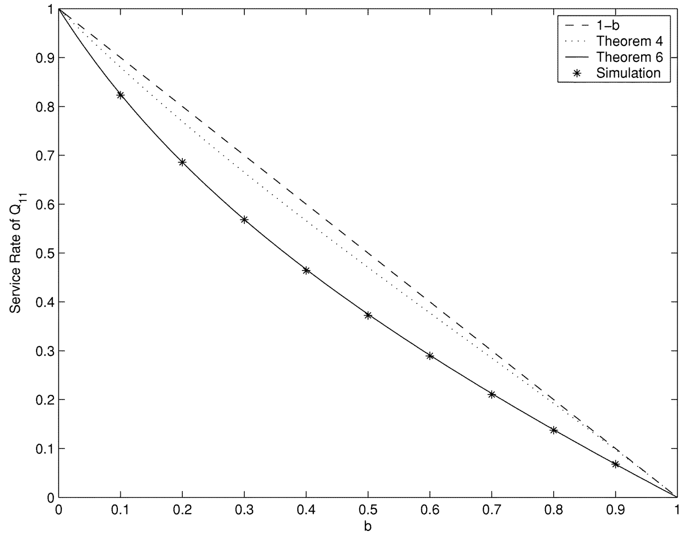

Fig. 1. Throughput of Q_{11} given by Theorem 4, by the exact formula in Theorem 6, and by simulation.

IV. Maximum Throughput of a 2×2 MSM Switch

Even though the objective of this letter was only to prove that the 2×2 MSM switch is unstable, it is actually possible to determine its exact throughput, as shown by the following theorem. The proof of a slightly more general form of the theorem can be found in [4].

Theorem 6: A 2×2 maximum size matching switch under admissible Bernoulli i.i.d. traffic of rate $\begin{pmatrix} a & b \\ b & 0 \end{pmatrix}$ with $a+b<1$ has less than 100% throughput if and only if

$$a > \frac{1-b^2}{1+2b}. \tag{7}$$

For example, when $b = 0.5$, the switch is unstable if and only if $a > 3/8 = 0.375$, compared to the sufficiency condition $a > 0.471$ given by Theorem 4.

Fig. 1 shows the service rate of Q_{11} when $a = 1 - b$, given by Theorem 4, by the exact formula in Theorem 6, and by simulation. As shown in the figure, the theoretical values from Theorem 6 agree extremely well with the simulations.

REFERENCES

[1] N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1260–1267, Aug. 1999.

[2] A. Mekkittikul and N. McKeown, "A practical scheduling algorithm to achieve 100% throughput in input-queued switches," in *Proc. IEEE Infocom*, vol. 2, San Francisco, CA, Apr. 1998, pp. 792–799.

[3] J. G. Dai and B. Prabhakar, "The throughput of data switches with and without speedup," in *Proc. IEEE INFOCOM*, vol. 2, Tel Aviv, Israel, Mar. 2000, pp. 556–564.

[4] R. Zhang-Shen, I. Keslassy, and N. McKeown, "Maximum size matching is unstable for any packet switch," Stanford Univ. High-Performance Networking Group, Stanford, CA, TR03-HPNG-030 100, 2003.