Any Work-conserving Policy Stabilizes the Ring with Spatial Reuse*

Leandros Tassiulas†
Dept. of Electrical Engineering and
Institute for Systems Research
University of Maryland
College Park, MD 20742
leandros@eng.umd.edu

Leonidas Georgiadis
Dept. of Electrical Engineering
University of Thessaloniki
Thessaloniki, GREECE
leonid@eng.auth.gr

March 10, 1996

Abstract

We consider the ring network with spatial reuse. Traffic streams may enter and exit the network at any node. We adopt an arrival traffic model with deterministic constraints on its sample paths, which conforms to the output traffic of a leaky bucket rate control mechanism. A transmission policy specifies at each time which traffic stream will be transmitted at the outgoing link by each node. We provide an upper bound on the asymptotic backlog of the ring that holds for all work-conserving policies and is independent of the initial conditions. This bound remains finite as long as the maximum load of every link is less than one. The latter condition is also necessary for the existence of an asymptotic bound that is independent of the initial conditions.

1 Introduction

We consider a uni-directional ring with spatial reuse, i.e., a ring in which multiple simultaneous transmissions are allowed as long as they take place over different links. A node can transmit at the outgoing link at the same time that it receives traffic from the incoming link. Initially it is assumed that the ring has *cut-through* capabilities, i.e., that a node receiving traffic from its incoming link may retransmit

*This work was presented in part at the INFOCOM'94 Conference, Toronto, Canada, June 1994.
†The research of the first author was supported in part by NSF under grant NCR-9406415, by the AFOSR under contract FA9620-95-1-0061 and by the New York State Center for Advanced Technology in Telecommunications, Polytechnic University.

the traffic immediately at the outgoing link. Results are obtained later for store and forward rings as well. The ring with spatial reuse, special implementation of which is the register insertion ring [2], has attracted a lot of attention recently since it allows a considerable increase in the throughput over the classical ring. Two classes of transmission control policies have been proposed and analyzed. The first class attempts to achieve the fair treatment of the nodes [7], [4], [11], [8], while the second guarantees the maximal stability region for the ring [1], [9].

In this paper we address the issue of stability of the ring under work-conserving policies. We adopt an arrival traffic model where certain deterministic constraints are satisfied by all sample paths. This model was introduced by Cruz [5] and and conforms to the output traffic of a leaky bucket rate control mechanism. When the ring is operated under distributed policies it may exhibit very complicated behavior as it has been demonstrated in recent studies. Cruz in [6] has obtained a sufficient stability condition for the ring operated under a simple distributed policy. The stability condition was strengthened later by [17]. However, the condition in [17] is still stronger than the condition that all link loads are less than one, and the question whether the latter condition is sufficient for stability of the ring remained open [13], [10]. While this condition is intuitive, related studies [12], [16], [14], [15] in the same context have shown that simple distributed systems may exhibit unstable behavior even if the condition is satisfied and sophisticated policies may be needed for their stabilization. For a single class Jackson-type network, it has been shown in [3] that the system is stable when the load to each node is less than one. The method presented in [3], however, does not seem to extend to multi-class non-feedforward networks, a special case of which is the ring topology studied here.

In this work we show that *any work conserving policy* provides the maximal stability region for the ring. This result indicates that no sophisticated control is required for stabilizing a ring network. More specifically we provide an upper bound on the asymptotic backlog of the ring that holds for all work-conserving policies and is independent of the initial conditions, under the condition that the maximum load of every link is less than one. This condition is also necessary for the existence of an asymptotic bound that is independent of the initial conditions. Moreover, if the maximum load of a link is larger than one, then there exist input traffic patterns for which the backlog cannot be bounded under any initial conditions. Note that the results of this paper hold for work-conserving policies. It can be seen by simple examples that the stability conditions of policies like Metaring [4] and Orwell [7], which do not have the work-conserving property, are in general stricter than the condition that the maximum load of every link is less than one (see also [8] where the throughput properties of these policies are studied under the assumption of zero propagation delays).

2 The model and the transmission policies

Let M be the number of nodes and denote the set of nodes, $\mathcal{M} = \{0, \ldots, M-1\}$. The operations $i \oplus j$ and $i \ominus j$ denote respectively, addition and subtraction modulo M. Furthermore, when i, j refer to node indices we denote $\sum_{k=i}^{j} x_k := x_i + x_{i\oplus 1} + \ldots + x_{j\ominus 1} + x_j$. We assume that the nodes are arranged on the ring according to their index so that the outgoing link of node i is the incoming link of node $i \oplus 1$. The ith link is joining nodes i and $i \oplus 1$. Nodes $i \oplus k$ and $i \ominus k$ are called the kth node "downstream" and "upstream" from node i, respectively. For any two nodes i, j, there may be traffic entering the ring at node i with destination node j. Let $A_{ij}(t_1,\ t_2),\ t_1 < t_2$, be the amount of traffic that arrives to node i from the outside with destination node j, in the interval $[t_1,\ t_2)$ and $A_i(t_1, t_2) = \sum_{j\in\mathcal{M}}\ A_{ij}(t_1, t_2)$. If $i = j$, then the traffic from node i has to make a full circle around the ring and exit at node i again. We assume that the input traffic satisfies the following relation

$$A_{ij}(t_1,\ t_2) \leq \rho_{ij}(t_2 - t_1) + \sigma_{ij},\ \ i, j \in \mathcal{M}. \tag{1}$$

The transmission capacity of all links is the same, equal to one. Each link i has a propagation delay β_i (it may be zero). At any time the amount of traffic that exists in link i is upper bounded by β_i. We define $\beta = \sum_{i=1}^{M} \beta_i$.

Let

$$\rho_m = \sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i} \rho_{ij}.$$

The quantity ρ_m is an upper bound of the average rate of the total flow through link m. Our objective is to show that if the maximum link utilization on the ring is less than one, i.e., $\rho := \max\{\rho_i :\ i \in \mathcal{M}\} < 1$, then the ring is stable under any work-conserving policy. This is made precise in the theorem of the next section. Before we proceed we need some notation. Let $Q_{ij}(t)$ be the amount of traffic with destination node j stored at node i at time t, and $Q_i(t) = \sum_{j\in\mathcal{M}}\ Q_{ij}(t)$. The amount of traffic traveling at time t on link i with destination link j is denoted by $e_{ij}(t)$. Also, let $e_i(t) = \sum_{j\in\mathcal{M}} e_{ij}(t)$ be the total traffic at time t on link i. Note that $e_i(t) \leq \beta_i$ for all $t \geq 0$. The degree of node m at time t, $d_m(t)$, is defined as the total amount of traffic that exists on the ring at time t, that has to cross node m in order to reach its destination. This traffic may be either queued at any of the nodes on the ring, or it may be traveling on any of the links. By definition,

$$\begin{aligned} d_m(t) &= \textstyle\sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i} Q_{ij}(t) + \sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i} e_{ij}(t) \\ &= \textstyle Q_m(t) + \sum_{i\in\mathcal{M},\, i\neq m} \sum_{j=m\oplus 1}^{i} Q_{ij}(t) + \sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i} e_{ij}(t). \end{aligned} \tag{2}$$

For the sake of generality we allow more than one packet to be transmitted simultaneously as long as the total transmission rate of all packets is less than 1 at all times. Let $\mu_{ij}(t)$ be the transmission rate of traffic with destination j through link i at time t. Then

$$\sum_{j\in\mathcal{M}} \mu_{ij}(t) \leq 1, \quad t \geq 0.$$

A transmission policy is any rule for selecting the transmission rates such that, in addition to the above capacity condition the following holds

$$\mu_{ij}(t) \leq \mu_{i\ominus 1 j}(t), \quad \text{if} \quad Q_{ij}(t) = 0,$$

which is necessary for flow conservation at every node. The backlog of stream j at node i evolves according to the equation

$$Q_{ij}(t_2) = Q_{ij}(t_1) + \int_{t_1}^{t_2} \mu_{i\ominus 1j}(t - \beta_{i\ominus 1})dt - \int_{t_1}^{t_2} \mu_{ij}(t)dt, \quad t_2 \geq t_1$$

In this paper we are interested on the class of work-conserving policies $\mathcal{W}$. It contains those policies for which each link transmits in full capacity whenever the backlog in its origin node is not empty, that is $\mu_i(t) = 1$ if $Q_i(t) > 0$.

3 The bound on the backlog

A bound of the backlog over all work-conserving policies is given in the following theorem.

Theorem 1 *If the traffic constraints satisfy the condition* $\max\{\rho_i : i \in \mathcal{M}\} < 1$, *then for any initial condition* $Q_{ij}(0)$, $e_{ij}(0)$, *and under any work conserving policy,*

$$\limsup_{t\to\infty} (\max\{Q_i(t) : i \in \mathcal{M}\}) \leq \frac{M\beta + M^2\sigma}{1-\rho} + \beta,$$

where $\sigma = \sum_{i,j\in\mathcal{M}} \sigma_{ij}$.

Proof. We will show the stronger result

$$\limsup_{t\to\infty} \left(\max\{d_i(t) :\ i \in \mathcal{M}\}\right) \le \frac{M\beta + M^2\sigma}{1-\rho} + \beta,$$

which by (2) implies the theorem. The proof will be by contradiction. The basic idea is to show that if for an arbitrarily large time the degree of some node is above the specified bound, then necessarily at some earlier time there will be a node with even higher degree. Following this backward in time procedure we will conclude that at the beginning of time at least one node has degree higher than the specified initial conditions, a contradiction.

Assume that the previous inequality is not true. Then there exists $\delta > 0$, such that for any time t there is a time $T \ge t$ at which the following holds.

$$\max\{d_i(T) :\ i \in \mathcal{M}\} > \frac{M\delta + M\beta + M^2\sigma}{1-\rho} + \beta. \tag{3}$$

Denote the right-hand side of equation (3) by B. Equation (3) implies that there is a node m_1, such that $d_{m_1}(T) > B$. Let $\tau = (B-\beta)/M$ and assume that t is fixed with $t > \tau$ so that $T > \tau$. By (2) there is at least one node, k, such that $Q_k(T) > (B-\beta)/M$. We will show next that the queue of node k is nonempty during the interval $[T-\tau, T)$. To see this, let τ_0 be the length of time since the last instant before T that node k was empty (if the node was nonempty in the interval $[0,\ T)$, the claim is true.) Then since the employed policy is work conserving we have that

$$Q_k(T) = A_k(T-\tau_0, T) + R_{k\ominus 1}(T-\tau_0, T) - \tau_0,$$

where $R_{k\ominus 1}(T-\tau_0, T)$ is the traffic received from link $k\ominus 1$ by node k in the interval $[T-\tau_0,\ T)$, with destination a node other than node k. Since the transmission rate of all nodes is 1 and the nodes have cut-through capability, we have that

$R_{k\ominus 1}(T-\tau_0, T) \leq \tau_0$ and therefore, $Q_k(T) \leq A_k(T-\tau_0,\ T) \leq \rho\tau_0 + \sigma$. Taking into account that $Q_k(T) > (B-\beta)/M = \tau$, we conclude

$$\tau_0 > \frac{\tau - \sigma}{\rho}.$$

It remains to observe that by the definition of τ and B, we have that $(\tau-\sigma)/\rho > \tau$.

Returning to node m_1, define τ_1 as follows. If the queue of node m_1 is nonempty in the interval $[T-\tau,\ T)$, set $\tau_1 = \tau$. Otherwise, let τ_1 be the length of time since the last instant before T that the queue of node m_1 was empty. Note that $\tau_1 = 0$ if $Q_{m_1}(T) = 0$.

Consider now the sequence of nodes m_i and time lengths τ_i, $1 \leq i \leq I$, obtained by the following procedure.

1. If $\tau_1 = \tau$, then $I = 1$; stop. Else, $i = 2$.
2. $m_i = m_{i-1} \ominus 1$
3. If the queue of node m_i is nonempty in the interval $[T-\tau,\ T-\sum_{j=1}^{i-1}\tau_j)$, set $\tau_i = \tau - \sum_{j=1}^{i-1}\tau_j$; $I = i$; stop.
4. else, let τ_i be the length of time since the last instant before $T - \sum_{j=1}^{i-1}\tau_j$ that the queue of node m_i was empty; $i \leftarrow i+1$; go to step 2.

Note that the fact that node k is nonempty in the interval $[T-\tau, T)$ implies that a) $I \leq M$ and b) $\sum_{j=1}^{I}\tau_j = \tau$.

Let $A^m(t_1,\ t_2)$ be the traffic that arrives in the interval $[t_1,\ t_2)$ from the outside to any node on the ring and has to be transmitted through node m in order to reach its destination. That is,

$$A^m(t_1,\ t_2) = \sum_{i\in\mathcal{M}} \sum_{j=m\oplus 1}^{i} A_{ij}(t_1,\ t_2). \tag{4}$$

¿From (1) we conclude that

$$A^m(t_1, t_2) \leq \rho_m(t_2 - t_1) + \sigma.$$

Since the queue of node $m_i,\ 1 \leq i \leq I$, is nonempty in the interval $[T - \sum_{j=1}^{i} \tau_j,\ T - \sum_{j=1}^{i-1} \tau_j)$ and the policy is work-conserving, node m_i transmits at the link transmission capacity in this interval and therefore,

$$\begin{aligned} d_{m_i}\left(T - \textstyle\sum_{j=1}^{i-1} \tau_j\right) &= d_{m_i}\left(T - \textstyle\sum_{j=1}^{i} \tau_j\right) + A^{m_i}\left(T - \textstyle\sum_{j=1}^{i} \tau_j,\ T - \textstyle\sum_{j=1}^{i-1} \tau_j\right) - \tau_i \\ &\leq d_{m_i}\left(T - \textstyle\sum_{j=1}^{i} \tau_j\right) + \rho\tau_i + \sigma - \tau_i, \end{aligned} \tag{5}$$

where we adopt the standard convention that $\sum_j^i = 0$ when $i < j$.

Note next that the ring topology implies the following relation for the node degrees

$$d_i(t) \leq d_{i\ominus 1}(t) + e_{i\ominus 1}(t), \quad \text{if} \ \ Q_i(t) = 0. \tag{6}$$

Assume now that $I \geq 2$. By (6) and the definition of the indices m_i we have that

$$d_{m_{i-1}}\left(T - \sum_{j=1}^{i-1} \tau_j\right) \leq d_{m_i}\left(T - \sum_{j=1}^{i-1} \tau_j\right) + \beta_{m_i}, \quad 2 \leq i \leq I. \tag{7}$$

¿From (5) and (7) we conclude that

$$d_{m_{i-1}}\left(T - \sum_{j=1}^{i-1} \tau_j\right) \leq d_{m_i}\left(T - \sum_{j=1}^{i} \tau_j\right) - (1-\rho)\tau_i + \sigma + \beta_{m_i}, \ \ 2 \leq i \leq I. \tag{8}$$

Summing the inequalities in (8) for $i = 2, \ldots, I$ together with the inequality in (5) corresponding to $i = 1$, canceling identical terms and taking into account the fact that $\sum_{j=1}^{I} \tau_j = \tau$, we get

$$d_{m_1}(T) \leq d_{m_I}(T - \tau) - (1-\rho)\tau + I\sigma + \beta. \tag{9}$$

¿From (5) we see that (9) holds for $I = 1$ as well. Replacing the value of τ and recalling that $I \leq M$, we finally conclude

$$\begin{aligned} d_{m_I}(T-\tau) &\geq d_{m_1}(T) + \delta + \beta + M\sigma - I\sigma - \beta \\ &\geq d_{m_1}(T) + \delta. \end{aligned}$$

If $T \geq 2\tau$, we can repeat the same argument at time $T - \tau$. In general, repeating the previous procedure we conclude that there are node indices $l_n,\ 1 \leq n \leq \left\lfloor \frac{T}{\tau} \right\rfloor + 1$, such that $l_1 = m_1$ and

$$d_{l_{n+1}}(T - n\tau) \geq d_{l_n}(T - (n-1)\tau) + \delta.$$

Therefore, for $n = \left\lfloor \frac{T}{\tau} \right\rfloor$, we have that

$$d_{l_{n+1}}\left(T - \left\lfloor \frac{T}{\tau} \right\rfloor \tau\right) \geq d_{m_1}(T) + \left\lfloor \frac{T}{\tau} \right\rfloor \delta \geq B + \left\lfloor \frac{T}{\tau} \right\rfloor \delta. \tag{10}$$

Since t can be chosen arbitrarily large and $T \geq t$, we conclude that there is a node whose degree can become arbitrarily large in the interval $[0, \tau)$. This is impossible however, since the maximum degree of any node in this interval is less than $\max\{d_i(0) :\ i \in \mathcal{M}\} + \rho\tau + \sigma + \beta$. □

The necessity for stability of the condition that the load of each link is less than one, follows from the following fact. If $\max\{\rho_i : i \in \mathcal{M}\} = 1$, then there are traffic patterns for which no asymptotic bound, independent of the initial conditions can be found. Consider for example a single queue where $\rho = 1$ and consider the input stream $A(t_1, t_2) = t_2 - t_1$. Clearly, the queue size in this case will be equal to the initial conditions. This example can be easily extended to any ring network. If $\max\{\rho_i : i \in \mathcal{M}\} > 1$, then there are traffic patterns such that the backlog on the ring cannot be bounded under any initial conditions. For example, consider the case $A_{ij}(t_1, t_2) = \rho_{ij}(t_2 - t_1)$ and let m be a node such that $\rho_m > 1$. Since in this case

$A^m(t_1, t_2) = \rho_m(t_2 - t_1)$, it is clear that the backlog of the traffic that has to cross node m will increase indefinitely.

4 Remarks-Discussion

There are several interesting consequences of the analysis in the proof of theorem 1. One of them is a bound on the elapsed time until the backlog will cross below a certain threshold. Let $\delta > 0$, and define T_δ^ℓ as the last time after which the node degrees will never exceed the threshold

$$B_\delta = \frac{M\delta + M\beta + M^2\sigma}{1-\rho} + \beta.$$

According to theorem 1 this time is well defined. Using a slight modification of the proof of this theorem, it can be shown that

$$T_\delta^\ell \leq \left(\frac{d(0) + \rho\tau_\delta + \sigma + \beta - B_\delta}{\delta} + 1 \right)^+ \tau_\delta, \tag{11}$$

where $a^+ := \max(a, 0)$, $\tau_\delta = (B_\delta - \beta)/M$ and $d(0) = \max\{d_i(0) : \ i \in \mathcal{M}\}$. Denote the right-hand side of equation (11) by U_δ. Indeed, if $T_\delta^\ell > U_\delta$, then for some $T \geq U_\delta$ we should have $d_m(T) > B_\delta$. Following the same argument as in the proof of theorem 1 we conclude from (10) that $d_{l_{n+1}}\left(T - \left\lfloor \frac{T}{\tau} \right\rfloor \tau\right) > d(0) + \rho\tau_\delta + \sigma + \beta$, which is impossible. Note that for $M \geq 2$, if $d(0) = 0$, then $U_\delta = 0$, and therefore the bound in theorem 1 is never exceeded.

The bound on T_δ^l readily gives an upper bound for the node degrees that holds at all times and not only asymptotically. For general initial condition we have

$$d_i(t) \leq \max\left\{ d(0) + U_\delta^\ell \rho + \sigma, \ B_\delta \right\}.$$

The results are easily generalizable when the ring has store-and-forward capabilities instead of cut-through. In this case, a node cannot process a packet until it is completely received. Let L be the maximum packet length. Minor modifications of the proof of the theorem can be used to show that the system is stable in this case as well and the corresponding bound is now

$$\frac{M\beta + M^2 L + M^2\sigma}{1-\rho} + \beta.$$

It is not clear whether the bound of theorem 1 is tight within the class of work conserving policies. Nevertheless it provides a basis for comparison. When there is no propagation delay, the bound of theorem 1 becomes $M^2\sigma/(1-\rho)$. For the adaptive quota policy considered in [9], the bound on the asymptotic backlog is $2(\sigma+1)/(1-\rho)$. Hence the latter policy performs better than the worst possible performance of a work conserving policy by a factor of the order of the square of the number of nodes in the network.

References

[1] N. Bambos and A. Nguyen. Optimal flow on ring network with spatial reuse. In *Proc. 31-st Conference on Decision and Control*, pages 2360–2363, Tuscon, AZ, December 1992.

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

[3] C.S. Chang. Stability, queue length and delay, part i: Deterministic queueing networks. Technical Report RC 17708, IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, 1992.

[4] I. Cidon and Y. Ofek. Metaring - a full-duplex ring with fairness and spatial reuse. pages 969–981, San Fransisco, CA, June 1990. IEEE INFOCOM.

[5] R. Cruz. A calculus of network delay, part i: Network elements in isolation. *IEEE Trans. on Information Theory*, IT-37:114–131, January 1991.

[6] R. Cruz. A calculus of network delay, part ii: Network analysis. *IEEE Trans. on Information Theory*, (IT-37):132–141, January 1991.

[7] 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.

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

[9] L. Georgiadis, W. Szpankowski, and L. Tassiulas. A scheduling policy with maximal stability region for ring networks with spatial reuse. *Queueing Systems: Theory and Applications*, 19:131–148, 1995.

[10] B. Hajek. Private communication.

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

[12] P.R. Kumar and T.I. Seidman. Distributed instabilities and stabilization methods in distributed real-time scheduling of manufacturing systems. *IEEE Trans. on Automat. Contr*, AC-35:289–298, March 1990.

[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] L. Tassiulas. Adaptive back pressure congestion control based on local information. *IEEE Trans. on Automatic Control*, 40(2):236–250, February 1995.

[15] L. Tassiulas. Push forward link-level scheduling for network-wide performance. *IEEE/ACM Transactions on Networking*, TR(CATT-93-56), 1995. submitted.

[16] W. Whitt. Large fluctuations in a deterministic multiclass network of queues. *Management Science*, 1992. to appear.

[17] O. Yaron and M. Sidi. Performance and stability of communication networks via robust exponential bounds. pages 372–385. IEEE/ACM Transactions on Networking, 1(3), June 1993.