



Delay Analysis of All-Optical Packet-Switching Ring and Bus Communications Networks

Izhak Rubin, Jing Ling

Department of Electrical Engineering, UCLA, Los Angeles, CA 90095

E-mail: rubin@ee.ucla.edu, jing@ee.ucla.edu

Received October 4, 2000; Revised October 30, 2000

Abstract. We study the delay performance of all-optical packet communication networks configured as ring and bus topologies employing cross-connect switches (or wavelength routers). Under a cross-connect network implementation, a packet experiences no (or minimal) internal queueing delays. Thus, the network can be implemented by high speed all-optical components. We further assume a packet-switched network operation, such as that using a slotted ring or bus access methods. In this case, a packet's delay is known before it is fed into the network. This can be used to determine if a packet must be dropped (when its end-to-end delay requirement is not met) at the time it accesses the network. It also leads to better utilization of network capacity resources. We also derive the delay performance for networks under a store-and-forward network operation. We show these implementations to yield very close average end-to-end packet queueing delay performance. We note that a cross-connect network operation can yield a somewhat higher queueing delay variance levels. However, the mean queueing delay for all traffic flows are the same for a cross-connect network operation (under equal nodal traffic loading), while that in a store-and-forward network increases as the path length increases. For a ring network loaded by a uniform traffic matrix, the queueing delay incurred by 90% of the packets in a cross-connect network may be lower than that experienced in a store-and-forward network. We also study a store-and-forward network operation under a nodal round robin (fair queueing) scheduling policy. We show the variance performance of the packet queueing delay for such a network to be close to that exhibited by a cross-connect (all-optical) network.

Keywords: all-optical, packet-switching, WDM, queueing delay, simulation

1 Introduction

We investigate the queueing delay performance of packet-switching ring and bus communication networks that employ cross-connect switches. Such a system can be implemented as an all-optical network using wavelength routers (as cross-connect switches) in a WDM operation. It can also operate as an ATM network using cross-connect (such as VP cross-connect) switches. Under a properly configured cross-connect network operation, packets that have accessed the network incur no internal random queueing delays. Packet can experience random delays while waiting at their access nodes to access the network. Also, no packet losses are incurred within the network. A station accesses the ring network by attaching to an access switch port. By using cross-connect switches, very little buffering is

needed in the network. This property allows this network to be synthesized by high speed optoelectronic and optical components. Other optical cross-connect networks are described in Watanabe et al. [1], Iannone and Sebella [2]. As no random internal queueing delays are incurred, a packet's end-to-end delay is known at the time its access node is ready to feed it into the network. This can be used to determine if a packet must be dropped (when its end-to-end delay requirement is not met) before it is fed into the network. This also leads to better utilization of network capacity resources. We compare the queueing delay performances of cross-connect networks to that of store-and-forward networks.

Our study is motivated by the study of cross-connect meshed-ring networks that are evolved as follows. Ring architectures have been used by many network systems such as token ring, FDDI, SONET

ring [3], MetaRing [4], and ATMR [5] networks and are shown to be highly efficient and survivable. For a ring network with spatial reuse, more than one transmission can be initiated simultaneously in time. Under a uniform traffic matrix, a spatial-reuse factor of 4 is achievable and the throughput efficiency is increased to 400% of that of a ring network without spatial reuse [6]. By meshing the ring, even better performance is achieved. The throughput performance of cross-connect meshed-ring networks has been studied in Rubin and Ling [7–9] and its predecessor, the SMARTNet (scaleable multichannel adaptable ring terabit network) [10,11]. In these studies, the network is divided into multiple subnets for simpler routing decisions. Users communicate to each other via subnets only. We observe from these study results that subnets have the topology of either a ring or a bus. To start our study of queueing delay performance of cross-connect meshed-ring networks, we first study the queueing delay performance of ring and bus networks. Thus, the objective of this paper is to carry out end-to-end queueing delay analysis for packet-switching cross-connect networks that are configured as ring or bus topologies. We further provide comparison with the end-to-end queueing delay performance of corresponding store-and-forward networks. The performance of meshed-ring networks will be provided in a later study.

The rest of the paper is organized as follows. Section 2 describes the network architecture and its concept of operation. In Section 3, we derive the waiting time performance for ring networks under Poisson and on-off traffic processes and under a uniform traffic matrix. In Section 4, we study networks with bus topology under equal and unequal link traffic loading levels. When unequal link loading levels are employed, we assign a link bandwidth capacity that is either proportional to the link's loading level or is equal for all links. Conclusions are drawn in Section 5.

2 System Description

2.1 Network Topology and Operation

Assumptions

We investigate ring and bus network topologies for cross-connect network operation. The ring and bus networks can operate in a directional or bi-directional manner. However, since the study procedure is the

same for each direction, we concentrate on a directional network. The loading of the network is characterized by the terminal loading matrix $\underline{\sigma} = \sigma_{ij}$, where σ_{ij} denotes the external traffic flow rate from (source) terminals attached to switch node i to (destination) terminals attached to switch node j , $0 \leq i, j \leq K - 1, i \neq j$. We note that for $\sigma_{ii} = 0$, for all i . For ring networks, a uniform traffic matrix is assumed, so that $\sigma_{ij} = \sigma$, for all i and j . We note that a node carries traffic flows originating by its directly attached terminals (“external” flows) as well as traffic flows that are received across network links from other nodes (“internal” flows).

This network employs a spatial-reuse packet switching method. The information unit generated by a traffic source is called a packet. The transmission time of a maximum length packet is equal to a slot duration. For simplicity, we assume here that packets are of fixed length (equal to the maximum packet length). The access algorithm is described as follows. First, we describe a synchronous or a slotted access method [5]. The channel is slotted. Each slot includes a tag that identifies it as being idle or busy. When a station inserts a packet into an idle slot, it marks the slot as being in a busy state. When a station removes a packet from its slot, it marks the slot as idle. By using this access mechanism, there are multiple slots that traverse the network. For a network employing cross-connect switches, we assume that at each node, internal traffic has higher priority than external traffic. This ensures that the cross-connect switches operate correctly. A destination removal procedure is used, so that once an internal packet reaches its destination, it is removed from the network. An external packet can be inserted into the network at its source (originating) node when that node receives an empty slot or when the internal packet in the slot is removed. An asynchronous access mechanism (e.g., by employing a buffer insertion scheme [6]) can also be employed. A similar operation is achieved through the use of a maximum packet length buffer and requiring each station to defer access to a current incoming packet (if any).

We also study store-and-forward networks as a comparison benchmark. When an internal packet arrived at a node, it is either removed by the node (if it is destined to this node) or it is inserted into the shared queueing facility at the node. This facility buffers both external and internal packets.

We use the term waiting time to represent the waiting time in the queue or the queueing delay

(denoted as W) of a packet. The term delay (denoted as D) is used to indicate the waiting time plus the (transmission) service time for a packet.

When a packet arrives at a node server, its arrival time is recorded at the beginning of a slot. Two traffic models are used in this paper to characterize the node-to-node flow: Poisson traffic (which is constant “on”), and an on-off traffic process. For an on-off process, “on” periods and “off” periods are generated alternately. Traffic is generated only during an “on” period and it then follows a Poisson distribution. Both “on” and “off” period lengths are exponentially distributed.

2.2 Switch Structure

Each node is outfitted with a switch that has two types of ports: link ports and terminal ports. Each port actually represents a pair: an input port and an output port. The link ports are used to connect the switch node with other switch nodes. Terminals (or stations, hosts, users) are connected to a switch node and access the network via terminal ports. Each switch can have multiple terminal ports and thus be connected to multiple users. A traffic flow (or a packet) whose destination is a terminal attached to switch node k , will be directed at switch node k to the corresponding output terminal port. An output queueing mechanism is assumed, where an incoming packet from a terminal is switched to its corresponding output queueing facility with zero input queueing delay.

For a cross-connect network, when internal traffic enters a node, it is passed directly (without experiencing a random queueing delay) to the output port (by the nodal cross-connect switch). At such a node, external packets arriving from an attached terminal are switched to (and queued at) an output buffer that corresponds to every output port. Packets from external traffic is queued in the buffer corresponds to their intended output port. For a store-and-forward network, all packets (from both internal and external traffic flows) wait (for transmission service) in a shared queueing facility. When a first come, first serve (FCFS) service discipline is employed, each node allocates one output queue for each output port. Under a round robin service discipline (Section 3.4), each node assigns one output queue for each (internal and external) traffic flow that traverses this node.

For cross-connect networks, each switch node acts as a cross-connect switch in handling internal traffic flows. Switching is performed in accordance with a

pre-scheduled switching matrix. For a ring or bus topology, each switch has only two ports. Each port connects to a bi-directional link. An incoming packet, not destined to the underlying node, arriving at one port is always switched to the other port.

2.3 Network Routing Assumption

For a bi-directional ring network, there are two paths available for each packet. A packet is always routed by selecting the shortest path. For a bi-directional bus network, there is a single path available to choose. Since internal traffic flows have higher priority than external traffic flows, by assigning sufficient capacity to each subnet, we ensure that an internal traffic flow that arrives at a switch node is guaranteed to have sufficient capacity available to it so that it can be switched to its desired outgoing link without incurring queueing delays. In some cases, instead of assigning sufficient capacity, by checking if sufficient bandwidth is available on links before insert an external packet into the network, we ensure that an internal packet just experiencing minimal predictable queueing delay. As a result, no packet loss takes place internally in the network.

3 Ring Networks

We study the delay performance under cross-connect and store-and-forward operations and compare them. For an K -node ring network, we assume its nodes are numbered from 0 to $K - 1$.

3.1 Simulation Results for Ring Networks with Poisson Arrivals

We consider a network with a uniform traffic matrix, $\sigma_{ij} = \sigma$, for all i, j and $i \neq j$. Each traffic flow follows a Poisson arrival process with intensity σ . Every packet is routed over its shortest path across the ring. We only consider the waiting time (i.e., queueing delay) performance since the transmission delay is a fixed value of one slot for all cases. For a ring network with a uniform traffic matrix, due to symmetry, the traffic intensities on all links are the same. For a network with K nodes, if K is odd, then there are $\frac{K^2-1}{8}$ distinct traffic flows that share every link in each direction. For example, when $K = 49$, the number of latter traffic flows is 300. Let λ denote the total arrival rate to each node, then $\lambda = 300\sigma$. Let μ denote the service rate of each node, and set $\mu = 1$ (noting that the mean packet

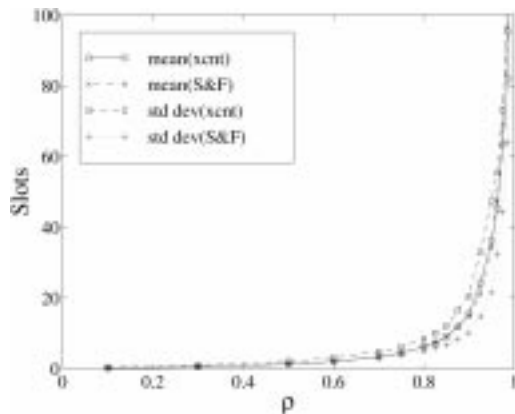


Fig. 1. Waiting time of a 49-node ring network.

service time, μ^{-1} , is equal to 1 slot). The normalized loading of each node, ρ , is equal to $\rho \triangleq \lambda/\mu = \lambda = 300\sigma$. We simulate the above described network for both cross-connect operation (where internal traffic has higher priority than external traffic) and store-and-forward operation with a FCFS nodal service policy, for $\rho < 1$. The mean and standard deviation values of the packet waiting time for both cases are plotted in Fig. 1, where *xcnt* denotes cross-connect networks and *S&F* denotes store-and-forward networks. For a cross-connect network, each packet experiences a queueing delay at the access node, while incurring no queueing delays at subsequent nodes along its route. For a store-and-forward network, all the nodal queueing delays experienced by a packet along its path to its destination are summed. The resulting value represents the (end-to-end) waiting time for the packet.

As we observe from Fig. 1, the mean waiting time values for both a cross-connect network and a store-and-forward network are almost the same, over the whole loading range. This can be explained intuitively as follows. Since both network operations employ work conserving service discipline and every packet uses the same path in both networks, any reduction in the waiting time value experienced by one packet (in comparing the two network operations) must be borne out by another packet. Therefore, the sum of incurred waiting time values (for all served packets) and consequently the average waiting time value, are the same for both network operations. Hence, $E(W_{SF}) = E(W_{XCNT})$. Thus, by employing a cross-connect network operation, which provides additional (network internal) service guarantee, we could still

achieve the same mean packet waiting time that attained by a store-and-forward operation. However, the waiting time standard deviation value incurred in a cross-connect network is higher than that experienced in a store-and-forward network. For a cross-connect network, consider a node, during a period in which the internal arriving traffic at the node is bursting for a long duration. In this case, external packets are blocked for an extended period of time. This is not necessarily the case for a store-and-forward network, where external packets could still access the network and be served on a FIFO basis, rather than be treated as lower priority packets. This phenomenon leads to higher waiting time variance performance for cross-connect networks. The effect of traffic burstiness on the waiting time variance is further accentuated when the network is loaded by the burstier on-off traffic processes (Section 3.4).

Although a cross-connect network operation leads to higher waiting time variance values when compared to a store-and-forward network operation, it offers the following advantages. In Fig. 2, we plot the average waiting time values for traffic flows of a given path length, under a loading rate of $\rho = 0.6$. For a store-and-forward network, the mean waiting time level increases as path length increases. In turn, under a cross-connect network operation, the mean waiting time level remains constant. This behavior repeats under any given loading level. Thus, if we want to guarantee a prescribed maximum average queueing delay level (for any flow in the network) regardless of the flow's path length, a cross-connect network would be a better choice. We have also plotted the standard

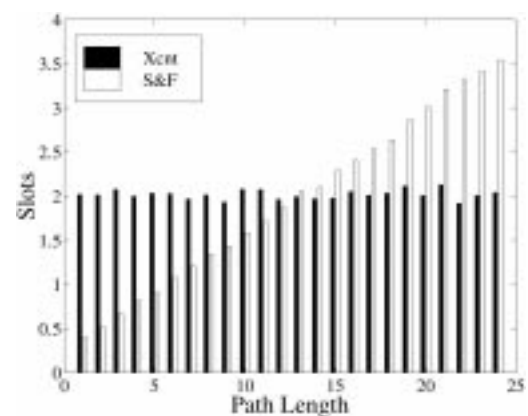


Fig. 2. Mean waiting time of a 49-node ring network with $\rho = 0.6$.

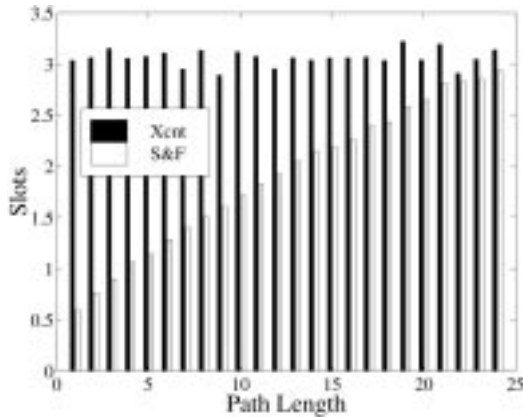


Fig. 3. Waiting time standard deviation of a 49-node ring network with $\rho = 0.6$.

deviation values of waiting time versus path length for $\rho = 0.6$, for each network operation, in Fig. 3. Although the average waiting time standard deviation values for a cross-connect network are higher than those attained in a store-and-forward network, the maximum standard deviation values over all path lengths for these two cases are close to each other, at $\rho = 0.6$. As ρ increases, the latter difference increases but is still less than the observed difference between average standard deviation values. We conclude that if a targeted quality-of-service (QoS) objective involves the guarantee of a maximum waiting time variance value for every traffic flow, instead of guaranteeing an average waiting time standard deviation level, a cross-connect network could provide similar performance to that attained by a store-and-forward network operation, while also providing for no internal random queueing delays and thus accommodating an all-optical operation.

3.2 Analytical Results for Cross-Connect Ring Networks with Poisson Arrivals

Since external traffic is only inserted into the network when there is no internal traffic, at each nodal processor, internal traffic is granted a higher priority than external traffic. To calculate the waiting time behavior of packets queued in a nodal output buffer, we thus form a single server queueing system model. This server process transmits the packets switched at the nodal to a given output port. Arriving packets belong to two priority classes, each obtaining a fixed service time (equal to 1 slot). We model the input

queue as a two-class priority discrete-time queueing system using a nonpreemptive priority service policy to schedule packets at the output queue (i.e., the transmission of a lower priority packet is not interrupted by the arrival of a higher priority packet). The external traffic follows a Poisson arrival pattern. The internal stream of packet arrivals is not governed by the statistics of a Poisson process. It consists of packets that departed from upstream nodes. At most one packet arrival per slot can take place. Therefore, we approximate internal traffic by a Bernoulli arrival process. The waiting time distribution of the lower priority class in this model (denoted as W_2) provides an approximation to the waiting time distribution experienced by external packets. This analytical result does not characterize the waiting time distribution of the external traffic exactly since the internal arrival process is modeled in an approximate manner. For cross-connect networks, the waiting time of internal packets is equal to zero. By using results from Rubin and Tsai [12], we derive approximate values for the mean and variance values of the packet waiting time for external traffic.

Using Equation (3.14) in Rubin and Tsai [12], we obtain the mean waiting time for external packets to be given by

$$E(W_2) = (1 - \rho_1)^{-1} \left\{ \frac{\overline{N^2} - \overline{N}}{2(1 - \rho)} + \overline{N}_1 + \frac{1}{2} \left(\frac{\overline{N}_2^2}{\overline{N}_2} - 1 \right) \right\}, \quad (1)$$

for $\rho < 1$. The notations in Equation (1) are explained in Rubin and Tsai [12] and are described below. We let $N_n^{(j)}$ denote the number of priority- j packet arrivals during slot n , $n \geq 1$, $j = 1, 2$. For each j , the class j arrival process is represented by $N^{(j)} \triangleq \{N_n^{(j)}, n \geq 1\}$, which is a sequence of independent identically distributed (i.i.d.) random variables, with moments

$$\overline{N}_j^i = E \left(\left[N_n^{(j)} \right]^i \right), \overline{N}_j = \overline{N}_j^1, \quad (2)$$

and moment-generating function

$$N_j^*(z) = \sum_{k=0}^{\infty} z^k P(N_n^{(j)} = k), |z| \leq 1. \quad (3)$$

Priority-1 packets are granted higher priority than priority-2 packets. The arrival processes $N^{(1)}$ and $N^{(2)}$ are statistically independent. The i.i.d. arrival process N is defined as $N \triangleq \{N_n = N_n^{(1)} + N_n^{(2)}, n \geq 1\}$,

with moments $\overline{N^i}$, $\overline{N} = \overline{N^1}$, and moment generating function $N^*(z)$, where $N^*(z) = N_1^*(z)N_2^*(z)$. Then, $\rho_1 \triangleq \overline{N_1}\mu_1^{-1}$, $\rho_2 \triangleq \overline{N_2}\mu_2^{-1}$, and $\rho \triangleq \rho_1 + \rho_2$, where μ_1^{-1} and μ_2^{-1} , the mean service time of class 1 and 2 packets, are both equal to 1. We also know that $\overline{N} = \rho$ and $\overline{N^2} = \overline{N_1^2} + \overline{N_2^2} + 2\rho_1\rho_2$ [12].

Using Equation (3.13) in Rubin and Tsai [12], we obtain

$$\eta^*(z) = \frac{(1-\rho)(z-1)}{z-N^*(z)} N_1^*(z) \frac{1-N_2^*(z)}{\overline{N_2}(1-z)}, \quad (4)$$

for $\rho < 1$. Let $A(z) = \frac{(1-\rho)(z-1)}{z-N^*(z)}$, $B(z) = N_1^*(z)$, and $C(z) = \frac{1-N_2^*(z)}{\overline{N_2}(1-z)}$, then

$$\begin{aligned} \overline{\eta^2} - \overline{\eta} &= \lim_{z \rightarrow 1} \frac{d^2}{dz^2} \eta^*(z) \\ &= \lim_{z \rightarrow 1} A''(z) + B''(z) + C''(z) \\ &\quad + 2A'(z)B'(z) + 2A'(z)C'(z) \\ &\quad + 2B'(z)C'(z). \end{aligned} \quad (5)$$

We then calculate the limit of each term.

$$\lim_{z \rightarrow 1} A'(z) = \frac{\overline{N^2} - \overline{N}}{2(1-\rho)}, \quad (6)$$

$$\lim_{z \rightarrow 1} B'(z) = \overline{N_1}, \quad (7)$$

$$\lim_{z \rightarrow 1} C'(z) = \frac{1}{2} \left(\frac{\overline{N_2^2}}{\overline{N_2}} - 1 \right), \quad (8)$$

$$\lim_{z \rightarrow 1} A''(z) = \frac{2(1-\overline{N})(\overline{N^3} - 3\overline{N^2} + 2\overline{N}) + 3(\overline{N^2} - \overline{N})^2}{6(1-\overline{N})^2} \quad (9)$$

$$\lim_{z \rightarrow 1} B''(z) = \overline{N_1^2} - \overline{N_1}, \quad (10)$$

$$\lim_{z \rightarrow 1} C''(z) = \frac{2\overline{N_2^3} - 6\overline{N_2^2} + 4\overline{N_2}}{6\overline{N_2}}. \quad (11)$$

We also know that $\overline{\eta} = \lim_{z \rightarrow 1} A'(z) + B'(z) + C'(z)$, then the value of $\overline{\eta^2}$ can be calculated. From $G_1^*(z) = zN_1^*(G_1^*(z))$ (using Equation (3.4) in Rubin and Tsai [12]), where $G_1^*(z)$ is the z -transform of the busy-period duration (denoted as G_1 and measured in slots) in a regular discrete-time queueing system with only priority-1 packets, we derive

$$\begin{aligned} \overline{G_1^2} &= (1-\rho)^{-1} \left((1 + 2\overline{N_1}G_1 + \overline{N_1^2}G_1^2) - 1 \right. \\ &\quad \left. - \rho_1\overline{G_1} - \rho_1\overline{N_1}G_1^2 + \overline{N_1^2}G_1^2 - \rho_1\overline{G_1^2} + \overline{G_1} \right), \end{aligned} \quad (12)$$

where $\overline{G_1} = (1-\rho)^{-1}$. From $W_2^*(z) = \eta^*[zN_1^*(G_1^*(z))]$ (Equation (3.3) in Rubin and Tsai [12]), where $W_2^*(z)$ is the z -transform of W_2 , we derive

$$\begin{aligned} \overline{W_2^2} &= \overline{\eta^2} (1 + 2\overline{N_1}G_1 + \overline{N_1^2}G_1^2) \\ &\quad - \overline{\eta} (1 + \overline{N_1}G_1 + \overline{N_1^2}G_1^2 - \overline{N_1^2}G_1^2 + \overline{N_1}G_1^2 \\ &\quad - \overline{N_1^2}G_1^2) + \overline{W_2}, \end{aligned} \quad (13)$$

for $\rho < 1$, where $\overline{W_2} = \overline{\eta}(1-\rho)^{-1}$ and is also given in Equation (1). We thus can calculate for value for $\overline{W_2^2}$. The standard deviation of the waiting time of the external traffic is approximated by $\sqrt{\overline{w_2^2} - \overline{w_2}^2}$.

Assume the arrival processes are Poisson streams. Let σ_{ij} denote the rate of the traffic flow from node i to node j . We assume a uniform traffic matrix. Then $\sigma_{ij} = \sigma$, for all i and j and $i \neq j$. Let the number of internal traffic flows to each node be M_1 , and the number of external traffic flows be M_2 . The number of internal traffic flows represents the number of upstream flows that pass this node. Due to the symmetry in the network, the values of M_1 or M_2 are same for each node. Let λ_2 denote the traffic rate of priority-2 traffic for each node. Since the superposition of Poisson traffic streams is itself a Poisson process, and priority-2 traffic represents the sum total of all external (Poisson) flows, the priority-2 traffic flow is a Poisson stream with intensity $\lambda_2 = M_2\sigma$. The internal traffic loading at a node is the sum total of M_1 upstream flows. We approximate it as a Bernoulli process with intensity $\lambda_1 = M_1\sigma$. Then, $\overline{N_1} = \overline{N_1^2} = \overline{N_1^3} = \lambda_1$, and $\overline{N_2} = \lambda_2$, $\overline{N_2^2} = \lambda_2^2 + \lambda_2$, $\overline{N_2^3} = \lambda_2^3 + 3\lambda_2^2 + \lambda_2$. We also have

$$\overline{N} = \lambda_1 + \lambda_2, \quad (14)$$

$$\overline{N^2} = \lambda_2 + \lambda_2^2 + 2\lambda_1\lambda_2 + \lambda_1, \quad (15)$$

$$\begin{aligned} \overline{N^3} &= \lambda_2^3 + 3\lambda_2^2 + \lambda_2 + 3\lambda_1\lambda_2^2 \\ &\quad + 6\lambda_1\lambda_2 + \lambda_1. \end{aligned} \quad (16)$$

By using these values in Equations (1)–(16), $\overline{W_2}$ and $\overline{W_2^2}$ can be calculated, for $\rho < 1$.

For the case of a 49-node ring network, we have: $M_1 = 276$, $M_2 = 24$. Then, $\rho = \overline{N_1} + \overline{N_2} = \lambda_1 + \lambda_2 = 276\sigma + 24\sigma = 300\sigma$. The simulation and analy-

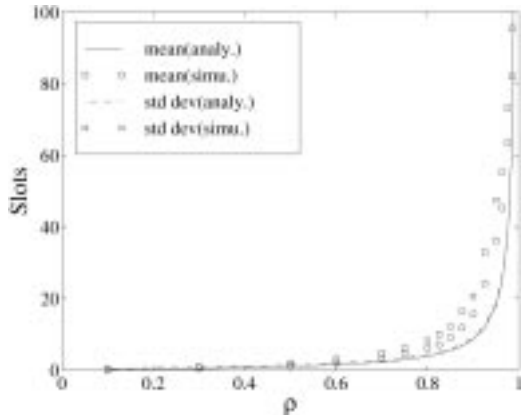


Fig. 4. Waiting time of a 49-node cross-connect ring.

tical results are plotted in Fig. 4. The exhibited simulation results are the average over all nodes in the network, while the analytical results correspond to a single node, noting that queueing delays at all nodes have the same distribution. The difference between the analytical and simulation results comes from the approximation of the internal traffic process used in our analytical model. We observe the analytical results for the mean queueing delay to follow closely the corresponding simulation results. Only small differences are noted between the standard deviation values. Hence, we conclude that the analytical result provides a good approximation for evaluating the system's performance behavior. We further note that the analytical performance evaluation approach described above is also applicable when other traffic matrices are assumed.

3.3 Analytical Results for Store-and-Forward Ring Networks with Poisson Arrivals

For a store-and-forward network with Poisson arrivals and exponential service times, it has been shown by Jackson that "the network acts as if each node can be viewed as an independent $M/M/1$ queue" (See, for example, p. 230 of Gross and Harris [13]). For our network, since the service time is deterministic, Jackson's result does not hold. To obtain an approximate performance evaluation, we nevertheless assume each node to act as an independent $M/D/1$ queue. Let λ denote the arrival rate (at a nodal switch output buffer) of the sum of all internal and external flows. Let μ denote the service rate of packets queued in the nodal output buffer. We have $\mu = 1$, as the

service time of each packet is equal to 1 slot. Let $W_{M/D/1}$ denotes the packet waiting time at an $M/D/1$ queue. Then, the mean waiting time (for $\lambda < \mu$) is equal to

$$E(W_{M/D/1}) = \frac{\lambda}{2\mu(\mu - \lambda)}. \quad (17)$$

Let $D_{M/D/1}$ denote the total delay incurred by a packet in its nodal output buffer, including its queueing and transmission delay, for an $M/D/1$ queue. For $\lambda < \mu$, its first moment, second moment, and variance values for this delay are given by

$$E(D_{M/D/1}) = \frac{\lambda}{2\mu(\mu - \lambda)} + \frac{1}{\mu}, \quad (18)$$

$$E(D_{M/D/1}^2) = \frac{\lambda^2 - 4\lambda\mu + 6\mu^2}{6(\mu - \lambda)^2\mu^2}, \quad (19)$$

$$\begin{aligned} \sigma^2(D_{M/D/1}) &= E(D_{M/D/1}^2) - E(D_{M/D/1})^2 \\ &= \frac{4\lambda_2\mu - \lambda_2^2}{12(\mu - \lambda_2)^2\mu^2}, \end{aligned} \quad (20)$$

using results from Gross and Harris [13]. The variance of $W_{M/D/1}$ is the same as that of $D_{M/D/1}$ since the packet service time is fixed.

We assume a uniform traffic matrix. Then $\sigma_{ij} = \sigma$, for all i and j and $i \neq j$. Let M denotes the total number of flows that traverse each node. Then $\lambda = M\sigma$ for every node. We assume that $\lambda < \mu$. For a 25-node ring network ($K = 25$), $M = \frac{K^2-1}{8} = 78$ and $\lambda = 78\sigma$. Let $\rho = \lambda/\mu$. For analytical results, the mean and variance values of the end-to-end waiting time are calculated by summing the values of the packet nodal waiting times over all nodes in the path of each traffic flow. Note that summing the nodal waiting time variances provides only an approximate evaluation of the end-to-end waiting time variance, since a packet waiting time in distinct nodes across its path may be correlated. For simulation results, the end-to-end waiting time value of each packet is calculated by summing the nodal waiting times over all nodes along the packet's flow path. We plot the simulation and analytical results in Fig. 5. We note that the analytical result for the mean and standard deviation of the waiting time is somewhat higher than the simulation result. Yet, they follow the same trend. The analytical result then provides a reasonably good upper bound. The analytical method serves as a numerically simple

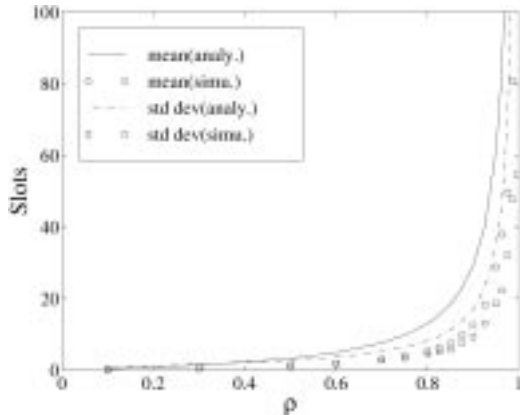


Fig. 5. Waiting time of a 25-node store-and-forward ring.

technique for estimating the system queueing delay performance. Since each node does not act as an independent $M/D/1$ queue, we expect the derived analytical result to overestimate the underlying performance levels. The analytical technique presented above is also applicable to other traffic matrix configurations.

3.4 Ring Networks with On-Off Traffic Sources

As mentioned in Section 3.1, under a cross-connect network, the burstiness of the internal traffic that loads a node tends to cause larger difference in the queueing delay of the external traffic from packet to packet than that observed in a store-and-forward network. Hence, under a cross-connect operation, packets experience higher waiting time variance levels. To investigate this phenomenon further, we assume that in this section that the network is loaded by terminals that generate traffic streams that are modeled as on-off processes, rather than as Poisson processes. To attain a better comparison of the performance features offered by the studied switching mechanisms, we further aim to reduce the internally incurred path queueing delays. This objective is automatically satisfied by cross-connect networks in which no internal queueing delays are experienced. In turn, for store-and-forward networks, we hereby employ a fair queueing scheduling algorithm at each node. We show the latter to lead to the reduction of internal packet queueing delays. Consequently, a comparison of the delay-throughput performance of these two network systems is carried out.

For an on-off traffic process, let the traffic rate be λ_p

during any “on” period. Let r denote the peak-to-average traffic rate ratio. The duration of each ‘on’ (“off”) period is exponentially distributed with average T_B (T_I). Then $T_C \triangleq T_B + T_I$ is the average length of a cycle (including an “on” period and an “off” period). For a uniform traffic matrix, $\sigma = \sigma_{ij}$ (for all i and j , and $i \neq j$) is the average traffic intensity of each i - j flow. Therefore, $\sigma = \lambda_p \frac{T_B}{T_C}$ and

$$r = \frac{\lambda_p}{\sigma} = \frac{T_C}{T_B}. \quad (21)$$

In our simulation, we set the value of T_B to be equal to $T_B = \frac{100}{\lambda_p}$ (measured in slots). For a 25-node network, given the normalized network loading level, ρ , we could calculate all corresponding values, for $\rho < 1$. With $\mu = 1$, $\rho = 78\sigma$ (Section 3.3) and $\rho = 78\lambda_p/r$, then $\lambda_p = \rho r/78$. Also, $T_B = \frac{100}{\lambda_p} = \frac{7800}{\rho r}$ and $T_C = \frac{7800}{\rho}$.

In previous sections, under a store-and-forward network operation, a FCFS service policy has been employed. By using such a service policy, with equal traffic intensity at all nodes (due to symmetry), the mean packet waiting time values at all nodes are approximately the same. This is different from the performance results obtained under a cross-connect network operation, where packets only experience queueing delays at the access node. The advantage of incurring no internal queueing delay has been discussed in Section 1. To provide a similar service guarantee in a store-and-forward network, a round robin service (fair queueing type) policy is employed. Under the round robin service policy, each traffic flow uses a different output queue, and each queue is polled by the server consecutively. When an empty queue is found, the server immediately skip it and proceeds to poll the next queue. Since all traffic flows are of the same intensity, using this scheduling policy offers each queue an equal chance to access the server. As shown in Parekh and Gallager [14], for a network employing generalized processor sharing (GPS) servers with leaky bucket input control, the end-to-end packet delay level can be upper bounded by the delay value incurred at the most congested node along the route. Although we do not employ a leaky bucket input flow regulation mechanism, nor use a GPS scheduling scheme, our results leads to a similar conclusion. That is, the end-to-end packet queueing delay incurred by a flow’s packets is dominated by the queueing delay experienced at the most congested node along the flow’s path. Hence, as demonstrated in

our simulations, we obtain the following property. Under a store-and-forward network operation, although we cannot reduce internal queueing delay levels to zero (as achieved under a cross-connect network operation), we can reduce internal packet queueing delay values to a small fraction of the end-to-end packet delay, for a higher fraction of the packets.

We study the waiting time performance of a 25-node ring network using the following operation assumptions: (1) a cross-connect network operation (*xcnt*), (2) a store-and-forward network operation employing a FCFS service policy (*FCFS*), and (3) a store-and-forward network operation using a round robin scheduling policy (*RR*). We also use two values of r , $r = 5$ and $r = 60$, to realize different levels of bustiness. We compare the end-to-end packet waiting time performance over different loading levels (ρ , for $\rho < 1$). The mean and standard deviation values of the packet waiting time are plotted in Figs. 6 and 7, respectively. The mean end-to-end packet waiting time value is approximately the same for all three cases, as explained in Section 3.1. The end-to-end packet waiting time standard deviation level achieved under a cross-connect network operation is higher than that attained under a store-and-forward network operation. This is induced by the bustiness features of the internal traffic processes, as noted above. By using a round robin service policy (rather than FCFS), the packet waiting time performance observed under a store-and-forward network operation is brought closer to that obtained under a cross-connect network operation. We show later in this section that the internal queueing delay attained under a store-and-

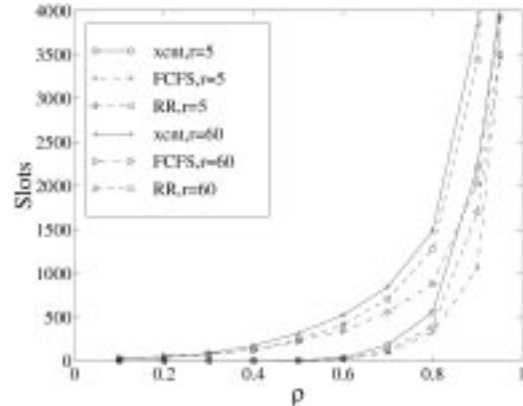


Fig. 7. Waiting time standard deviation of a 25-node ring with on-off source.

forward network operation using a round robin service policy is lower than that achieved under the use of a FCFS service policy. We also compare the 90-percentile and the 99-percentile end-to-end waiting time levels (i.e., 90% or 99% of packets experience queueing delay values that are lower than or equal to the latter levels). In Tables 1 and 2, we exhibit the waiting time percentile performance vs. the loading level for 25-node ring networks with bustiness ratios of $r = 5$ and $r = 60$, respectively. We observe that store-and-forward network operations generally lead to lower 99-percentile waiting time levels as compared to those experienced in a cross-connect network. In turn, the 90-percentile waiting time levels are realized under cross-connect network operations are close to (and sometimes lower than) those experienced under store-and-forward network operations. The relatively low 90-percentile waiting time values observed under a cross-connect network operation are

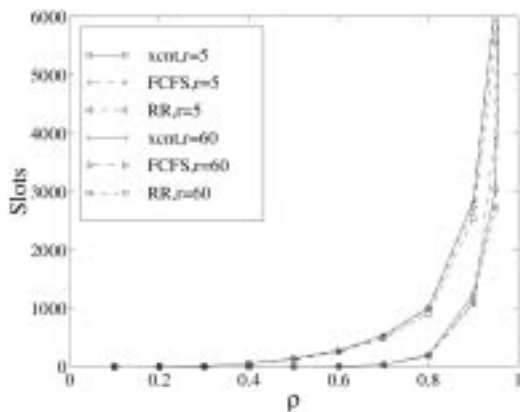


Fig. 6. Mean waiting time of a 25-node ring with on-off source.

Table 1. Waiting time of a 25-node ring with $r = 5$.

ρ	90-Percentile			99-Percentile		
	FCFS	RR	xcnt	FCFS	RR	xcnt
0.1	1	0	0	1	2	2
0.2	1	1	1	2	3	3
0.3	2	2	2	4	4	5
0.4	2	3	3	6	6	7
0.5	4	4	4	10	10	14
0.6	9	8	8	64	61	54
0.7	93	74	35	498	566	840
0.8	567	552	603	1612	1925	2841
0.9	2709	3017	3796	5231	8349	10375

Table 2. Waiting time of a 25-node ring with $r = 60$.

ρ	90-Percentile			99-Percentile		
	FCFS	RR	xcnt	FCFS	RR	xcnt
0.1	1	1	1	2	2	3
0.2	2	2	2	12	12	11
0.3	27	23	10	215	214	307
0.4	155	148	146	544	578	845
0.5	376	359	434	999	1177	1540
0.6	719	684	829	1665	2017	2539
0.7	1295	1274	1526	2782	3464	3993
0.8	2350	2306	2726	4483	6252	7147
0.9	6057	6331	7320	11221	16776	17669

induced by the attained equal end-to-end packet queueing delay levels for all traffic flows, as explained in Section 3.1 and as illustrated in Fig. 2. The packet waiting time performance curves achieved under a round robin service policy are generally between the corresponding curves characterizing the other two cases.

For store-and-forward network using a round robin scheduling policy, in observing the queueing delays incurred by flow's packets across the flow's path, we expect the bulk of the delay to be experienced at the bottleneck node (i.e., the node across the path that experiences the highest loading level). To study this behavior, we have measured, along each flow's path, the fraction of the queueing delay incurred at the bottleneck node (in relation to the overall end-to-end packet delay) for each flow. In Fig. 8, we show the percent level of the end-to-end queueing delay incurred by a packet at the bottleneck node for different burstiness (r) and loading (ρ) values. At low network loading, a high fraction of the overall end-to-end queueing delay is concentrated at a single node. The probability that the latter node is the access node is very high (it is equal to 0.95 when $\rho = 0.1$). As the network loading intensity increases, a large portion of the end-to-end packet queueing delay still occurs at a single node, even for a path length value of 12. The probability that the latter node is the access node decreases as the loading level (ρ) increases. Yet, even for high loading value of $\rho = 0.995$, the latter probability level is still equal to about 0.7. In contrast, under a store-and-forward network operation that employs a FCFS scheduling policy, the end-to-end packet queueing delay is spread evenly along the path. Thus, by using a round robin service policy, a store-and-forward network operation leads to lower internal

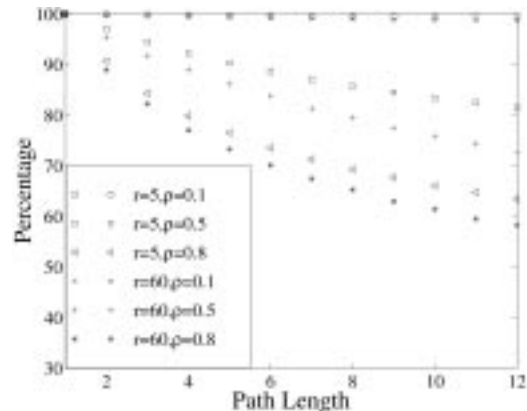


Fig. 8. Percent of end-to-end queueing delay incurred at the bottleneck node.

queueing delay levels. In comparison with a FCFS policy, the packet waiting time performance is closer to that achieved under a cross-connect network operation.

4 Bus Networks

We study the queueing delay performance of networks with a unidirectional bus topology layout. For a K -node bus network, nodes are sequentially numbered from 0 to $K - 1$. We consider traffic loading conditions that involve equal and unequal link traffic loading levels. For the latter case, the capacity of each link can be assigned in two ways: it is either set equal for all links or it is selected in proportion to the link's loading level.

4.1 Bus Networks Under Equal Link Loading

Assume the traffic loading intensity to be the same for every node of the bus network. We then assign equal link capacity values. We study a 7-node bus (with 6 links) under the following traffic matrix.

$$\underline{\sigma} = \begin{bmatrix} 0 & \sigma & \sigma & \sigma & \sigma & 0 & 0 \\ 0 & 0 & \sigma & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & \sigma & \sigma & 0 & \sigma \\ 0 & 0 & 0 & 0 & \sigma & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 & \sigma & 2\sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & 2\sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (22)$$

This traffic matrix characterizes a single direction of a bi-directional bus network. Consequently, $\sigma_{ij} = 0$,

for $i > j$. The traffic flow from node i to j follows a Poisson arrival process with intensity σ_{ij} . Let λ denote the total traffic rate loading to each node. By employing the traffic matrix represented by Equation (22), we conclude that $\lambda = 7\sigma$ for every node. Let μ denote the service rate at each node. We set $\mu = 1$, by assuming the service time of each packet to be equal to one slot. Then, $\rho \triangleq \lambda/\mu = 7\sigma$. The mean and standard deviation values of the end-to-end packet waiting time incurred under cross-connect and store-and-forward network operations, for $\rho < 1$, are plotted in Fig. 9. The waiting time performance results are similar to those observed for ring networks with Poisson arrival processes. That is, the incurred mean end-to-end packet waiting time values are almost the same under both network operations (i.e., under cross-connect and store-and-forward) and a cross-connect network operation leads to higher packet waiting time standard deviation values. Analytical results for this bus network configuration can be obtained by using the equations derived in Sections 3.2 and 3.3.

4.2 Bus Networks Under Unequal Link Loading

In this section, we study a network configuration with unequal link loading levels. To demonstrate bus network performance under an unequal link loading pattern, we select the following two loading cases. (We note that the following traffic matrices have been attained from the 49-node meshed ring network presented in Section 1. It is loaded according to a uniform traffic matrix. They represent the traffic flows

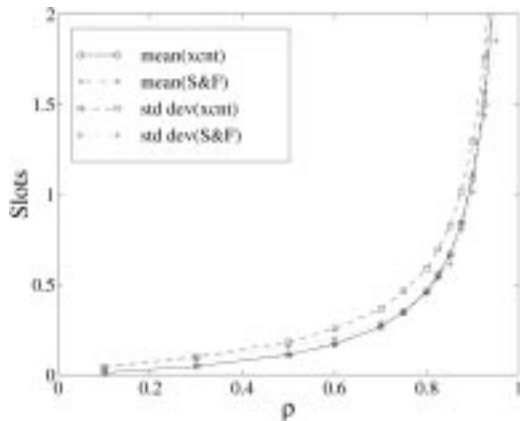


Fig. 9. Waiting time of a 7-node bus network with equal loading.

across two selected 8-node and 7-node bus subnetworks of this 49-node network.)

$$\underline{\sigma}_1 = \begin{bmatrix} 0 & \sigma & \sigma & \sigma & \sigma & \sigma & \sigma & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma & \sigma & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma & \sigma & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma & \sigma & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (23)$$

$$\underline{\sigma}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & \sigma & \sigma & \sigma \\ 0 & 0 & 0 & 0 & \sigma & \sigma & \sigma \\ 0 & 0 & 0 & 0 & \sigma & \sigma & \sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (24)$$

First, we assign each link the same capacity level. Under a cross-connect operation, by using such an assignment and granting higher priority to internal flows, when an internal packet enters a node, it can always be immediately served without experiencing any queuing delay. This simplifies the network implementation. In this case, let ρ denote the normalized loading of the most heavily loaded node. Each traffic flow follows a Poisson process. We again assume that the packet service time, μ^{-1} , is equal to one slot, for all nodes. Then, $\rho = 9\sigma$ for both traffic matrices (for the corresponding 8-node and 7-node bus network). For $\rho < 1$, the mean and standard deviation values of the end-to-end queuing delay achieved under traffic matrices $\underline{\sigma}_1$ and $\underline{\sigma}_2$ are plotted in Figs. 10 and 11, respectively. We again observe that the incurred mean packet waiting time values are similar for both cross-connect and store-and-forward network operations. Under a cross-connect network operation, the waiting time standard deviation values are somewhat higher. By using the equations derived in Sections 3.2 and 3.3, we obtain the corresponding analytical results. Note that for the first node in a cross-connect bus network, the internal traffic flow rate is zero. Therefore, the waiting time performance

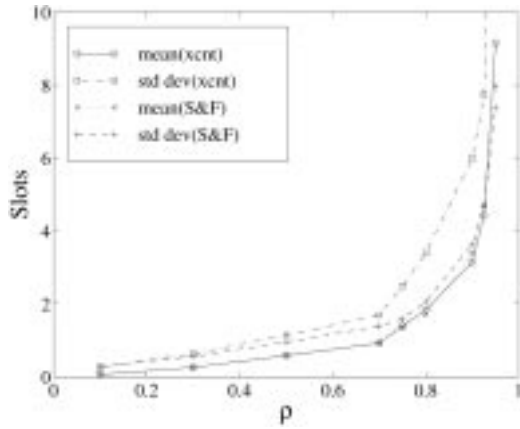


Fig. 10. Waiting time of a bus network with equal bandwidth (σ_1).

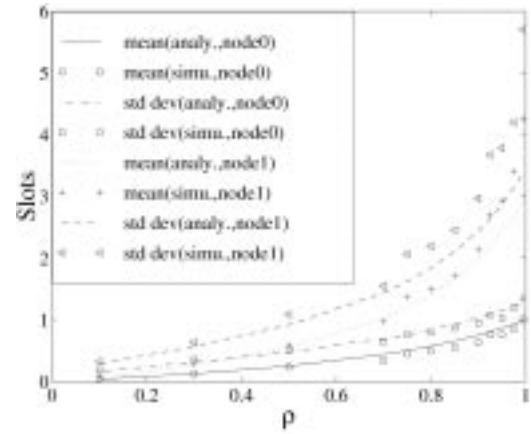


Fig. 12. Nodal waiting time of a cross-connect bus network with equal bandwidth (σ_1).

results for an $M/D/1$ queue can be applied. With traffic matrix $\underline{\sigma}_1$, we plot the packet waiting time performance results for node 0 and node 1 under a cross-connect network operation in Fig. 12, and those for paths of length 1 and 8 under a store-and-forward network operation in Fig. 13. We observe that the derived analytical results provide a good approximation for evaluating the system's performance behavior. The reasons for the differences noted between analytical and simulation results are explained in Sections 3.2 and 3.3.

Next, we assign link capacity levels that are proportional to the traffic intensity on each link. Hence, the fractional loading level of each link and each node is equal to a constant value ρ . For given total capacity, this assignment method leads to a higher maximum throughput when compared to the

previous assignment method. However, this assignment method requires a more complicated network implementation under a cross-connect network operation. Under a cross-connect network operation, we note that the packet size is the same in a network and a slot duration for a link is equal to the time required to transmit one packet at this link. When two links are assigned different capacity levels, these two links employ time slots of different sizes. Thus, to transmit a packet (received from the first link) using the server at the second link, the latter packet has to be buffered. Therefore, this operation induces some internal waiting times. The induced internal time values are small and are deterministically predictable. They lead to minor changes in the end-to-end network waiting time performance. An all-optical implementa-

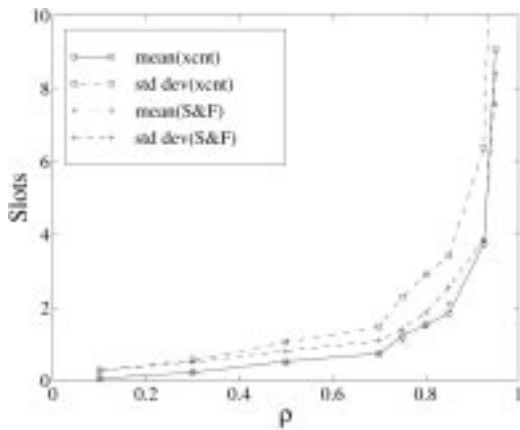


Fig. 11. Waiting time of a bus network with equal bandwidth (σ_2).

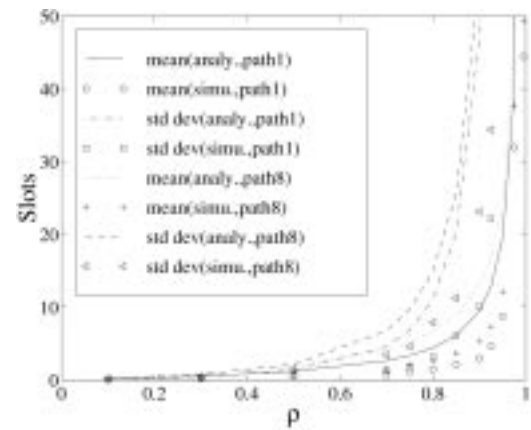


Fig. 13. End-to-end flow waiting time of a store-and-forward bus network with equal bandwidth (σ_1).

tion is still feasible. When equal link capacity is assigned (using the first assignment method), a packet can be inserted into the network whenever an empty slot is detected. When unequal link capacity values are assigned (using the second assignment method), we have to make sure that a packet will be granted sufficient capacity in all downstream nodes before we insert it into the network. The latter check can be done based solely on the information known at the current node.

For $\rho < 1$, we plot the mean and standard deviation values of the packet queueing delay incurred under traffic matrices σ_1 and σ_2 , using the proportional bandwidth assignment method, in Figs. 14 and 15, respectively. Under traffic matrix σ_2 , the packet waiting time performance results are similar to those obtained using the first assignment method. Under a cross-connect network operation employing σ_1 , when we do not regulate the input traffic, a noticeable decrease in the throughput capacity is observed. This is caused by terminal flows (such as the flow from node 4 to node 7) that are at high probability precluded from accessing the network due to the preferred access provided to upstream flows. A simple solution to this problem is to dedicate part of the bandwidth along the path of the latter flow to such a blocked flow. Using such a bandwidth reservation scheme, the queueing delay performance results realized under a cross-connect network operation (the case indicated by *mod* in Fig. 14) is again noted to be close to that attained under a store-and-forward network operation. Corresponding analytical results

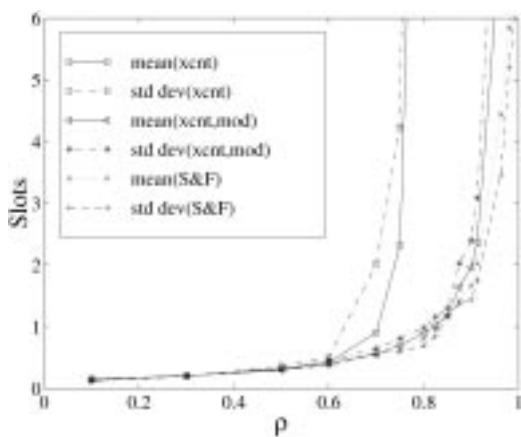


Fig. 14. Waiting time of a bus network with proportional bandwidth (σ_1).

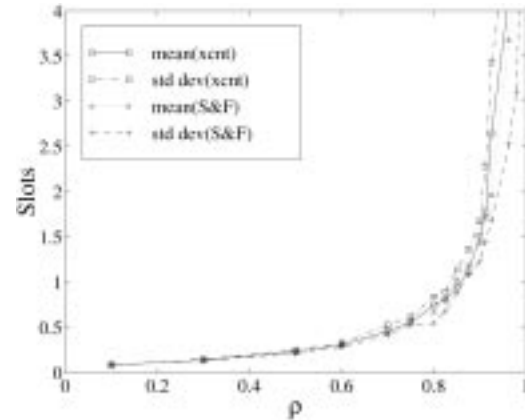


Fig. 15. Waiting time of a bus network with proportional bandwidth (σ_2).

for waiting time performance have been obtained (not shown here) by using the equations derived in Sections 3.2 and 3.3.

5 Conclusions

In this paper, we study the waiting time performance of cross-connect networks and compare the results to those obtained from store-and-forward networks. We study two types of network topologies: ring and bus. For a ring network, a uniform traffic matrix is used. The mean and standard deviation of the packet end-to-end waiting time (queueing delay) is derived for both cross-connect and store-and-forward network by using simulation and analytical methods, under a Poisson arrival process. To access the operation of the network under bursty traffic processes, we also load the network by an on-off type traffic stream. We obtain the simulation results for the above mentioned performance measures for cross-connect networks and for store-and-forward networks, under operated under FCFS and round robin scheduling policies. We also calculated the 90-percentile and 99-percentile waiting time under both modes of network operation.

For a network with bus topology, two traffic loading cases, corresponding to equal and unequal link loading conditions, are studied. When the traffic loading levels are equal for all links, equal bandwidth (link capacity) values are assigned to every link. Under unequal link loading levels, either equal link capacity is assigned to every link or a link capacity

proportional to the link loading level is assigned. While the former case results in a simpler network implementation, the latter case can accommodate higher throughput (when the total capacity value is the same in both cases). Simulation results representing the mean and standard deviation levels of the packet end-to-end waiting time are derived for a network that is configured as a bus. Analytical results are also obtained for such topologies using the performance equations obtained for a ring network topology.

We note the cross-connect network to exhibit the following advantages. It can accommodate an all-optical network since packets experience no internal queueing delays. Such a network offers high link transmission rate levels. As no random internal queueing delays are incurred, a packet's end-to-end delay is known at the time its access node is ready to feed it into the network. This can be used to determine if a packet must be dropped (when its end-to-end delay requirement is not met) before it is fed into the network. This also leads to better utilization of network capacity resources. In this paper, we have shown that under a cross-connect implementation the resulting mean packet queueing delay is approximately the same as that experienced in a store-and-forward network. Although a cross-connect network results in higher standard deviation value of the queueing delay, the difference is generally minor, especially at low network loading levels. This is true for both ring and bus networks. Also, as we observe for a ring network with a uniform traffic matrix, for 90% of all packets, the queueing delay experienced in a cross-connect network is either similar or less than that incurred in a store-and-forward network. When a round robin scheduling policy is employed by the nodes in a store-and-forward network, the internal packet queueing delay is reduced and the end-to-end queueing delay performance becomes closer to that exhibited by a cross-connect network.

Acknowledgments

This work was supported by ARO Grant No. DAAG55-98-1-0338 and by Pacific Bell and University of California MICRO Grant No. 98-131

and by office of Naval Research (ONR) Contract No. N00014-01-C-0016.

References

- [1] A. Watanabe, S. Okamoto, K. Sato, Optical path cross-connect system architecture suitable for large scale expansion, *IEEE Journal of Lightwave Technology*, vol. 14, no. 10, (Oct. 1996), pp. 2162–2172.
- [2] E. Iannone, R. Sabella, Optical path technologies: A comparison among different cross-connect architectures, *IEEE Journal of Lightwave Technology*, vol. 14, no. 10, (Oct. 1996), pp. 2184–2196.
- [3] J. Drake, A review of the four major SONET/SDH rings, *Proc. of ICC'93*, (Geneva, Switzerland, May 1993), vol. 2, pp. 878–884.
- [4] Y. Ofek, Overview of the MetaRing architecture, *Computer Networks and ISDN Systems*, vol. 26, no. 6–8, (March 1994), pp. 817–829.
- [5] K. Imai, T. Ito, H. Kasahara, N. Morita, ATMR: asynchronous transfer mode ring protocol, *Computer Networks and ISDN Systems*, vol. 26, no. 6–8, (March 1994), pp. 785–798.
- [6] I. Cidon, Y. Ofek, MetaRing—a full-duplex ring with fairness and spatial reuse, *IEEE Transaction on Communications*, vol. 41, no. 1, (Jan. 1993), pp. 110–120.
- [7] I. Rubin, J. Ling, Survivable all-optical cross-connect meshed-ring communications networks, *Proc. of SPIE*, (Dallas, TX, Nov. 1997), vol. 3228, pp. 280–291.
- [8] I. Rubin, J. Ling, All-optical cross-connect meshed-ring communications networks using a reduced number of wavelengths, *Proc. IEEE INFOCOM'99*, (New York, NY, March 1999), vol. 2, pp. 924–931.
- [9] I. Rubin, J. Ling, All-optical WDM cross-connect meshed-ring communications networks, *European Transactions on Telecommunications*, vol. 11, no. 1, (Jan. 2000), pp. 83–90.
- [10] I. Rubin, H.-K. Hua, An all-optical wavelength-division meshed-ring packet-switching network, *Proc. IEEE INFOCOM'95*, (Boston, MA, April 1995), vol. 3, pp. 969–976.
- [11] I. Rubin, H.-K. Hua, SMARTNet: An all-optical wavelength-division meshed-ring packet-switching network, *Proc. IEEE GLOBECOM'95*, (Singapore, Malaysia, Nov./Dec. 1995), vol. 3, pp. 1756–1760.
- [12] I. Rubin, Z.-H. Tsai, Message delay analysis of multiclass priority TDMA, FDMA, and discrete-time queueing systems, *IEEE Transactions on Information Theory*, vol. 35, no. 3, (May 1989), pp. 637–647.
- [13] D. Gross, C. M. Harris, *Fundamentals of Queueing Theory*, second edition, (Wiley, 1985).
- [14] A. K. Parekh, R. G. Gallager, A generalized processor sharing approach to flow control in integrated services networks: the multiple node case, *IEEE/ACM Transactions on Networking*, vol. 2, no. 2, (April 1994), pp. 137–150.

Izhak Rubin received the B.Sc. and M.Sc. from the Technion - Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree from Princeton University, Princeton, NJ, all in Electrical Engineering. Since 1970, he has been on the faculty of the UCLA School of Engineering and Applied Science where he is currently a Professor in the Electrical Engineering Department.



Dr. Rubin has had extensive research, publications, consulting, and industrial experience in the design and analysis of commercial and military computer communications and telecommunications systems and networks. At UCLA, he is leading a large research group. He also serves as President of IRI Computer Communications Corporation, a leading team of computer communications and telecommunications experts engaged in software development and consulting services.

Professor Rubin served as co-chairman of the 1981 IEEE International Symposium on Information Theory; as program chairman of the 1984 NSF-UCLA workshop on Personal Communications; as program chairman for the 1987 IEEE INFOCOM conference; and as program co-chair of the IEEE 1993 workshop on Local and Metropolitan Area networks. Dr.

Rubin has been elected as a Fellow of IEEE for his contributions to the analysis and design of computer communications networks. He has served as editor of the IEEE Transactions on Communications, of the ACM/Baltzer journal on *Wireless Networks*, of the Kluwer journal on *Photonic Network Communications* and of the Wiley InterScience *International Journal on Communications Systems*.

Jing Ling received her B.S. degree in 1995 from University of Wisconsin, Madison and the M.S. and Ph.D. degree in 1996 and 2000, respectively, from University of California, Los Angeles, all in Electrical Engineering. Her research topic is about optical and ATM cross-connect networks. During summer 1999, she worked at HRL on Medium Access Control for satellite networks. During summer 1998, she worked at NORTEL/BayNetworks in the area of virtual private networks. In summer 1996, she worked at Motorola on the design of test equipment for telecommunication networks.

