Maximum Size Matching Is Unstable for Any Packet Switch

Isaac Keslassy, Rui Zhang-Shen, Nick McKeown
{keslassy, rzhang, nickm}@stanford.edu
Computer Systems Laboratory
Stanford University
Stanford, CA 94305-9030

***Abstract*— Input-queued packet switches use a matching algorithm to configure a non-blocking 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 in [1] 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*— Maximum Size Matching, switching algorithms, instability.**

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).* 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 counter-example 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.

II. PROBLEM STATEMENT

We will consider a packet switch with 2 inputs and 2 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 non-empty. As a consequence, for $t \geq 1$, $L_{ij}(t)$ satisfies the following equation:

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

Arrivals - For our counter-example, 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:

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} a & b \\ c & 0 \end{bmatrix}, \qquad (2)$$

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

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 assume $b = c$, then 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:*

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

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 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 $t+1$.

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

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* $t \geq 1$,

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

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 non-empty. 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 non-empty. 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 non-empty, 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* $t \geq 1$,

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

Proof: Let $t \geq 1$. We will show that $\Pr(L_{12}(t) + L_{21}(t) > 0) \geq \frac{a \cdot b}{2}$, 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 non-empty 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*

$$\frac{1-b}{1+b(1-b)/4} < a < 1-b \qquad (6)$$

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 Equation (6), which is true over a non-empty 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. ■

IV. Maximum Throughput of a 2×2 MSM Switch

In this section we will derive a closed form formula for the maximum throughput of a 2×2 switch under traffic of the form given in (2). We will do so by comparing two systems, later called the *original* system and the *new* one.

We call the 2×2 maximum size matching switch the *original* system, and denote it by parameters with a superscript (o). By contrast, in the *new* system, all parameters are the same except that a packet generator is attached to Q_{11}. The packet generator generates a packet for Q_{11} at the very beginning, and whenever Q_{11} is empty. Thus, in the new system, Q_{11} is always occupied. This enables us to analyze the evolution of Q_{12} and Q_{21} without considering Q_{11}.

A. The Single Unstable Queue Lemma

First we will prove the following useful lemma which applies to *any* admissible traffic:†

Lemma 6 *Under Bernoulli admissible traffic, a* 2×2 *MSM switch can have at most one unstable queue.*

Proof: Let the doubly strictly substochastic matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ represent the input traffic rate. Suppose that more than one queue is unstable. Without loss of generality, let one of them be Q_{11}.

If Q_{12} is also unstable, let us consider $L_{11} + L_{12}$. After a finite time, both queues are always occupied, so together, they have departures at rate 1. But the incoming traffic rate is $a + b < 1$, so at least one of Q_{11} and Q_{12} should eventually drain to zero. Therefore, Q_{12} cannot be unstable. Similarly, Q_{21} cannot be unstable.

Finally, if both Q_{11} and Q_{22} are unstable, both queues will be permanently occupied after a finite time. Then Q_{12} and

†Note that the new system may not be admissible due to the packet generator, and hence more queues can be unstable than shown in the lemma.

TABLE I

TRANSITIONS OF (L_{12}, L_{21}), THEIR PROBABILITIES, AND THE SERVICE PROBABILITY OF Q_{11}

Arrival (A_{12}, A_{21})	Probability	Initial and final values of (L_{12}, L_{21}): $(0,0)$		$(1,0)$		$(0,1)$	
$(0,0)$	$(1-b)(1-c)$	$(\mathbf{0},\mathbf{0})$		$(0,0)$	$(\mathbf{1},\mathbf{0})$	$(0,0)$	$(\mathbf{0},\mathbf{1})$
$(0,1)$	$(1-b)c$	$(0,0)$	$(\mathbf{0},\mathbf{1})$	$(0,0)$		$(0,1)$	$(\mathbf{0},\mathbf{2})$
$(1,0)$	$b(1-c)$	$(0,0)$	$(\mathbf{1},\mathbf{0})$	$(1,0)$	$(\mathbf{2},\mathbf{0})$	$(0,0)$	
$(1,1)$	bc	$(0,0)$		$(1,0)$		$(0,1)$	
prob. Q_{11} is served		$1-\frac{1}{2}(b+c)$		$\frac{1}{2}(1-c)$		$\frac{1}{2}(1-b)$	

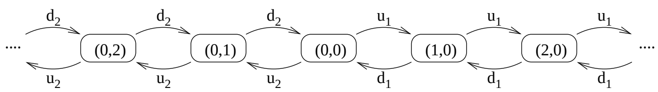

Fig. 1. Markov chain of (L_{12}, L_{21}) in the new system

Q_{21} can only be served if both queues have packets, i.e., at a maximum rate of $\min(b,c)$. Then Q_{11} and Q_{22} are served at a rate no less than $1-\min(b,c) > \max(a,d)$. So Q_{22} cannot be unstable because we've assume Q_{11} to be unstable.

Thus we conclude that at most one queue can be unstable in a 2×2 MSM switch with Bernoulli admissible traffic. ■

The following corollary will enable us to link the new and the original systems.

Corollary 7 *In a 2×2 maximum size matching switch under admissible traffic* $\left(\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right)$*, only* Q_{11} *can be unstable.*

Proof: Assume that Q_{12} is unstable, and thus that after a finite time with probability one its occupancy never returns to zero. Then Q_{11} has to be stable due to Lemma 6. Thus Q_{11} is served at rate a and Q_{12} at rate $1-a$ because it is backlogged. But this is impossible because the incoming traffic rate to Q_{12} is $b < 1-a$. Similarly Q_{21} cannot be unstable either. Thus, only Q_{11} can be unstable. ■

B. Throughput of the New System

We will now study the *new* system by analyzing the Markov chain with state (L_{12}, L_{21}) under the traffic matrix in Equation (2). Table I summarizes the transitions of (L_{12}, L_{21}) over a time-slot as a result of different arrivals. This table only mentions three initial states, because MSM only requires the binary information of whether a queue is occupied or not. Note that there are two possible configurations of the crossbar: either $S_{11} = S_{22} = 1$ and $S_{12} = S_{21} = 0$, or $S_{11} = S_{22} = 0$ and $S_{12} = S_{21} = 1$. When both configurations have an equal number of packets that can depart, i.e. when there is one packet for each configuration, one of them is chosen at random with equal probabilities, resulting in two equally likely final states. The bold entries in the table represent the transitions during which Q_{11} is served. The last row summarizes the service probability of Q_{11} given the initial state (L_{12}, L_{21}).

With at least one of L_{12} and L_{21} always being zero (Lemma 1), the Markov chain of (L_{12}, L_{21}) becomes one dimensional, as shown in Figure 1, where u_1 (respectively u_2) and d_1 (resp. d_2) are the probabilities that Q_{12} (resp. Q_{21}) increases or decreases by 1 during a time-slot. (We have omitted the probability of staying in a state from the figure.) Thus, we have

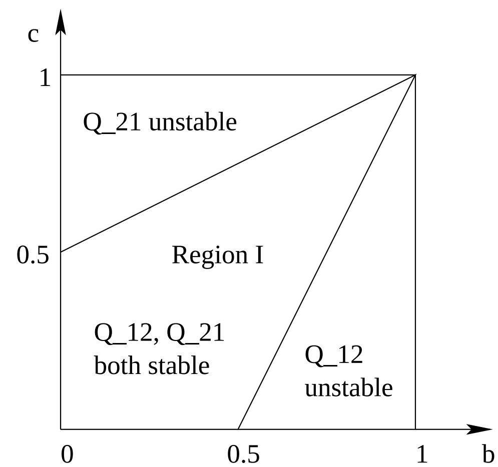

Fig. 2. Stability regions for the *new* system

$$\rho_1 \stackrel{\text{def}}{=} \frac{u_1}{d_1} = \frac{\frac{1}{2}b(1-c)}{\frac{1}{2}(1-b)(1+c)} = \frac{b(1-c)}{(1-b)(1+c)} \tag{7}$$

$$\rho_2 \stackrel{\text{def}}{=} \frac{u_2}{d_2} = \frac{\frac{1}{2}(1-b)c}{\frac{1}{2}(1+b)(1-c)} = \frac{c(1-b)}{(1-c)(1+b)} \tag{8}$$

We will only consider cases when both Q_{12} and Q_{21} are stable, which means $\rho_1 \leq 1$ and $\rho_2 \leq 1$. As shown in Figure 2, solving the two inequalities gives the region:

$$I = \{b \leq (1+c)/2, c \leq (1+b)/2\}. \tag{9}$$

We will now derive the service rate of Q_{11} in region I. Let $\{\pi_{ij}, i \cdot j = 0\}$ be the stationary distribution of the Markov chain for (L_{12}, L_{21}). Solving the normalization equation $\sum \pi_{ij} = 1$, we get:

$$\pi_{00} = \frac{(1-2b+c)(1-2c+b)}{(1-b)(1-c)(1+b+c)} \tag{10}$$

$$\pi_{i0} = \rho_1^i \pi_{00} \tag{11}$$

$$\pi_{0i} = \rho_2^i \pi_{00}. \tag{12}$$

Then the service rate of Q_{11} is then given by:

$$\begin{aligned}\sigma_{11} &= \frac{1}{2}(1-c)\sum_{i=1}^{\infty}\pi_{i0} + \frac{1}{2}(1-b)\sum_{i=1}^{\infty}\pi_{0i} \\ &\quad +(1-\frac{1}{2}(b+c))\pi_{00} \\ &= \frac{1+bc-b^2-c^2}{1+b+c}. \quad (13)\end{aligned}$$

It is easy to verify that $\sigma_{11} < \min(1-b, 1-c)$ in the interior of I. If $a < \sigma_{11}$, the packet generator generates packets at an average rate of $\sigma_{11} - a$. On the other hand, if $a > \sigma_{11}$, packets in Q_{11} will build up at the rate $a - \sigma_{11}$, and with probability one the packet generator will only generate a finite number of packets before the queue becomes permanently occupied.[‡]

C. Throughput of the Original System

In this section we will show that the stability condition for the original system is the same as the new system. We will use *throughput* to denote the expected number of packets leaving the system in each time slot.

Lemma 8 *The throughput of the original system is no greater than the new system.*

Proof: We have shown that only Q_{11} can be unstable in the original system, so after running a long enough time, either all queues are stable, or Q_{11} becomes permanently backlogged with probability one and the other two queues are stable. In either case, (L_{12}, L_{21}) form a Markov chain with the same transitions paths as the new system (as in Figure 1), but with possibly different transition probabilities. Let $p = (1-a)\Pr\{L_{11} = 0\}$ be the probability that Q_{11} is empty after arrivals in a time slot. If Q_{11} is unstable then $p = 0$, otherwise $p > 0$.

Now let us look at the one-step transitions of (L_{12}, L_{21}) and distinguish two cases.

Case 1: With probability $1-p$, Q_{11} is occupied after arrivals, and the transition probabilities will be the same as in the new system.

Case 2: With probability p, Q_{11} is empty. Denote by the superscript *empty* the transitions probabilities in this case. Then $u_1^{empty} = 0 < u_1$, $u_2^{empty} = 0 < u_2$; and $d_1^{empty} = 1-b > d_1$, $d_2^{empty} = 1-c > d_2$.

Therefore, considering both cases leads to $u_i^{(o)} = (1-p)u_i + pu_i^{empty} \le u_i$ and $d_i^{(o)} = (1-p)d_i + pd_i^{empty} \ge d_i$. Thus, by properties of birth-death chains, we have:

$$\pi_{00} \le \pi_{00}^{(o)} \quad (14)$$

$$\sum_{i=1}^{\infty}\pi_{i0} \ge \sum_{i=1}^{\infty}\pi_{i0}^{(o)} \quad (15)$$

$$\sum_{i=1}^{\infty}\pi_{0i} \ge \sum_{i=1}^{\infty}\pi_{0i}^{(o)} \quad (16)$$

Now let $q_i, i = 0, 1, 2, \sum q_i = 1$ denote the probability that i packets are are served in the 2×2 switch during a time slot. Then the expected number of packets served in a time-slot can be calculated by $\eta = q_1 + 2q_2$. Given the queue occupancy at the beginning of a time slot, the probability that two packets are served is c if Q_{12} is occupied (with probability $\sum_{i=1}^{\infty}\pi_{i0}$), b if Q_{21} is occupied (with probability $\sum_{i=1}^{\infty}\pi_{0i}$), and $b \cdot c < b, c$ if both queues are empty (with probability π_{00}). Therefore, the inequalities (14-16) give $q_2 \ge q_2^{(o)}$. In the new system, there is always a non-empty queues, therefore $q_0 = 0 \le q_0^{(o)}$. We can rewrite η as $\eta = q_1 + 2q_2 = (1-q_0) + q_2$. Then it is easy to see that the last two inequalities imply $\eta \ge \eta^{(o)}$, i.e., the original system has a throughput no greater than the new system. ■

[‡]If $a = \sigma_{11}$, the Markov chain for Q_{11} is null recurrent. The queue is stable because the probability of L_{11} returning to zero is one.

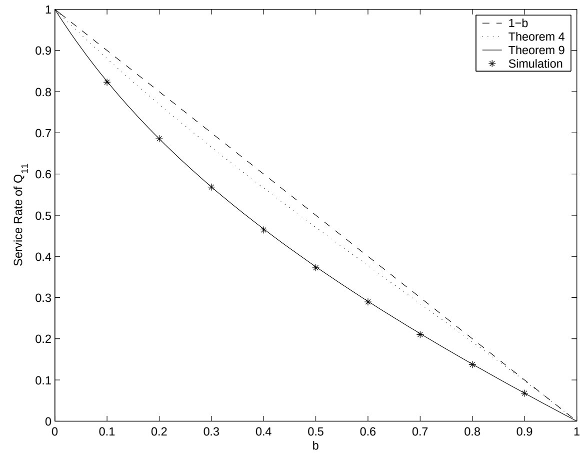

Fig. 3. Throughput of Q_{11} given by formula and by simulation

Theorem 9 *A 2×2 maximum size matching switch under admissible traffic* $\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}$ *is unstable if and only if* $b \le (1+c)/2$, $c \le (1+b)/2$, *and* $a > \sigma_{11} = \frac{1+bc-b^2-c^2}{1+b+c}$.

Proof: In region I (Equation (9)), Q_{12} and Q_{21} are stable in both systems. (They are also stable in the original system because $u_i^{(o)} \le u_i$, $d_i^{(o)} \le d_i$.) By Lemma 8, the original system has a throughput no greater than the new system, therefore we must have

$$\sigma_{11}^{(o)} \le \sigma_{11} = \frac{1+bc-b^2-c^2}{1+b+c}. \quad (17)$$

In the 2×2 MSM switch, if $a > \sigma_{11}$, Q_{11} will have a positive drift, and with probability 1, Q_{11} is never empty after a finite time. Then the switch behaves the same as the one with a packet generator. Thus, the maximum throughput for Q_{11} is σ_{11}. On the other hand, if $a \le \sigma_{11}$, Q_{11} has to be stable. Otherwise, if it is unstable, it must have a service rate less than a, and the queue is never empty after a finite time. By comparison with the new system, the service rate must be σ_{11}, which is a contradiction.

Outside the region I, if Q_{11} is unstable, we again have the two systems having the same throughput for Q_{11}. But one of Q_{12} and Q_{21} is also unstable, which contradicts Corollary 7. So Q_{11} is stable outside of region I. ■

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. Figure 3 shows the service rate of Q_{11} when $b = c$ and $a = 1 - b$, given by Theorem 4, by the exact formula in Theorem 9, and by simulation. As shown in the figure, the theoretical values from Theorem 9 agree extremely well with the simulations.

ACKNOWLEDGMENTS

We wish to thank Balaji Prabhakar, Devavrat Shah, Persi Diaconis, and Arnab Chakraborty for their helpful comments regarding Section IV. We would also like to thank Shang-Tse Chuang, Gireesh Shrimali and Nandita Dukkipati for their comments.

REFERENCES

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

[2] A. Mekkittikul, N. McKeown, "A Practical Scheduling Algorithm to Achieve 100% Throughput in Input-Queued Switches", *Proceedings of IEEE Infocom*, Vol 2, pp. 792-799, San Francisco, April 1998.

[3] J.G. Dai, B. Prabhakar, "The throughput of data switches with and without speedup", *Proceedings of IEEE INFOCOM*, 2:556-564, Tel Aviv, Israel, March 2000.