Greedy weighted matching for scheduling the input-queued switch

Andrew Brzezinski and Eytan Modiano
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
{brzezin,modiano}@mit.edu

***Abstract*— We consider greedy maximal weighted matching based scheduling for input-queued switches. We present simulation results that demonstrate the attractive throughput and delay performance achievable under maximal weighted matching. For the 2×2 input-queued switch we subsequently prove for i.i.d. Bernoulli arrival processes that greedy maximal weighted matching achieves 100% throughput.**

I. INTRODUCTION

We consider the scheduling problem for the $N \times N$ input-queued switch. It is widely held that the $O(N^3)$ computational complexity of maximum weighted matching is overly burdensome for implementation on a slot-by-slot basis in practical systems operating at high rates [1]. Many practitioners resort to suboptimal matching algorithms in conjunction with speedup to provide optimal throughput performance. In this paper, we consider greedy maximal weighted matching as a suboptimal matching algorithm, and we make no use of speedup. We conduct numerical and analytical studies to demonstrate the attractive throughput and delay performance properties of greedy matching based scheduling.

The switch scheduling literature often takes advantage of the simple fact that a greedy matching on a weighted bipartite graph provides a 2-approximation to the weight of the maximum weighted matching, and thus that at least 50% throughput is achievable. Consequently, it is simple to demonstrate that 100% throughput is achievable under greedy matching in conjunction with a speedup of two. Less can be found on the topic of greedy matchings with no speedup [2].

In this paper, we pursue two important goals:

1) To develop numerical simulations that attest to the attractive throughput and delay properties of greedy weighted matching based schedulers; and
2) To prove the throughput optimality of greedy weighted matching based scheduling in the 2×2 input-queued switch under i.i.d. Bernoulli arrival processes.

II. GREEDY MATCHING BASED SCHEDULING

A greedy maximal weighted matching on a complete weighted bipartite graph (S, T, E) with $|S| = |T| = N$ and weight function $w : E \to \mathbb{R}_+$ selects sequentially the maximum weighted edge allowed to be included in the matching. For edge $e \in E$ let $s(e) \in S$ and $t(e) \in T$ denote the S and T nodes corresponding to edge e, respectively. Thus, a greedy maximal weighted matching $M \subset E$ under weight function w is given as follows.

Repeat N times, starting with $M = \{\}$:

$$M \leftarrow M \cup e^*, \tag{1}$$

where

$$e^* \in \underset{\{e: s(e) \neq s(\bar{e})\, \forall \bar{e} \in M,\, t(e) \neq t(\bar{e})\, \forall \bar{e} \in M\}}{\arg\max} w(e). \tag{2}$$

Since the bipartite graph (S, T, E) is complete, the resulting matching clearly contains N edges.

The scheduling scenario consists of an input-queued switch employing virtual output queues (VOQs) for each source-destination pair. Each VOQ contains fixed-size cells awaiting transmission to a particular output port of the switch. We consider slotted time, with index n, and assume that one time slot is required for transmission of any cell across the switch fabric. For $n \geq 0$, we define by $X(n)$ the queue occupancy matrix at time n, with $X_{i,j}(n)$ equal to the number of cells in $\text{VOQ}_{i,j}$ at time n. The cumulative arrival process is defined as $A(n)$, with $A_{i,j}(n)$ equal to the total number of cell arrivals to input port i for destination port j by time slot n. We assume that each arrival process is i.i.d. Bernoulli, with $\text{VOQ}_{i,j}$ having arrivals at rate $\lambda_{i,j}$ for all i, j. The rate matrix λ gathers each of these rates together, and is called *admissible* if it is strictly doubly substochastic:

$$\sum_j \lambda_{i,j} < 1\, \forall i, \; \sum_i \lambda_{i,j} < 1\, \forall j. \tag{3}$$

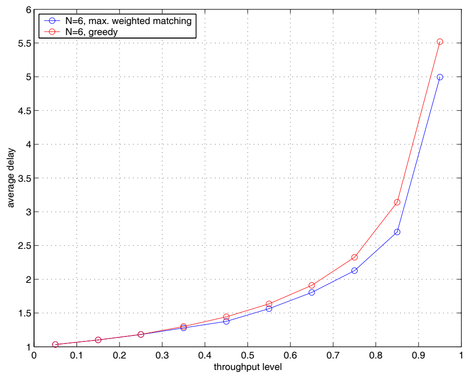
The greedy maximal weighted matching based scheduler creates at each time $n \geq 0$ the complete weighted $N \times N$ bipartite graph (S, T, E) with $w(e) = X_{s(e),t(e)}(n)$ for all $e \in E$ and configures the switch according to a greedy maximal weighted matching on the bipartite graph.

III. NUMERICAL STUDY

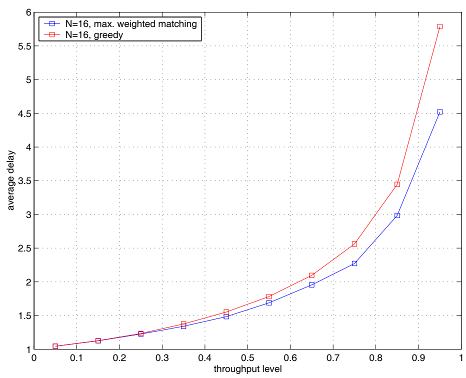
Here we report the results of our numerical simulations of greedy matching based scheduling. In addition to demonstrating the attractive throughput properties of the greedy scheduler, we also observe delay performance quite similar to that achievable under maximum weighted matching based scheduling.

This work was supported by NSF under grants CCR-0325401 and ANI-0335217 and by ONR grant number N000140610064.

1-4244-0350-2/06/$20.00 ©2006 IEEE



(a) $N = 6$



(b) $N = 16$

Fig. 1. Average delay performance over a range of throughput levels for both greedy and maximum weighted matching based scheduling.

Our simulation scenario considers $N = 6, 16$, and a range of throughput levels for simulation. Each throughput level is given by the maximum row/column sum of the arrival rate matrix. At each throughput level, 25 arrival rate matrices are generated randomly and for each rate matrix, a sample path is simulated over 2.5×10^5 time slots, starting at initial VOQ occupancies of zero. The average queueing delay over each sample path is averaged over the 25 sample paths to generate an individual data point representing the average delay at that throughput level. We present the simulation results in Fig. 1.

Note above that maximum weighted matching and the greedy algorithm maintain a close level of delay over the entire range of throughput considered. Additionally, we observe that greedy scheduling never suffers instability over the range of throughput levels. This points to significantly improved throughput performance over the 50% level that can be trivially shown to be sufficient (though clearly not necessary) under any 2-approximation algorithm to maximum weighted matching.

Given these attractive delay and throughput performance properties of greedy weighted matching based scheduling, we next analytically pursue the maximum throughput properties of the switch under the greedy algorithm. Specifically, we consider the 2×2 switch.

IV. Greedy Matching Achieves 100% Throughput in the 2×2 Input-Queued Switch

For the 2×2 switch, there are only two configurations that can be selected as greedy matchings. In matrix form, they are given by

$$v_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{4}$$

The main result of this paper is provided next.

Theorem 1: For the 2×2 input-queued switch, under Bernoulli arrivals, greedy maximal weighted matching based scheduling achieves 100% throughput.

Our notion of stability is referred to as *weak stability*, and is defined as follows.

Definition 1: The Markov Chain $\{X(n)\}_{n=0}^{\infty}$ is *weakly stable* if there exists a Lyapunov function V such that for any $\epsilon > 0$, there exists $B > 0$ such that

$$\lim_{n \to \infty} P[V(X(n)) > B] < \epsilon. \tag{5}$$

□

We make use of the following Lyapunov function, $V : \mathbb{R}_+^{2\times 2} \to \mathbb{R}_+$. Let X be a 2×2 matrix. Then

$$V(X) = \max_{i,j} X_{i,j}. \tag{6}$$

For the entire proof, we assume without loss of generality that $\text{VOQ}_{1,1}$ has the maximum number of cells: $X_{1,1} \geq X_{i,j}, \forall i, j = 1, 2$. Any other case can be trivially converted to this scenario by relabeling the input/output ports. We divide the proof into two key cases: the first case has v_1 strictly dominating v_2 under greedy matching, and the second case has v_1 equivalent to v_2 as greedy choices.

A. Case 1: v_1 dominates v_2 under greedy weighted matching

Lemma 1: When the queue occupancy matrix $X(n)$ satisfies $X_{1,1}(n) - 1 \geq X_{1,2}(n)$, $X_{1,1}(n) - 1 \geq X_{2,1}(n)$, and $X_{1,1}(n) \geq X_{2,2}(n)$,

$$\begin{aligned} &E\left[V(X(n+1)) - V(X(n)) | X(n)\right] \\ &\qquad \leq -(1-\lambda_{1,1})(1-\lambda_{1,2})(1-\lambda_{2,1})(1-\lambda_{2,2}). \end{aligned} \tag{7}$$

Proof: Assume $V(X(n)) \geq 1$, with $X_{1,1}(n) - 1 \geq X_{1,2}(n)$, $X_{1,1}(n) - 1 \geq X_{2,1}(n)$, and $X_{1,1}(n) \geq X_{2,2}(n)$. Designating $a_{i,j}(n)$ as the number of arrivals to $\text{VOQ}_{i,j}$ at the beginning of time slot n, we have

$$X_{1,1}(n+1) = X_{1,1}(n) - 1 + a_{1,1}(n+1) \tag{8}$$
$$X_{1,2}(n+1) = X_{1,2}(n) + a_{1,2}(n+1) \tag{9}$$
$$X_{2,1}(n+1) = X_{2,1}(n) + a_{2,1}(n+1) \tag{10}$$
$$X_{2,2}(n+1) = \max\{X_{2,2}(n) - 1, 0\} + a_{2,2}(n+1) \tag{11}$$

We are interested in finding an upper-bound for the expression $E\left[V(X(n+1)) - V(X(n)) | X(n)\right]$. Thus, we must find the minimum probability of the event $\{V(X(n+1)) - V(X(n)) = -1 \mid X(n)\}$. Clearly, this occurs when $X_{1,1}(n) - 1 = X_{1,2}(n) = X_{2,1}(n) = X_{2,2}(n) - 1$, since an arrival to any VOQ at the beginning of time slot $n + 1$ results in $V(X(n+1)) = V(X(n))$. The following bounds are then evident, under the assumed Bernoulli-distributed arrivals.

$$P\left[V(X(n+1)) = V(X(n)) | X(n)\right] \leq 1 - (1-\lambda_{1,1})(1-\lambda_{1,2})(1-\lambda_{2,1})(1-\lambda_{2,2}) \quad (12)$$

$$P\left[V(X(n+1)) = V(X(n)) - 1 | X(n)\right] \geq (1-\lambda_{1,1})(1-\lambda_{1,2})(1-\lambda_{2,1})(1-\lambda_{2,2}). \quad (13)$$

Since (12) and (13) completely characterize the difference $V(X(n+1)) - V(X(n))$ in this scenario, we have

$$\begin{aligned} E&\left[V(X(n+1)) - V(X(n)) | X(n)\right] \\ &= 0 \cdot P[V(X(n+1)) - V(X(n)) = 0 | X(n)] \\ &\quad - 1 \cdot P[V(X(n+1)) - V(X(n)) = -1 | X(n)] \quad (14) \\ &\leq -(1-\lambda_{1,1})(1-\lambda_{1,2})(1-\lambda_{2,1})(1-\lambda_{2,2}). \quad (15) \end{aligned}$$

□

The remainder of this section treats the case where the maximum element of $\{X_{1,1}(n), X_{2,2}(n)\}$ equals the maximum element of $\{X_{1,2}(n), X_{2,1}(n)\}$.

B. Drift analysis for a simple two queue system

Consider a queueing system consisting of two queues, with respective backlogs at time n equal to $Y_1(n), Y_2(n)$. These queues are subject to Bernoulli arrivals with possibly *time-varying rates*. Let $\tilde{A}_1(n), \tilde{A}_2(n)$ be the cumulative arrivals to queues Y_1, Y_2 respectively, up to and including time n. There is a single server that is only able to serve one of the two queues at a time, with one unit of service at a queue resulting in a reduction in the queue's backlog by one cell at the end of the time slot. The scheduling policy employed is *longest queue first* (LQF), where the queue having maximum backlog is serviced at each slot, with queue 1 chosen as the default queue for service in the event of equal queue backlogs. Suppose at time n, $Y_1(n) = Y_2(n) = a \geq 0$, and there exist τ, k, such that the arrival rates to Y_1 and Y_2 are upper-bounded by r_1 and r_2, respectively, over time slots $n+\tau, \ldots, n+\tau+k$. We are interested in understanding the expected *drift* of this queueing system from time slot n through time slot $n+\tau+k$, given by

$$\begin{aligned} d(n,\tau,k,a) = E[&\max\{Y_1(n+\tau+k), Y_2(n+\tau+k)\} \\ &- \max\{Y_1(n), Y_2(n)\} \mid Y(n) = (a,a)]. \quad (16) \end{aligned}$$

The following lemma demonstrates that when $|Y_1(n) - Y_2(n)| \leq 2$, then after any number of time slots, prior to the arrivals but after service, Y_1 and Y_2 are within 1 unit of each other. For convenience, we denote the queue backlogs prior to arrivals but after service by

$$\hat{Y}_1(n) = Y_1(n) - \tilde{A}_1(n) \quad (17)$$

$$\hat{Y}_2(n) = Y_2(n) - \tilde{A}_2(n). \quad (18)$$

Lemma 2: When $|Y_1(n) - Y_2(n)| \leq 2$, for $k \geq 1$,

$$\left|\hat{Y}_1(n+k) - \hat{Y}_2(n+k)\right| \leq 1. \quad (19)$$

Proof: Our proof is by induction. Assume $|Y_1(n) - Y_2(n)| \leq 2$. Without loss of generality, we assume that $\max\{Y_1(n), Y_2(n)\} = Y_1(n)$, from which we can assume without loss of generality that under LQF, queue 1 is selected for service at time slot n. Then, $|\hat{Y}_1(n+1) - \hat{Y}_2(n+1)| \leq 1$. For the inductive step, assume that $|\hat{Y}_1(n+k) - \hat{Y}_2(n+k)| \leq 1$. Thus it must be true that the arrivals at time slot $n+k$ are such that $|Y_1(n+k) - Y_2(n+k)| \leq 2$. Under LQF, the service over time slot k is applied to queue i, where $Y_i(n+k) \geq Y_j(n+k)$, $i \neq j$. This implies $|\hat{Y}_1(n+k+1) - \hat{Y}_2(n+k+1)| \leq 1$, which completes the induction. □

The following lemma bounds the $(\tau + k)$-slot drift in our two-queue system, when both queue occupancies are initially equal.

Lemma 3: Consider the case $Y_1(n) = Y_2(n) = a$. Suppose there exist integers $\tau, k \geq 0$ such that over time slots $n+\tau+1, \ldots, n+\tau+k$, the arrival rates to queues $1, 2$ are upper-bounded by r_1, r_2, respectively. Then, when $a \geq \tau + k$,

$$d(n,\tau,k,a) \leq \frac{3+\tau}{2} - \frac{r_1+r_2}{2} - \frac{k(1-r_1-r_2)}{2} \quad (20)$$

Proof: Let $Y_1(n) = Y_2(n) = a \geq \tau + k$. For time slots $n, \ldots, n+\tau$, a maximum of 2τ arrivals can occur in total to both queues. Over τ slots, τ total cells must be serviced from the queues under LQF. By Lemma 2, $|Y_1(n+\tau) - Y_2(n+\tau)| \leq 2$. Thus, the total number of cells in both queues at time $n+\tau$ is at most $2a+\tau$, which implies $\max\{Y_1(n+\tau), Y_2(n+\tau)\} \leq a + \tau/2 + 1$.

Subsequent to time $n + \tau$, assume arrival rates upper-bounded by r_1 and r_2 (Bernoulli arrivals) to queues Y_1 and Y_2, respectively. We consider the sum total number of arrivals to both queues from time slot $n+\tau+1$ through time $n+\tau+k-1$. Define

$$\begin{aligned} \rho(i,\tau,k) \triangleq P[&\tilde{A}_1(n+\tau+k-1) - \tilde{A}_1(n+\tau) \\ &+ \tilde{A}_2(n+\tau+k-1) - \tilde{A}_2(n+\tau) = i \mid Y(n)]. \quad (21) \end{aligned}$$

From Lemma 2, it is clear that after $k-1$ arrival opportunities and k service opportunities to the queues Y_1, Y_2, irrespective of the manner in which the arrivals and service occurred, the queue backlogs must be within 1 cell of one another. We can then conclude that if i total arrivals occurred to the queues over time slots $n+\tau+1, \ldots, n+\tau+k$, the queue backlogs (in no particular order) are upper-bounded by

$$a + \left\lceil \frac{\tau+i-k}{2} \right\rceil, \quad a + \left\lfloor \frac{\tau+i-k}{2} \right\rfloor. \quad (22)$$

To complete the characterization of the drift, we must account for the arrivals at time slot $n+\tau+k$. A sufficient bound is to assign probability 1 to the occurrence of an additional arrival to the maximum-valued queue. The drift is then upper-bounded by

$$\begin{aligned}
&d(n,\tau,k,a) \\
&\quad \le \sum_{i=0}^{2k-2}\left(1+\frac{1}{2}+\frac{\tau+i-k}{2}\right)\rho(i,k), && (23)\\
&\quad = \frac{3+\tau-k}{2}+\frac{1}{2}\sum_{i=0}^{2k-2} i\rho(i,\tau,k), && (24)\\
&\quad = \frac{3+\tau-k}{2}+\frac{1}{2}E[\tilde{A}_1(n+\tau+k-1)-\tilde{A}_1(n+\tau) \\
&\qquad\qquad +\tilde{A}_2(n+\tau+k-1)-\tilde{A}_2(n+\tau)]. && (25)
\end{aligned}$$

The limits of the sum in (23) account for up to $2(k-1)$ total arrivals to the queues over $k-1$ time slots. The 1 in (23) corresponds to the additional cell whose arrival occurs with probability 1 at time $n+\tau+k$, and the $1/2$ term in (23) provides a bound on the ceiling of (22). The expectation at right in (25) is effectively upper-bounded as follows,

$$\begin{aligned}
&E[\tilde{A}_1(n+\tau+k-1)-\tilde{A}_1(n+\tau) \\
&\quad +\tilde{A}_2(n+\tau+k-1)-\tilde{A}_2(n+\tau)\mid Y(n)] && (26)\\
&= \sum_{i=0}^{k-2}(E[\tilde{A}_1(n+\tau+i+1)-\tilde{A}_1(n+\tau+i)] \\
&\qquad + E[\tilde{A}_2(n+\tau+i+1)-\tilde{A}_2(n+\tau+i)]), && (27)\\
&\le (k-1)(r_1+r_2). && (28)
\end{aligned}$$

Above, (28) follows because at each time slot, we have assumed that the Bernoulli arrival rates at queues Y_1, Y_2 are upper-bounded by r_1, r_2 respectively. (20) follows immediately. □

C. Case 2: v_1 and v_2 are both valid greedy choices

In this section, we assume without loss of generality that $X(n)$ has the form

$$X(n)=\begin{bmatrix} a & a\\ b & c\end{bmatrix}, \qquad (29)$$

where $a \ge b \ge 0$ and $a \ge c \ge 0$.

We study the probability of an increase in either $Z_1(n+m)=\max\{X_{1,1}(n+m), X_{2,2}(n+m)\}$ or $Z_2(n+m)=\max\{X_{1,2}(n+m), X_{2,1}(n+m)\}$ due to cell arrivals. Here, Z_1, Z_2 may be considered as a pair of induced queues corresponding to v_1, v_2, respectively. Given $X(n)$ of the form (29), the arrival rate at Z_1 at time $n+m$ is given by

$$\begin{aligned}
&r_1(n+m)= \\
&(\lambda_{1,1}+\lambda_{2,2}-\lambda_{1,1}\lambda_{2,2})P[X_{1,1}(n+m)=X_{2,2}(n+m)|X(n)] \\
&\quad +\lambda_{1,1}P[X_{1,1}(n+m)>X_{2,2}(n+m)|X(n)] \\
&\quad +\lambda_{2,2}P[X_{1,1}(n+m)<X_{2,2}(n+m)|X(n)]. \qquad (30)
\end{aligned}$$

Similarly, the arrival rate at Z_2 at time $n+m$ is given by

$$\begin{aligned}
&r_2(n+m)= \\
&(\lambda_{1,2}+\lambda_{2,1}-\lambda_{1,2}\lambda_{2,1})P[X_{1,2}(n+m)=X_{2,1}(n+m)|X(n)] \\
&\quad +\lambda_{1,2}P[X_{1,2}(n+m)>X_{2,1}(n+m)|X(n)] \\
&\quad +\lambda_{2,1}P[X_{1,2}(n+m)<X_{2,1}(n+m)|X(n)]. \qquad (31)
\end{aligned}$$

Define

$$\gamma=\frac{1}{4}\left(1-\max\{\lambda_{1,1},\lambda_{2,2}\}-\max\{\lambda_{1,2},\lambda_{2,1}\}\right). \qquad (32)$$

If we can obtain the bound

$$\begin{aligned}
&(\lambda_{1,1}+\lambda_{2,2}-\lambda_{1,1}\lambda_{2,2})P[X_{1,1}(n+m)=X_{2,2}(n+m)|X(n)] \\
&\qquad <\gamma, \qquad (33)
\end{aligned}$$

then we can conclude

$$\begin{aligned}
&r_1(n+m) \\
&\quad <\frac{1}{4}\left(1-\max\{\lambda_{1,1},\lambda_{2,2}\}-\max\{\lambda_{1,2},\lambda_{2,1}\}\right) \\
&\qquad +\max\{\lambda_{1,1},\lambda_{2,2}\} && (34)\\
&\quad =\frac{1}{4}+\frac{3}{4}\max\{\lambda_{1,1},\lambda_{2,2}\}-\frac{1}{4}\max\{\lambda_{1,2},\lambda_{2,1}\} && (35)\\
&\quad \triangleq \bar{r}_1. && (36)
\end{aligned}$$

Similarly for $r_2(n+m)$, if

$$\begin{aligned}
&(\lambda_{1,2}+\lambda_{2,1}-\lambda_{1,2}\lambda_{2,1})P[X_{1,2}(n+m)=X_{2,1}(n+m)|X(n)] \\
&\qquad <\gamma, \qquad (37)
\end{aligned}$$

then we can conclude

$$\begin{aligned}
&r_2(n+m) \\
&\quad <\frac{1}{4}\left(1-\max\{\lambda_{1,1},\lambda_{2,2}\}-\max\{\lambda_{1,2},\lambda_{2,1}\}\right) \\
&\qquad +\max\{\lambda_{1,2},\lambda_{2,1}\} && (38)\\
&\quad =\frac{1}{4}-\frac{1}{4}\max\{\lambda_{1,1},\lambda_{2,2}\}+\frac{3}{4}\max\{\lambda_{1,2},\lambda_{2,1}\} && (39)\\
&\quad \triangleq \bar{r}_2. && (40)
\end{aligned}$$

In this section we will establish the rate upper-bounds $\bar{r}_1, \bar{r}_2$, implying that the induced queues Z_1, Z_2 will fit into the drift analysis of Section IV-B. Note for λ admissible, $\bar{r}_1+\bar{r}_2<1$.

Define

$$\xi_0=\left\lceil\frac{3-\bar{r}_1-\bar{r}_2}{1-\bar{r}_1-\bar{r}_2}\right\rceil+1. \qquad (41)$$

Lemma 4: If $a-c>2\xi_0$ and $a-b>2\xi_0$, and $a\ge 2\xi_0+1$, then

$$\begin{aligned}
&P[X_{1,1}(n+m)=X_{2,2}(n+m)|X(n)]=0, \\
&\qquad\qquad m=1,2,\ldots,\xi_0 \qquad (42)\\
&P[X_{1,2}(n+m)=X_{2,1}(n+m)|X(n)]=0, \\
&\qquad\qquad m=1,2,\ldots,\xi_0 \qquad (43)
\end{aligned}$$

Proof: Over time slots $n,\ldots,n+\xi_0-1$, the greedy matching algorithm ensures $X_{1,1}$ decreases by at most ξ_0 cells. The backlog $X_{2,2}$ cannot increase by more than ξ_0 cells over time slots $n,\ldots,n+\xi_0$. Then we must have that $X_{1,1}(n+m)\ge a-\xi_0$ and $X_{2,2}(n+m)<a-\xi_0$ for $m=1,2,\ldots,n+\xi_0$, from which (42) follows. The proof for (43) follows identically and is omitted. □

Lemma 4 ensures that (33) and (37) are satisfied over slots $n,\ldots,n+\xi_0$. Thus, Lemma 3 can be applied to induced queues Z_1, Z_2, with values $\tau = 0, k = \xi_0, a \geq 2\xi_0$. Simple algebraic manipulation provides $d(n, 0, \xi_0, a) < 0$. Thus the drift of induced queues Z_1, Z_2 in any switch state satisfying the assumptions of Lemma 4 is negative after ξ_0 time slots.

Lemma 5: If $a - c \leq 2\xi_0$, then there exist τ_1, ξ_1 such that if $a - b > 2(\tau_1 + \xi_1)$ and $a \geq 2\xi_0 + 2\xi_1 + 2\tau_1$ then

$$P[X_{1,1}(n+m) = X_{2,2}(n+m)|X(n)] < \gamma, \quad m = \tau_1 + 1, \ldots, \tau_1 + \xi_1 \quad (44)$$

$$P[X_{1,2}(n+m) = X_{2,1}(n+m)|X(n)] = 0, \quad m = 1, \ldots, \tau_1 + \xi_1 \quad (45)$$

Proof: For any m such that over time slots $n, \ldots, n+m$, neither $\text{VOQ}_{1,1}$ or $\text{VOQ}_{2,2}$ reaches zero occupancy, we have

$$X_{1,1}(n+m) - X_{2,2}(n+m) = a - c + (A_{1,1}(n+m) - A_{1,1}(n)) - (A_{2,2}(n+m) - A_{2,2}(n)). \quad (46)$$

The expression $(A_{1,1}(n+m) - A_{1,1}(n)) - (A_{2,2}(n+m) - A_{2,2}(n))$ can be regarded as a summation of m i.i.d. random variables, each of which take values from the set $\{-1, 0, 1\}$. Using (46), we have for any m such that over time slots $n, \ldots, n+m$, neither $\text{VOQ}_{1,1}$ or $\text{VOQ}_{2,2}$ reaches zero occupancy,

$$P[X_{1,1}(n+m) = X_{2,2}(n+m)|X(n)] = P[(A_{1,1}(n+m) - A_{1,1}(n)) - (A_{2,2}(n+m) - A_{2,2}(n)) = c - a]. \quad (47)$$

We throw out the case of $\lambda_{1,1} = \lambda_{2,2} = 0$, since in this case, $X_{1,1}(n) = 0, X_{2,2}(n) = 0, \forall n$ almost surely. For any other $\lambda_{1,1}, \lambda_{2,2}$ values, a simple normal approximation to the i.i.d. summation guarantees the existence of τ_1^{a-c} such that for all $m > \tau_1^{a-c}$,

$$P[(A_{1,1}(n+m) - A_{1,1}(n)) - (A_{2,2}(n+m) - A_{2,2}(n)) = c - a] < \gamma. \quad (48)$$

Taking the maximum over all $a - c$ values considered in this lemma, we obtain a common value τ_1 such that after time $n + \tau_1$, (48) is satisfied:

$$\tau_1 = \max_{a-c \in \{0, \ldots, 2\xi_0\}} \tau_1^{a-c}. \quad (49)$$

Note τ_1 is a constant derived only from the constants $\lambda_{1,1}, \lambda_{2,2}$.

Define

$$\xi_1 = \frac{3 + \tau_1 - \bar{r}_1 - \bar{r}_2}{1 - \bar{r}_1 - \bar{r}_2} + 1. \quad (50)$$

Suppose $a \geq 2\xi_0 + 2\xi_1 + 2\tau_1$. Then since $a - c \leq 2\xi_0$, we have $c \geq 2(\tau_1 + \xi_1)$. Thus, for the first $\tau_1 + \xi_1$ services to configuration v_1, both $X_{1,1}$ and $X_{2,2}$ are reduced by one cell at each service (since neither queue reaches zero occupancy). Since $c \geq 2(\tau_1 + \xi_1)$ guarantees that $X_{1,1}(n) > 0$ and $X_{2,2}(n) > 0$ for time slots $n, \ldots, n + \tau_1 + \xi_1$, we conclude that (44) is satisfied. Finally, assuming $a - b > 2(\tau_1 + \xi_1)$, then following in a similar manner to the proof of Lemma 4, there is no sample path on which the queue backlogs $X_{1,2}$ and $X_{2,1}$ coincide over time slots $n, \ldots, n + \tau_1 + \xi_1$, giving (45) as desired. □

Lemma 5 provides that (33) and (37) are satisfied over slots $n + \tau_1, \ldots, n + \tau_1 + \xi_1$. Thus, applying Lemma 3 with $\tau = \tau_1, k = \xi_1, a \geq 2\xi_0 + 2\xi_1 + 2\tau_1$, we have $d(n, \tau_1, \xi_1, a) < 0$.

Corollary 1: If $a - b \leq 2\xi_0$, then there exist τ_2, ξ_2 such that if $a - c > 2(\tau_2 + \xi_2)$ and $a \geq 2\xi_0 + 2\xi_2 + 2\tau_2$ then

$$P[X_{1,1}(n+m) = X_{2,2}(n+m)|X(n)] = 0, \quad m = 1, \ldots, \tau_2 + \xi_2 \quad (51)$$

$$P[X_{1,2}(n+m) = X_{2,1}(n+m)|X(n)] < \gamma, \quad m = \tau_2 + 1, \ldots, \tau_2 + \xi_2 \quad (52)$$

Proof: The proof follows identically to the proof of Lemma 5. □

Denote $\tau_3 = \max\{\tau_1, \tau_2\}$ and $\xi_3 = \max\{\xi_1, \xi_2\}$.

Lemma 6: If $a - b \leq 2(\tau_3 + \xi_3)$ and $a - c \leq 2(\tau_3 + \xi_3)$ and $a \geq 3(\tau_3 + \xi_3)$ then

$$P[X_{1,1}(n+m) = X_{2,2}(n+m)|X(n)] < \gamma, \quad m = \tau_3 + 1, \ldots, \tau_3 + \xi_3 \quad (53)$$

$$P[X_{1,2}(n+m) = X_{2,1}(n+m)|X(n)] < \gamma, \quad m = \tau_3 + 1, \ldots, \tau_3 + \xi_3 \quad (54)$$

Proof: From Lemma 5 and Corollary 1, it is clear that τ_3 slots are sufficient to ensure convergence of arrival rates to induced queues Z_1, Z_2 to less than $\bar{r}_1, \bar{r}_2$, respectively. This entire discussion is valid so long as no queue reaches zero occupancy over time slots $n, \ldots, n + \tau_3 + \xi_3 - 1$. Clearly, restricting $a \geq 3(\tau_3 + \xi_3)$, $b, c \geq \tau_3 + \xi_3$ is sufficient to guarantee this condition. Thus (53) and (54) follow as desired. □

Note in Lemma 6 that ξ_3 is the number of time slots required to ensure $d(n, \tau_3, \xi_3, a) < 0$, when $a \geq 3\tau_3 + 3\xi_3$.

We have organized the results of this section in Table I. Note that the full set of possible $a - b, a - c$ values is covered by the results of this section.

TABLE I

REQUIRED VALUES FOR a, τ, k FOR DIFFERENT POSSIBLE $a - b, a - c$ VALUES, SUCH THAT WHEN $X(n)$ IS AS IN (29), THEN THE DRIFT OF THE INDUCED QUEUES Z_1, Z_2 IS NEGATIVE: $d(n, \tau, k, a) < 0$.

$a - b$	$a - c$	$a \geq$	τ	k
$> 2\xi_0$	$> 2\xi_0$	$2\xi_0$	0	ξ_0
$> 2\tau_1 + \xi_1$	$\leq 2\xi_0$	$2\xi_0 + 2\xi_1 + 2\tau_1$	τ_1	ξ_1
$\leq 2\xi_0$	$> 2\tau_2 + 2\xi_2$	$2\xi_0 + 2\xi_2 + 2\tau_2$	τ_2	ξ_2
$\leq 2(\tau_3 + \xi_3)$	$\leq 2(\tau_3 + \xi_3)$	$3(\tau_3 + \xi_3)$	τ_3	ξ_3

We now have all the tools necessary to prove Theorem 1.

D. Proof of Theorem 1

The proof is through a Lyapunov drift argument. We will determine a sequence of time slot indices $\{\zeta_i\}_{i=0}^{\infty}$ such that $\zeta_0 = 0$, there exists $M < \infty$ such that $\zeta_{i+1} - \zeta_i < M$, and

$$E[V(X(\zeta_{i+1})) - V(X(\zeta_i))|X(\zeta_n)] < 0,\ V(X(\zeta_i)) > \nu \tag{55}$$

$$E[V(X(\zeta_{i+1}))|X(\zeta_i)] < \infty,\ V(X(\zeta_i)) \leq \nu. \tag{56}$$

for all $i \geq 0$.

Consider any integer $i \geq 0$. At time ζ_i, suppose $V(X(\zeta_i)) \geq 2\xi_0 + 3(\tau_3 + \xi_3) \triangleq \nu$. The system state $X(\zeta_n)$ is accounted for by one of Lemmas 1, 4, 5, 6, and Corollary 1. For Lemma 1 to be valid, we require $V(X(\zeta_n)) \geq 1$, which is satisfied by ν. For Lemmas 4, 5, 6 and Corollary 1, the values are listed as restrictions on the variable a in Table I. Each of these are satisfied by ν. From (20), the condition for eventually attaining $d(n, \tau, k, a) < 0$ is that $r_1 + r_2 < 1$. Thus, so long as $\bar{r}_1 + \bar{r}_2 < 1$, Lemma 3 combined with each of Lemmas 4, 5, 6, and Corollary 1 imply that the values of k in Table I will yield negative drift, with the values for ζ_{i+1} listed in Table II. This completes the characterization of (55).

TABLE II
VALUES OF ζ_{i+1} USED UNDER THE DIFFERENT POSSIBLE SYSTEM STATES AT TIME ζ_i.

Case covered by	$\zeta_{i+1} - \zeta_i$
Lemma 1	1
Lemma 4	ξ_0
Lemma 5	$\tau_1 + \xi_1$
Corollary 1	$\tau_2 + \xi_2$
Lemma 6	$\tau_3 + \xi_3$

The only remaining case is when $V(X(\zeta_n)) \leq \nu$. In this case, we use $\zeta_{n+1} = \zeta_n + 1$. Then it is clear under Bernoulli arrivals that $|V(X(\zeta_{n+1})) - V(X(\zeta_n))| \leq 1$, implying (56).

The above application of Lemma 1 requires only that $0 \leq \lambda_{i,j} < 1, \forall i, j$. Further, the above application of Lemmas 3, 4, 5, 6, and Corollary 1 holds for any $\bar{r}_1 + \bar{r}_2 < 1$, which implies that $\max\{\lambda_{1,1}, \lambda_{2,2}\} + \max\{\lambda_{1,2}, \lambda_{2,1}\} < 1$. This is equivalent to requiring that λ is strictly doubly substochastic.

In order to conclude that Foster's Criteria [3, Ch I, Prop. 5.3] are satisfied for positive recurrence of the embedded Markov Chain $\{X(\zeta_n)\}_{n=0}^{\infty}$, it is trivial to demonstrate that the embedded chain has a single irreducible class.

Since $\{X(\zeta_n)\}_{n=0}^{\infty}$ is irreducible, $\inf_X V(X) = 0$, and by (55), (56), [3, Ch I Prop. 5.3] implies $\{X(\zeta_n)\}_{n=0}^{\infty}$ is positive recurrent. As explained in [4], we may then conclude that for any $\epsilon > 0$, there exists $B_1 > 0$ such that

$$\lim_{n\to\infty} P[V(X(\zeta_n)) > B_1] < \epsilon. \tag{57}$$

Finally, we turn our attention to the weak stability of the queue backlog process, $\{X(n)\}_{n=0}^{\infty}$. Define $\kappa(n)$ as the maximum index m such that $\zeta_m \leq n$:

$$\kappa(n) = \max\{m : \zeta_m \leq n\}. \tag{58}$$

Also, define

$$M = \max\{1, \xi_0, \tau_1 + \xi_1, \tau_2 + \xi_2, \tau_3 + \xi_3\}. \tag{59}$$

Clearly M is a finite constant providing an upper bound on $\zeta_{n+1} - \zeta_n$ for any $n \geq 0$. Then

$$V(X(n)) = V(X(\zeta_{\kappa(n)})) + \big(V(X(n)) - V(X(\zeta_{\kappa(n)}))\big) \tag{60}$$

$$\leq V(X(\zeta_{\kappa(n)})) + M. \tag{61}$$

Above, (61) follows by the fact that the maximum queue backlog in the system can only increase by 1 cell at each slot, and that there are at most M slots between times $\kappa(n)$ and n. Then it immediately follows that for $B_2 > 0$,

$$P[V(X(n)) > B_2] \leq P[V(X(\zeta_{\kappa(n)})) + M > B_2]. \tag{62}$$

Then we have the following series of equations.

$$\lim_{n\to\infty} P[V(X(n)) > B_2]$$

$$\leq \lim_{n\to\infty} P[V(X(\zeta_{\kappa(n)})) + M > B_2] \tag{63}$$

$$= \lim_{n\to\infty} P[V(X(\zeta_n)) > B_2 - M] \tag{64}$$

Above, (64) follows since $\kappa(n) \to \infty$ as $n \to \infty$. Using (57), we conclude that if $B_2 \geq B_1 + M$, then

$$\lim_{n\to\infty} P[V(X(n)) > B_2] < \epsilon \tag{65}$$

Thus, we have the weak stability of the queue backlog process as desired, for all strictly doubly substochastic arrival rate matrices λ. Since the set of substochastic arrival rate matrices is the maximum set of admissible rates in a 2×2 switch, and is identical to the set of strictly substochastic matrices up to a set of measure zero, we conclude that greedy maximal matching achieves 100% throughput.

V. CONCLUSION

We have considered greedy maximal weighted matching based scheduling for the input-queued switch. Our numerical studies attest to the attractive throughput and delay performance of the greedy algorithm next to the throughput-optimal maximum weighted matching based scheduler. Additionally, for the case of the 2×2 switch under i.i.d. Bernoulli arrivals, we have proven the throughput optimality of the greedy matching based scheduler.

REFERENCES

[1] P. Krishna, N.S. Patel, A. Charny, and R.J. Simcoe, "On the speedup required for work-conserving crossbar switches," *IEEE J. Select. Areas Commun.*, vol. 17, no. 6, Jun 1999.

[2] S. Kumar, P. Giaccone, and E. Leonardi, "Rate stability of stable marriage scheduling algorithms in input-queued switches," in *Proceedings Allerton Conf. Communication, Control, and Computing*, Oct 2002.

[3] S. Asmussen, *Applied Probability and Queues*, Springer, 2000.

[4] E. Leonardi, M. Mellia, F. Neri, and M. Ajmone Marsan, "Bounds on average delays and queue size averages and variances in input-queued cell-based switches," in *IEEE Proc. INFOCOM*, 2001.