Throughput Properties of Fair Policies in Ring Networks

Leonidas Georgiadis, Roch Guérin and Israel Cidon

We consider a slotted ring in which simultaneous transmissions of messages by different stations is allowed, a property referred to as spatial reuse. Ring networks with spatial reuse can achieve significantly higher throughput than standard token rings but they also introduce the possibility of starvation for some nodes on the ring. To alleviate this problem, various policies have been suggested in the literature. Our objective is to characterize the node throughputs achievable by general transmission policies in ring networks with spatial reuse and then to evaluate the throughput trade-off for a class of policies that has been proposed in the literature in order to avoid starvation. Specifically, we study a policy that is based on the idea of allocating transmission quotas to the nodes. Each node is guaranteed transmission of his quota within a specified interval. We show that by appropriately allocating the quotas, policies that satisfy general optimality criteria-in particular criteria related to fairness -can be designed. We also study the asymptotic behavior of the quota policy when either the quotas or the number of nodes increase. Ring Networks, Spatial Reuse, Scheduling Policies, Fair Policies, Asymptotic Analysis

The authors are with IBM T. J. Watson Research Center, P. O. Box 704, Yorktown Heights, NY 10598. I. Cidon is also with the Department of Electrical Engineering, Technion, Haifa 32000, Israel

1 Introduction

In recent years, the dramatic increase in transmission speeds has drastically altered many of the operating assumptions of communication networks. In the local area (LAN) environment, the effort towards defining new approaches that take better advantage of the available technology has resulted in a number of new standards and architecture proposals [1, 2, 6, 5] and led to a renewed interest in rings networks that employ *spatial reuse* [6, 5, 10] as an attractive alternative for new high speed LANs. Spatial reuse, also termed destination release, has the potential to significantly improve network throughput by allowing multiple simultaneous transmissions on the LAN as long as they take place on different links. The significance of this advantage has been recognized, and proposals have even been made to upgrade some of the early high speed LAN standards so that they can support this feature, e.g., erasure nodes in DQDB, etc. [14, 11].

When applied to ring networks, spatial reuse does increase throughput but it also introduces the possibility of *starvation* for nodes on the ring. Specifically, it is possible that certain nodes be denied access to the ring for extended periods of time, while others enjoy uninterrupted transmissions. This results in an unfair allocation of ring bandwidth between the nodes that are sharing the same ring. It is, therefore, necessary to provide mechanisms that will enforce a fair allocation of network resources while preserving their efficient sharing.

The importance of this problem has been recognized, not only in the context of LANs but also for wide area networks, e.g., see [8, 13]. Considerable work has been done to address this issue in rings that employ spatial reuse and several algorithms have been proposed that attempt to preserve fairness without significantly impacting ring throughput [6, 15, 5]. It is beyond the scope of this paper to compare the respective merits of all these algorithms. Rather, our focus is on understanding and sizing the loss of network throughput incurred when attempting to guarantee fairness. This can then be applied to the analysis of existing policies and the design of new and improved ones. Since the focus in this paper is on node throughputs, we assume that the nodes are full with packets always waiting for transmission. The study of the access queue at a node as a function of the nodal arrival processes and destinations distributions, clearly a topic of interest, is beyond the scope of this paper. We note, however, that our model also provides insight to the operation of systems with finite arrival rates under appropriate time scales. Indeed, even in such systems, there may be relatively long time intervals of congestion, during which a large number of nodes have nonempty queues, although in the long run the queues will become empty again. During those intervals of congestion, it is important to assure the proper operation of the system and it is on this aspect that the current paper concentrates on.

Our objective is to characterize the node throughputs achievable by general transmission policies in ring networks with spatial reuse and then to evaluate the throughput trade-off for a class of policies that has been proposed in the literature in order to avoid starvation. Specifically, we study a policy that is based on the idea of allocating transmission quotas to the nodes, see [6, 5]. Each node is guaranteed transmission of his quota within a specified interval. We show that by appropriately allocating the quotas, policies that satisfy general optimality criteria-in particular criteria related to fairness -can be designed. We also study the node throughputs as either the quotas or the size of the network increase. It is found that for a fixed number of nodes, as the quotas increase proportionally, the node throughputs approach the optimal values exponentially fast when there is only one bottleneck node in the ring. When more than one bottleneck nodes exist, the convergence speed is $O(\nu^{-1/2})$, where ν is the node quota . When the number of nodes increases, for a wide class of message destination probabilities the optimal node throughput is achieved even when the quota sizes remain constant.

The rest of the paper is organized as follows. Section 2 precisely defines the ring model and the relevant notations. The throughput space of the ring is then characterized in Section 3. Section 4 contains the results on the asymptotic behavior of the policy as either the quotas or the number of nodes increase.

2 System Model

We first introduce some notation. The symbols $\ominus$, $\oplus$ refer respectively to subtraction and addition modulo M, the number of nodes in the ring. If the index k refers to nodes we denote $\sum_{k=i}^{m} x_k := x_i + x_{i\oplus 1} + \ldots + x_{m\ominus 1} + x_m$. We define $\mathcal{M} = \{0, \ldots, M-1\}$ and for a sequence $\{X(k)\}_{k=1}^{\infty}$, we denote $X := X(1)$.

System Model. We consider a unidirectional ring with M nodes. The nodes are numbered starting from zero, so that the node next (downstream) to node i in the direction of message transmission is node $i \oplus 1$. The round trip delay of the ring is 0. The last assumption is made in order to simplify the discussion. As we will see in Section 4 our methodology can also be applied to models with nonzero round trip delays such as the one studied in [12]. The system is slotted, that is, time is divided in slots of length 1 and the nodes can transmit messages at the beginning of each slot. Slot T corresponds to the time interval $[T, T+1)$. We assume that the length of each message is one slot. There is an infinite queue of messages in each node and the destination of the kth message in the queue of node i is denoted by $D_i(k)$. The sequence $\{D_i(k)\}_{k=1}^{\infty}$ consists of i.i.d. random variables independent of the destination sequences in the other nodes. We set $q_{ij} := \Pr(D_i(k) = j)$ (note that we do not exclude the possibility that $q_{ii} > 0$, that is, the message has to travel around the ring - broadcast message). Multiple messages can be transmitted at the same time on the ring. Assume that at the beginning of a slot, node $i \ominus 1$ transmits a message with destination node j. If $j = i$, node i can transmit one of its messages in the same time slot, while it receives the message from node $i \ominus 1$. If $j \neq i$, then, *in the same slot*, node i can either retransmit the upstream message or it can send one of its own messages and store the upstream message for later transmission. Although restrictions are imposed in practice on the number of upstream messages that a node can hold, we consider the operation in this generality initially, in order to describe the throughput space of the system under general scheduling policies (see Section 3).

We now specify the class of admissible message transmission policies. Let $S_{ij}(T)$ be the number of messages sent up to time T, $T \geq 1$, by node i with destination node j and $S_i(T) := \sum_{j=0}^{M-1} S_{ij}(T)$. Let also $R_{ij}(T)$ be the number of messages received by node j with origin node i and $R_i(T) := \sum_{j=0}^{M-1} R_{ij}(T)$. We denote by $\mathbf{\Pi}$ the class of policies that satisfy the following properties.

1. The long term transmission rate of $S_{ij}(T)$, $V_{ij} := \lim_{T\to\infty} S_{ij}(T)/T$, exists for all i, j. Note that we do not require that the limit be the same for all sample paths. That is, V_{ij} can be a random variable.

2. The long term reception rate (throughput) of $R_{ij}(T)$ exists and is equal to the transmission rate V_{ij}.

3. The long term proportion of messages originated at node i with destination node j is equal to the proportion of messages with destination j that are in the queue of node i when the system starts. That is, if $\lim_{T\to\infty} S_i(T) = \infty$, then
$$\lim_{T\to\infty} \frac{S_{ij}(T)}{S_i(T)} = q_{ij}.$$

These assumptions are mainly technical and do not impose any significant restrictions on the class of practical transmission policies. For the system under consideration, however, additional constraints on the admissible policies are imposed. First, the buffer space available at a node for the storage of messages originated in other nodes is limited to a single message. This buffer space is referred to as the "ring buffer." Second, each node provides priority to the messages originated in other stations. Due to these two requirements, a single ring buffer assures the delivery of all messages without the possibility of buffer overflow. It this way, the use of expensive fast buffers is avoided and the fast delivery of messages entering the network is guaranteed. A ring that employs a policy $\pi \in \Pi$ that has the above two additional properties will be called a "buffer insertion ring". For fairness reasons a third constraint is often imposed on the admissible policies. Specifically, it is required that a policy guarantees a finite upper bound on the "channel access time", that is, the time elapsed between two successive transmissions of messages from the same node. The class of policies from Π which have the three additional properties described above will be denoted by Π_0.

3 Throughput Space

Let $V_i = \sum_{j=0}^{M-1} V_{ij}$ be the throughput of node i under a policy $\pi \in \Pi$. We are interested in finding the set of values that the vector $\mathbf{V} := \{V_0, \ldots, V_{M-1}\}$ can take when π is employed. We call this set of values the "throughput space" of the ring and denote it by $\mathcal{V}$.

Assume first that $V_i > 0,\ i \in \mathcal{M}$. Since $\pi \in \Pi$, this implies that $\lim_{T\to\infty} S_i(T) = \lim_{T\to\infty} R_i(T) = \infty$. Let $L_{im}(T)$ be the number of messages with origin node i that passed through node m up to time T ($L_{ii}(T)$ denotes the number of messages originated by node i up to time T). Clearly,
$$\sum_{j=m\oplus 1}^{i\oplus M} R_{ij}(T) \le L_{im}(T),$$

and since $\sum_{i\in\mathcal{M}} L_{im}(T) \le T$, we have that

$$\sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i\oplus M} \frac{R_{ij}(T)}{R_i(T)} \frac{R_i(T)}{T} \le 1. \tag{1}$$

The properties of π imply that $\lim_{T\to\infty} R_{ij}(T)/R_i(T) = \lim_{T\to\infty} S_{ij}(T)/S_i(T) = q_{ij}$ and therefore, using (1) we have that

$$\sum_{i\in\mathcal{M}} a_{im} V_i \le 1, \quad m \in \mathcal{M}, \tag{2}$$

where $a_{im} = \sum_{j=m\oplus 1}^{i\oplus M} q_{ij}$; a_{im} is the probability that a message generated by node i is destined to a node that is downstream from node m, in particular, $a_{ii} = 1$. It is easy to see that (2) continues to hold when $V_i = 0$ for some nodes. From the above discussion we see that

$$\mathcal{V} \subset \mathcal{W} := \{\mathbf{v} \in I\!\!R^M : v_m \ge 0, \ \sum_{i=0}^{M-1} a_{im} v_i \le 1, \ \ m \in \mathcal{M}\}. \tag{3}$$

It can also be shown (see [7]) that $\mathcal{W}^o \subset \mathcal{V}$, where $\mathcal{W}^o$ is the interior of the set $\mathcal{W}$.

4 Study of a Policy in Π_0

In the rest of this paper we will study the following policy.

(π_1^*) Each node has a ring buffer for storing a single message originated in other nodes. In addition, with node $m \in \mathcal{M}$ there is an associated preassigned integer number $\nu_m \ge 0$ called "quota" and a variable $Q_m(T)$ which is initialized to ν_m, $Q_m(1) = \nu_m$. At time T node m performs the following actions.

1. If there is a message in its ring buffer, the node transmits this message. A message transmitted by node $m \ominus 1$ (if any) in the time slot $[T, T+1)$, is stored in the ring buffer of node m.
2. Otherwise,
   (a) if $Q_m(T) > 0$, the node transmits the first message in its queue and sets
   $$Q_m(T) := Q_m(T) - 1.$$
   A message transmitted by node $m \ominus 1$ (if any) in the time slot $[T, T+1)$, is stored in the ring buffer of node m.

(b) if $Q_m(T) = 0$ the node does not transmit any of the messages in its queue. If node $m \ominus 1$ is transmitting a message with destination a node other than m, node m is retransmitting the same message *in the same slot* (therefore, no message is stored in the ring buffer of node m in this case).

At the time instant T at which $Q_m(T) = 0$ for all $m \in \mathcal{M}$ and the ring buffers of all nodes are empty (i.e., all the node quotas have been delivered to their destination), we reset $Q_m(T) = \nu_m$. In this case we say that node m is "allocated new quota".

Policy π_1^* is a synchronous version of the algorithm proposed in [5]. The main difference is that in the distributed algorithm of [5], node m resets its quota whenever it receives a circulating token, while under π_1^* all nodes reset $Q_i(T)$ at the same time. The modification is introduced here to simplify the analysis. Also, π_1^* has similarities to the policy considered in [12], which models the operation of the Orwell protocol [6]. The main difference is that in [12] the round trip delay is considered to be one slot, while in our model the round trip delay is zero. Although there are also differences on the order in which the nodes are given permission to transmit on the ring, as will be explained shortly (see the remark after Proposition 1), our results can be directly applied to the model in [12]. Finally, we note that the distributed implementation of policy π_1^* requires a mechanism that informs the nodes when the quotas of all the nodes on the ring have reached their destination. One such mechanism is provided in [6] using the TRIAL and RESET slots. The implementation of this mechanism will increase by two slots the "evacuation time" (for the definition see two paragraphs below) of the ring and the results of this paper can be easily adapted to incorporate this increase.

The policy described above uses quotas that are integer numbers. In the following we will be interested in policies that allocate quotas according to predetermined ratios. Since any ratios can be expressed either exactly or arbitrarily close with integer numbers, integer quotas are sufficient for most applications. However, if desired, π_1^* can be easily modified to accommodate any ratios (e.g. $\sqrt{2} : \sqrt{3}$). In this case ν_m is interpreted as *average* quota and the policy is implemented as follows. The kth quota allocated to node m is a random variable, $\boldsymbol{\nu}_m(k)$, that takes the values $\lfloor \nu_m \rfloor$ and $\lfloor \nu_m \rfloor + 1$ with probabilities $1 - \nu_m + \lfloor \nu_m \rfloor$ and $\nu_m - \lfloor \nu_m \rfloor$ respectively. The random variables $\boldsymbol{\nu}_m(k),\ k = 1, \ldots$ are i.i.d and independent of the rest of the processes in the system. It is easy to see that $E\boldsymbol{\nu}_m(k) = \nu_m$. Note that when integer quotas are allocated, $\boldsymbol{\nu}_m(k) \equiv \nu_m$, that is, the modified policy reduces to π_1^*. For the sake of generality, we will study the modified policy in the following and we will also denote it by π_1^*.

From the description of the algorithm we see that under π_1^* the system operates in cycles. At the beginning of cycle k, node $m \in \mathcal{M}$ has quota $\boldsymbol{\nu}_m(k)$ and the end of the cycle, the quotas of all nodes are zero and their ring buffers contain no messages. Formally let $\tilde{B}_m(T)$ be the number of messages in the ring buffer of node m at time T. Let also $T_0 = 1$ and define for $k \geq 0$,

$$\begin{aligned} T_{k+1} &= \min\{T > T_k : \ Q_m(T) = 0, \\ &\qquad \tilde{B}_m(T) = 0, \ m \in \mathcal{M}\}, \\ Q_m(T_{k+1}) &= \boldsymbol{\nu}_m(k+1), \ \ m \in \mathcal{M}. \end{aligned}$$

Provided that there are no messages in the ring buffers at time $T = 1$, the sequence $\{T_{k+1} - T_k\}_{k=0}^{\infty}$ consists of i.i.d random variables. Using the terminology in [12] we will call the random variable $T_e := T_1 - T_0$ the "evacuation time" of the ring. As we will see, the evacuation time is crucial in determining the performance of π_1^* in terms of node throughput. We derive next an expression for T_e. Let K_{ij} be the number of messages originated from node i with destination node j during T_e. Let also N_m be the number of messages that are originated or have to be retransmitted by node m during T_e. Clearly,

$$\sum_{j \in \mathcal{M}} K_{ij} = \boldsymbol{\nu}_i$$

and

$$N_m = \sum_{i \in \mathcal{M}} \sum_{j=m\oplus 1}^{i} K_{ij}. \tag{4}$$

Since at least one node is transmitting in each slot in the interval $[1, T_e + 1)$ and the quotas are finite, T_e is also finite. Also, since a node can transmit at most one message in each slot, we have

$$T_e \geq U_e := \max_{m \in \mathcal{M}} N_m. \tag{5}$$

The next proposition shows that equality holds in (5).

Proposition 1 *Provided that $\nu_i > 0$ for some $i \in \mathcal{M}$,*

$$T_e = \max_{m \in \mathcal{M}} N_m.$$

Proof. We use induction on the number of nodes, M. For $M = 1$, the proposition is clearly true. Assume its validity for M and consider a ring with $M + 1$ nodes. Let $T_a + 1$ be the first time at which we have $Q_i(T_a + 1) = 0, \ \tilde{B}_i(T_a + 1) = 0$ for at least one node $i \in \mathcal{M}$ (note that if $\nu_j = 0$ for some $j \in \mathcal{M}$, then $T_a = 0$). Clearly, $T_a \leq T_e$. Let also $\mathcal{A}$ be the set of nodes $i \in \mathcal{M}$ with the property $Q_i(T_a + 1) = 0, \ \tilde{B}_i(T_a + 1) = 0$. By the definition of π_1^*, every node in $\mathcal{M}$ is busy in the interval $[1, T_a + 1)$. Therefore, If $\mathcal{A} = \mathcal{M}$ then $T_a = T_e = N_m, \ m \in \mathcal{M}$ and the proposition is true. Assume now that $\mathcal{A}$ is a strict subset of $\mathcal{M}$ and consider the operation of the ring after time $T_a + 1$. The actions taken by the nodes in $\mathcal{M} - \mathcal{A}$ are the same as the actions taken by the corresponding nodes in a ring where the following modifications are made.

1. If $\tilde{B}_i(T_a + 1) = 1$, the message in the ring buffer of node i in the original ring is moved to the head of the queue of node i in the modified ring. The rest of the messages in node i in the modified ring are identical to the messages in the queue of node i at time $T_a + 1$ in the original ring.

2. The quota of node $i \in \mathcal{M} - \mathcal{A}$ is $\bar{\nu}_i := Q_i(T_a + 1) + \tilde{B}_i(T_a + 1)$.

3. The ring buffer of node $i \in \mathcal{M} - \mathcal{A}$ is empty.

4. If a node $i \in \mathcal{M} - \mathcal{A}$ sends a message to a node $j \in \mathcal{A}$, in the modified system the same message is sent to the first node in $\mathcal{M} - \mathcal{A}$ that is downstream from j.

Denote by $\overline{T}_e$, $\overline{N}_m$, the quantities corresponding to T_e, N_m, in the modified ring. In both rings, in the interval $[T_a + 1, T_e + 1)$, the nodes in $\mathcal{M} - \mathcal{A}$ are transmitting exactly the same messages at the same time. Therefore, the cycles for both rings will end at the same time and

$$\overline{T}_e = T_e - T_a$$

From the definition of the modified ring and the fact that all nodes in the original ring were busy in the interval $[1, T_a + 1)$, it follows that

$$\overline{N}_m = N_m - T_a, \quad m \in \mathcal{M} - \mathcal{A}. \tag{6}$$

The modified ring contains at most M nodes and from the induction hypothesis and (6) we have that

$$\overline{T}_e = \max_{m \in \mathcal{M} - \mathcal{A}} \overline{N}_m = \max_{m \in \mathcal{M} - \mathcal{A}} N_m - T_a.$$

Therefore,

$$T_e = \overline{T}_e + T_a = \max_{m \in \mathcal{M} - \mathcal{A}} N_m \leq U_e,$$

which together with (5) shows the validity of the proposition for $M + 1$. □

Remark. In the model considered in [12], assume that the slot starts from a specific point O on the ring and let τ_e be the first time the slot reaches O and all nodes have completed their quotas. Following essentially the same approach, it can be shown that $\tau_e = \max_{m \in \mathcal{M}} N_m + 1$ and therefore, the evacuation time, $\widehat{T}_e$, of the ring studied in [12] can be expressed as

$$\widehat{T}_e = \max_{m \in \mathcal{M}} N_m + 1 - Y,$$

where $0 \leq Y \leq 1$, is a random variable that expresses the time it takes for the slot to travel from the node that last finished its quota to the point O. Since the asymptotic properties of $\widehat{T}_e$ depend again on $\max_{m \in \mathcal{M}} N_m$, our analysis can be applied to the model in [12] as well.

Since T_e is bounded and $ET_e \leq \sum_{i\in\mathcal{M}}(\nu_i+1) < \infty$, using regenerative arguments it is easy to establish that $\pi_1^* \in \Pi_0$ and that the throughput of node $i \in \mathcal{M}$ is

$$v_i = \frac{E\boldsymbol{\nu}_i}{ET_e} = \frac{\nu_i}{ET_e}. \tag{7}$$

In the next section we will investigate the asymptotic properties of this algorithm when either the quotas or the number of nodes becomes large.

4.1 Limiting Behavior of the Algorithm

4.1.1 Node Throughputs for Large Quotas

We consider first the asymptotic behavior of $v_i,\ i \in \mathcal{M}$, when the quotas increase proportionally. Let real numbers $r_i \geq 0,\ \ i \in \mathcal{M}$, be given. To avoid the trivial case we assume that $r_i > 0$ for at least one $i \in \mathcal{M}$. We set $\nu_i = \nu r_i,\ i \in \mathcal{M},\ \nu \geq 1$. Whenever needed, to explicitly denote the dependence of a quantity on ν, say quantity X, we write $X(\nu)$. We also denote by $\mathbf{1}_A$ the indicator of the event A. The next proposition provides the node throughputs when the quotas become large while maintaining the proportions $r_i,\ i \in \mathcal{M}$.

Proposition 2 *For all $i \in \mathcal{M}$,*

$$\lim_{\nu\to\infty} v_i(\nu) = \frac{r_i}{\max_{m\in\mathcal{M}}\left\{\sum_{j\in\mathcal{M}} r_j a_{jm}\right\}}.$$

Proof. Recalling the definition of $D_i(k)$ from Section 2, we can write $K_{ij} = \sum_{k=1}^{\boldsymbol{\nu}_i} \mathbf{1}_{\{D_i(k)=j\}}$. Using the independence of $D_i(k),\ k = 1,\ldots,$ we have that

$$\begin{aligned}
r_i q_{ij} &= r_i \lim_{\nu\to\infty} \frac{\sum_{k=1}^{\lfloor \nu r_i \rfloor} \mathbf{1}_{\{D_i(k)=j\}}}{\nu r_i} \\
&\leq \lim_{\nu\to\infty} \frac{K_{ij}(\nu)}{\nu} \\
&\leq r_i \lim_{\nu\to\infty} \frac{\sum_{k=1}^{\lfloor \nu r_i \rfloor+1} \mathbf{1}_{\{D_i(k)=j\}}}{\nu r_i} \\
&= r_i q_{ij}, \quad a.e.
\end{aligned} \tag{8}$$

Therefore, taking into account (4) and Proposition 1 we have that

$$\begin{aligned}
\lim_{\nu\to\infty} \frac{T_e(\nu)}{\nu} &= \lim_{\nu\to\infty} \max_{m\in\mathcal{M}} \left\{ \frac{N_m(\nu)}{\nu} \right\} \\
&= \max_{m\in\mathcal{M}} \left\{ \sum_{j\in\mathcal{M}} r_j a_{jm} \right\}, \quad a.e.,
\end{aligned} \tag{9}$$

where a_{im} are as defined in (2). Since

$$0 \leq \frac{T_e(\nu)}{\nu} \leq \sum_{i \in \mathcal{M}} \frac{\boldsymbol{\nu}_i}{\nu} \leq \sum_{i \in \mathcal{M}} r_i + M,$$

the process $T_e(\nu)/\nu,\ \nu \geq 1$ is uniformly integrable, and we conclude from (9) that

$$\lim_{\nu \to \infty} \frac{ET_e(\nu)}{\nu} = \max_{m \in \mathcal{M}} \left\{ \sum_{j \in \mathcal{M}} r_j a_{jm} \right\}. \tag{10}$$

Using finally (7) we have that

$$\lim_{\nu \to \infty} v_i(\nu) = \lim_{\nu \to \infty} \frac{\nu r_i}{ET_e(\nu)} = \frac{r_i}{\max_{m \in \mathcal{M}} \left\{ \sum_{j \in \mathcal{M}} r_j a_{jm} \right\}}.$$

□

Let us now see how Proposition 2 can be used in the design of ring access policies. A design objective for such policies is usually associated with an optimization problem in $\mathcal{W}$. For example, in ring networks with spatial reuse, fairness is the main issue. The simplest optimization problem associated with this criterion is to maximize the minimum node throughput in the network

$$\max_{\mathbf{v} \in \mathcal{M}} \min_{m \in \mathcal{M}} v_m. \tag{11}$$

When preferential treatment of the nodes is desirable, the slightly more general problem

$$\max_{\mathbf{v} \in \mathcal{M}} \min_{m \in \mathcal{M}} \frac{v_m}{g_m},\ g_m > 0,\ m \in \mathcal{M}, \tag{12}$$

is appropriate. It is not difficult to see that the vector with coordinates

$$v_i^* = \frac{g_i}{\max_{m \in \mathcal{M}} \{ \sum_{j \in \mathcal{M}} a_{jm} g_j \}},\ i \in \mathcal{M}, \tag{13}$$

solves the problem in (12). A stronger optimization problem related to fairness and widely used in the literature is to find the *max-min optimal* vector in $\mathcal{W}$ (see [3, Section 6.5.2] for a description of the fairness properties of max-min optimal vectors). The max-min optimal vector in $\mathcal{W}$ can be easily determined by slightly modifying the algorithm in [3]. Both for problem (11) and the max-min optimization, it is also easy to solve the more general problem obtained by replacing v_m with a reward $f_m(v_m)$, where $f_m(\cdot)$ is a non-decreasing function. Depending on the application other criteria may be desirable. However, one main characteristic of most of these problems (including those mentioned before) is that the resulting optimal point $\mathbf{v}^*$ lies on the "upper" boundary of $\mathcal{W}$, that is, for some $k \in \mathcal{M}$, the equality $\sum_{j \in \mathcal{M}} a_{jk} v_j^* = 1$ holds. Given such a vector we can pick appropriate quotas so that π_1^* provides node throughputs arbitrarily close to $\mathbf{v}^* = \{v_0^*, \ldots, v_{M-1}^*\}$. To see this, let $v^* = \min\{v_i^* : v_i^* > 0,\ i \in \mathcal{M}\}$ and define $r_i^* = v_i^*/v^*,\ i \in \mathcal{M}$. Then, since $\sum_{j\in\mathcal{M}} a_{jm} r_j^* v^* \leq 1,\ m \in \mathcal{M}$ and $\sum_{j\in\mathcal{M}} a_{jk} r_j^* v^* = 1$, we have that

$$v^* = \frac{1}{\max_{m\in\mathcal{M}}\left\{\sum_{j\in\mathcal{M}} r_j^* a_{jm}\right\}}$$

and therefore,

$$v_i^* = \frac{r_i^*}{\max_{m\in\mathcal{M}}\left\{\sum_{j\in\mathcal{M}} r_j^* a_{jm}\right\}}. \tag{14}$$

According to Proposition 2 if we pick $r_i = r_i^*$, as ν becomes large, the vector of node throughputs induced by π_1^* converges to the desired vector $\mathbf{v}^*$. Therefore, once the optimal vector $\mathbf{v}^*$ is determined, it is easy to design the desired policy. For problem (12) as well as the max-min optimization problem, $\mathbf{v}^*$ can be computed very simply, while for other optimization problems, $\mathbf{v}^*$ can be computed by standard numerical procedures. The main disadvantage with this approach is that for the determination of r_i, knowledge of the statistical parameters q_{ij} that determine the performance space is required. As such, the approach can be useful in environments in which these parameters do not change fast. Observe, however, that for problem (12) no such knowledge of statistical parameters is required. Indeed, in this case we can simply pick $r_i = g_i,\ i \in \mathcal{M}$. Once $r_i,\ i \in \mathcal{M}$, have been determined, the desired optimal point can be approached arbitrarily close by increasing ν. In the next two sections we will examine the dependence of the node throughputs on the quota.

Remark. In [8], it was shown that when sessions are established, which corresponds to the case $q_{ij} = 0$ or 1 in our framework, the round robin scheduling policy is max-min optimal. However, besides the fact that this policy cannot be implemented in buffer insertion rings, if q_{ij} can take values other than 0 or 1, it can be seen from simple examples that the round robin policy does not solve even the weaker problem (11).

4.1.2 Bounds on the Rate of Convergence for Large Quota

In this section, we provide bounds on the rate of convergence to the optimal point under policy π_1^*, as the quotas increase proportionally. Since larger quotas imply larger channel access times, it is important to know whether the node throughput approaches the optimal throughput quickly as the quotas increase. The main result of this section is Theorem 1 which states that the rate of convergence is exponentially fast when there is one bottleneck node on the ring and of order $O(\nu^{1/2})$ when the are at least two bottleneck nodes. Before considering the technical details, we provide an outline of the basic arguments that lead to this conclusion. Let m_0 be one of the *bottleneck nodes*, that is,

$$\sum_{j \in \mathcal{M}} r_j a_{jm_0} = \max_{m \in \mathcal{M}} \left\{ \sum_{j \in \mathcal{M}} r_j a_{jm} \right\}$$

Observe from (4) that

$$EN_m = \nu \sum_{j \in \mathcal{M}} r_j a_{jm} \tag{15}$$

and therefore, we can write for the limit values in Proposition 2 for $i \in \mathcal{M}$,

$$v_i^* := \frac{r_i}{\max_{m \in \mathcal{M}} \left\{ \sum_{j \in \mathcal{M}} r_j a_{jm} \right\}} = \frac{\nu r_i}{\max_{m \in \mathcal{M}} \{EN_m\}}.$$

On the other hand, the throughput of the system when $\nu_i = r_i \nu$ is

$$v_i(\nu) = \frac{r_i \nu}{E\left(\max_{m \in \mathcal{M}} \{N_m\}\right)}.$$

Therefore, the loss in throughput is due to the fact

$$E\left(\max_{m \in \mathcal{M}} \{N_m\}\right) \geq \max_{m \in \mathcal{M}} \{EN_m\},$$

i.e., the cycles are elongated since the message destinations are random. However, when the quotas are becoming large the law of large numbers takes effect and as a result, the maximum of $N_m,\ m \in \mathcal{M}$, is attained at one of the bottleneck nodes most of the time. If there is only one bottleneck node on the ring the maximum will almost always be located at that node, and the effect of randomness will disappear quickly. However, when at least two bottleneck nodes exist, the effect of randomness tends to persist for larger quotas since the maximum may be attained at different bottleneck nodes at different times.

For simplicity in the exposition, we assume in this section that both ν and $r_i, i \in \mathcal{M}$, are positive integers and therefore, no randomization of the allocated quota is needed. We also assume without loss of generality that node 0 is a bottleneck. Let $\overline{H}_{jm}(l) = \sum_{k=1}^{m} \left(E\mathbf{1}_{\{D_j(l)=k\}} - \mathbf{1}_{\{D_j(l)=k\}}\right)$. Let also $I_{jm}(x) = \sup_{\theta \in I\!R} \left\{\theta x - \log E\ \left(e^{\theta \overline{H}_{jm}}\right)\right\}$ be the rate function of $\overline{H}_{jm}$. Denote by $\mathcal{M}_0$ the set of bottleneck nodes and $\mathcal{M}_0^c = \mathcal{M} - \mathcal{M}_0$

Theorem 1 *When ν and $r_i,\ i \in \mathcal{M}$ are positive integers,*

$$\begin{aligned} 0 &\leq \frac{v_i^* - v_i(\nu)}{v_i^*} \\ &\leq \frac{v_i^*}{r_i} \left(\frac{1}{\nu^{1/2}} \sum_{m \in \mathcal{M}_0 - \{0\}} \sum_{j \in \mathcal{M}} r_j^{1/2} (E\overline{H}_{jm}^2)^{1/2} \right) \end{aligned}$$

$$+\frac{v_i^*}{r_i}\left(\sum_{m\in\mathcal{M}_0^c}\sum_{j\in\mathcal{M}} r_j m e^{-r_j\nu I_{jm}(\gamma_m/(r_j M))}\right).$$

Proof. According to Proposition 2 the maximum throughput achievable by node i is given by

$$v_i^* = \frac{r_i}{\max_{m\in\mathcal{M}}\{\sum_{j\in\mathcal{M}} r_j a_{jm}\}}.$$

The evacuation time can be written as follows.

$$T_e = N_0 + \max_{m\in\mathcal{M}}\{N_m - N_0\}. \tag{16}$$

Observe that since node 0 is a bottleneck,

$$EN_0/\nu = \max_{m\in\mathcal{M}}\{\sum_{j\in\mathcal{M}} r_j a_{jm}\} := r_i/v_i^*.$$

Setting $\Delta_m = N_m - N_0$, we have for $i \in \mathcal{M}$.

$$\begin{aligned}
0 &\leq \frac{v_i^* - v_i(\nu)}{v_i^*} \\
&= 1 - \frac{r_i}{v_i^*(EN_0/\nu) + v_i^* E\left(\max_{m\in\mathcal{M}}\{\Delta_m/\nu\}\right)} \\
&= \frac{v_i^* E\left(\max_{m\in\mathcal{M}}\{\Delta_m/\nu\}\right)}{r_i + v_i^* E\left(\max_{m\in\mathcal{M}}\{\Delta_m/\nu\}\right)} \\
&\leq \frac{v_i^*}{r_i} E\left(\max_{m\in\mathcal{M}}\{\Delta_m/\nu\}\right). \qquad (17)
\end{aligned}$$

From (17) we see that to have bounds on the rate of convergence of the throughput of node i to the maximal throughput v_i^*, it is sufficient to develop bounds on the expectation of $\max_{m\in\mathcal{M}}\{\Delta_m/\nu\}$. Let F_m be the number of messages with destination node m that are transmitted during a cycle, that is, $F_m := \sum_{j\in\mathcal{M}}\sum_{l=1}^{\nu r_j} \mathbf{1}_{\{D_j(l)=m\}}$. Then, by definition

$$N_{m\oplus 1} = N_m - F_{m\oplus 1} + \boldsymbol{\nu}_{m\oplus 1}, \quad m \in \mathcal{M}$$

and therefore for $m \geq 1$,

$$\Delta_m = \sum_{k=1}^{m} \boldsymbol{\nu}_k - \sum_{k=1}^{m} F_k. \tag{18}$$

Let $\gamma_m := \sum_{j\in\mathcal{M}} r_j a_{j0} - \sum_{j\in\mathcal{M}} r_j a_{jm} \geq 0,\ m \in \mathcal{M}$. Taking into account (15) we have

$$E\Delta_m = EN_m - EN_0 = -\nu\gamma_m$$

and therefore, we can write Δ_m as follows (note that $\boldsymbol{\nu}_k = \nu_k$ since ν_k is integer),

$$
\begin{aligned}
\Delta_m &= \Delta_m - E\Delta_m - \nu\gamma_m \\
&= \sum_{k=1}^{m} (\boldsymbol{\nu}_k - \nu_k) - \sum_{k=1}^{m} (F_k - EF_k) - \nu\gamma_m \\
&= \sum_{k=1}^{m} (EF_k - F_k) - \nu\gamma_m \\
&= \sum_{j\in\mathcal{M}} \left(\sum_{l=1}^{r_j\nu} \overline{H}_{jm}(l) - r_j\nu \frac{\gamma_m}{r_j M} \right). \qquad (19)
\end{aligned}
$$

Denoting $x^+ = \max\{0, x\}$, and observing that $\Delta_0 = 0$ and $\gamma_m = 0,\ m \in \mathcal{M}_0$, we have

$$
\begin{aligned}
E\left(\max_{m\in\mathcal{M}} \left\{\frac{\Delta_m}{\nu}\right\}\right) &\leq E\left(\max_{m\in\mathcal{M}_0-\{0\}} \left\{\frac{\Delta_m^+}{\nu}\right\}\right) \qquad (20) \\
&\quad + E\left(\max_{m\in\mathcal{M}_0^c} \left\{\frac{\Delta_m^+}{\nu}\right\}\right) \\
&\leq E\left(\frac{1}{\nu} \sum_{m\in\mathcal{M}_0-\{0\}} |\Delta_m|\right) \qquad (21) \\
&\quad + E\left(\frac{1}{\nu} \sum_{m\in\mathcal{M}_0^c} \Delta_m^+\right) \\
&\leq \sum_{m\in\mathcal{M}_0-\{0\}} \sum_{j\in\mathcal{M}} r_j E\left|\frac{\sum_{l=1}^{r_j\nu} \overline{H}_{jm}(l)}{r_j\nu}\right| \\
&\quad + \sum_{m\in\mathcal{M}_0^c} \sum_{j\in\mathcal{M}} r_j E\left(\widehat{H}_{jm}\right)^+, \qquad (22)
\end{aligned}
$$

where

$$
\widehat{H}_{jm} := \frac{\sum_{l=1}^{r_j\nu} \overline{H}_{jm}(l)}{r_j\nu} - \frac{\gamma_m}{r_j M}.
$$

Observe that for fixed $j, m, \in \mathcal{M}$, the variables $\overline{H}_{jm}(l),\ l = 1, \ldots$ are i.i.d. with zero mean and therefore,

$$
\begin{aligned}
E\left|\frac{\sum_{l=1}^{r_j\nu} \overline{H}_{jm}(l)}{r_j\nu}\right| &\leq \left(E\left(\frac{\sum_{l=1}^{r_j\nu} \overline{H}_{jm}(l)}{r_j\nu}\right)^2\right)^{1/2} \\
&= \frac{\left(E\overline{H}_{jm}^2\right)^{1/2}}{(r_j\nu)^{1/2}}. \qquad (23)
\end{aligned}
$$

Since

$$\widehat{H}_{jm} \le \frac{1}{r_j\nu}\sum_{l=1}^{r_j\nu}\sum_{k=1}^{m} E\mathbf{1}_{\{D_j(l)=k\}} \le m,$$

we have,

$$\begin{aligned} E\left(\widehat{H}^{+}_{jm}\right) &= \int_{\{\ \hat{H}_{jm}\ >\ 0\ \}} \widehat{H}_{jm} dP \\ &\le m \Pr\left(\frac{\sum_{l=1}^{r_j\nu}\overline{H}_{jm}(l)}{r_j\nu} > \frac{\gamma_m}{r_j M}\right). \qquad (24) \end{aligned}$$

(25)

Since for fixed $j, m, \in \mathcal{M}$, the random variables $\overline{H}_{jm}(l),\ l = 1, \dots$ are i.i.d., using the Chernoff bound [4] we have

$$\Pr\left(\frac{\sum_{l=1}^{r_j\nu}\overline{H}_{jm}(l)}{r_j\nu} > \frac{\gamma_m}{r_j M}\right) \le e^{-r_j\nu I_{jm}(\gamma_m/(r_j M))}. \qquad (26)$$

The theorem follows from (17), (22), (23), (25) and (26). □

Since $\gamma_m > 0$, $m \in \mathcal{M}_0^c$, and $\overline{H}_{jm}$ is bounded, it can be seen that $I_{jm}(\gamma_m/(r_i M)) > 0$. Therefore, when $\mathcal{M}_0 = \{0\}$, that is when there is only one bottleneck node in the ring, the convergence is exponentially fast. Otherwise, the upper bound in Theorem 1 decreases as the square root of ν. The following example shows that the bound on the rate cannot be improved in general when there is more than one bottleneck node in the ring.

Example 1 Consider a ring with two nodes, $q_{01} = q_{10} = q,\ 1 > q > 0$, and $r_0 = r_1 = 1$. In this case, both nodes are bottleneck nodes. From (19) we have,

$$\Delta_1 = \sum_{l=1}^{\nu} \overline{H}_{01}(l) + \overline{H}_{11}(l) = \sum_{l=1}^{\nu} \overline{H}_1(l),$$

where $\overline{H}_1(l) = \overline{H}_{01}(l) + \overline{H}_{11}(l)$. Simple calculations show that $\left(E\overline{H}_1^2\right)^{1/2} = (2q(1-q)))^{1/2} =: \sigma$. Now,

$$\max\{0, \Delta_1/\nu\} = \frac{\sigma}{\nu^{1/2}} \max\left(0, \frac{\sum_{l=1}^{\nu}\overline{H}_1(l)}{\sigma\nu^{1/2}}\right).$$

By the Central Limit Theorem,

$$S(\nu) := (\sigma\nu^{1/2})^{-1}\left(\sum_{l=1}^{\nu}\overline{H}_1(l)\right)$$

converges in distribution to a normally distributed random variable, W, with zero mean and variance 1. Since the function $f(x) = \min\{K, \max(0,x)\}$, $K > 0$ is continuous and bounded, we have that $\lim_{\nu\to\infty} Ef(S(\nu)) = Ef(W) > 0$. Since $\max\{0, S(\nu)\} \geq f(S(\nu))$, for ν large enough,

$$E(\max\{0, \Delta_1/\nu\}) = \frac{\sigma}{\nu^{\frac{1}{2}}} E(\max\{0, S(\nu)\}) \geq \frac{\sigma}{\nu^{1/2}} \frac{Ef(W)}{2}.$$

We conclude that for ν large enough and $i = 0, 1$,

$$\frac{v_i^* - v_i(\nu)}{v_i^*} = \frac{v_i^* E(\max\{0, \Delta_1/\nu\})}{r_i + v_i^* E(\max\{0, \Delta_1/\nu\})} \geq \frac{C}{\nu^{1/2}}, \quad C > 0.$$

∎

Remark. To simplify the discussion we assumed that ν and r_i, $i \in \mathcal{M}$, are positive integers. Of course, the case $r_i = 0$ for some $i \in \mathcal{M}$ is trivial, since in this case $v_i^* = v_i(\nu) = 0$. In the general case, that is, when ν and r_i, $i \in \mathcal{M}$, are nonnegative real numbers, the exposition is more complicated, however, the essential steps are the same. The main difference is that in the second equation in (19), the term $\sum_{k=1}^{m} \boldsymbol{\nu}_k - \nu_k$ may not be zero. This term is due solely to the discrete nature of the messages. The main effect is that when there is more than one bottleneck node on the ring, an additional term appears in the upper bound in Theorem 1 that is of order $O(1/\nu)$. Therefore, this term does not alter the asymptotic behavior of the policy.

From the previous discussion we see that the quotas needed to reach within a small percentage of the optimal node throughput can be relatively large if there is more than one bottleneck node in the ring. However there is another factor, namely the number of nodes on the ring, M, that affects favorably the quota size. Note that the maximum channel access time is upper bounded by the sum of the quota allocated to the nodes. Therefore, for rings with small number of nodes, a proportional increase of the allocated quotas may be acceptable. However, for rings with large number of nodes, the same increase will result in much longer channel access times in the worst case. However, as we will see in the next section, for rings with large number of nodes increasing the allocated quotas may be unnecessary, since for a wide family of message destination probabilities, small quotas can provide node throughputs close to the optimal.

4.1.3 Limiting Behavior for Large Number of Nodes

We now consider the situation in which the number of nodes, M, becomes large. As in the previous section, when necessary, we use the notation $X(M)$ to make the dependence of a quantity X on M explicit. It was found in [12] that when

$\nu_i = 1,\ i \in \mathcal{M}$ and $q_{ij} = 1/M,\ i,j\ \in \mathcal{M}$, then $\lim_{M\to\infty} M/ET_e(M) = 2$. For these parameters the maximum throughput of node i is

$$v_i^u(M) = \frac{1}{\sum_{i=0}^{M-1} \frac{M-i}{M}} = \frac{2}{M+1},$$

independent of i. We conclude that

$$\lim_{M\to\infty} \frac{v_i(M)}{v_i^u(M)} = \lim_{M\to\infty} \frac{M+1}{ET_e(M)2} = 1. \tag{27}$$

Therefore, as the number of nodes increases, the throughput of every node approaches the maximal throughput even with quotas of one slot. We will investigate this property further under general message destination probabilities and arbitrary but fixed quotas.

According to Proposition 2, given $r_i,\ i \in \mathcal{M}$, the maximal throughput for node i is

$$v_i^* = \frac{r_i}{\max_{m\in\mathcal{M}} \left\{\sum_{j=0}^{M-1} r_j a_{jm}\right\}}.$$

In Section 4.1.1 we showed that the throughput of node i can be arbitrarily close to v_i^*, however, this can be achieved by increasing the quotas ($\nu_i = r_i\nu$) of all nodes proportionally. We will show next that for a wide class of message destination probabilities, and as the number of nodes increases, there is no need to increase the quotas in order to approach the optimal node throughputs. We provide first an example that shows that the result cannot hold under arbitrary message destination probabilities.

Example 2 Consider a ring with $M = 3K,\ K > 1,\ \nu_i = 1$, and for $m = 0,\ldots,K-1$, $q_{3m,3m\oplus1} = q,\ q_{3m,3m\oplus2} = 1-q,\ q_{3m\oplus1,3m\oplus2} = q_{3m\oplus2,3m\oplus3} = 1$. In this case the ring contains K identical non-interfering sub-rings. The evacuation time takes only the values $1,2$. Since $T_e = 1$ if and only if all nodes with indices $3m,\ m = 0,\ldots,K-1$, transmit to their neighbors, we have $ET_e = 2 - q^K$ and therefore, $\lim_{M\to\infty} v_i(M) = \lim_{K\to\infty}(2-q^K)^{-1} = 1/2$. The maximal node throughputs under the specified quotas, are $v_i^*(M) = 1/(2-q),\ i \in \mathcal{M}$, and since $0 \le q < 1$, we see that in this case the node throughput can be reduced by 50 percent relative to the optimal node throughput if the quota remain small. ∎

Assume that we are given a sequence of real numbers $r_i,\ i = 0,1,\ldots,$ such that

$$\sup_i r_i < \infty.$$

When the number of nodes is M, let policy π_1^* operate with quota $\nu_i = r_i,\ i \in \mathcal{M}$. We will identify a class of message destination probabilities, $q_{ij}(M),\ i,j \in$

$\mathcal{M},\ M = 1, 2, \cdots$, for which policy π_1^* induces node throughputs arbitrarily close to the maximal as the number of nodes increases. As in the previous section, the issue is to identify conditions under which $E\left(\max_{m\in\mathcal{M}} N_m(M)\right) \approx \max_{m\in\mathcal{M}} EN_m(M)$. While the strong law of large numbers was in effect in that section, as will be seen in the following, in the current situation the main reason for the above approximation is that under the appropriate condition on the message destination probabilities, the variability of $N_m(M)$ relative to $\max_{m\in\mathcal{M}} EN_m(M)$ is small.

Theorem 2 *Let* $\sup_i r_i < \infty$ *and assume that when the number of nodes in the ring is* M*, policy* π_1^* *operates with quota* $\nu_i = r_i,\ i \in \mathcal{M}$*. If for some* $\delta > 0$*,*

$$\lim_{M\to\infty} \frac{M^\delta}{\max_{m\in\mathcal{M}} \left\{\sum_{i=0}^{M-1} r_i a_{im}(M)\right\}} = 0,$$

then

$$\lim_{M\to\infty} \frac{v_i(M)}{v_i^*(M)} = 1.$$

A few remarks before proceeding with the proof of this theorem. When $q_{ij} = 1/M,\ i \in \mathcal{M}$ and $r_i(M) = 1$, then $\max_{m\in\mathcal{M}} \left(\sum_{i=0}^{M-1} a_{im}(M)\right) = (M+1)/2$ and Theorem 2 holds. This result was shown in [12] using different methods. The question arises whether the condition on the destination probabilities in Theorem 2 can be weakened. The next example shows that the numerator M^δ cannot be replaced with $(\ln M)^\beta,\ \beta < 1$. Therefore, the condition in the theorem is close to being necessary.

Example 3 Let $I\!N_0$ denote the set of nonnegative integers, $\alpha > 0$ and define $K_M = \max\{k \in I\!N_0 :\ k^{1+\alpha} 2^k \leq M\}$, $L_M = \left\lfloor K_M^\alpha 2^{K_M} \right\rfloor$. Pick M large enough so that $K_M \geq 2$. Let $\nu_i = 1$ and consider a ring with the following message destination probabilities.

$$\begin{aligned} q_{(m-1)K_M+i,mK_M} &= q_{(m-1)K_M+i,(m-1)K_M+i+1} \\ &= \frac{1}{2}, m = 1, \ldots, L_M,\ i = 0, \ldots, K_M - 2, \\ q_{mK_M-1,mK_M} &= 1, m = 1, \ldots, L_M, \\ q_{i-1,i} &= 1, i = L_M K_M, \ldots, M. \end{aligned}$$

The ring contains L_M identical non-interfering sub-rings and it is easy to see that

$$\max_{m\in\mathcal{M}} N_m = \max_{1\leq m\leq L_M} N_{I_m},\quad m = 1, \ldots, L_M,$$

where $I_m = mK_M - 1$. Setting $Y_{mi} = \mathbf{1}_{\left\{D_{(m-1)K_M+i} = mK_M\right\}}$, $m = 1, \ldots, L_M$, $i = 0, \ldots, K_M - 2$, we can write

$$N_{I_m} = 1 + \sum_{i=0}^{K_M-2} Y_{im}.$$

Clearly, the random variables $Y_{im},\ m = 1, \ldots, L_M,\ i = 0, \ldots, K_M - 2$, are i.i.d. and

$$\max_{m \in \mathcal{M}} \left\{ \sum_{i=0}^{M-1} a_{im}(M) \right\} = EN_{I_1}(M) = 1 + \frac{K_M - 1}{2}$$

which implies that

$$\lim_{M\to\infty} \frac{\ln M}{\sum_{i=0}^{M-1} a_{im}(M)} = 2 \lim_{M\to\infty} \frac{\ln M}{K_M} = 2\ln 2.$$

Define now

$$Z_m = K_M \mathbf{1}_{\{N_{I_m} = K_M\}},\ m = 1 \ldots L_M.$$

Since $\max_{1\le m\le L_M} Z_m = 0$ if and only if $N_{I_m} < K_M,\ m = 1, \ldots, L_M$, it is easy to see that

$$E\left(\max_{0\le m\le L_M} Z_m\right) = K_M\left(1 - (1 - 2^{1-K_M})^{L_M}\right),$$

Since $Z_m \le N_{I_m}$ and $\lim_{M\to\infty}(1 - 2^{1-K_M})^{L_M} = 0$ we conclude that

$$\begin{aligned}
\lim_{M\to\infty} \frac{v_i(M)}{v_i^*(M)} &= \lim_{M\to\infty} \frac{1 + \frac{K_M-1}{2}}{E\left(\max_{1\le m\le L_M} N_{I_m}\right)} \\
&\le \lim_{M\to\infty} \frac{1 + \frac{K_M-1}{2}}{E\left(\max_{1\le m\le L_M} Z_m\right)} = \frac{1}{2}.
\end{aligned}$$

∎

The proof of Theorem 2 is based on two lemmas and the next theorem due to Rosenthal [9, Theorem2.12].

Theorem 3 [Rosenthal] *If $\{Y_k, \mathcal{F}_k, 0 \le k \le M-1\}$ is a martingale and $2 \le \rho < \infty$, then there is a constant C depending only on ρ such that*

$$\begin{aligned}
E\left|Y_{M-1}\right|^{\rho} &\le CE\left(\left(\sum_{i=1}^{M-1} E\left(X_i^2 | \mathcal{F}_{i-1}\right) + EX_0^2\right)^{\rho/2}\right) \\
&\quad + C\sum_{i=0}^{M-1} E|X_i|^{\rho},
\end{aligned}$$

where $X_0 = Y_0,\ X_i = Y_i - Y_{i-1}, i \ge 1$.

Let $G_{im} = \sum_{j=m\oplus 1}^{i} (K_{ij} - EK_{ij})$ and let $m_0(M)$ be one of the bottleneck nodes when the number of nodes is M. For the rest of this section we assume without loss of generality that $m_0(M) = 0$.

Lemma 1 *For $\rho \geq 2$ and for $m \in \mathcal{M}$ there is a constant C_m that depends only on ρ such that*

$$\left(E \left| \sum_{i=0}^{M-1} G_{im} \right|^{\rho} \right)^{\frac{1}{\rho}} \leq C_m \left(\left(\sum_{i=0}^{M-1} r_i a_{im} \right)^{\frac{1}{2}} (2R)^{\frac{1}{2}} + M^{\frac{1}{\rho}} R \right),$$

where $R := \sup_i r_i + 1$.

Proof. Observe that for fixed m the random variables $G_{im},\ i \in \mathcal{M}$, are independent (not identically distributed in general) with zero mean. Therefore, for fixed $m \in \mathcal{M}$, the process $Y_k := \sum_{i=0}^{k} G_{im},\ k = 0, \ldots, M-1$ is a martingale.

We will apply Theorem 3 with $Y_k = \sum_{i=0}^{k} G_{im}$ and $\mathcal{F}_i = \mathcal{F}_{im}$, where $\mathcal{F}_{im}$ is the σ-field generated by the random variables $G_{km},\ k = 0, \ldots, i$. Since for fixed m the random variables $G_{im},\ i \in \mathcal{M}$, are independent,

$$\begin{aligned} E(G_{im}^2 | \mathcal{F}_{i-1,m}) &= EG_{im}^2 \leq E\left(\left(\sum_{j=m\oplus 1}^{i} K_{ij} \right)^2 \right) \\ &= E\left(\left(\sum_{l=1}^{\boldsymbol{\nu}_i} \sum_{j=m\oplus 1}^{i} \mathbf{1}_{\{D_i(l)=j\}} \right)^2 \right) \\ &\leq E\left(\boldsymbol{\nu}_i \sum_{l=1}^{\boldsymbol{\nu}_i} \left(\sum_{j=m\oplus 1}^{i} \mathbf{1}_{\{D_i(l)=j\}} \right)^2 \right) \end{aligned}$$

Observe now that

$$\sum_{j=m\oplus 1}^{i} \mathbf{1}_{\{D_i(l)=j\}} = \mathbf{1}_{B_{im}(l)},\ B_{im}(l) = \cup_{j=m\oplus 1}^{i} \{D_i(l) = j\}.$$

Since $E\mathbf{1}_{B_{im}(l)} = a_{im}$ and the random variable $\boldsymbol{\nu}_i$ is independent of $\mathbf{1}_{B_{im}(l)},\ l = 1, \ldots,$ we conclude that

$$\begin{aligned} E(G_{im}^2 | \mathcal{F}_{i-1,m}) &\leq E\left(\boldsymbol{\nu}_i \sum_{l=1}^{\boldsymbol{\nu}_i} \mathbf{1}_{B_{im}(l)} \right) \\ &= a_{im} E\boldsymbol{\nu}_i^2 \end{aligned} \tag{28}$$

If $r_i \leq 1$, then by definition $\boldsymbol{\nu}_i$ takes the values 0 or 1 and therefore, $E\boldsymbol{\nu}_i^2 = E\boldsymbol{\nu}_i = r_i$. If $r_i > 1$, then again by definition, $E\boldsymbol{\nu}_i^2 \leq (\lfloor r_i \rfloor + 1)^2$. Therefore, taking into account (28), we have for general r_i,

$$E(G_{im}^2 | \mathcal{F}_{i-1,m}) \leq 2(r_i+1) r_i a_{im} \leq 2R r_i a_{im}. \tag{29}$$

Since $|G_{im}| \leq \lfloor r_i \rfloor + 1 \leq R$ we have,

$$\sum_{i=0}^{M-1} E|G_{im}|^\rho \leq M R^\rho. \tag{30}$$

From Theorem 3, (29) and (30) we conclude that for $\rho \geq 2$ and for $m \in \mathcal{M}$ there is a constant C_m that depends only on ρ such that

$$\begin{aligned}
\left(E \left| \sum_{i=0}^{M-1} G_{im} \right|^\rho \right)^{\frac{1}{\rho}} &\leq C_m \left(\left(\sum_{i=0}^{M-1} 2R r_i a_{im} \right)^{\frac{\rho}{2}} + M R^\rho \right)^{\frac{1}{\rho}} \\
&\leq C_m \left(\sum_{i=0}^{M-1} r_i a_{im} \right)^{\frac{1}{2}} (2R)^{\frac{1}{2}} \\
&\quad + C_m M^{\frac{1}{\rho}} R
\end{aligned}$$

□

Let $\widetilde{\Delta}_m := N_m - EN_0$. We are now in a position to prove the following lemma.

Lemma 2 *For any $\bar{\delta} > 0$, there is a constant $\widehat{C}$ that depends only on $\bar{\delta}$ and R, such that*

$$\begin{aligned}
0 &\leq E\left(\max_{m \in \mathcal{M}} \widetilde{\Delta}_m \right) \\
&\leq 1 + \widehat{C} \left(\left(\max_{m \in \mathcal{M}} \left\{ \sum_{i=0}^{M-1} r_i a_{im} \right\} \right)^{\frac{1}{2}} M^{\bar{\delta}} + M^{\bar{\delta}} \right).
\end{aligned}$$

Proof. Let $\overline{G}_m = \sum_{i \in \mathcal{M}} G_{im}$, and consider the event $F_k = \{\max_{m \in \mathcal{M}} |\overline{G}_m| \geq k\}$. We can express F_k as the union of the disjoint events $F_{kn} = \{|\overline{G}_n| \geq k, |\overline{G}_j| < k,\ j = 0, \ldots, n-1\},\quad n = 0, \ldots, M-1$. Since $|\overline{G}_n| \geq k$ on the set F_{kn}, using Holder's inequality and Lemma 1, we have for $\rho \geq 2$,

$$k \Pr(F_k) \leq \sum_{n=0}^{M-1} E(|\overline{G}_n| \mathbf{1}_{F_{kn}})$$

$$
\begin{aligned}
&\leq \sum_{n=0}^{M-1} (E|\overline{G}_n|^\rho)^{1/\rho} (\Pr(F_{kn}))^{\frac{\rho-1}{\rho}} \\
&\leq C\Phi \sum_{n=0}^{M-1} \Pr(F_{kn})^{\frac{\rho-1}{\rho}}, \qquad (31)
\end{aligned}
$$

where $C = \max_{m\in\mathcal{M}} C_m$ and

$$
\Phi = \left(\max_{m\in\mathcal{M}} \left\{ \sum_{i=0}^{M-1} r_i a_{im} \right\} \right)^{\frac{1}{2}} (2R)^{\frac{1}{2}} + M^{\frac{1}{\rho}} R.
$$

Using the inequality

$$
\sum_{i=0}^{M-1} \beta_i^{\frac{1}{\alpha}} \leq M^{(\alpha-1)/\alpha} \left(\sum_{i=0}^{M-1} \beta_i \right)^{1/\alpha}, \ \beta_i \geq 0, \ \alpha \geq 1,
$$

we have

$$
\begin{aligned}
\sum_{n=0}^{M-1} (\Pr(F_{kn}))^{\frac{\rho-1}{\rho}} &\leq M^{1/\rho} \left(\sum_{n=0}^{M-1} \Pr(F_{kn}) \right)^{1-1/\rho} \\
&= M^{1/\rho} \Pr(F_k)^{1-1/\rho}. \qquad (32)
\end{aligned}
$$

From (31) and (32) it follows that

$$
\begin{aligned}
\Pr(F_k) &\leq \Pr(F_k)^{\frac{1}{\rho}} \\
&\leq \frac{C\Phi M^{\frac{1}{\rho}}}{k}. \qquad (33)
\end{aligned}
$$

Since $|\overline{G}_m| \leq MR$, using the inequality

$$
E|X| \leq 1 + \sum_{k=1}^{\infty} \Pr(|X| \geq k),
$$

and taking into account (33) we find that

$$
\begin{aligned}
E\left(\max_{m\in\mathcal{M}} |\overline{G}_m| \right) &\leq 1 + \sum_{k=1}^{MR} \Pr(F_k) \\
&\leq 1 + C\Phi M^{\frac{1}{\rho}} \sum_{k=1}^{MR} \frac{1}{k} \\
&\leq 1 + C\Phi M^{\frac{1}{\rho}} (1 + \ln(MR)).
\end{aligned}
$$

Since $EN_0 \geq EN_m, \ m \in \mathcal{M}$, recalling (4) we have that

$$\widetilde{\Delta}_m \leq N_m - EN_m = \sum_{i\ \in \mathcal{M}} G_{im} = \overline{G}_m. \tag{34}$$

Using (34) and the fact that $E\left(\max_{m\in\mathcal{M}} \widetilde{\Delta}_m\right) \geq E\widetilde{\Delta}_0 = 0$, we have for $\rho \geq 2$,

$$\begin{aligned} 0 \leq E\left(\max_{m\in\mathcal{M}} \tilde{\Delta}_m\right) &\leq E\left(\max_{m\in\mathcal{M}} |\overline{G}_m|\right) \\ &\leq 1 + C\Phi M^{\frac{1}{\rho}}(1+\ln(MR)). \end{aligned}$$

For $0 < \bar{\delta} \leq 2$, the lemma follows by picking $\rho = 4/\bar{\delta}$ and

$$\hat{C} = C \max\left\{(2R)^{1/2}, R\right\} \sup_{M\geq 1} \left\{\frac{1+\ln(MR)}{M^{\bar{\delta}}/2}\right\}.$$

For $\bar{\delta} > 2$, we can simply pick the constant $\hat{C}$ that corresponds to $\bar{\delta} = 2$. □

Proof of Theorem 2. Write $T_e = EN_0 + \max_{m\in\mathcal{M}} \widetilde{\Delta}_m$. Since node 0 is a bottleneck node,

$$v_i^*(M) = \frac{r_i}{(\sum_{j=0}^{M-1} r_j a_{j0}(M))} = \frac{r_i}{EN_0(M)}$$

and therefore, the ratio of the throughput of node $i \in \mathcal{M}$ to the maximal throughput achievable under the specified r_i is

$$\frac{v_i(M)}{v_i^*(M)} = \frac{EN_0(M)}{ET_e(M)} = \frac{EN_0(M)}{EN_0(M) + E\left(\max_{m\in\mathcal{M}}\{\widetilde{\Delta}_m(M)\}\right)}. \tag{35}$$

We see from (35) that

$$\lim_{M\to\infty} \frac{v_i(M)}{v_i^*(M)} = 1 \ \text{ iff } \ \lim_{M\to\infty} \frac{E\left(\max_{m\in\mathcal{M}}\{\widetilde{\Delta}_m(M)\}\right)}{EN_0(M)} = 0 \tag{36}$$

and therefore, it suffices to show that $E\left(\max_{m\in\mathcal{M}}\{\widetilde{\Delta}_m(M)\}\right)$ becomes arbitrarily small relative to $EN_0(M)$ as M increases. In Lemma 2 pick $\bar{\delta} = \delta/2$. Recalling that $EN_0(M) = \max_{m\in\mathcal{M}} \left\{\sum_{i=0}^{M-1} r_i a_{im}(M)\right\}$, we then have,

$$\begin{aligned} 0 &\leq \lim_{M\to\infty} \frac{E\left(\max_{m\in\mathcal{M}} \widetilde{\Delta}_m(M)\right)}{EN_0(M)} \\ &\leq \lim_{M\to\infty}\left(\frac{M^\delta}{EN_0(M)}\frac{1}{M^\delta}\right) \\ &\quad + \lim_{M\to\infty} \hat{C}\left(\frac{M^{\delta/2}}{(EN_0(M))^{1/2}}\frac{M^\delta}{EN_0(M)}\frac{1}{M^{\frac{\delta}{2}}}\right) \\ &= 0 \end{aligned}$$

and the right hand side in (36) is true. □

5 Conclusions

We studied a slotted ring in which simultaneous transmissions of messages by different stations is allowed. Under the assumption that the stations have infinite access queues, we described the space of achievable node throughputs under general transmission policies. Next, we studied a policy that is based on the idea of allocating transmission quotas to the nodes as described in [6] and [5]. While we considered a synchronous version of the policy, there is a simple relation between the node throughputs of the policy studied in this paper and the system studied in [12] which models the ring in [6]. This relation permits us to apply our results to that system as well. We conjecture that our results also hold for the asynchronous version of the policy proposed in [5]. We do not have a proof in this case, however, the following heuristic argument can be given to support the conjecture. In the asynchronous version, a node may start transmission of its new quota before the upstream nodes complete the transmission of their quotas from the previous cycle. On the other hand, when the same node finishes its new quota, it cannot start new transmissions until all the upstream nodes complete the transmission of their quotas from the previous cycle. Therefore, we expect that the net effect is as if the quotas of all nodes have been increased proportionally, but not more than twice the original quota. In fact, it is easy to establish that for given $r_i,\ i \in \mathcal{M}$, the node throughputs induced by the asynchronous version are proportional to r_i, as is the case with the policy studied in this paper.

Acknowledgements. We are grateful to Ilan Kessler and Yuval Shavitt for many helpful discussions and comments.

References

[1] FDDI token ring - media access control. ANSI X3.139, American national Standard, November 1986.

[2] Distributed queue dual bus (DQDB) subnetwork of a metropolitan area network (MAN). Standard document P802.6/d14, IEEE Project 802.6, July 1990.

[3] D. Bertsekas and R. Gallager. *Data Networks.* Prentice Hall, second edition, 1992.

[4] J. A. Bucklew. *Large Deviation Techniques in Decision, Simulation, and Estimation.* Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, 1991.

[5] I. Cidon and Y. Ofek. Metaring – a full duplex ring with fairness and spatial reuse. *IEEE Trans. Commun.*, 41(1):110–120, January 1993.

[6] R. M. Falconer and J. L. Adams. Orwell: A protocol for an integrated service local network. *Br. Telecom Technol. J.*, 3(4):21–35, April 1985.

[7] L. Georgiadis, R. Guerin, and I. Cidon. Throughput properties of fair policies in ring networks. Technical Report RC 18293(80185), IBM, September 1992.

[8] E. L. Hahne. Round-robin scheduling for max-min fairness in data networks. *IEEE Journal on Selected Areas in communications*, 9(7):1024–1039, September 1991.

[9] P. Hall and C. C. Heyde. *Martingale Limit Theory and its Applications*. Academic Press, New York, 1980.

[10] P. Heinzmann, H. R. Müller, D. A. Pitt, and H. R. van As. Buffer-insertion cell-synchronized multiple access (BCMA) on a slotted ring. In *Proc. Second Intl. Conf. Local Commun. Sys.: LAN and PABX*, Palma, Balearic Islands, Spain, June 26–28 1991.

[11] M. J. Karol and R. D. Gitlin. High-performance optical local and metropolitan area networks: Enhancements of FDDI and IEEE 802.6 DQDB. *IEEE J. Select. Areas Commun*, 8(8):1439–1448, October 1990.

[12] I. Kessler and A. Krishna. On the cost of fairness in ring networks. *IEEE/ACM Trans. Network.*, 1(3):306–313, June 1993.

[13] A. K. J. Parekh. *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks*. PhD thesis, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, February 1992. No. LIDS-TH-2089.

[14] M. A. Rodrigues. Erasure node: Performance improvements for the IEEE 802.6 MAN. In *Proc. INFOCOM'90*, pages 636–643, San Francisco, CA, June 1990.

[15] H. J. F. Sabourin. Offset flow protocol. In *Proc. Second IEEE Workshop Future Trends Distrib. Comput. Sys.*, pages 185–189, 1990.

Leonidas Georgiadis (M'86) received the Diploma degree in electrical engineering from the University of Thessaloniki, Greece, in 1979, and his M.S. and Ph.D. degrees both in electrical engineering from the University of Connecticut, in 1981 and 1986 respectively. From 1981 to 1983 he has been with the Greek army.

From 1986 to 1987 he was Research Assistant Professor at the University of Virginia, Charlottesville. In 1987 he joined IBM T. J. Watson Research

Center, Yorktown Heights, where he is presently a Research Staff Member in the Network System Design group. His interests are in the area of local and wide area networks, scheduling, distributed systems, modeling and performance analysis.

Dr. Georgiadis is a member of IEEE Communications Society. In 1992 he received the IBM Outstanding Innovation Award for his work on Goal-Oriented Workload Management for Multi-class systems. His e-mail address is leonid@watson.ibm.com Roch Guérin Roch Guérin (S'84, M'86, SM'91) received the "Diplôme d'Ingénieur" from the École Nationale Supérieure des Télécommunications, Paris, France, in 1983, and his M.S. and Ph.D. from the California Institute of Technology, both in Electrical Engineering, in 1984 and 1986 respectively. Since August 1986 he has been with IBM at the Thomas J. Watson Research Center, Yorktown Heights, New York, where he now manages the Network System Design group in the Advanced Networking Laboratory department. His current activities include analysis and architecture of high-speed networks and their use by applications. His research interests are in the area performance analysis and modeling of high-speed networks, with emphasis on congestion control, bandwidth management, and dynamic routing.

Dr. Guérin is a member of Sigma Xi and the IEEE Communications Society, and is an editor for the IEEE Communications Magazine and was an editor for the IEEE Transactions on Communications. His email address is guerin@watson.ibm.com. Israel Cidon Israel Cidon (M'85-SM'90) received the B.Sc. (summa cum laude) and the D.Sc. degrees from the Technion - Israel Institute of Technology in 1980 and 1984, respectively, both in electrical engineering. From 1984 to 1985 he was on the faculty with the Electrical Engineering Department at the Technion. In 1985 he joined the IBM T.J. Watson Research Center, NY, where he has been a Research Staff Member and a manager of the Network Architectures and Algorithms group involved in various broadband networking projects. Since 1990 he is also with the Department of Electrical Engineering at the Technion. His research interests include high-speed local and wide area networks and distributed algorithms.

He currently serves as an editor for the IEEE/ACM Transactions on Networking. Previously he served as the Editor for Network Algorithms for the IEEE Transactions on Communications and as a guest editor for Algorithmica. In 1989 and 1993 he received the IBM Outstanding Innovation Award for his work on the PARIS high speed network and topology update algorithms respectively. His research interests are high-speed networks, distributed algorithms and wireless networks.