Message Transmission with Timing Constraints in Ring Networks*

Ching-Chih Han
Department of Electrical Engineering
The Ohio State University
Columbus, OH 43210–1272
cchan@ee.eng.ohio-state.edu

Kang G. Shin
Real-Time Computing Laboratory
Dept. of Elect. Engin. and Computer Science
University of Michigan, Ann Arbor, MI 48109
kgshin@eecs.umich.edu

Abstract

We consider the message transmission problem in unidirectional slotted ring networks. We consider two operation modes: 1) the evacuation mode, *in which all messages arrive before system initialization and no new message arrives during system operation; 2) the* continuation mode, *in which new messages may arrive after system initialization.*

We study the performance of several message scheduling policies with respect to three performance measures: meeting all message deadlines, minimizing the maximum delay or the total length of busy periods, and minimizing the average delay. We show that the Least-Slack-time-First *(LSF) scheduling policy is optimal in evacuation mode operation with respect to meeting all message deadlines, while no optimal scheduling policy can possibly exist for continuation mode operation.*

For the other two performance measures, we show that in the case when messages may be of variable lengths, 1) the Farthest-Destination-First *(FDF) policy is optimal in terms of minimizing the maximum delay and minimizing the total length of busy periods for evacuation mode and for continuation mode operation, respectively, and 2) no optimal scheduling policy can possibly exist in terms of minimizing the average delay under either evacuation mode or continuation mode operation.*

1. Introduction

Message transmission is an important issue in both communication networks and multiprocessor interconnection systems. Effective transmission scheduling policies may greatly reduce the queueing delay of messages and improve the utilization of the overall communication network or enhance the performance of parallel computation. Ring is a simple, effective, and commonly-used network architecture for local area networks (LANs) [5, 9, 12] and for interconnecting the processors in a multiprocessor. To reduce message delay and to improve the system throughput of a ring network, several ring architectures, *medium access control* (MAC) protocols, and message transmission scheduling policies have been proposed to exploit spatial slot reuse [4, 5, 7, 8]. A ring network that employs spatial slot reuse allows multiple simultaneous transmissions in the network as long as the transmissions take place on different links.

To study the interaction of different traffic streams and their effects on the performance of a ring network that employs spatial slot reuse, Tassiulas and Joung [11] adopted a simple ring network model. They considered a unidirectional slotted ring in which each node can transmit to its downstream neighbor and receive from its upstream neighbor simultaneously, and studied the performance of several scheduling policies for transmitting unit-length messages, called *packets*, under two operation modes: 1) the *evacuation mode*, in which all the messages arrive before system initialization and no new message arrives during system operation; 2) the *continuation mode*, in which new messages may arrive after system initialization. They showed that 1) the *Farthest-Destination-First* (FDF) policy minimizes the maximum delivery time (or called *evacuation time*) in evacuation mode operation and minimizes the total length of busy periods in continuation mode operation, and 2) the *Closest-Destination-First* (CDF) policy minimizes the average delay of the packets in evacuation mode operation.

Several similar problems have been studied in the context of data transmission in multiprocessor systems with respect to different initial conditions of the data/messages and with the objective of minimizing the evacuation time [3, 6, 10]. These problems include: 1) *scattering*: one processor has messages for some other processors, 2) *gathering*: some processors have messages for a specific processor, 3) *broadcasting*: a processor has one specific message for all the other processors, and 4) *gossiping*: each processor has one specific message for all the other processors. Saad and Schultz [10] studied the above problems for different parallel architectures, including broadcast bus model, shared memory model, ring architecture, two and three-dimensional mesh-connected arrays, hypercubes, and switch connected model. Fraigniaud *et al.* [6] focused on the scattering problem in a ring architecture and

*The work reported in this paper was supported in part by the ONR under Grant N00014-94-1-0229 and by the NSF under Grant MIP-9203895.

1052-8725/96 $5.00 © 1996 IEEE

showed that the FDF policy is optimal in terms of minimizing the evacuation time. Bhatt *et al.* [3] studied the scattering and gathering problems with variable-length messages in a tree network, and showed that the FDF policy is optimal for scattering and proposed a near-optimal scheduling policy for gathering in general trees.

All of the aforementioned work is concerned with minimizing either the evacuation time or the average delay. To our best knowledge, the issue of guaranteeing the timely delivery of isochronous (real-time) messages with delivery deadlines has not been addressed in ring networks with spatial slot reuse. However, for embedded real-time applications [1] such as air-traffic controls, automated factories, and industrial process controls, hard real-time messages need to be delivered by their respective deadlines for the timely completion of diverse real-time applications. Moreover, in some hard real-time systems, failure to meet the deadlines of the messages may even lead to catastrophic consequences. Conventional performance objectives, such as maximizing the throughput or minimizing the maximum delivery time, are not of the most important concern to hard real-time systems. Instead, predictable and dependable timing guarantees must be provided.

The main intent of this paper is, thus, to study message transmission problems with timing constraints in ring networks. We adopt the same ring network model as that used in [11], consider the same operation modes, i.e., the evacuation mode and the continuation mode, but relax the assumption that messages are of unit lengths, i.e., we consider messages of variable lengths. Messages are fragmented into fixed-length cells/packets at their source nodes, and are reassembled at their destination nodes. Each cell needs one time slot to transmit from a node to its downstream neighbor, and may be forwarded one hop per slot independently as long as the transmission sequence of all the cells of a message is preserved at each intermediate node. We show that the *Least-Slack-time-First* (LSF) scheduling policy is optimal in evacuation mode operation with respect to meeting all message deadlines, while no optimal scheduling policy can possibly exist for continuation mode operation.

Another salient contribution of this paper is that we generalize the results in [11] and study the performance of message transmission scheduling policies with respect to minimizing the evacuation time or minimizing the average delay for the case that messages may be of variable lengths. We show that for the case of variable-length messages, 1) the *Farthest-Destination-First* (FDF) policy is still optimal in terms of minimizing the maximum delay and minimizing the total length of busy periods under evacuation mode and continuation mode operations, respectively, and 2) no optimal scheduling policy can possibly exist in terms of minimizing the average delay under either evacuation mode or continuation mode operation. For the cases in which no optimal scheduling policy exists, we then conduct simulations and empirically compare the performance of several commonly-used scheduling policies, including FIFO, FDF, CDF (*Closest-Destination-First*), SMF (*Shortest-Message-First*), LSF, and EDF (*Earliest-Deadline-First*).

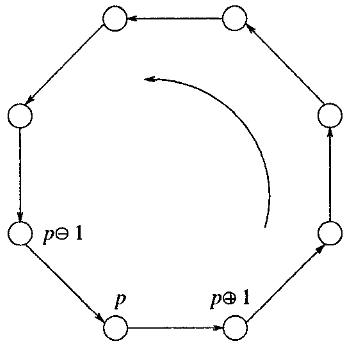


Figure 1. A unidirectional slotted ring network with 8 nodes.

The rest of the paper is organized as follows. In Section 2, we formally define the message transmission problem to be addressed here. In Section 3, we study the message transmission problem in which each message is associated with a delivery deadline, and the objective of a scheduling policy is to meet all message deadlines. In Section 4, we comment on the impact of variable-length messages on the message transmission problem with the two performance measures used in [11]. For the cases in which no optimal scheduling policy can possibly exist, we first give formal non-existence proofs, and then conduct simulations to compare the performance of several scheduling policies. The results of these simulations are presented in Section 5. The paper concludes with Section 6.

2. Problem formulation

The network under consideration is a unidirectional *slotted* ring (Fig. 1) with N nodes numbered 0, 1, ..., $N-1$. Define $p \oplus q = (p+q) \bmod N$ and $p \ominus q = (p-q) \bmod N$. The transmission link between two neighboring nodes p and $p \oplus 1$ is denoted as $L(p, p \oplus 1)$, or simply $(p, p \oplus 1)$ if no ambiguity may arise. Each node p may transmit to its downstream neighbor $p \oplus 1$ and receive from its upstream neighbor $p \ominus 1$ simultaneously. Moreover, all nodes may transmit at the same time to exploit spatial reuse, and they are synchronized to begin transmission in the beginning of each slot (the slot length is the same for all nodes). Without loss of generality, we assume time is measured in slots and each slot begins at an integral time instant, i.e., each time interval $[t, t+1]$, $t \geq 0$, is a slot and is denoted as $S[t, t+1]$, or simply $[t, t+1]$. Messages are fragmented into fixed-length cells/packets at their source nodes, and are reassembled at their destination nodes. Each cell needs one time slot to transmit from a node to its downstream neighbor, i.e., the actual transmission time plus the propagation delay and communication startup overhead is equal to one slot. Each cell of a message may be forwarded one hop per slot independently as long as the transmission sequence of all the cells of a message is preserved at each intermediate node (and hence, the cells arrive at the destination in their original order). That is, message (but not cell) transmission is preemptable at cell boundaries—the transmissions of the cells of different messages on a link can interleave with each other.

A message M_i in the ring network is characterized by $(a_i, \ell_i, N_i^s, N_i^d, d_i)$, where

- a_i is the arrival time of M_i, i.e., the time M_i is ready for transmission at its source node,
- ℓ_i is the length of M_i (measured in cells/packets),
- N_i^s is the source node id of M_i,
- N_i^d is the destination node id of M_i, and
- d_i is the *absolute* deadline of M_i, i.e., (all the cells of) M_i must be delivered to its destination node N_i^d no later than time d_i.

Note that if we need not consider individual message deadlines, we set $d_i = \infty$ for all i. We denote the j-th cell of the i-th message M_i as M_{ij}, $1 \leq j \leq \ell_i$. If M_i has only one cell (i.e., $\ell_i = 1$), we denote the only cell as M_i and interchangeably call M_i a message or a cell.

A transmission scheduling policy determines the transmission priorities of the cells at each node. We consider two operation modes: *evacuation* (or *static*) and *continuation* (or *dynamic*). In evacuation (static) mode operation, there is a set, $\mathbf{M} = \{M_1, M_2, \ldots, M_n\}$, of n messages in the ring network at system initialization and after system initialization, no new message may arrive at any node in the ring network, i.e., $a_i = 0$ for all i. In continuation (dynamic) mode operation, new messages may arrive at the nodes in the ring network during system operation, i.e., $a_i \geq 0$ for all i.

To facilitate the presentation, we define the following notation. Let π be a scheduling policy. The transmission schedule generated by π acting on a set, $\mathbf{M}$, of messages is denoted by a two-dimensional array (table), $\pi^{\mathbf{M}}$. The (t,p)-th (t-th row, p-th column) entry $\pi^{\mathbf{M}}(t,p)$, $t \geq 0$ and $0 \leq p \leq N-1$, in the transmission table describes the status (activity) at node p (link $(p, p \oplus 1)$) at time t (slot $[t, t+1]$). Specifically, $\pi^{\mathbf{M}}(t,p) = (M, \mathbf{A}; M')$ denotes that cell M is transmitted from node $p \ominus 1$ to node p during slot $[t-1, t]$ and M's destination is p, cells in $\mathbf{A}$ are residing (waiting for transmission) at node p at time t, and cell M' is transmitted by node p to its downstream neighbor $p \oplus 1$ during slot $[t, t+1]$. Note that when a cell M is delivered to its destination p in slot $[t-1, t]$, M is passed onto the higher communication transport layer at node p and exits from the ring network. This is the reason why M is not included in the $\mathbf{A}$ component of an entry $\pi^{\mathbf{M}}(t,p)$. Furthermore, M will henceforth not appear in any entry $\pi^{\mathbf{M}}(t',p)$, for $t' > t$ and $0 \leq p \leq N-1$. Also, $\pi^{\mathbf{M}}(t,p) = (M, \mathbf{A}; M')$ (and M' is non-null) implies $M' \in \mathbf{A}$. Note that in the context of ring networks, the transmission table actually wraps around at the left and right ends. For notational convenience, the superscript will be dropped in the following discussion as long as it doesn't cause any ambiguity, i.e., π denotes both the scheduling policy and the transmission table (schedule) of π acting on a message set.

Example 1 Table 1 shows the transmission schedule of the *Farthest-Destination-First* (FDF) policy acting on the set of messages, $\mathbf{M} = \{M_1 = (0,2,0,2,\infty), M_2 = (0,2,1,3,\infty), M_3 = (0,1,0,4,\infty)\}$. (As will be discussed later, the FDF policy gives priority to the message with longest way to go at each node, and ties can be broken arbitrarily by a specific rule.) A "-" in the corresponding component of an entry denotes a null/empty cell/set. □

Each row of a transmission table is called a *state*. Each cell M_{ij} will appear in at most one entry in each row of the transmission table, and if M_{ij} appears in the entry $\pi(t,p)$, then M_{ij} will only appear in either $\pi(t+1, p \oplus 1)$ (if M_{ij} is transmitted on link $(p, p \oplus 1)$ in slot $[t, t+1]$) or $\pi(t+1, p)$ (if M_{ij} is not transmitted on link $(p, p \oplus 1)$ in slot $[t, t+1]$). If we draw a line on the transmission table to trace the progress of the delivery of a cell M_{ij}, the line will be piecewise linear and each linear segment goes either from north to south or from northwest to southeast. We call the line the *transmission path* of M_{ij} (under policy π). Note that the transmission path of M_{ij} is different from the unique physical path, $(N_i^s, N_i^s \oplus 1, \ldots, N_i^d)$, through which M_{ij} is delivered. The transmission paths of all the cells in Example 1 are shown in Fig. 2, in which the two solid lines are the transmission paths of the two cells of M_1, respectively, the two dashed lines are the transmission paths of the two cells of M_2, respectively, and the dotted line is the transmission path of M_3.

	0; $(0,1)$	1; $(1,2)$	2; $(2,3)$	3; $(3,4)$	4; $(4,5)$
0; $[0,1]$	-, $\{M_{11},M_{12},M_3\}$; M_3	-, $\{M_{21},M_{22}\}$; M_{21}	-, -; -	-, -; -	-, -; -
1; $[1,2]$	-, $\{M_{11},M_{12}\}$; M_{11}	-, $\{M_{22},M_3\}$; M_3	-, $\{M_{21}\}$; M_{21}	-, -; -	-, -; -
2; $[2,3]$	-, $\{M_{12}\}$; M_{12}	-, $\{M_{22},M_{11}\}$; M_{22}	-, $\{M_3\}$; M_3	M_{21}, -; -	-, -; -
3; $[3,4]$	-, -; -	-, $\{M_{11},M_{12}\}$; M_{11}	-, $\{M_{22}\}$; M_{22}	-, $\{M_3\}$; M_3	-, -; -
4; $[4,5]$	-, -; -	-, $\{M_{12}\}$; M_{12}	M_{11}, -; -	M_{22}, -; -	M_3, -; -
5; $[5,6]$	-, -; -	-, -; -	M_{12}, -; -	-, -; -	-, -; -

Table 1. Transmission schedule generated by the FDF policy in Example 1.

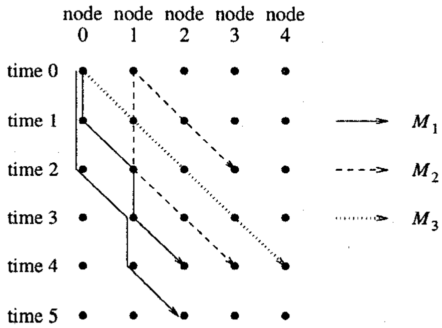


Figure 2. The transmission paths of $M_1 = (0,2,0,2,\infty)$, $M_2 = (0,2,1,3,\infty)$, and $M_3 = (0,1,0,4,\infty)$ under the FDF scheduling policy.

The transmission paths of two cells M_{ij} and M_{kl} are said to *intersect* each other at time t at node p (i.e., at the (t,p)-th entry of π) if one cell, say M_{kl}, arrives at p later than the other, M_{ij}, but M_{kl} is transmitted to node $p\oplus 1$ in slot $[t,t+1]$ before M_{ij} is transmitted to $p\oplus 1$. Note that intersection is a symmetric relation, i.e., if path 1 intersects path 2 then path 2 intersects path 1. For example, in Fig. 2, the *only* two transmission paths that intersect each other are those of M_3 and M_{22} and they intersect each other at FDF(1, 1). The transmission paths of M_{11} and M_{12} do not intersect each other. Nor does the transmission path of M_{11} intersect that of M_{22}.

Let D_i^π denote the *delivery time* of message M_i under scheduling policy π, i.e., D_i^π is the time (the last cell of) M_i is delivered to N_i^d under scheduling policy π. The *delay* of M_i under π is defined to be $D_i^\pi - a_i$. For evacuation mode operation, we consider the following performance measures:

- the *evacuation time* $V(\pi)$ (also called *schedule length* or *makespan*), which is defined to be the *maximum delivery time*, i.e., $V(\pi) = \max_i\{D_i^\pi\}$, and
- the *average delay* $A(\pi)$, which is defined as $A(\pi) = \frac{1}{n}\sum_{i=1}^{n}(D_i^\pi - a_i)$.

For continuation mode operation, in addition to the average delay $A(\pi)$, we consider the total length of *busy periods*, $B(\pi)$, where a busy period is defined as a time interval $[t,t']$ during which the ring network is nonempty and in slots $[t-1,t]$ and $[t',t'+1]$ the ring network is empty.

Tassiulas and Joung [11] studied message transmission in a ring network in which each message is of unit length, i.e., $\ell_i = 1$ for all i, and showed that under evacuation mode operation, the FDF policy, which gives priority to the message with longest way to go at each node (ties can be broken arbitrarily by a specific rule), minimizes the evacuation time, and the *Closest-Destination-First* (CDF) policy, which gives priority to the message with shortest way to go at each node, minimizes the average delay. They also showed that under continuation mode operation, the FDF policy maximizes the fraction of the time at which the ring is empty (i.e., minimizes the total length of busy periods) for any message arrival pattern.

In what follows, we study the problem of transmitting variable-length messages with timing constraints in a ring network. In addition, we also study the impact of variable-length messages on the performance of message transmission with respect to the performance measures mentioned above, i.e., minimizing $V(\pi)$ or $B(\pi)$, and minimizing $A(\pi)$.

3. Message transmission with deadline constraints

3.1. Evacuation mode operation

In this section, we consider message transmission with deadline constraints in evacuation mode operation, i.e., the arrival times, a_i's, of all messages are 0. We first consider the case when all messages have unit length, i.e., $\ell_i = 1$ for all i, and show that the Least-Slack-time-First (LSF) policy, rather than the Earliest-Deadline-First (EDF) policy, is optimal in terms of meeting message/cell deadlines. We then relax the assumption and consider the case when messages may have variable lengths.

A commonly-used scheduling policy for messages/cells with hard deadlines is the EDF policy which gives priority to the message/cell with closest deadline.

However, EDF is not optimal in the ring network model (even for the case of $\ell_i = 1$ for all i).

The reason why EDF is not optimal is that EDF does not take into account the distance from the current position, p, of a cell M_i at time t to its destination, N_i^d. To take this into account, we define the slack time, $s_i(t,p)$, of M_i at node p at time t as

$$\begin{aligned} s_i(t,p) &= (d_i - t) - (N_i^d \ominus p), \\ &= \begin{cases} (d_i - t) - (N_i^d - p) & \text{if } 0 \le p \le N_i^d, \\ (d_i - t) - (N_i^d + N - p) & \text{if } N_i^d < p < N. \end{cases} \end{aligned}$$

In the above definition, $d_i - t$ is the *relative* deadline (relative to the current time t) of M_i and $N_i^d \ominus p$ is the remaining transmission time (remaining distance to the destination) of M_i. Note that with the above definition, the slack time of a cell changes dynamically with time and with the current position of the cell. Note, however, that two cells compete for transmission on a link only if they are residing at the same node at the same time. Therefore, to ease the calculation of cell slack times, we further define $s_i(p)$, $1 \le i \le n$ and $0 \le p \le N-1$, as follows:

$$s_i(p) = \begin{cases} d_i - N_i^d & \text{if } 0 \le p \le N_i^d, \\ d_i - N_i^d - N & \text{if } N_i^d < p < N. \end{cases} \tag{1}$$

It is easy to see that $s_i(t,p) - s_j(t,p) = s_i(p) - s_j(p)$. Moreover, the difference between the slack times of two cells does not change with time and cell positions as long as the two cells are residing at the same node at the time of comparison. Formally, we have the following theorem.

Theorem 1 Suppose at time t_1, both cells M_i and M_j are residing at node p_1. If at time t_2, both M_i and M_j are residing at node p_2, then $s_i(t_1,p_1) - s_j(t_1,p_1) = s_i(t_2,p_2) - s_j(t_2,p_2)$ (and hence, $s_i(p_1) - s_j(p_1) = s_i(p_2) - s_j(p_2)$).

Proof: Assume, without loss of generality, that the order of the three nodes p_1, N_i^d, and N_j^d in the direction of the ring is $p_1 \to N_i^d \to N_j^d$. It is easy to see that the order of the three nodes p_2, N_i^d, and N_j^d in the direction of the ring is $p_2 \to N_i^d \to N_j^d$ since if the order is $N_i^d \to p_2 \to N_j^d$, then $t_1 < t_2$ ($t_1 > t_2$) implies that M_i (M_j) is transmitted beyond its destination N_i^d (N_j^d) and reaches node p_2 (p_1) at time t_2 (t_1). Therefore, we have

$$(N_j^d \ominus p_1) - (N_i^d \ominus p_1) = (N_j^d \ominus N_i^d) = (N_j^d \ominus p_2) - (N_i^d \ominus p_2), \tag{2}$$

and

$$\begin{aligned} & s_i(t_1,p_1) - s_j(t_1,p_1) \\ = \;& (d_i - t_1) - (N_i^d \ominus p_1) - ((d_j - t_1) - (N_j^d \ominus p_1)) \\ = \;& d_i - (N_i^d \ominus p_1) - (d_j - (N_j^d \ominus p_1)) \\ = \;& d_i - (N_i^d \ominus p_2) - (d_j - (N_j^d \ominus p_2)) \\ = \;& s_i(t_2,p_2) - s_j(t_2,p_2). \end{aligned}$$

□

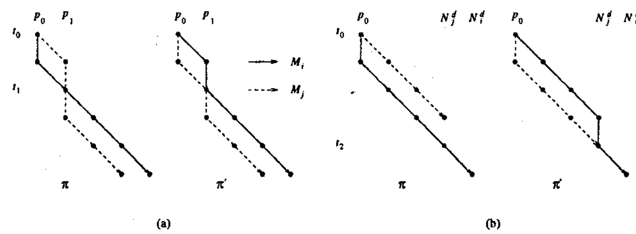


Figure 3. The construction of π' in Lemma 1.

As mentioned earlier, two cells will compete for transmission on a link only if they are residing at the same node at the same time. Therefore, Theorem 1 implies that it suffices for us to use Eq. (1) as the definition of cell slack times, in which case each cell's slack time needs to be updated at most once at the time when the cell passes through node 0. The resulting implementation of the LSF policy (to be discussed next) is greatly simplified. We will henceforth use Eq. (1) as the definition of slack times.

Since the slack time serves as an index of how tight the message/cell laxity is, the LSF policy, which gives priority to the cell with the smallest slack time[1] (i.e., each node p always transmits first the cell with the smallest slack time among all the cells currently residing at p), should give the best performance. We formally prove in Theorem 2 that the LSF policy is optimal in the sense that if a set of messages is schedulable (i.e., the deadlines of all the message cells can be met) then the LSF policy is guaranteed to generate a feasible schedule for the set.

The following two lemmas are needed for the proof of the optimality of the LSF policy (Theorem 2).

Lemma 1 Given a policy π, suppose the schedule produced by π acting on a set of messages is feasible, and suppose at time t_0, both M_i and M_j are residing at node p_0. If the destination, N_i^d, of M_i is the same as or farther than that, N_j^d, of M_j from node p_0, i.e., $N_i^d \ominus p_0 \ge N_j^d \ominus p_0$, and the slack time, $s_i(p_0)$, of M_i is smaller than that, $s_j(p_0)$, of M_j, but policy π transmits M_j in the slot $[t_0, t_0+1]$ from node p_0 to node $p_0 \oplus 1$, then there exists a policy π' such that $\pi'(t,p) = \pi(t,p)$ for all $t \le t_0$ and all $0 \le p \le N-1$, except that p_0 transmits M_i, instead of M_j, to node $p_0 \oplus 1$ in slot $[t_0, t_0+1]$ under π', and the resulting schedule generated according to π' is also feasible.

Proof: There are two cases to consider.

Case 1: The transmission paths of M_i and M_j in π intersect each other at node p_1 at time $t_1 > t_0$.

Policy π' is constructed as follows (see Fig. 3(a)). π' is exactly the same as π except in the time interval $[t_0, t_1]$, in which whenever a node transmits M_j in

[1]Ties can be broken arbitrarily by a specific rule.

π, the node transmits M_i in π', and whenever a node transmits M_i in π, the node transmits M_j in π'. It is easy to see that in the resulting schedule generated according to policy π', every cell is delivered to its destination at the same time as in the schedule generated by policy π. Therefore, if π is feasible, so is π'.

Case 2: The transmission paths (after time t_0) of M_i and M_j in π do not intersect each other.

Policy π' is constructed as follows (see Fig. 3(b)). Let t_2 be the time when M_i is transmitted to N_j^d under π. π' is exactly the same as π except in the time interval $[t_0, t_2]$, in which whenever a node transmits M_j in π, the node transmits M_i in π', and whenever a node transmits M_i in π, the node transmits M_j in π'. It is easy to see that $D_i^{\pi'} \leq D_i^{\pi} \leq d_i$, and hence M_i meets its deadline under policy π'. Since $D_j^{\pi'} = t_2$ and $D_i^{\pi} - t_2 \geq N_i^d \ominus N_j^d$, we have

$$D_j^{\pi'} \leq D_i^{\pi} - (N_i^d \ominus N_j^d),$$

and since $s_i(p_0) \leq s_j(p_0)$, $s_i(p_0) - s_j(p_0) = d_i - d_j - (N_i^d \ominus N_j^d) \leq 0$. Therefore,

$$D_j^{\pi'} \leq D_i^{\pi} - (N_i^d \ominus N_j^d) \leq D_i^{\pi} - (d_i - d_j) \leq d_j,$$

and hence M_j meets its deadline under policy π'. Moreover, since all the other cells are delivered to their destination at the same time as in the schedule generated according to policy π, if π is feasible, so is the resulting schedule generated according to policy π'.

The lemma follows from **Case 1** and **Case 2**. □

In Lemma 1, the original policy π transmits cell M_j with a closer destination N_j^d before cell M_i with a farther destination N_i^d at node p_0, while policy π' switches the transmission order at node p_0 and transmits M_i before M_j. A dual lemma of Lemma 1 can be similarly proved in which the original policy π transmits cell M_j with a farther destination before cell M_i with a closer destination at a node, while policy π' switches the transmission order at the node and transmits M_i before M_j.

Lemma 2 Given a policy π, suppose the schedule produced by π acting on a set of messages is feasible, and suppose at time t_0, both M_i and M_j are residing at node p_0. If the destination, N_i^d, of M_i is closer than that, N_j^d, of M_j, from node p_0, i.e., $N_i^d \ominus p_0 < N_j^d \ominus p_0$, and the slack time, $s_i(p_0)$, of M_i is smaller than that, $s_j(p_0)$, of M_j, but policy π transmits M_j in the slot $[t_0, t_0+1]$, from node p_0 to node $p_0 \oplus 1$, then there exists a policy π' such that $\pi'(t,p) = \pi(t,p)$ for all $t \leq t_0$ and all $0 \leq p \leq N-1$, except that p_0 transmits M_i, instead of M_j, to node $p_0 \oplus 1$ in slot $[t_0, t_0+1]$ under π', and the resulting schedule generated according to π' is also feasible. □

The proof of Lemma 2 is omitted due to space limitations.

Theorem 2 For the case of $\ell_i = 1$ for all i, the LSF policy is optimal in terms of meeting all deadlines under evacuation mode operation.

Proof: Given a set of messages/cells, suppose there is a feasible schedule π for this set but π does not follow the LSF policy. Let t_0 be the earliest time and p_0 be the smallest node id such that π does not follow the LSF policy, i.e., in slot $[t_0, t_0+1]$, node p_0 does not transmit the least-slack-time cell, say M_i, residing at p_0 to its downstream neighbor $p_0 \oplus 1$.

There are three cases to consider. First, if in slot $[t_0, t_0+1]$, the link $(p_0, p_0 \oplus 1)$ is idle, then we construct a schedule π' which is exactly the same as π except that node p_0 transmits M_i to node $p_0 \oplus 1$ in slot $[t_0, t_0+1]$ and p_0 is idle (under π') in the slot that p_0 transmits M_i to $p_0 \oplus 1$ under policy π. It is easy to see that $D_i^{\pi'} \leq D_i^{\pi}$ and $D_j^{\pi'} = D_j^{\pi}$ for all $j \neq i$. Second, if in slot $[t_0, t_0+1]$, node p_0 transmits another cell M_j whose destination N_j^d is not farther from p_0 than M_i's destination N_i^d (i.e., $N_j^d \ominus p_0 \leq N_i^d \ominus p_0$), then by Lemma 1, there exists a feasible policy π' such that $\pi'(t,p) = \pi(t,p)$ for all $t \leq t_0$ and all $0 \leq p \leq N-1$, except that π' follows the LSF policy and transmits M_i, instead of M_j, from p_0 to $p_0 \oplus 1$ in slot $[t_0, t_0+1]$. Finally, if in slot $[t_0, t_0+1]$, node p_0 transmits another cell M_j which has a farther destination N_j^d from p_0 than M_i (i.e., $N_j^d \ominus p_0 > N_i^d \ominus p_0$), then by Lemma 2, there exists a feasible policy π' such that $\pi'(t,p) = \pi(t,p)$ for all $t \leq t_0$ and all $0 \leq p \leq N-1$, except that π' follows the LSF policy and transmits M_i, instead of M_j, from p_0 to $p_0 \oplus 1$ in slot $[t_0, t_0+1]$.

In all of the above three cases, schedule π' is feasible and follows the LSF policy for all $t \leq t_0$ and all $p \leq p_0$. By repeating the above process, eventually we will get a feasible schedule which follows the LSF policy at all times and at all nodes. □

For the case when ℓ_i's may not be all equal to 1, since each cell of a message may be transmitted independently, we need to define the slack time for each cell. Since it is required that the cells of a message are transmitted to their destination in their original order, the ℓ_i-th cell of M_i, M_{i,ℓ_i}, must be delivered to its destination no later than time d_i, the $(\ell_i - 1)$-th cell of M_i, M_{i,ℓ_i-1}, must be delivered to its destination no later than time $d_i - 1$, and so on. Therefore, we define the deadline, d_{ij}, of each cell M_{ij} of message M_i as $d_{ij} = d_i - (\ell_i - j)$, and the slack time, $s_{ij}(p)$, of M_{ij} at node p as follows:

$$s_{ij}(p) = \begin{cases} d_{ij} - N_i^d & \text{if } 0 \leq p \leq N_i^d \\ d_{ij} - N_i^d - N & \text{if } N_i^d < p < N \end{cases} \quad (3)$$

With the above definition of slack times, we show in the following theorem that the LSF policy is still optimal for the case when ℓ_i's may be arbitrary positive integers.

Theorem 3 The LSF policy is optimal (in terms of meeting all message deadlines) for evacuation mode operation.

Proof: Since the LSF policy treats all the cells of a message independently, by Theorem 2, we know that if a set of cells is schedulable, then the LSF policy will produce a feasible schedule for the set of cells, i.e., each cell M_{ij} will meet its deadline d_{ij}, for all i and j. Moreover, if two cells, M_{ij} and M_{ik}, of M_i are residing at the same node p at the same time t, then $j < k$ implies $s_{ij}(p) < s_{ik}(p)$, and hence the LSF policy will transmit the cells of a message in the correct order at any time at any node. □

3.2. Continuation mode operation

In this section, we prove in Theorem 4 by using the *adversary argument* (a detailed account of which can be found in [2]) that there does not exist *any* optimal scheduling policy (in terms of meeting cell/message deadlines) under continuation mode operation.

Theorem 4 There does not exist any optimal scheduling policy (in terms of meeting cell/message deadlines) under continuation mode operation even for the case of $\ell_i = 1$ for all i.

Proof: Suppose there are two messages, $M_1 = (0,1,0,3,5)$ and $M_2 = (0,1,0,1,2)$, that arrived at node 0 at time 0. If a scheduling policy chooses M_1 to transmit first, we assume that there is another message, $M_3 = (1,1,0,1,2)$, that arrived at node 0 at time 1. It is easy to see from the schedule of Table 4(a) that either M_2 or M_3 will miss its deadline. However, as shown in the schedule of Table 4(b), if we transmit M_2 first during slot $[0,1]$, then all the three messages can meet their deadlines.

On the other hand, if the scheduling policy chooses M_2 to transmit first in slot $[0,1]$. We assume that there are two other messages $M_3 = (2,1,1,2,3)$ and $M_4 = (3,1,1,2,4)$. As shown in the schedule of Table 4(a), at least one of M_1, M_3, and M_4 cannot meet its deadline. Note that after transmitting M_2 to node 1 in slot $[0,1]$, node 0 transmits M_1 to node 1 in slot $[1,2]$. At time 2 (i.e., before the beginning of slot $[2,3]$), M_3 arrives at node 1. Since the deadline of M_3 is 3, node 1 must transmit M_3 during slot $[2,3]$ in order to meet its deadline. At time 3, M_4 arrives at node 1. Since M_4 has a deadline 4, node 1 must transmit M_4 to M_4's destination (node 2) during slot $[3,4]$ (transmitting M_1 to node 2 in slot $[3,4]$ will cause M_4 to miss its deadline). Then it will be too late for M_1 to meet its deadline. However, the transmission schedule of Table 4(b) is feasible for the same message set.

From the above adversary argument, we conclude that there cannot exist any optimal scheduling policy (without knowing the future message arrivals) for continuation mode operation even for the case of $\ell_i = 1$ for all i. □

Since there does not exist any optimal scheduling policy for continuation mode operation, we resort to simulations and compare the performance of several commonly-used scheduling policies in Section 5.

4. Minimizing evacuation time and average delay

In this section, we discuss the impact of variable-length messages on the performance of several scheduling policies with respect to the two performance measures used in [11], i.e., minimizing the evacuation time (or total length of busy periods) and minimizing the average delay.

4.1. Minimizing evacuation time or total length of busy periods

As mentioned earlier, if each message has unit length, i.e., $\ell_i = 1$ for all i, Tassiulas and Joung [11] proved that for evacuation mode operation, the FDF scheduling policy minimizes the evacuation time. Moreover, they also showed that the FDF policy maximizes the fraction of the time at which the ring is empty, i.e., the FDF policy minimizes the total length of busy periods, in continuation mode operation. By a similar argument as we generalize the optimality (in terms of meeting message/cell deadlines) of the LSF policy from the case of $\ell_i = 1$ for all i to the case of $\ell_i \geq 1$ for all i (Theorem 3), we can also show that both of the above results still hold in the case that messages may be of variable lengths since a message of length ℓ_i may be viewed as ℓ_i messages of unit-length.

4.2. Minimizing average delay

Tassiulas and Joung [11] show that if all messages have unit length, the *Closest-Destination-First* (CDF) policy is optimal under evacuation mode operation in the sense that the average (total) delay is minimized. However, the CDF policy is no longer optimal if messages may have variable lengths. Note that the delay of a message is defined to be the delivery time of the (last cell of the) message minus the arrival time of the message. Again, using the adversary argument, we can show that there does not exist *any* optimal schedul-

	0; (0, 1)	1; (1, 2)	2; (2, 3)	3; (3, 4)
0; [0, 1]	-, $\{M_1, M_2\}$; M_1	-, -; -	-, -; -	-, -; -
1; [1, 2]	-, $\{M_2, M_3\}$; M_2 or M_3	-, $\{M_1\}$; M_1	-, -; -	-, -; -
2; [2, 3]	-, $\{M_3\}$ or $\{M_2\}$; deadline miss	M_2 or M_3, -; -	-, $\{M_1\}$; M_1	-, -; -

(a)

	0; (0, 1)	1; (1, 2)	2; (2, 3)	3; (3, 4)
0; [0, 1]	-, $\{M_1, M_2\}$; M_2	-, -; -	-, -; -	-, -; -
1; [1, 2]	-, $\{M_1, M_3\}$; M_3	M_2, -; -	-, -; -	-, -; -
2; [2, 3]	-, $\{M_1\}$; M_1	M_3, -; -	-, -; -	-, -; -
3; [3, 4]	-, -; -	-, $\{M_1\}$; M_1	-, -; -	-, -; -
4; [4, 5]	-, -; -	-, -; -	-, $\{M_1\}$; M_1	-, -; -
5; [5, 6]	-, -; -	-, -; -	-, -; -	M_1, -; -

(b)

Table 2. Illustrative schedules used in the proof of Theorem 4.

	0; (0, 1)	1; (1, 2)	2; (2, 3)	3; (3, 4)
0; [0, 1]	-, $\{M_1, M_2\}$; M_2	-, -; -	-, -; -	-, -; -
1; [1, 2]	-, $\{M_1\}$; M_1	M_2, -; -	-, -; -	-, -; -
2; [2, 3]	-, -; -	-, $\{M_1, M_3\}$; M_3	-, -; -	-, -; -
3; [3, 4]	-, -; -	-, $\{M_1, M_4\}$; M_1 or M_4	M_3, -; -	-, -; -
4; [4, 5]	-, -; -	-, $\{M_4\}$ or $\{M_1\}$; deadline miss	M_1 or M_4, -; -	-, -; -

(a)

	0; (0, 1)	1; (1, 2)	2; (2, 3)	3; (3, 4)
0; [0, 1]	-, $\{M_1, M_2\}$; M_1	-, -; -	-, -; -	-, -; -
1; [1, 2]	-, $\{M_2\}$; M_2	-, $\{M_1\}$; M_1	-, -; -	-, -; -
2; [2, 3]	-, -; -	M_2, $\{M_3\}$; M_3	-, $\{M_1\}$; M_1	-, -; -
3; [3, 4]	-, -; -	-, $\{M_4\}$; M_4	M_3, -; -	M_1, -; -
4; [4, 5]	-, -; -	-, -; -	M_4, -; -	-, -; -

(b)

Table 3. Illustrative schedules used in the proof of Theorem 4.

ing policy under evacuation mode operation if messages may be of variable lengths.

Theorem 5 There does not exist any optimal scheduling policy (in terms of minimizing the average/total delay) under evacuation mode operation if messages may have variable lengths. □

The proof of the theorem is omitted due to space limitations.

The above theorem also implies that there cannot exist any optimal scheduling policy in terms of minimizing the average delay under continuation mode operation if $\ell_i \geq 1$ for all i since evacuation mode operation can be considered as a special case of continuation mode operation. By a similar argument, we can also show that under continuation mode operation, there cannot exist any optimal scheduling policy in terms of minimizing the average delay even in the case of $\ell_i = 1$ for all i. The proof is left for the interested reader.

5. Performance evaluation

Table 4 summarizes all the findings derived in Sections 3–4 and in [11], i.e., the optimal scheduling policies, if any, for meeting all message deadlines, for minimizing the evacuation time or total length of busy periods, and for minimizing the average delay, in both the cases of unit-length and variable-length messages.

For the cases in which no optimal scheduling policy can possibly exist, we conduct simulations and compare the performance of several commonly-used scheduling policies, including FIFO, FDF, CDF, SMF (*Shortest-Message-First*), LSF, and EDF. In the FIFO policy,

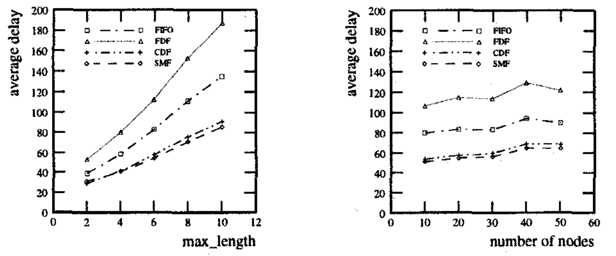
Evacuation mode $a_i = 0, \forall i$	meeting deadlines	minimizing evacuation time	minimizing average delay
$\ell_i = 1, \forall i$	LSF [§3.1]	FDF [11]	CDF [11]
$\ell_i \geq 1, \forall i$	LSF [§3.1]	FDF [§4.1]	not exist [§4.2]

(a) Evacuation mode operation.

Continuation mode $a_i \geq 0, \forall i$	meeting deadlines	minimizing length of busy periods	minimizing average delay
$\ell_i = 1, \forall i$	not exist [§3.2]	FDF [11]	not exist
$\ell_i \geq 1, \forall i$	not exist [§3.2]	FDF [§4.1]	not exist [§4.2]

(b) Continuation mode operation.

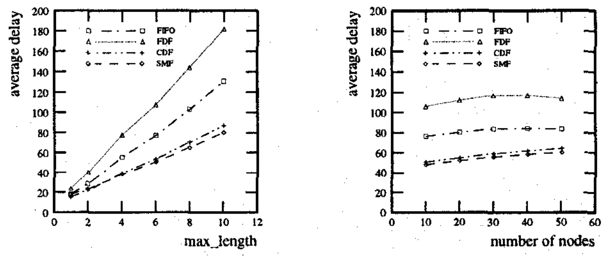
Table 4. Optimal scheduling policies for different performance measures.



(a) Average delays for different maximum message lengths.

(b) Average delays for different network sizes.

Figure 4. Average delays under evacuation mode operation ($a_i = 0$) with $\ell_i \geq 1$ for all i.



(a) Average delays for different maximum message lengths.

(b) Average delays for different network sizes.

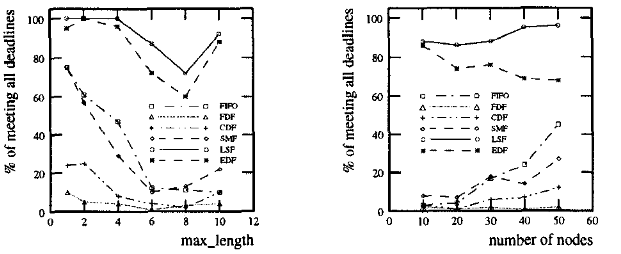
Figure 5. Average delays under continuation mode operation ($a_i \geq 0$) with $\ell_i \geq 1$ for all i.

each node gives priorities to the cells (not messages) based on the arrival times of the cells at the node. In the SMF policy, each node gives priorities to the messages based on the lengths of the messages.

Fig. 4 shows the average delays for the case of $a_i = 0$ (i.e., evacuation mode operation) and $\ell_i \geq 1$ for all i. In Fig. 4(a), for each value of *max_length* = 2, 4, 6, 8, and 10, we generate 100 message sets, each of which contains messages whose lengths are randomly generated between 1 and *max_length*. The results show that the SMF policy outperforms (in terms of minimizing the average delay) the other policies, except when the maximum message length is small (i.e., *max_length* = 2), in which case the CDF policy outperforms the others (note that if all messages are of unit length, the CDF policy is optimal). In general, SMF performs slightly better than CDF (except when the maximum message length is small), and FDF has the worst performance among these four policies. In Fig. 4(b), the maximum message length, *max_length*, is fixed at 6 and the number of nodes, N, in the ring network ranges from 10 to 50. 100 message sets are again generated for each $N =$ 10, 20, 30, 40, and 50. The results coincide with those in Fig. 4(a).

Fig. 5 depicts the average delays for the case of $a_i \geq 0$ (i.e., continuation mode operation) and $\ell_i \geq 1$ for all i. The input data (message sets) are similarly generated as in the evacuation mode case (Fig. 4). The results coincide with those in the evacuation mode case: SMF slightly outperforms CDF (except when the maximum message length is small, i.e., *max_length* = 1 or 2), and FDF has the worst performance among these four policies.

Fig. 6 gives the percentage of the 100 input message sets that all message deadlines can be met under different scheduling policies. The results in Fig. 6(a) show that the LSF policy outperforms (in terms of meeting all message deadlines) the other policies (FIFO, FDF, CDF, SMF, and EDF). In general, LSF performs slightly better than EDF, and both LSF and EDF perform much better than the other policies (FIFO, FDF, CDF, and SMF). The results in Fig. 6(b) show that in



(a) Percentage of meeting all deadlines for different maximum message lengths.

(b) Percentage of meeting all deadlines for different network sizes.

Figure 6. Percentage of meeting all deadlines under continuation mode operation ($a_i \geq 0$) with $\ell_i \geq 1$ for all i.

general, LSF outperforms EDF, and the larger the network the more significant the difference between LSF and EDF. This is perhaps due to the fact that when the network size gets larger, the distance factor becomes more pronounced and only the deadline itself cannot reflect the degree of the urgency of a message.

In conclusion, the FDF policy is optimal in all four cases (i.e., in evacuation or continuation mode operation and for unit-length or variable-length messages) in terms of minimizing the evacuation time or total length of busy periods. In terms of minimizing the average delay, the CDF policy performs better in the case when message lengths are small, but the SMF policy performs better in the case when message lengths are large. In terms of meeting message deadlines, the LSF policy always performs better than the other policies.

6. Conclusion

We treated the message transmission problem in unidirectional slotted ring networks with spatial slot reuse under two operation modes, i.e., evacuation mode and continuation mode, and with respect to three performance measures, i.e., meeting all message deadlines, minimizing the evacuation time or total length of busy periods, and minimizing the average delay. We showed that the LSF scheduling policy is optimal (in terms of meeting all message deadlines) under evacuation mode operation, while no optimal scheduling policy can possibly exist under continuation mode operation. We generalized the results reported in [11] (in which only unit-length messages are considered) and showed that the FDF policy is optimal in terms of minimizing the evacuation time and minimizing the total length of busy periods under evacuation mode operation and continuation mode operation, respectively, even in the case when messages may have variable lengths. We also showed that in terms of minimizing the average delay, no optimal scheduling policy can possibly exist if messages may have variable lengths for either evacuation mode or continuation mode operation.

For those cases in which no optimal scheduling policy can possibly exist, we conducted simulations and compare the performance of several commonly-used scheduling policies, including FIFO, FDF, CDF, SMF, LSF, and EDF. The simulation results show that LSF outperforms the others in terms of meeting all message deadlines, and SMF outperforms the others in terms of minimizing the average delay in both evacuation mode and continuation mode operations, except when maximum message length is small (in which case CDF outperforms the others).

References

[1] C. M. Aras, J. F. Kurose, D. S. Reeves, and H. Schulzrinne. Real-time communication in packet-switched networks. *Proceedings of the IEEE*, 82(1):122–139, January 1994.

[2] S. Baase. *Computer Algorithms: Introduction to Design and Analysis.* Addison-Wesley, 2nd edition, 1991.

[3] S. N. Bhatt, G. Pucci, A. Renade, and A. L. Rosenberg. Scattering and gathering messages in networks of processors. *IEEE Trans. on Computers*, 42:398–949, August 1993.

[4] S. Breuer and T. Meuser. Enhanced throughput in slotted rings employing spatial slot reuse. In *Proc. of IEEE INFOCOM'94*, pages 1120–1129, 1994.

[5] I. Cidon and Y. Ofek. Metaring—A full-duplex ring with fairness and spatial reuse. *IEEE Trans. on Communications*, 41:110–120, 1993.

[6] P. Fraigniaud, S. Miguet, and Y. Robert. Scattering on a ring of professors. *Parallel Computing*, 13:377–383, 1990.

[7] L. Georgiadis, R. Guérin, and I. Cidon. Throughput properties of fair policies in ring networks. *IEEE/ACM Trans. on Networking*, 1(6):718–728, Dec. 1993.

[8] A. E. Kamal and V. C. Hamacher. Utilizing bandwidth sharing in the slotted ring. *IEEE Trans. on Computers*, 39(3):289–299, March 1990.

[9] P. J. B. King and I. Mitrani. Modeling a slotted ring local area network. *IEEE Trans. on Computers*, C-36(5):554–561, May 1987.

[10] Y. Saad and M. H. Schultz. Data communication in parallel architectures. *Parallel Computing*, 11:131–150, 1989.

[11] L. Tassiulas and J. Joung. Performance measures and scheduling policies in ring networks. *IEEE/ACM Trans. on Networking*, 3(5):576–584, October 1995.

[12] M. Zafirovic-vukotic, I. G. Niemegeers, and D. S. Valk. Performance analysis of slotted ring protocols in HSLAN's. *IEEE J. on Selected Areas in Communications*, 6(6):1011–1024, July 1988.