

**AN ANALYSIS OF A PRIORITIZED POLLING MECHANISM  
FOR CYCLIC SERVER SYSTEMS**

**M. M. SRINIVASAN  
and  
HYO-SEONG LEE**

**Department of Industrial & Operations Engineering  
The University of Michigan  
Ann Arbor, Michigan 48109, USA**

**Technical Report 87-6  
January 1987**

## An Analysis of a Prioritized Polling Mechanism for Cyclic Server Systems

### Abstract

The cyclic server system has been the subject of considerable research over the last few years. Interest in analyzing such systems has gained momentum in recent years due to their application in the performance analysis of token ring networks. In this paper we analyze a cyclic server system where the server services the requests at one of the nodes exhaustively while servicing all other nodes in the system on a non-exhaustive basis. Such a service mechanism finds application in the operation of interconnected ring networks. Exact analysis is presented here for the case where there are two nodes. A conservation law is developed which enables an accurate approximation technique for larger systems.

## 1. Introduction:

A cyclic server system is a system in which a single server attends, in a cyclic manner, to a number of centers at which requests queue up for service. Interest in analyzing such systems has gained momentum owing to their direct application in the analysis of token ring networks [4]. In a token ring network, a number of computers (nodes) communicate with each other through a common transmission medium. Messages are generated at each node for transmission to other nodes in the network through the transmission medium. Each message is broken up into one or more 'packets', and these form the unit of transmission. To transmit these packets, the node must gain control of the medium, and this control is arbitrated by a mechanism known as a token, which polls each node in some manner to determine whether it wishes to transmit some information.

The analysis of these token ring networks is usually done by modeling the token as a single server. The polling disciplines that are commonly modelled are the exhaustive, non-exhaustive and gated disciplines [29]. In all these disciplines, the server cycles around the system, visiting each node exactly once in each cycle. In the exhaustive service discipline, the node visited by the server transmits all the packets queued up there, including any that may arrive during the time the server is present at the node. In the gated service discipline, transmission is restricted to only those packets present when the server arrives at that node. In the non-exhaustive service discipline, each time the server visits a node, exactly one packet is transmitted. Although the non-exhaustive service mechanism is the most difficult to analyze, it finds the maximum application owing to its perceived 'fairness' in servicing the nodes, and has been implemented in a number of systems, notable among them being the Apollo token ring [21].

This paper proposes a variant of these polling disciplines which would be appropriate in certain situations. In this proposed scheme, one node alone receives exhaustive service, while all other nodes receive non-exhaustive service. One instance where such a scheme would be appropriate is when there is a system of interconnected ring networks. In these systems, one of the nodes in the ring serves as a 'bridge' node which does not generate any messages on its own, but merely serves to

forward traffic from other rings bound for this local ring [16,17]. In this case, such a prioritized form of service to the bridge node may be required, especially when the inter-ring traffic is large relative to the intra-ring traffic. Another situation where such a prioritized system could be considered is when a node is doing a video transmission, where the non-exhaustive service mechanism is clearly inappropriate.

The organization of the rest of this paper is as follows: section 2 reviews existing work in the analysis of cyclic service systems, section 3 presents the description of the proposed model, motivating the problem situation where this model would be appropriate. The exact analysis for the system with two nodes is developed in section 4 which also presents a conservation law for such systems, section 5 develops an approximation algorithm for the case with more than 2 nodes, and section 6 presents the conclusions.

## 2. Previous work:

A comprehensive survey on the analysis of cyclic service systems for the three well-known disciplines (exhaustive, gated, and non-exhaustive) is given by Watson [29] and Takagi [25]. Most of the analytic models treat these three disciplines. Some work has also been done on the analysis of some prioritized polling mechanisms [15,22,26,27].

The seminal work on cyclic service systems is due to Cooper [8,9], who considered systems with exhaustive and gated service disciplines, with Poisson arrival processes and general service time distributions at each node. The work of Cooper assumed that the switchover time, which is the time taken by the server to switch from one node to another, was zero. Further work on analyzing such systems using the exhaustive and gated service disciplines has since been reported. The paper by Ferguson and Aminetzah [12] presents exact results for non-symmetric cyclic service systems, with non-zero switchover times. Approximate analytical models for non-symmetric systems have recently been analyzed by Bux and Truong [5], and Carsten et al. [6].

The analysis of token ring networks using the non-exhaustive service mechanism is, however, extremely difficult. In general, even obtaining the mean waiting times presents difficulties, and exact results on these mean waiting times have been obtained only for a very limited set of configurations. The analysis for two queues, without switchover times has been obtained by Eisenberg [11]. Exact results for mean waiting times are also available in a simple closed form expression for the symmetric cyclic service system [13,29]. The intractability of the general model has led to several techniques being proposed to obtain approximate mean waiting times [1,2,18,19,23,28]. These approaches are then validated through extensive simulations.

### 3. Model description:

We consider a system with  $N$  nodes, where nodes 1 through  $N-1$  are served in a non-exhaustive manner, while node  $N$  is served in an exhaustive manner. As one application for such a 'prioritized' polling scheme, consider the system of interconnected computers as depicted in Figure 1. Suppose that a remote workstation needs to send a message to a workstation in ring 3. The packets generated by this message ultimately queue up for service at the bridge node,  $B$ , of ring 3. If this ring would have operated in isolation, then every node could conceivably be treated in the same manner. However, in the case of an interconnected ring, it seems appropriate to give the bridge node some preferential treatment, considering the fact that the packets arriving at this node have travelled over the network, while the packets waiting at any other node in the ring have been generated at that point itself. It is remarked that the paper by Takine et al. [27] has considered a somewhat similar type of service discipline for interconnected networks.

We make the assumption that the arrival process to the node with the exhaustive service discipline is Poisson. In light of the above application, this assumption can be justified by the fact that the arrival process obtained by merging a large number of independent renewal processes is approximately Poisson [20]. In general packets are generated (i.e., requests arrive) to node  $n$  of this system according to an independent Poisson process with rate  $\lambda_n$ . The service time demands of

customers arriving at node  $n$ ,  $H_n$ , are independent, identically distributed random variables with an arbitrary distribution, having mean  $h_n$ , and second moment  $h_n^{(2)}$ . The utilization of the server at node  $n$ ,  $\rho_n$ , is defined as

$$\rho_n = \lambda_n h_n \quad (3.1)$$

The switchover time between node  $n$  and node  $n + 1$  is an independent, identically distributed random variable,  $S_n$ , with mean  $s_n$  and second moment  $s_n^{(2)}$ . Let  $s$  denote the sum of the mean switchover times experienced in one complete scan of the system, and let  $\rho$  denote the sum of the utilizations at all nodes. In the ensuing discussion, unless otherwise stated, the index for any summation is assumed to range over 1 through  $N$ . We thus have,

$$s = \sum_n s_n \quad (3.2)$$

and

$$\rho = \sum_n \rho_n \quad (3.3)$$

It can be shown (refer [19], for example) that the following conditions are necessary and sufficient for the stability of this system:

$$\rho < 1, \quad (3.4)$$

and

$$\max_n \lambda_n s < 1 - \rho: \quad n = 1, \dots, N-1 \quad (3.5)$$

Let  $c$  denote the mean time taken by the server to complete one scan of the ring. This quantity, termed the cycle time, is given for any cyclic server system as [19]

$$c = s / (1 - \rho) \quad (3.6)$$

In this paper, we are interested in obtaining the mean waiting times,  $W_n$ , experienced by an arriving packet at node  $n$ . The exact analysis of this system for these mean waiting times, in the general case, presents considerable difficulties. Here, we present an exact analysis for the case with two nodes. A conservation law is also developed for the type of systems considered here, which gives one equation in the  $N$  unknowns,  $W_n$ ,  $n = 1, \dots, N$ . For the general case with  $N$  nodes, with nodes 1 through  $N-1$  having identical characteristics, an approximation scheme is presented which gives very good results.

#### 4. Exact analysis for a system with two nodes:

For the exact analysis in the two-queue case, we require that the switchover time from node 2 to node 1 be deterministic. The analysis for the  $W_n$ 's is then based on two conditional cycle times,  $C_1$  and  $C_2$ , where  $C_1$  is the conditional cycle time given there is no service at node 1 and  $C_2$  is the conditional cycle time given there is a service at node 1. Let  $U_n(\cdot)$  and  $U_n^*(\cdot)$  respectively be the distribution function and the Laplace-Stieltjes Transform (LST) of  $C_n$ ,  $n = 1, 2$ . Define the corresponding inter-visit times to node 2 as  $V_1$  and  $V_2$ , where

$$V_1 = S_1 + S_2. \quad (4.1)$$

and

$$V_2 = S_1 + S_2 + H_1, \quad (4.2)$$

and let  $R_n(\cdot)$  and  $R_n^*(\cdot)$  respectively denote the distribution function and LST of  $V_n$ ,  $n = 1, 2$ . The following result, stated as Lemma 4.1, expresses the conditional cycle times in terms of the LST's of the  $V_n$ 's.

Lemma 4.1:

$$U_n^*(x) = R_n^*(x + \lambda_2 - \lambda_2 B_2^*(x)), \quad n = 1, 2. \quad (4.3)$$

where  $B_2^*(x)$  is the LST of a busy period at node 2.

Proof:

We obtain, here, the LST for the distribution function  $U_1(\cdot)$ . The derivation for the LST of  $U_2(\cdot)$  is quite similar.

The conditional cycle time,  $C_1$ , will not exceed a value  $t$  if (and only if) the intervisit time,  $V_1$ , lasts a length  $\omega \leq t$ , and the  $k \geq 0$  customers who arrive at node 2 during this time generate a  $k$  busy period of length at most  $t - \omega$ . Let  $B_2^k(\cdot)$  and  $B_2^{*(k)}(\cdot)$  denote, respectively, the distribution function and the LST of a  $k$  busy period at node 2. Then,

$$U_1(t) = \sum_{k=0}^{\infty} \int_{\omega=0}^t e^{-\lambda_2 \omega} \frac{(\lambda_2 \omega)^k}{k!} B_2^k(t - \omega) dR_1(\omega), \quad (4.4)$$

and taking Laplace Stieltjes transforms on both sides, noting that

$$(B_2^{*(k)}(x)) = (B_2^*(x))^k,$$

we get

$$U_1^*(x) = \int_{t=0}^{\infty} e^{-xt} dU_1(t). \quad (4.5)$$

Substituting the expression for  $U_1(t)$  in (4.4), and after some manipulation, we get

$$\begin{aligned} U_1^*(x) &= \int_{t=0}^{\infty} \sum_{k=0}^{\infty} e^{-\lambda_2 t} \frac{(\lambda_2 t)^k}{k!} (B_2^*(x))^k e^{-xt} dR_1(t) \\ &= \int_{t=0}^{\infty} e^{-(\lambda_2 + x)t} e^{\lambda_2 B_2^*(x)t} dR_1(t) \\ &= R_1^*(x + \lambda_2 - \lambda_2 B_2^*(x)). \end{aligned}$$

□

Now, let  $\psi^2_n$  denote the variance of the switchover time from node  $n$  to node  $(n + 1) \bmod N$ .

Theorem 4.1 obtains an expression for the mean waiting time,  $W_1$ , at node 1.



Theorem 4.1:

$$W_1 = \frac{1}{2(1-\lambda_1 C)(1-\rho_2)} \left[ \frac{\sum_{n=1}^2 \lambda_n h_n^{(2)}}{1-\rho} + c(1-\rho_1+\rho_2) + \frac{\Psi_1^2}{s} \right] \quad (4.6)$$

Proof:

Let  $q_1(k)$  be the probability that there are  $k$  requests waiting at node 1 at the instant the server arrives there. Then,

$$q_1(k) = \sum_{j=1}^{k+1} q_1(j) \int_0^\infty \frac{(\lambda_1 t)^{k-j+1}}{(k-j+1)!} e^{-\lambda_1 t} dU_2(t) + q_1(0) \int_0^\infty \frac{(\lambda_1 t)^k}{k!} e^{-\lambda_1 t} dU_1(t) \quad (4.7)$$

The probability generating function (p.g.f) of the number of requests at node 1 at server arrival

instant,  $Q_1(z) = \sum_{k=0}^{\infty} q_1(k) z^k$ , is then obtained after some straightforward manipulation as

$$Q_1(z) = q_1(0) \frac{zR_1^*(F(z)) - R_2^*(F(z))}{z - R_2^*(F(z))} \quad (4.8)$$

where

$$F(z) = \lambda_1(1-z) + \lambda_2 - \lambda_2 B_2^*(\lambda_1(1-z)) \quad (4.9)$$

Let  $Q_n$  denote the expected number of requests at node  $n$  at server arrival instant,  $n = 1, 2$ . Using the fact that the distribution of waiting requests as seen by a departing request is identical to the distribution seen by an arriving request, and that Poisson arrivals see time averages [30], the following expression is easily obtained:

$$W_1 = \frac{\overline{Q_1}}{\lambda_1(1-q_1(0))} - \frac{1}{\lambda_1} \quad (4.10)$$

The probability that there are zero customers waiting at node 1 at the instant the server arrives there is given by [19]–

$$q_1(0) = 1 - \lambda_1 c.$$

Hence, from equations (4.8) and (4.10), noting that

$$\bar{Q}_1 = \frac{d}{dz} Q_1(z) \Big|_{z=1},$$

and after some manipulation, we obtain,

$$(1 - \lambda_1 c) W_1 = \frac{1}{2(1 - \rho_2)} \left[ \frac{\sum_{n=1}^2 \lambda_n h_n^{(2)}}{1 - \rho} + c(1 + \rho_1 - \rho_2) + \frac{\Psi_1^2}{s} \right],$$

from which the result follows. □

In [24], a conservation law is presented for cyclic server systems with  $N$  nodes, where nodes 1 through  $N-1$  are serviced non-exhaustively and node  $N$  is serviced on an exhaustive basis. This law gives one equation in the  $N$  unknowns  $W_1$  through  $W_N$ . In general, for a system with  $N$  nodes, having nodes 1 through  $m$  serviced non-exhaustively, and nodes  $m + 1$  through  $N$  serviced exhaustively, this law is expressed as Theorem 4.2.

Theorem 4.2:

$$\begin{aligned} \sum_{n=1}^m \rho_n (1 - \lambda_n c) W_n + \sum_{n=m+1}^N \rho_n W_n &= \frac{\rho}{2(1 - \rho)} \sum_{n=1}^N \lambda_n h_n^{(2)} + \frac{\rho}{2s} \sum_{n=1}^N \Psi_n^2 \\ &+ \frac{c}{2} \left[ \sum_{n=1}^m \rho_n (1 + \rho_n) + \sum_{n=m+1}^N \rho_n (1 - \rho_n) \right]. \end{aligned} \quad (4.11)$$

This law is obtained in a straightforward manner, although with a great deal of algebraic manipulation, using an approach similar to that used by Watson for the non-exhaustive case [29].

The proof is omitted. The above law has also been derived independently by Boxma and Gronendijk [3].

From Theorems (4.1) and (4.2), the mean waiting time,  $W_2$ , at node 2 is easily obtained, and this is expressed as Theorem 4.3:

Theorem 4.3:

$$W_2 = \frac{1}{2(1-\rho_2)} \left[ \sum_{n=1}^2 \lambda_n h_n^{(2)} + s(1+\rho_1-\rho_2) + \frac{\Psi_1^2}{c} \right]. \quad (4.12)$$

Proof:

For the two node case we are considering here, equation (4.11) reduces to

$$\rho_1(1-\lambda_1c)W_1 + \rho_2W_2 = \frac{\rho}{2(1-\rho)} \sum_{n=1}^2 \lambda_n h_n^{(2)} + \frac{\rho}{2s} \Psi_1^2 + \frac{c}{2} [\rho_1(1+\rho_1) + \rho_2(1-\rho_2)]. \quad (4.13)$$

The result follows from equations (4.6) and (4.13), after some straightforward algebra.

□

Comparing equations (4.10) and (4.13), we note the following simple relationship between  $W_1$  and  $W_2$ :

$$(1-\lambda_1c)(1-\rho)W_1 = W_2. \quad (4.14)$$

It must be noted that  $W_2$  can be easily obtained directly, using the fact that the variance of the inter-visit time to node 2 is the sum of the variance of the residence time at node 1 plus the variance of the switchover time from node 1. Then, an application of the conservation law given by equation (4.12) gives  $W_1$ . However, the approach detailed above has the advantage that it gives more information about the system, since it also calculates the p.g.f. of the number of customers at node 1 at server arrival instant.

5. Extension to systems with more than 2 nodes:

In this section we will develop an approximation method to analyze the system with more than two nodes. We consider a system with N nodes where all the non-exhaustive nodes (nodes 1 through N-1) have same arrival rates, service distributions and switchover times. This configuration of the system is quite realistic since nodes within a ring network frequently have similar characteristics for arrival rates, service and switchover times. In the following discussion, unless specified otherwise, the index for any summation will be over the range 1 through N.

Let  $V_n$  now denote the intervisit time for node n, namely, the time interval from the instant the server departs from node n until it returns to the same node, and  $T_n$  be the residence time of server at node n during one cycle. To analyse the system, let's consider the exhaustive node N first. It is well known that the mean waiting time at the exhaustive node, N, can be expressed as [10].

$$W_N = \frac{E(V_N)}{2} + \frac{Var(V_N)}{2E(V_N)} + \frac{\lambda_N h_N^{(2)}}{2(1-\rho_N)}. \quad (5.1)$$

Hence the mean waiting time of the exhaustive node can be obtained if the mean and variance of the intervisit times are determined. The mean of intervisit time can be obtained straightforwardly as follows.

$$E(V_N) = c(1-\rho_N). \quad (5.2)$$

In fact equation (5.2) can be applied to any type of node whether it is exhaustive or non-exhaustive.

The variance of the intervisit time is very difficult to obtain and we use an approximation scheme here. To get the estimate of variance of intervisit time to node N, we establish our first heuristic assumption.

<Assumption 1>

The residence times at the non-exhaustive nodes and the switchover times are stochastically independent of each other.

This independence assumption obviously underestimates the variance  $\text{Var}(V_N)$  and therefore  $W_N$  since it ignores the positive correlations among the residence times at the non-exhaustive nodes and the switchover times. Nevertheless, as extensive simulation results show, the error in this approximation of the mean waiting time at node N is very small.

Using the fact that the probability that there is at least one request waiting at a non-exhaustive node n at server arrival instant is  $\lambda_n c$ , we can determine the means and variances of residence times at non-exhaustive nodes. Therefore under independence assumption, we get the following expression for the variance of intervisit time at node N

$$\begin{aligned} \text{Var}(V_N) &\approx \sum_{n \neq N} \text{Var}(T_n) + \sum_n \text{Var}(S_n) \\ &= \sum_{n \neq N} \left[ c\lambda_n h_n^{(2)} - (cp_n)^2 \right] + \sum_n \psi_n^2, \end{aligned} \quad (5.3)$$

where  $\psi_n^2$  is, again, the variance of switchover time from node n. Substituting the expressions for  $E(V_N)$  and  $\text{Var}(V_N)$  given by equations (5.2) and (5.3) into equation (5.1), we obtain the approximate value for the mean waiting time,  $W_N$ , at node N. The approximate values for  $W_1$  through  $W_{N-1}$  is then obtained based on the following assumption.

<Assumption 2>

The mean waiting times at all the non-exhaustive nodes are same.

In fact the locations of nodes slightly affect the mean waiting times since the variance of cycle time at one node could be different from that at another node. But extensive simulation results show that these differences are extremely small and this fact justifies our assumption 2.

Let  $W_{ave}$  denote the average of the mean waiting times  $W_1$  through  $W_{N-1}$ . To obtain  $W_{ave}$ , we use the conservation law derived in the previous section to get

$$(N-1)\rho_1(1-c\lambda_1)\overline{W}_{ave} + \rho_N W_N = \frac{\rho}{2(1-\rho)} \sum_n \lambda_n h_n^{(2)} + \frac{\rho}{2s} \sum_n \psi_n^2 + \frac{c}{2} \left[ (N-1)\rho_1(1+\rho_1) + \rho_N(1-\rho_N) \right]$$

Since we have already estimated the value of  $W_N$ ,  $W_{ave}$  can be obtained directly from this conservation equation. It must be noted that since  $W_N$  is always underestimated, we can see from the conservation equation that  $W_{ave}$  should then always be overestimated in the approximation.

### 5.1. Numerical results:

The approximation method was tested for its accuracy in estimating the mean waiting times by comparing these estimates with simulation results for a number of cases with networks having from 3 to 10 nodes.. Some of these test problems, chosen at random are reported in tables 1 through 4. In these tables the mean waiting estimates obtained by the approximations are listed as also the results of the simulation. In addition, assuming that simulation results are the closest to the exact mean values, the differences in the estimates of the approximations with respect to the simulation results are also expressed as a percentage of the simulation values in these tables. In these tables, the approximation and simulation methods are expressed as App. and Sim. respectively.

Each table has three test problems where the arrival rates at node N are respectively one, two and three times higher than those at the non-exhaustive nodes when the system has 3 or 4 nodes. For the system consisting of 10 nodes, the arrival rates of exhaustive nodes for the three test problems are respectively chosen to be one, three and five times higher than those at the non-exhaustive nodes. Among the four tables table 3 treats the examples with exponential switchover times while the other tables assume deterministic switchover times.

In some cases, the simulation could take a long time to stabilize. Equation (3.4) and (3.5) would identify such instances. All the simulation experiments that we conducted were designed such that

and

$$\rho \leq 0.85 ,$$

$$\gamma = \max_{n \leq N-1} (\lambda_n s) + \rho \leq 0.96$$

For larger values of  $\rho$  or  $\gamma$ , we found that the simulation results did not stabilize even after a very long time. In our experiments, while simulation results of the cases for  $\rho < 0.8$  seem to be quite reliable, the results for higher traffic intensities appear to be a little less accurate.

The results of the numerical experiments can be summarized as follows.

- As we expect, comparison with simulation shows that the value  $W_N$  obtained by the approximation is always an underestimate. However, the tables may indicate that the value of  $W_{ave}$  obtained by the approximation is sometimes less than that obtained by simulation. These exceptions occur at high values of  $\rho$  ( $\geq 0.8$ ), and are due to the fact that the simulation results are not completely accurate.

- The relative error of the approximation method is usually within 10% for  $W_N$  and within 3% for  $W_{ave}$ .

- The error in estimating  $W_N$  decreases with increasing values of  $\rho$ . This can be explained by the fact that at higher traffic intensities, the switchover times has less impact on the residence times at the non-exhaustive nodes. Hence the approximation assumption 1 becomes more accurate.

- The relative error of the approximation method is not sensitive to the number of nodes. As noted above, the relative error appears to decrease with increasing traffic intensities.

## 6. Conclusion:

This paper presents and analyzes a new polling mechanism for cyclic server systems which is appropriate for interconnected token ring networks, where one node receives exhaustive service while all other nodes receive non-exhaustive service. An exact analysis for a system with one

exhaustive service node and one non-exhaustive node was made. For the system with more than one non-exhaustive service node, an approximation scheme was developed for the special case where these non-exhaustive service nodes had identical characteristics. This approximation scheme is extremely simple and, as noted from the comparison with simulation results, it is very accurate. The conservation law used in this approximation gives a very good insight into the behavior of such ring networks.

For further study, it is desirable that the approximation method be extended for more general systems having more than one exhaustive service node, and with non-exhaustive service nodes having different characteristics.



Results for Expected Waiting Times

<Table 1>

N = 3 queues  
 Exponential service times  
 $\lambda_1 = \lambda_2 = 0.2$ ,  $h_1 = h_2 = h_3 = 0.8$   
 All switchover times are 0.25, deterministic.

	$\lambda_3 = \lambda_1$ ( $\rho = 0.48$ )		$\lambda_3 = 2\lambda_1$ ( $\rho = 0.64$ )		$\lambda_3 = 3\lambda_1$ ( $\rho = 0.80$ )	
	$W_{ave}$	$W_3$	$W_{ave}$	$W_3$	$W_{ave}$	$W_3$
Sim.	2.440*	1.058	5.722*	1.427	34.801*	2.050
App.	2.442	1.019	5.791	1.383	34.421	2.021
E(%)	0.1	-3.7	1.2	-3.1	-1.1	-1.5

<Table 2>

N = 4 queues  
 Exponential service times  
 $\lambda_1 = \lambda_2 = \lambda_3 = 0.2$ ,  $h_1 = h_2 = h_3 = h_4 = 0.7$ .  
 All switchover times are 0.15, deterministic.

	$\lambda_4 = \lambda_1$ ( $\rho = 0.56$ )		$\lambda_4 = 2\lambda_1$ ( $\rho = 0.70$ )		$\lambda_4 = 3\lambda_1$ ( $\rho = 0.84$ )	
	$W_{ave}$	$W_3$	$W_{ave}$	$W_4$	$W_{ave}$	$W_4$
Sim.	2.462*	1.076	5.634*	1.412	33.216*	1.974
App.	2.514	0.996	5.771	1.319	34.656	1.911
E(%)	2.1	-7.4	2.4	-6.6	4.3	-3.2

the results represent mean waiting times averaged over the non-exhaustive nodes.

< Table 3 >

N = 4 queues

Exponential service times

$\lambda_1 = \lambda_2 = \lambda_3 = 0.2$ ,  $h_1 = h_2 = h_3 = h_4 = 0.7$ .

All switchover times are 0.15, exponential.

	$\lambda_4 = \lambda_1$ ( $\rho = 0.56$ )		$\lambda_4 = 2\lambda_1$ ( $\rho = 0.70$ )		$\lambda_4 = 3\lambda_1$ ( $\rho = 0.84$ )	
	$W_{ave}$	$W_4$	$W_{ave}$	$W_4$	$W_{ave}$	$W_4$
Sim.	2.580*	1.129	5.884*	1.466	35.582*	2.000
App.	2.634	1.034	5.945	1.350	35.172	1.932
E(%)	2.1	-8.4	1.0	-7.9	-1.2	-3.4

< Table 4 >

N = 10 queues

Exponential service times

$\lambda_1 = \lambda_2 = \dots = \lambda_9 = 0.1$ ,  $h_1 = h_2 = \dots = h_{10} = 0.6$

All switchover times are 0.12, deterministic.

	$\lambda_{10} = \lambda_1$ ( $\rho = 0.60$ )		$\lambda_{10} = 3\lambda_1$ ( $\rho = 0.72$ )		$\lambda_{10} = 5\lambda_1$ ( $\rho = 0.84$ )	
	$W_{ave}$	$W_{10}$	$W_{ave}$	$W_{10}$	$W_{ave}$	$W_{10}$
Sim.	3.631*	1.930	7.225*	2.410	33.302*	3.296
App.	3.647	1.741	7.317	2.199	34.287	3.171
E(%)	0.4	-9.8	1.3	-8.8	2.9	-3.8

the results represent mean waiting times averaged over the non-exhaustive nodes.

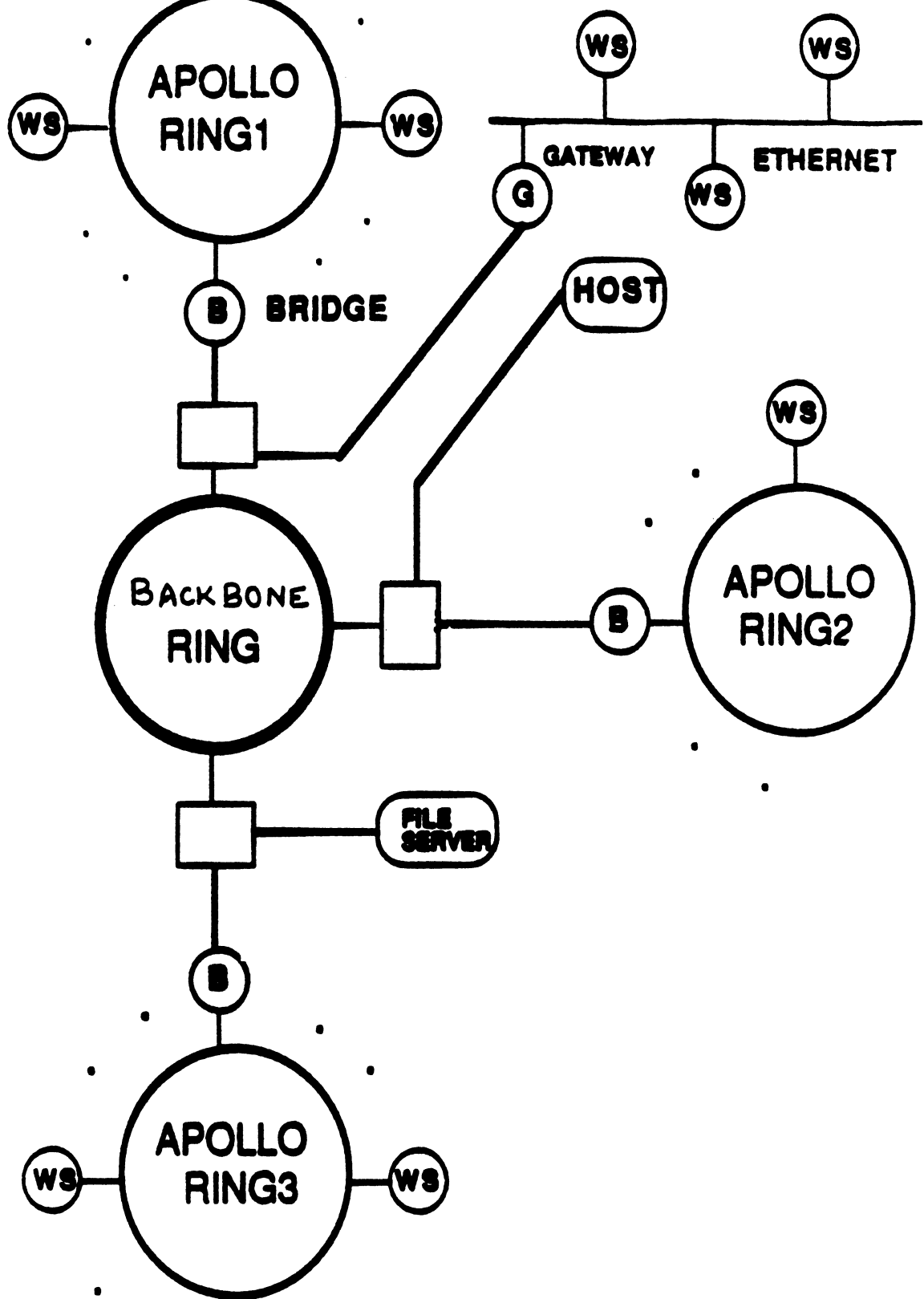


Figure 1: A system with 3 Apollo rings interconnected by a Backbone ring.

## REFERENCES

1. Berry, R., and Chandy, K.M., "Performance models of token ring local area networks", Performance Evaluation Review, (special issue), pp. 266-274, August 1983.
2. Boxma, O.J., and Meister, B., "Waiting-time approximations for cyclic- service systems with switch-over times", Performance '86 and ACM Sigmetrics 1986, Proceedings of the Joint Conference on Computer Performance modeling, Measurement and Evaluation, pp. 254-262, Raleigh, May 1986.
3. Boxma, O.J., and Gronendijk, W.P., "Pseudo-conservation laws in cyclic service systems", Journal of Applied Probability, vol 24, 1987 (to appear).
4. Bux, W., "Local area subnetworks: A performance comparison", IEEE Transactions on Communications, v 10, October 1981, pp. 1465-1473.
5. Bux, W., and Truong, H.L., "Mean-delay approximation for cyclic-service queueing systems", Performance Evaluation, v 3, 1983, pp. 187-196.
6. Carsten, R.T., Newhall, E.E., and Posner, M.J.M., "A simplified analysis of scan times in an asymmetrical Newhall loop with exhaustive service", IEEE Transactions on Communications, v COM-25, no 9, September 1977, pp. 951-957.
7. Cooper, R.B., Introduction to Queueing Theory, North-Holland, New York, 1981.
8. Cooper, R.B., and Murray, G., "Queues served in cyclic order", Bell System Technical Journal, v 48, no 3, March 1969, pp. 675-689.
9. Cooper, R.B., "Queues served in cyclic order: waiting times", Bell System Technical Journal, v 49, no 3, March 1970, pp. 399-413.
10. Eisenberg, M., "Queues with periodic service and changeover time", Operations Research, v 19, 1972, pp. 440-451.
11. Eisenberg, M., "Two queues with alternating service", SIAM Journal of Applied Mathematics, v 36, 1979, pp. 287-303.
12. Ferguson, J.J., and Aminetzah, Y.J., "Exact results for non-symmetric token-ring systems", IEEE Transactions on Communications, v COM-33, no. 3, pp. 223-231, March 1985.
13. Fuhrmann, S.W., "Symmetric queues served in cyclic order", Operations Research Letters, v 4, no 3, 1985, pp. 139-144.
14. Fuhrmann, S.W., and Cooper, R.B., "Stochastic decompositions in the M/G/1 queue with generalized vacations", Operations Research, v 33, no 5, September-October 1985, pp. 1117-1129.
15. Hayes, J.F., Modeling and Analysis of Computer Communications Networks, Plenum Press, New York, 1986.
16. Kandlur, D.D., Park, J.T., Srinivasan, M.M., and Teorey, T.J., "Modeling and simulation of a large interconnected local area network", Technical Report CITI-

TR-2-86, Center for Information Technology Integration, The University of Michigan, Ann Arbor, December 1986.

17. Kant, K., "Performance analysis of hierarchical ring networks", Technical Report, The Pennsylvania State University, 1986.
18. Kimura, G., and Takahashi, Y., "Diffusion approximation for a token ring system with non-exhaustive service", IEEE Journal on Selected Areas in Communication, v SAC-4, no 6, September 1986, pp. 794-801.
19. Kuehn, P.J., "Multiqueue systems with non-exhaustive service", Bell System Technical Journal, v 58, no 3, pp. 671-698, March 1979.
20. Lavenberg, S.S., Computer Performance Modeling Handbook, Academic Press, New York, 1983.
21. Leach, P.J., Levine, P.H., Douros, B.P., Hamilton, J.A., Nelson, D.L., and Stumpf, B.L., "The architecture of an integrated local network", IEEE Journal on Selected Areas in Communications, v SAC-1, no 5, November 1983, pp. 842-856.
22. Sevcik, K.C., and Johnson, M.J., "Cycle time properties of the FDDI token ring protocol", Technical Report CSRI-179, April 1986, Computer Systems Research Institute, University of Toronto, Canada M5S 1A1.
23. Srinivasan, M.M., "An approximation algorithm for cyclic server systems with non-exhaustive service", Technical Report 86-36, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor, MI, September 1986.
24. Srinivasan, M.M., and Lee, H.S., "A priority polling mechanism for token ring networks", ORSA/TIMS special interest conference on Queueing Networks and their applications, New Brunswick, N.J., January 1987.
25. Takagi, H., "Analysis of polling systems", The MIT Press, 1986.
26. Takagi, H., "Queueing analysis of non-preemptive reservation polling discipline", Performance '86 and ACM Sigmetrics 1986, Proceedings of the Joint Conference on Computer Performance Modeling, Measurement and Evaluation, pp. 237-244, Raleigh, May 1986.
27. Takine, T., Takahashi, Y., and Hasegawa, T., "Performance analysis of a polling system with single buffers and its application to interconnected networks", IEEE Journal on Selected Areas in Communication, v SAC-4, no 6, September 1986, pp. 802-812.
28. Tran-Gia, P., and Raith, T., "Multiqueue systems with finite capacity and non-exhaustive cyclic service", International Seminar on Computer Networking and Performance Evaluation, pp. 1-13, September 1985, Tokyo, Japan.
29. Watson, K.S., "Performance of cyclic service strategies - a survey", Performance '84, Elsevier Publishers, New York, 1984.
30. Wolff, R.W., "Poisson arrivals see time averages", Operations Research, v 30, 1982, pp. 223-231.