Exact Queueing Analysis of Discrete Time Tandems with Arbitrary Arrival Processes

Michael J. Neely
University of Southern California
mjneely@usc.edu — http://www-rcf.usc.edu/∼mjneely/

***Abstract*—We consider a discrete time tandem of queues serving fixed length packets, where each queue can serve a single packet during a timeslot. Arrivals and departures take place at each stage according to arbitrary stochastic processes. Using sample path techniques and stochastic coupling methods, we present an exact analysis of the queue occupancy distribution at each stage when all queues operate according to the *Furthest-to-Go* service discipline. Explicit expressions for average queue occupancies are provided in terms of the average occupancy in a single queue with a superposition of the original inputs. To our knowledge, this is the first analysis of a multi-input multi-output queueing network yielding exact solutions for general arrival processes.**

***Index Terms*—Packet Scheduling, Queueing Analysis, Network Calculus, Stochastic Coupling**

I. Introduction

In this paper we consider a discrete time tandem of K queueing nodes serving fixed length packets (Fig. 1). Time is slotted and slots are normalized to integral units. Packets enter the network according to a set of arbitrary arrival processes $\{A_{ij}(t)\}$, where $A_{ij}(t)$ represents the number of packets that exogenously arrive to source i at timeslot t which are destined for node j. Each node can serve a single packet during a timeslot. After service, packets either exit the network or are forwarded to the next node in the tandem, according to their destinations.

To analyze such a system, a *service policy* must be specified that determines which packet to serve when packets from different streams are waiting in the same queue. We consider the *Furthest-to-Go* (FTG) service policy, which serves the packet destined for the furthest downstream node of the tandem (ties are broken arbitrarily). Under this FTG policy, we develop a simple network decomposition that preserves steady state queue occupancy when inputs are independent and stationary. Explicit expressions for the average number of packets from each traffic class and in each node of the network are computed in terms of a single queue with a superposition of the original inputs. Furthermore, exact occupancy distributions can be obtained in terms of a simplified 2-queue equivalent model. We note that the FTG service policy is shown in [1] to be stable in all discrete time networks with fixed routing. Our results demonstrate that this policy is easily analyzable for all input traffic when the network has a tandem topology.

As queueing networks are non-linear systems driven by stochastic events, exact analytical results are largely limited to systems with the special structure of *reversibility* [2] [3] [4] [5]. Reversible networks have product-form occupancy distributions. The best known example of a reversible network is the classic M/M/1 Jackson network with Poisson inputs and independent, exponential service times [5]. Exact analysis of non-reversible queueing systems is usually confined to small networks (see [6] [7] [8] for analysis of a single discrete time queue with general inputs, and [9] for a moment generating function analysis of a two-queue tandem with i.i.d. arrivals every timeslot). Approximation methods are developed in [10] [11] for modeling discrete time tandems with arrivals and departures at each stage in the special case when inputs have a specified Markovian structure. Bounding techniques for general networks are developed in [12] [13] [14] using a calculus of network service curves.

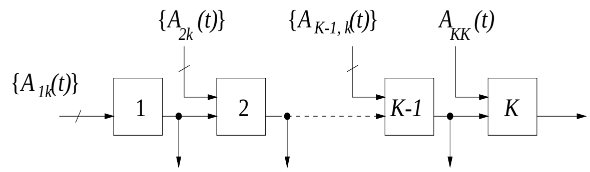

Fig. 1. A tandem of K queues with arrivals and departures at each stage.

Our approach uses the sample-path method of [15] [16], where tree networks of deterministic service time queues with general inputs are analyzed using a packet conservation property for multi-input single output systems. A similar approach is used in [17] and [18] to analyze discrete time trees and tandems with Poisson inputs. We show that a similar conservation property holds for discrete time tandems with general arrivals and departures at each stage when the FTG service policy is used. To our knowledge, this result is the first analysis of a multi-input multi-output queueing network yielding exact solutions for general inputs.

This paper is structured as follows: In the next section we describe the tandem and define the arrival parameters. In Section III we establish an equivalent model for each node of the tandem which preserves the exact occupancy distribution. In Section IV we provide expressions for the average occupancy within each node of the tandem. Finally, we compare the exact and simulated delay under the FTG strategy to the simulated delay under the *Nearest-to-Go* (NTG) service discipline.

II. Tandem Model

Consider the K node tandem of Fig. 1. We label the nodes in increasing order according to the integers $\{1, 2, \ldots, K\}$, so that node 1 is the first node of the tandem and node K is the last (furthest downstream) node of the tandem.

A. The Arrival Processes

Define:

$$A_{ij}(t) \triangleq \text{Number of packets arriving to node } i \text{ during slot } t \text{ that are destined for node } j$$

$$\lambda_{ij} \triangleq \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} A_{ij}(\tau)$$

We assume that the arrival streams $A_{ij}(t)$ are rate-convergent so that the above limits defining the arrival rates λ_{ij} exist with probability 1. Note that from the tandem structure of the network, $A_{ij}(t)\triangleq 0$ for all t whenever $i > j$. The stochastics of the arrival streams are otherwise arbitrary.

B. The Queueing Equation

Define the *class* of a packet as the destination of a packet, so that packets of class $n \in \{1, 2, \ldots, K\}$ are destined for node n. The current number of packets from each class and the current service decisions for each queue are defined as follows:

$$N_n^{(m)}(t) \triangleq \text{Number of packets destined for node } n \text{ that are currently in node } m$$

$$N^{(m)}(t) \triangleq \sum_{n=1}^{K} N_n^{(m)}(t)$$

$$S_n^{(m)}(t) \triangleq \text{Number of class } n \text{ packets served by node } m \text{ at timeslot } t$$

Note that the service variables $S_n^{(m)}(t)$ can take the values of either 0 or 1. Because each node can serve at most one packet during a timeslot, whenever $S_n^{(m)}(t) = 1$ for some node m and class n, it must be the case that $S_{\tilde{n}}^{(m)}(t) = 0$ for all other classes $\tilde{n} \neq n$ in that node.

All nodes of the tandem are assumed to be initially empty, and the input processes $A_{ij}(t)$ are applied to the tandem at time 0. Formally, we assume that all data streams satisfy $A_{ij}(t) = 0$ for all $t < 0$. The queueing dynamics for each node m and each class $n \geq m$ proceed as follows:

$$N_n^{(m)}(t+1) = [N_n^{(m)}(t) - S_n^{(m)}(t)] + A_{mn}(t) + S_n^{(m-1)}(t)$$

We assume that the only packets elligible for service in a particular queue are those within the queue at the beginning of a timeslot, and hence newly arriving packets cannot be served until the next slot. Under the *Furthest-to-Go* (FTG) service discipline, the service variable $S_n^{(m)}(t)$ is defined to be 1 if n is the largest class among all packets currently contained in node m, and 0 otherwise. Thus, every timeslot each node serves the packet having the largest number of future nodes to traverse among all packets waiting in its queue. New packets arrive to a given node from both the exogenous input stream and the endogenous input stream consisting of departures from the previous node.

III. Equivalent Models

In this section we develop a simple decomposition of the tandem which serves as an equivalent model for all analytical purposes. We begin by presenting the relevant results of [15].

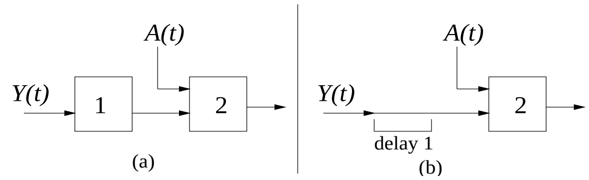

Fig. 2. A 2-queue system with two inputs and a single output, together with its equivalent model formed by replacing the first queue by a pure delay of 1 timeslot.

A. Previous Results for Single-Output Systems

Consider the special case tandem illustrated in Fig. 2a consisting of two nodes 1 and 2 and two general input processes $Y(t)$ and $A(t)$ delivering packets destined for the final node. We have the following fact from [15].

Fact 1: *(Delay Replacement — from [15]) The departure process of node 2 is unchanged if node 1 is replaced by a pure time delay of one time unit, as shown in the equivalent model of Fig. 2b, so that the total input stream delivered into the final node of the equivalent model is given by* $Y(t-1)+A(t)$. □

The proof of this fact is developed in [15]. We note that an independent proof was given in [19]. Intuitively, the result follows because the final node serves packets no faster than the first node can deliver them. Hence, the busy period of the final node in Fig. 2a cannot finish before the busy period of the final node in the equivalent model.

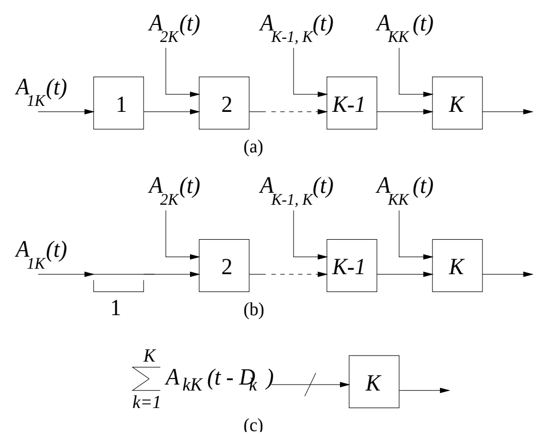

Fig. 3. Equivalently representing the departure process of the final node in a multi-input single output tandem with delayed versions of the original inputs, where the time delay for the k^{th} input stream is given by $D_k = (K - k)$.

Consider now the multi-input single output tandem of K queues of Fig. 3a with exogenous arrivals at each stage, all of which deliver packets destined for the end node K. The delay replacement result of Fact 1 can be used to iteratively replace all preliminary nodes with pure delays while preserving the output process of node K. Indeed, consider the 2-node system consisting of nodes 1 and 2. By Fact 1, node 1 can be replaced by a delay line without changing the output process of node 2, and hence without affecting the output process of any other nodes further downstream (see Fig. 3b). The result can be used again to replace node 2 with a delay line (where the sum process $A_{1K}(t-1) + A_{2K}(t)$ entering node 2 in Fig. 3b is treated as a single collective input when applying Fact 1). This simplifying procedure is iteratively repeated until we are left with a single node K with a superposition of inputs that are delayed versions of the originals (see Fig. 3c). Specifically, the input stream $A_{kK}(t)$ for node k is delayed by D_k timeslots (where $D_k \triangleq (K-k)$), so that the sum process $\sum_{k=1}^{K} A_{kK}(t-D_k)$ enters node K. We call such a sum process a *superposition* of the delayed input streams. The resulting departure process of node K is unchanged under this transformation.

Before proceeding with the general problem of a tandem with arrivals and departures at each stage, we provide a simple stochastic coupling argument that proves the intuitive result that, whenever the inputs are independent and stationary, all time delays can be removed without affecting the steady state occupancy distribution in the queues.

B. Removing the Delays Via Stochastic Coupling

Consider the 2-node tandem with nodes 1 and 2 and input streams $X(t)$, $Y(t)$, and $A(t)$ delivering packets destined for node 2, as shown in Fig. 4. We note that the packet occupancy in the final node of any single output tandem can always be described in terms of the occupancy in node 2 shown in the figure, where the input process $A(t)$ represents the actual process of packets directly entering the final node, and the processes $X(t)$ and $Y(t)$ represent superpositions of delayed versions of the remaining inputs of the tandem, with delays given as in the previous subsection so that the output process of the second-to-last node is unchanged.

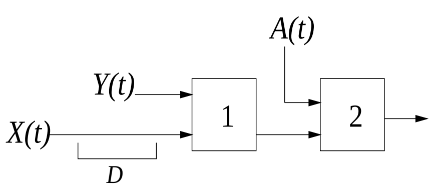

Fig. 4. A canonical 2-queue system with input $X(t)$ delayed by an integer number of timeslots.

Note that the $X(t)$ process is explicitly shown with a time delay of D timeslots (where D is an integer). We show that if $X(t)$ and $Y(t)$ are independent and stationary, the delay can be removed without affecting the steady state distribution of packets in node 1 or node 2. We first define the notions of steady state and stationarity.

Definition 1. *Let the stochastic process $N(t)$ represent the number of packets in a queue as a function of time. The* steady state distribution $F[n]$ *for the queue is defined:*

$$F[n] \triangleq \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} Pr[N(\tau) \leq n] \tag{1}$$

whenever the limit exists.

To define the notion of stationarity, for any arrival process $X(t)$ and any positive integer D, we define the process $\tilde{X}_D(t)$ as follows:

$$\tilde{X}_D(t) \triangleq \begin{cases} 0 & \text{if } t < D \\ X(t) & \text{if } t \geq D \end{cases}$$

Thus, $\tilde{X}_D(t)$ can be viewed as a version of the $X(t)$ data stream in which packets during the first D slots are thrown away.

Definition 2. *An arrival process $X(t)$ is* stationary *if for any positive delay D, the delayed process $X(t-D)$ is stochastically indistinguishable from $\tilde{X}_D(t)$.*

Note that for two stationary arrival processes $X(t)$ and $Y(t)$ that are also independent, the superposition $X(t-D)+Y(t)$ is stochastically equivalent to the superposition $\tilde{X}_D(t)+Y(t)$, and hence either superposition applied to a queue yields the same steady state distribution, provided that the distribution exists.

Theorem 1. *For any general inputs $X(t)$, $Y(t)$, $A(t)$ that are independent and stationary, the steady state occupancy distribution in node 1 of Fig. 4 exists if and only if the steady state occupancy distribution exists when the time delay on the $X(t)$ input stream is removed. If the distributions exist, they are identical.*

Similarly, the steady state distribution in node 2 is the same with or without the time delay on the $X(t)$ stream, provided the distribution exists.

Proof. The proof is given in the Appendix. □

This result for the simple 3 input model of Fig. 4 proves that all time delays from any superposition of multiple inputs can be removed, as the delay lines can be removed iteratively without affecting the steady state distribution, treating the stream for which the delay is removed as the $X(t)$ stream and the collection of all other (delayed or non-delayed) streams as $Y(t)$.

C. Multi-Input Multi-Output Tandems under FTG

Consider now the general tandem of K queues shown in Fig. 1, with arrivals and departures at each stage and operating under the Furthest-to-Go service discipline. We define arrival processes $Y_n^{(m)}(t)$ and their rates $\gamma_n^{(m)}$ as follows:

$$Y_n^{(m)}(t) \triangleq \sum_{i=1}^{m} \sum_{j=\max(n,m)}^{K} A_{ij}(t)$$

$$\gamma_n^{(m)} \triangleq \sum_{i=1}^{m} \sum_{j=\max(n,m)}^{K} \lambda_{ij}$$

For given integers n, m, the $Y_n^{(m)}(t)$ process represents the superposition of all input streams of packets destined for node n or higher that pass through node m, and $\gamma_n^{(m)}$ represents the aggregate data rate of this process. All packets from the $Y_n^{(m)}(t)$ stream are thus from class n or higher. By definition of the FTG policy, these packets have priority over all packets with class lower than n. Thus, to analyze the packet occupancy or departure processes for packets of class n or higher, all lower class packets can be ignored. It follows that the departure process of class n or higher packets from any node $m \leq n$ is identical to the departure process in a modified multi-input single output tandem of nodes $1,\ldots,m$ where all arrival streams for packets exiting the tandem before node n are removed. Using Fact 1 of the previous section, all preliminary nodes of this modified tandem can be replaced with pure delay lines. Thus, the departure process of class n or higher packets is identical to the departures in a single queue with an arrival process $\sum_{i=1}^{m}\sum_{j=\max(n,m)}^{K} A_{ij}(t-(m-i))$. Note that this arrival process is similar to the process $Y_n^{(m)}(t)$ with the exception that some of the component processes $A_{ij}(t)$ are timeshifted.

Note that $\sum_{k\geq n} N_k^{(m)}(t)$ represents the current number of packets in node m that have class greater than or equal to n.

Theorem 2. *If all arrival processes are independent and stationary, then for any node $m \in \{2,\ldots,K\}$, the steady state distribution of $\sum_{k\geq n} N_k^{(m)}(t)$ is identical to the steady state distribution in the second node of the 2 node tandem in Fig. 2a, with input process $Y(t)\triangleq Y_n^{(m-1)}(t)$ at the first stage and input process $A(t)\triangleq\sum_{k\geq n} A_{mk}(t)$ at the second stage.*

Proof. The exogenous inputs to node m consisting of packets of class n or higher are given by the process $\sum_{k\geq n} A_{mk}(t)$. The endogenous arrivals to node m consist of the departures of class n or higher packets from node $m-1$. This departure process is invariant when all preliminary nodes $\{1,2,\ldots,m-2\}$ are deleted and inputs to node $m-1$ are replaced by the stream $\sum_{i=1}^{(m-1)}\sum_{j=\max(n,m-1)}^{K} A_{ij}(t-(m-1-i))$. This transforms the tandem into the canonical 2-queue system of Fig. 2a, with inputs as specified. Because all processes $A_{ij}(t)$ are stationary and independent, by Theorem 1, all time delays can be removed without affecting the steady state occupancy distribution in the second node, and hence the input stream $\sum_{i=1}^{(m-1)}\sum_{j=\max(n,m-1)}^{K} A_{ij}(t-(m-1-i))$ can be replaced by $Y_n^{(m-1)}(t)$. □

IV. Mean Occupancy Analysis

Here we use the equivalent model theory of the previous section to express the mean occupancy in each node of the multi-input multi-output tandem under the FTG service policy in terms of the mean occupancy in a single queue. In particular, for a given input stream $X(t)$, we define $Q(X)$ as the average number of packets this input stream yields when applied as the sole input to a discrete time queue with unit service times. We assume throughout that such steady state averages exist, and that all inputs are stationary and independent.

A. Total Queue Occupancy

Consider a single node m of the tandem, where $m > 1$. Note that all packets in node m are of class m or higher, and hence the total number of packets in this node is given by $N^{(m)}(t)=\sum_{n=m}^{K} N_n^{(m)}(t)$. Thus, for the special case $n=m$, Theorem 2 implies that that the steady state occupancy distribution for node m is equal to the distribution in a modified tandem, where all nodes less than or equal to $m-2$ and all packets less than or equal to class m are deleted, and the arrival stream $Y_m^{(m-1)}$ enters node $m-1$ while the arrival stream $\sum_{k\geq m} A_{mk}(t)$ enters node m. The canonical representation of this 2-queue system is illustrated in Fig. 2a.

Let Θ represent the total number of packets in this equivalent 2-queue tandem. This value is equal to the sum of the average values in each of the two queues, and hence:

$$\Theta = Q\left(Y_m^{(m-1)}\right)+\overline{N}^{(m)} \tag{2}$$

However, by Fact 1, the first queue of this tandem can be replaced by a pure delay line without affecting the departure process from the second queue, and hence, without affecting the total occupancy in the tandem (see Fig. 2b). It follows that Θ can alternately be described as the sum of the average number of packets in the delay line and the average number of packets in the second queue, and we have:

$$\begin{aligned}\Theta &= \gamma_m^{(m-1)} + Q\left(Y_m^{(m-1)}(t-1)+\sum_{k\geq m} A_{mk}(t)\right)\\ &= \gamma_m^{(m-1)} + Q\left(Y_m^{(m-1)}(t)+\sum_{k\geq m} A_{mk}(t)\right) \quad (3)\\ &= \gamma_m^{(m-1)} + Q\left(Y_m^{(m)}\right) \quad (4)\end{aligned}$$

where (3) follows because, by Theorem 1, the average queue occupancy is not affected by a time delay on one of the inputs.

Combining (4) and (2), we have:

$$\overline{N}^{(m)} = \gamma_m^{(m-1)} + Q\left(Y_m^{(m)}\right) - Q\left(Y_m^{(m-1)}\right) \tag{5}$$

The above equality establishes a simple and exact formula for the average node occupancy in a multi-input multi-output tandem under FTG in terms of the average occupancy in a single queue with a superposition of the original inputs. We note that computational techniques for finding the average occupancy in single queues are given for general Markovian inputs in [6] [7], for periodic inputs in [8], and upper bounds on average occupancy for arbitrary rate-convergent inputs are given in [20].

Example: Suppose all inputs are Poisson, so that the function $Q(X)$ can be written as $Q(\lambda)$, a function only of the rate λ of the input stream X, and is given by the standard equation for average occupancy in an M/D/1 queue [5]:[1]

$$Q(\lambda) = \frac{\lambda^2}{2(1-\lambda)}+\lambda$$

Hence, the average number of packets in any node m can be written:

$$\begin{aligned}\overline{N}^{(m)} &= \gamma_m^{(m-1)} + Q\left(\gamma_m^{(m)}\right) - Q\left(\gamma_m^{(m-1)}\right)\\ &= \gamma_m^{(m)} + \frac{(\gamma_m^{(m)})^2}{2\left(1-\gamma_m^{(m)}\right)} - \frac{(\gamma_m^{(m-1)})^2}{2\left(1-\gamma_m^{(m-1)}\right)} \quad (6)\end{aligned}$$

Note that the above result for Poisson inputs cannot be derived using the theory of reversible Markov chains. Indeed, the state dynamics are not reversible, as no more than 1 packet can depart from any output line during a timeslot, while an arbitrarily large number of packets can arrive at the input ports during a slot.

[1] It can be shown that the queueing equations for a non-slotted M/D/1 queue are identical to those for a slotted M/D/1 queue where arrivals occur on slot boundaries.

B. Individual Class Occupancy

The average occupancy of packets from individual classes can likewise be analyzed. Define: $\overline{Z}_n^{(m)} \triangleq \sum_{k=n}^{K} \overline{N}_k^{(m)}$

We clearly have $\overline{N}_n^{(m)} = 0$ if $n < m$. Suppose $n \geq m$. By Theorem 2, the average occupancy of class n or higher packets in node m is the same as the average occupancy in the second node of a 2-queue tandem with input stream $Y_n^{(m-1)}(t)$ at the first queue and input stream $\sum_{k\geq n} A_{mk}(t)$ at the second queue. Using the same analysis as in the section above, we have [compare with (5)]:

$$\overline{Z}_n^{(m)} = \gamma_n^{(m-1)} + Q(Y_n^{(m)}) - Q(Y_n^{(m-1)}) \quad (n \geq m)$$

The average occupancy of class n packets in any queue $m \leq n$ is thus given by $\overline{N}_n^{(m)} = \overline{Z}_n^{(m)} - \overline{Z}_{n+1}^{(m)}$.

V. Comparison to NTG

Throughput properties of the FTG strategy and the *Nearest-to-Go* (NTG) strategy are compared in [1] [21] for general networks. Our results for tandem networks enable explicit occupancy and delay analysis of the FTG strategy, and simulation experiments (omitted for brevity) suggest that delay under NTG is similar, being noticeably different only under high loadings.

Appendix

Proof. (Theorem 1) It is useful to define $N_{[A(\tau)]}(t)$ as the number of packets at timeslot t in a discrete time queue that is initially empty with a general arrival process $A(\tau)$ applied at time 0, where $A(\tau)$ could represent a superposition of processes. Note that the number of packets $N_{[A(\tau)]}(t)$ is always greater than or equal to the number of packets in a queue with the same input process but with some of the arriving packets deleted. Hence, the following inequalities hold deterministically for all timeslots t:

$$N_{[\tilde{X}_D(\tau)+\tilde{Y}_D(\tau)]}(t) \leq N_{[\tilde{X}_D(\tau)+Y(\tau)]}(t) \leq N_{[X(\tau)+Y(\tau)]}(t) \quad (7)$$

The process $N_{[X(\tau)+Y(\tau)]}(t)$ on the right of the above inequality represents the number of packets in node 1 of Fig. 4 when the $X(\tau)$ and $Y(\tau)$ streams are applied directly with no time delay, while the middle term of the above inequality represents the corresponding number of packets when arrivals during the first D slots are deleted from the $X(\tau)$ stream. Likewise, the leftmost term considers the case when packets from both the $X(\tau)$ and $Y(\tau)$ streams are deleted during the first D slots.

However, note that the arrival process $\tilde{X}_D(\tau) + \tilde{Y}_D(\tau)$ is stochastically equivalent to the process $X(\tau - D) + Y(\tau - D)$ representing a delayed version of inputs $X(\tau)$ and $Y(\tau)$. Likewise, the arrival process $\tilde{X}_D(\tau) + Y(\tau)$ is stochastically equivalent to the process $X(\tau - D) + Y(\tau)$. We thus have the following stochastic equalities for all timeslots t:[2]

$$N_{[\tilde{X}_D(\tau)+\tilde{Y}_D(\tau)]}(t) \underset{st.}{=} N_{[X(\tau)+Y(\tau)]}(t - D)$$
$$N_{[\tilde{X}_D(\tau)+Y(\tau)]}(t) \underset{st.}{=} N_{[X(\tau-D)+Y(\tau)]}(t)$$

[2] The relation $A \underset{st.}{=} B$ denotes stochastic equality of A and B, and the relation $A \underset{st.}{\leq} B$ denotes stochastic inequality (see [22]).

Using these stochastic inequalities in (7) yields:
$$N_{[X(\tau)+Y(\tau)]}(t-D) \underset{st.}{\leq} N_{[X(\tau-D)+Y(\tau)]}(t) \underset{st.}{\leq} N_{[X(\tau)+Y(\tau)]}(t)$$

The upper and lower bounds in the above inequality are time delayed versions of the same process, namely, the process of packets in node 1 of Fig. 4 when $X(\tau)$ and $Y(\tau)$ are applied directly. It follows that their time average distributions (defined in (1)) are equal, and converge if and only if the middle term converges. The middle term represents the process of packets in node 1 when the $X(\tau)$ stream first passes through the D-slot delay. Thus, the steady state distribution in node 1 is unchanged if the time delay is removed.

The proof of the corresponding property for node 2 is similar and is omitted for brevity. □

References

[1] D. Gamarnik. Stability of adaptive and non-adaptive packet routing policies in adversarial queueing networks. *Proc. of 31st ACM Symposium on the Theory of Computing*, 1999.

[2] F. P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1984.

[3] H. Daduna. *Queueing Networks with Discrete Time Scale*. Springer, 2001.

[4] R. Gallager. *Discrete Stochastic Processes*. Kluwer Academic Publishers, Boston, 1996.

[5] D. P. Bertsekas and R. Gallager. *Data Networks*. New Jersey: Prentice-Hall, Inc., 1992.

[6] R. Jafari and K. Sohraby. Combined m/g/1-g/m/1 type structured chains: A simple algorithmic solution and applications. *IEEE Proceedings of INFOCOM*, Vol. 2, April 2001.

[7] J. He and K. Sohraby. A new analysis framework for discrete time queueing systems with general stochastic sources. *IEEE Proceedings of INFOCOM*, Vol. 2, April 2001.

[8] J. W. Roberts and J. T. Virtamo. The superposition of periodic cell arrival streams in an atm multiplexer. *IEEE Transactions on Communications*, Feb. 1991.

[9] J. A. Morrison. Two discrete-time queues in tandem. *IEEE Transactions on Communications*, Vol. COM-27, No. 3, March 1979.

[10] S. K. Walley and A. M. Viterbi. A tandem of discrete-time queues with arrivals and departures at each stage. *Queueing Systems*, Vol. 23, pp. 157-176, 1996.

[11] Q. Wang and V. Frost. Estimation of cell loss probabilities for tandem atm queues. *IEEE Proceedings of the International Conference on Communications*, 1994.

[12] R. L. Cruz. A calculus for network delay. i. network elements in isolation. *IEEE Transactions on Information Theory*, vol. 37:1, pp. 114-131, Jan. 1991.

[13] R. L. Cruz. A calculus for network delay. ii. network analysis. *IEEE Transactions on Information Theory*, vol. 37:1, pp. 132-141, Jan. 1991.

[14] A. K. Parekh and R. Gallager. A generalized processor sharing approach to flow control in integrated services networks: The single-node case. *IEEE/ACM Transactions on Networking*, Vol. 1, pp. 344-357, 1993.

[15] M. J. Neely and C. E. Rohrs. Equivalent models and analysis for multi-stage tree networks of deterministic service time queues. *Proceedings of the 38th Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2001.

[16] M. J. Neely and C. E. Rohrs. Inequality comparisons and traffic smoothing in multi-stage atm multiplexers. *IEEE Proceedings of the International Conference on Communications*, 2000.

[17] E. Modiano, J. E. Wielselthier, and A. Ephremides. A simple analysis of average queueing delay in tree networks. *IEEE Transactions on Information Theory*, Vol. IT-42, no. 2, pp. 660-664, March 1996.

[18] M. Shalmon and M. A. Kaplan. A tandem network of queues with deterministic service and intermediate arrivals. *Operations Research*, Vol. 32, No. 4, July-August 1984.

[19] J. A. Morrison. A combinatorical lemma and its applications to concentrating trees of discrete-time queues. *Bell Sys. Tech. J.*, May-June 1978.

[20] M. J. Neely. *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, Massachusetts Institute of Technology, Laboratory for Information and Decision Systems (LIDS), 2003.

[21] W. Aiello, R. Ostrovsky, E. Kushilevitz, and A. Rosén. Dynamic routing on networks with fixed-size buffers. *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, Jan. 2003.

[22] S. Ross. *Stochastic Processes*. John Wiley & Sons, Inc., New York, 1996.