NASA GRANT
IN-63-CR
7595
P30

ANALYSIS OF EXHAUSTIVE LIMITED SERVICE FOR TOKEN RING NETWORKS

(NASA-CR-188092) ANALYSIS OF EXHAUSTIVE LIMITED SERVICE FOR TOKEN RING NETWORKS (Houston Univ.) 30 p CSCL 09B

N91-21790

Unclas

G3/63 0007595

Jeffery H. Peden

Digital Technology

October 1990

Cooperative Agreement NCC 9-16
Research Activity No. SE.31

NASA Johnson Space Center
Engineering Directorate
Flight Data Systems Division

Research Institute for Computing and Information Systems
University of Houston - Clear Lake

T·E·C·H·N·I·C·A·L R·E·P·O·R·T

The RICIS Concept

The University of Houston-Clear Lake established the Research Institute for Computing and Information systems in 1986 to encourage NASA Johnson Space Center and local industry to actively support research in the computing and information sciences. As part of this endeavor, UH-Clear Lake proposed a partnership with JSC to jointly define and manage an integrated program of research in advanced data processing technology needed for JSC's main missions, including administrative, engineering and science responsibilities. JSC agreed and entered into a three-year cooperative agreement with UH-Clear Lake beginning in May, 1986, to jointly plan and execute such research through RICIS. Additionally, under Cooperative Agreement NCC 9-16, computing and educational facilities are shared by the two institutions to conduct the research.

The mission of RICIS is to conduct, coordinate and disseminate research on computing and information systems among researchers, sponsors and users from UH-Clear Lake, NASA/JSC, and other research organizations. Within UH-Clear Lake, the mission is being implemented through interdisciplinary involvement of faculty and students from each of the four schools: Business, Education, Human Sciences and Humanities, and Natural and Applied Sciences.

Other research organizations are involved via the "gateway" concept. UH-Clear Lake establishes relationships with other universities and research organizations, having common research interests, to provide additional sources of expertise to conduct needed research.

A major role of RICIS is to find the best match of sponsors, researchers and research objectives to advance knowledge in the computing and information sciences. Working jointly with NASA/JSC, RICIS advises on research needs, recommends principals for conducting the research, provides technical and administrative support to coordinate the research, and integrates technical results into the cooperative goals of UH-Clear Lake and NASA/JSC.

Preface

This research was conducted under auspices of the Research Institute for Computing and Information Systems by Jeffery H. Peden and Digital Technology. Dr. George Collins, Associate Professor of Computer Systems Design, served as RICIS technical representative for this activity.

Funding has been provided by the Engineering Directorate, NASA/JSC through Cooperative Agreement NCC 9-16 between NASA Johnson Space Center and the University of Houston-Clear Lake. The NASA technical monitor for this activity was Frank W. Miller, of the Systems Development Branch, Flight Data Systems Division, Engineering Directorate, NASA/JSC.

The views and conclusions contained in this report are those of the author and should not be interpreted as representative of the official policies, either express or implied, of NASA or the United States Government.

Analysis of Exhaustive Limited Service for Token Ring Networks

JEFFERY H. PEDEN

University of Virginia, Charlottesville, Virginia

Abstract. Token ring operation is well-understood in the cases of exhaustive, gated, gated limited, and ordinary cyclic service. There is no current data, however, on queueing models for the *exhaustive limited* service type. This service type differs from the others in that there is a preset maximum (ω) on the number of packets which may be transmitted per token reception, and packets which arrive after token reception may still be transmitted if the preset packet limit has not been reached. Exhaustive limited service is important since it closely approximates a timed token service discipline (the approximation becomes exact if packet lengths are constant).

In this paper we present a method for deriving the z-transforms of the distributions of the number of packets present at both token departure and token arrival for a system using exhaustive limited service. This will then allow us to derive a formula for mean queueing delay and queue lengths. The method used is theoretically applicable to any ω, but due to computational complexity, it becomes impractical for large values of ω. Fortunately, as the value of ω becomes large (typically values on the order of $\omega = 8$ are considered large), the exhaustive limited service discipline closely approximates an exhaustive service discipline.

Categories and Subject Descriptors: D.4.8 [Performance]: Queueing theory; C.2.5: Networks, local.

General Terms: Performance, Theory.

Additional Key Words and Phrases: Queueing theory, network performance measurement, network design, network protocol analysis.

1. *Introduction*

The general operation of a token ring follows. A circulating token arrives at the various ring stations in either logical or physical order, depending on whether the ring is a logical ring implemented on a bus (e.g., IEEE 802.4 Token Bus), or is an actual physical ring (e.g., IEEE

This research was supported by the Flight Data Systems Branch of NASA-Johnson Space Center.

Author's address: Jeffery H. Peden, Department of Computer Science, Thornton Hall, University of Virginia, Charlottesville, VA 22903.

802.5 Token Ring, FDDI Token Ring). If a station has packets enqueued, the reception of the token allows some or all of these packets to be transmitted. The token is then released so that it can visit the next station. If there are no packets enqueued at a station at token reception, the token is immediately sent to the next station.

The number of packets that may be transmitted at token reception is controlled by the service discipline. There are six possible service disciplines: (1) ordinary cyclic service, (2) exhaustive, (3) gated, (4) timed token, (5) gated limited (GL) service, and (6) exhaustive limited (EL) service. Ordinary cyclic service allows only one packet to be transmitted when the token is received. Exhaustive service allows packets to be transmitted until the station's queue is completely empty, that is, there is no upper limit to the amount of time a station may use the token. Gated service allows only those packets present at token arrival to be transmitted. GL service is identical to gated service with the exception that an upper limit is placed on the number of packets that may be transmitted, that is, the token serves the minimum of the number of packets present at token arrival, or the maximum number set by the limit (an excellent treatment of gated limited service can be found in [2]). Timed token service places a limit on the amount of *time* a station may hold the token before releasing it, with the provision that at least one packet may always be transmitted. The last service type, EL service, allows up to ω packets to be served, regardless of whether or not they were present at token arrival. It also includes the provision that the token departs the station as soon as the station's queue is empty, even if this event occurs before the ω packet limit has been reached.

The EL service discipline has obvious application in the analysis of token rings which use token holding timers (timed token service discipline), in that the ω packet limit may be used to approximate the maximum number of packets which may be transmitted during a service session. The limited gated discipline may also be used, although it is less accurate. Among token rings which use EL service are the IEEE 802.4 Token Bus [3], the IEEE 802.5 Token Ring [4], and the FDDI Token Ring [1]. Note that if single packet per token service is used, EL, GL, and ordinary cyclic service become identical.

In this paper we will develop the z-transform of the density function giving the number of packets present at token departure, using the embedded Markov chain approach. Since the transform is dependent upon the state probabilities of our Markov chain, we show how these probabilities may be obtained. We then present an example solution along with numerical results.

2. *Assumptions*

Our analysis is based on several assumptions, the first of which is that the stations on the ring are indistinguishable. That is, arrival processes are the same at each station, along with mean packet service time, station load, station latency, etc. We also assume that arrivals at each station follow a Markov process, and that all packets are offered at the same priority level, resulting in an M/G/1 model.

We assume *pseudo*-independence of successive token cycles, rather than complete independence. The complete independence of token cycles is an assumption frequently made for mathematical tractability, which states that a transmission has no effect on the length of the following token cycle. Our assumption of pseudo-independence assumes the departure from an equilibrium state on a single successive token cycle, and then a return to equilibrium. We show that valid results are derived using this assumption. The reasons for this are that (1) the dependence of any token cycle on the previous one is very small, and (2) the fact that we are using unconditional distributions based on the mean value of the token cycle time (which is unaffected by the number of packets per token allowed -- see Section 7) allows us to derive the unconditional probabilities we use in our analysis.

3. *Preliminaries*

Several terms and functions are used in our analysis. Following the standard definitions, the z-transform $G(z)$ of a function $G(t)$ is

$$G(z) \equiv E[z^X] = \sum_{k=0}^{\infty} g_k z^k \tag{1}$$

where $g_k \equiv P[X=k]$, and $E[X]$ is the mean or expected value of the random variable X. Similarly, the *Laplace-Stieltjes* transform (or LST) of a function is

$$F^*(s) \equiv E[e^{-sX}] = \int_{-\infty}^{\infty} e^{-sx} dF(x) \tag{2}$$

The expectation operator $E[X]$ for some random variable X is

$$E[X] \equiv \sum_{k=0}^{\infty} k\, P[X=k] \tag{3}$$

for the discrete case. For the continuous case, it is

$$E[X] \equiv \int_{-\infty}^{\infty} x\, dF(x) \tag{4}$$

We use the following notation for conditional expectation:

$$P[x \mid y] \equiv P[x \text{ given that } y \text{ has occurred}] \equiv \frac{P[x \cap y]}{P[y]} \tag{5}$$

4. *Fundamental Relationship*

This analysis is for EL service, where the number of packets per token may be any finite integer greater than or equal to one; the analysis implicitly includes a derivation for ordinary cyclic service.

The fundamental equation of our system is

$$R_{n+1} = R_n + A_n - \Omega_n \tag{6}$$

where

$R_n \equiv$ the number of packets present at the n^{th} token departure

$A_n \equiv$ the number of packets which arrive between the n^{th} and $n+1^{st}$ token departures

$\Omega_n \equiv$ the number of packets transmitted between the n^{th} and $n+1^{st}$ token departures

This is a time-dependent process. In order to remove the time dependency, we let $n \to \infty$, and take the expectation. This results in

$$E[R] = E[R] + E[A] - E[\Omega] \tag{7}$$

From this equation, we see that queue length at token departure is a natural point at which to embed a semi-Markov process.

5. *Probability Distribution of Packet Arrivals*

In order to define the probability distribution of packet arrivals, we first need certain terms. An *arrival slot* is defined as being either the token vacation or the duration of a packet transmission; there are at most $\omega + 1$ arrival slots in a token cycle, there is a minimum of one. The quantity $\vec{s}$ is defined as a vector of length $\omega + 1$. The elements of $\vec{s}$ are numbered from 0 to ω. The j^{th} element of $\vec{s}$ is an integer representing the number of arrivals during arrival slot j. The 0^{th} element represents the number of arrivals during the token vacation; successive elements represent arrivals during packet transmissions.

The inclusion function $I(\vec{s}, \Omega, A, R)$ is defined as

$$I(\vec{s}, \Omega, A, R) = 1, \quad R + A \geq \Omega$$

$$I(\vec{s}, \Omega, A, R) = 0, \quad \text{otherwise} \tag{8}$$

where Ω is the slot number of $\vec{s}$, A is the number of packets which have arrived previous to slot Ω given the current value of $\vec{s}$, and R is the state of the queue at the previous token departure. The purpose of this function is to control the inclusion of certain terms in the calculation of arrival probabilities.

THEOREM 1. *The probability $\pi_{i|j}$ of i packet arrivals occurring between the current token departure and the next token departure given state j at the current token departure is given by*

$$\pi_{i\,|\,j} = \sum_{\vec{s}=[0,\ldots,k]}^{[k,\ldots,0]} P[\vec{s}_0 = \vec{s}] \prod_{\Omega=0}^{\omega} I(\vec{s}, \Omega, A_\Omega, R) \tag{9}$$

with $\vec{s}$ *running through all canonical forms of all lengths of* $\vec{s}$ *from* $len(\vec{s}) = 1$ *to* $len(\vec{s}) = \omega + 1$, *beginning with* $\vec{s} = [0,\ldots,k]$ *and ending with* $\vec{s} = [k,\ldots,0]$ *such that the integer elements of* $\vec{s}$ *sum to k.*

PROOF. The calculation of $\pi_{j\,|\,i}$ is complicated by the fact that the arrivals are constrained to arrive early enough to make a contribution to service time. For example, it is evident that for any packets to be left behind at token departure, ω packets must have been transmitted (if less than ω packets were transmitted, the station was empty of packets at token departure). However, if only a single packet was present at the previous token departure, and the number left behind at the current token departure is greater, then at least $\omega + 1$ packets must have arrived during a total of ω packet service times plus the token vacation time. But if all ω arrived during the ω^{th} service time, there would only have been a single service time (a contradiction), since only the single packet present would have been served with the token then departing. Therefore, some of the packets had to have arrived prior to the ω^{th} service time for there to have been ω packets transmitted. All possible arrival forms are generated, whether or not they are admissible; the function $I(\vec{s}, \Omega, A_\Omega, R)$ excludes inadmissible forms. An example of possible arrival forms for $\omega = 2$ is given in Figure 1.

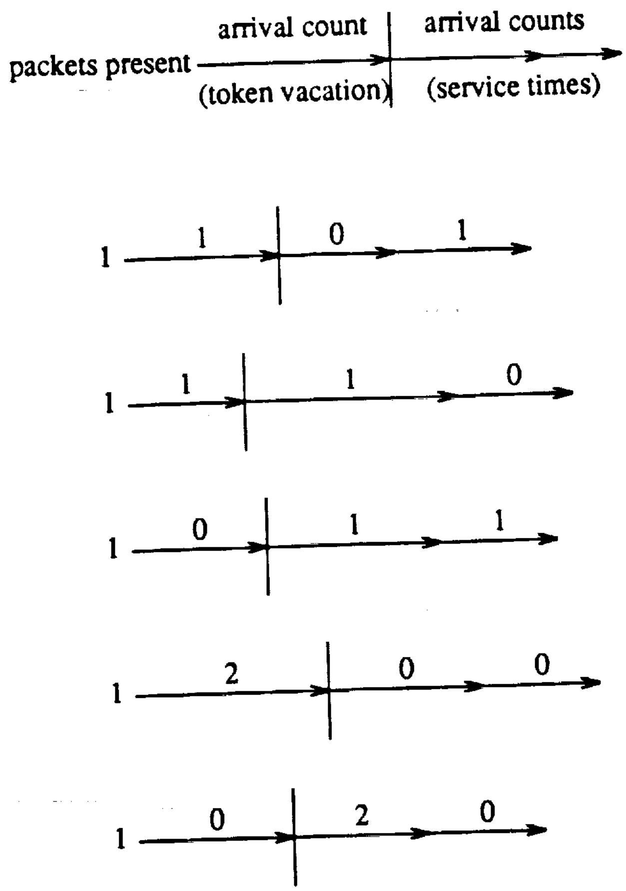

FIG. 1. Possible arrival patterns for two packet arrivals.

To complete the proof, we must show that

$$\sum_{i=0}^{\infty} \pi_{i|j} = 1 \tag{10}$$

If we represent the conditional state probabilities as $p_{0|j}, p_{1|j}, p_{2|j}, \ldots$, then we may propose the following lemma:

LEMMA 1.

$$
\begin{aligned}
p_{0|j} &= \pi_{0|j} + \pi_{1|j} + \cdots + \pi_{\omega-j|j} \\
p_{1|j} &= \pi_{\omega-j+1|j} \\
p_{2|j} &= \pi_{\omega-j+2|j}
\end{aligned} \tag{11}
$$

etc.

PROOF. If j packets are present at a station, then $\omega - j$ packet may still arrive and be served, thus leaving the station empty of packets at token departure. If i packets are to be present at token departure given j present at the previous token departure, then since up to $\omega - j$ packets may arrive for no packets to be present at token departure, $\omega - j + i$ packets must arrive for i packets to remain. □

We know that that $\sum_{i=0}^{\infty} p_{i\,|\,j} = 1$. From the above lemma we have $\sum_{i=0}^{\infty} p_{i\,|\,j} = \sum_{i=0}^{\infty} \pi_{i\,|\,j}$. Thus

$\sum_{i=0}^{\infty} \pi_{i\,|\,j} = 1$. □

COROLLARY 1. *The number of permissible arrival forms is always less than or equal to* $\binom{\omega + i}{i}$ *for i arriving packets.*

PROOF. Since arrivals are indistinguishable, the total number of arrival forms is given by the *Bose-Einstein* statistic. Due to the fact that some arrival forms do not meet the criteria given in Theorem 1 for timeliness of arrival, this statistic represents a maximum. □

Corollary 1 is an important measure since it shows the complexity of the algorithm used to compute arrival probabilities.

6. *The z-transform*

The method we use to derive the z-transform is to derive the conditional transforms which will thus allow us to express the unconditional transform in terms of unconditional state probabilities. We will show that the number of unconditional state probabilities needed is equal to ω, and then indicate a method for finding them.

THEOREM 2.

$$G(z \mid \varepsilon = \omega + \delta) = z^{\delta}\, V^{*}(\lambda - \lambda z \mid \varepsilon)\, [B^{*}(\lambda - \lambda z)]^{\omega} \tag{12}$$

where $V^*(s)$ and $B^*(s)$ are the LST's of the token vacation time and packet service time distributions, respectively.

PROOF. Any state $\varepsilon < \delta$ is unreachable since only ω packets may be transmitted. Since there are at least $\omega + \delta$ packets present at token arrival, exactly ω packets will be transmitted. Therefore, the z-transform is defined by the expression

$$G(z \mid \varepsilon = \omega + \delta) = \sum_{k=0}^{\infty} \pi_k z^{k+\delta}$$

$$= z^{\delta} \sum_{k=0}^{\infty} \pi_k z^k \tag{13}$$

where π_k is the probability of k packet arrivals during the token vacation time plus ω service times. We may now apply the well-known result from queueing theory that a system with Poisson arrivals which only experiences unit changes in state has the following relationship between its z-transform and the LST of its packet service time distribution:

$$H(z) = U^*(\lambda - \lambda z) \tag{14}$$

Defining the LST of the token vacation time distribution to be $V^*(s)$, and the LST of the packet service time distribution to be $B^*(s)$, we have

$$G(z \mid \varepsilon = \omega + \delta) = z^{\delta} V^*(\lambda - \lambda z)[B^*(\lambda - \lambda z)]^{\omega} \tag{15}$$

□

For any state $0 \le \varepsilon < \omega$, the construction of the z-transform is accomplished by realizing that the number of packets left behind in the queue at token departure is the number left behind at the previous token departure, plus the number which arrive to be served, minus the number actually served. Defining $G(z \mid \varepsilon)$ to be the z-transform of the state transition probabilities, and $H(z \mid \varepsilon)$ to be the z-transform of the packet arrival probabilities, the relationship between $g(z \mid \varepsilon)$ and $H(z \mid \varepsilon)$ is given by the following theorem.

THEOREM 3.

$$G(z \mid \varepsilon) = \frac{H(z \mid \varepsilon) z^{\varepsilon}}{\sum_{k=0}^{\omega-\varepsilon-1} \pi_k z^{k+\varepsilon} + \left[1 - \sum_{k=0}^{\omega-\varepsilon-1} \pi_k\right] z^{\omega}} \tag{16}$$

PROOF. The z-transform for the number left behind at the previous token departure is merely z^{ε}. Therefore, since addition and subtraction in the distribution domain corresponds to multiplication and division respectively in the transform domain, the numerator of $G(z \mid \varepsilon)$ is equal to $H(z \mid \varepsilon) z^{\varepsilon}$. The transform of the number delivered is given by the denominator of $G(z \mid \varepsilon)$. If no packets are left behind, i.e., $\varepsilon = 0$, then the number delivered is equal to the lesser of the number which arrive and ω. If a single packet is left behind, the number delivered is at least one, plus those that arrive, or ω, whichever is less, and so on. Using mathematical notation, this quantity is identical to the denominator of $G(z \mid \varepsilon)$. Since these packets are subtracted from those which are already present plus those which arrive, this transform is in the denominator of $G(z \mid \varepsilon)$. □

THEOREM 4. *For any state $0 \le \varepsilon < \omega$, the conditional z-transform for that state is given by*

$$\begin{aligned} G(z \mid \varepsilon) = {} & V^*(\lambda - \lambda z)\,[B^*(\lambda - \lambda z)]^{\omega} \\ & - \pi_{0\mid\varepsilon}[B^*(\lambda - \lambda z)]^{\omega-\varepsilon} + \pi_{0\mid\varepsilon} \\ & - \pi_{1\mid\varepsilon}\, z\,[B^*(\lambda - \lambda z)]^{\omega-\varepsilon-1} + \pi_{1\mid\varepsilon}\, z \\ & - \pi_{2\mid\varepsilon}\, z^2\,[B^*(\lambda - \lambda z)]^{\omega-\varepsilon-2} + \pi_{1\mid\varepsilon}\, z^2 \\ & - \cdots \\ & - \pi_{\omega-\varepsilon-1\mid\varepsilon}\, z^{\omega-\varepsilon-1} B^*(\lambda - \lambda z) + \pi_{\omega-\varepsilon-1\mid\varepsilon}\, z^{\omega-\varepsilon-1} \end{aligned} \tag{17}$$

PROOF. $G(z)$ is always less than or equal to $V^*(\lambda - \lambda z)\,[B^*(\lambda - \lambda z)]^{\omega}$. The forms which must be omitted consist of those arrival formats such that the number of arrivals is less than $\omega - \varepsilon$. These, however, are given by $\pi_{0\mid\varepsilon}[B^*(\lambda - \lambda z)]^{\omega-\varepsilon} + \pi_{0\mid\varepsilon}$, $\pi_{1\mid\varepsilon}[B^*(\lambda - \lambda z)]^{\omega-\varepsilon-1} + \pi_{1\mid\varepsilon}$, through $\pi_{\omega-\varepsilon-1|\varepsilon} B^*(\lambda - \lambda z) + \pi_{\omega-\varepsilon-1|\varepsilon}$. These forms serve to eliminate the first $\omega - \varepsilon - 1$ terms from the z-transform. However, these terms are replaced by $\pi_{0|\varepsilon} + \pi_{1|\varepsilon} z + \cdots + \pi_{\omega-\varepsilon-1|\varepsilon} z^{\omega-\varepsilon-1}$, which are the correct $\omega - \varepsilon - 1$ first terms of the transform. □

Once all the conditional transforms have been obtained, the final unconditional transform $G(z)$ for any ω is given by

$$G(z) = \sum_{\varepsilon=0}^{\infty} P_\varepsilon G(z \mid \varepsilon) \tag{18}$$

THEOREM 5. *The final unconditional z-transform is dependent upon* ω *unconditional state probabilities* $P_0, P_1, \ldots, P_{\omega-1}$.

PROOF. From Theorem 2, we have

$$G(z) = \sum_{\varepsilon=0}^{\omega-1} P_\varepsilon G(z \mid \varepsilon) + \sum_{\varepsilon=\omega}^{\infty} P_\varepsilon z^{\varepsilon-\omega} V^*(\lambda - \lambda z)[B^*(\lambda - \lambda z)]^\omega \tag{19}$$

We recognize that

$$\sum_{\varepsilon=\omega}^{\infty} P_\varepsilon z^{\varepsilon-\omega} V^*(\lambda - \lambda z)[B^*(\lambda - \lambda z)]^\omega$$

$$= \frac{V^*(\lambda - \lambda z)[B^*(\lambda - \lambda z)]^\omega \left(G(z) - \sum_{\varepsilon=0}^{\omega-1} P_\varepsilon z^\varepsilon\right)}{z^\omega} \tag{20}$$

$G(z)$ ultimately resolves to

$$G(z) = \frac{V^*(\lambda-\lambda z)[B^*(\lambda-\lambda z)]^\omega \sum_{\varepsilon=0}^{\omega-1} P_\varepsilon z^\varepsilon - z^\omega \sum_{\varepsilon=0}^{\omega-1} P_\varepsilon G(z \mid \varepsilon)}{V^*(\lambda-\lambda z)[B^*(\lambda-\lambda z)]^\omega - z^\omega} \tag{21}$$

Thus $G(z)$ is dependent only on $P_0, P_1, \ldots, P_{\omega-1}$. □

The state probabilities $P_0, P_1, \ldots, P_{\omega-1}$ are obtained using the method in [2]. The unconditional transform has a degree ω polynomial in z in the denominator, and a degree $\omega - 1$ polynomial in the numerator. Since the transform is certainly analytic on and inside the unit disk, Rouche's Theorem guarantees that the numerator and denominator will have $\omega - 1$ identical roots inside the unit disk (from inspection, the ω^{th} root is 1). The denominator polynomial, which involves only z, is used to obtain $\omega - 1$ of the zeros of the numerator polynomial, resulting in $\omega - 1$ equations in ω unknowns. The necessary ω^{th} equation is obtained using the fact that $G(z)\big|_{z=1} = 1$. These zeros are now used in the numerator to solve for $P_0, P_1, \ldots, P_{\omega-1}$.

Using $G(z)$ we can now obtain the moments $E[R]$ and $E[R^2]$ of queue length at token departure via differentiation of $G(z)$:

$$E[R] = \frac{d}{dz} G(z) \Bigg|_{z=1} \tag{22}$$

$$E[R^2] = \frac{d^2}{dz^2} G(z) \Bigg|_{z=1} + E[R] \tag{23}$$

The z-transform derived above is not directly invertible, thus making it impossible to derive a closed-form expression for the probability distribution function. However, it is possible to numerically invert the z-transform to any degree desired, that is, it is theoretically possible to compute any state probability; however, the computation of P_k increases quickly in difficulty as k grows large. Fortunately, the state probabilities quickly become numerically insignificant as k grows large.

THEOREM 6. *Given unconditional state probabilities* $P_0, P_1, \ldots, P_{k-1}$,

$$P_k = \frac{P_{k-\omega} - \sum_{i=0}^{k-1} p_{k-\omega|i} P_i}{p_{k-\omega|k}} \tag{24}$$

PROOF. State $\varepsilon = k - \omega$ can be reached only from states $P_0, P_1, \ldots, P_k$. Therefore

$$P_{k-\omega} = \sum_{i=0}^{k-1} p_{k-\omega|i} P_i + p_{k-\omega|k} P_k \tag{25}$$

with the only unknown being P_k. Solving for P_k, we have

$$P_k = \frac{P_{k-\omega} - \sum_{i=0}^{k-1} p_{k-\omega|i} P_i}{p_{k-\omega|k}} \tag{26}$$

□

COROLLARY 2. *Given unconditional state probabilities $P_0, P_1, \ldots, P_{\omega-1}$, it is possible to find any unconditional state probability.*

PROOF. The proof is by induction on k, where the state probability to be found is $P_{\omega+k}$. It is evident that we can find P_ω, given that we have $P_0, P_1, \ldots, P_{\omega-1}$, thus demonstrating validity for $k = 1$. Assuming the statement to be true for $k = n - \omega - 1$, we can easily find P_n, so the corollary is valid for $k = n - \omega$. Therefore, by induction, the corollary is valid for all k. □

7. *Token Cycle Time*

It has been proven elsewhere that the mean token cycle time is unaffected by the choice of ω [8]. It is necessary to qualify this, however, by stating that if the network load is outside the region of stability for any ω (that is, a mean queue length exists), then the statement applies only to a choice of ω' such that the network load is within the region of stability for ω'. Given this, we may now state:

THEOREM 7. *Given nondeterministic network operation, the mean number of packets present at token arrival is less than ω if the mean exists.*

PROOF. Given stability, the mean token cycle time is unaffected by the choice of ω (given suitable constraints on the maximum load). It is easily demonstrated that for any ω, the denominator of the equation giving the mean number of packets present at token arrival will contain the factor

$$\omega - (\omega B^{*(1)}(\lambda - \lambda z) + V^{*(1)}(\lambda - \lambda z)) \tag{27}$$

which is multiplicative of all non-zero terms in the denominator. The term

$$\omega B^{*(1)}(\lambda - \lambda z) + V^{*(1)}(\lambda - \lambda z) \tag{28}$$

represents the mean number of packets arriving at a fully utilized station during a full token cycle. For the term in Equation (28) to remain positive, the number of arriving packets during a token cycle must be strictly less than ω. Since the number of arriving packets must be less than ω, a mean queue length of greater than or equal to ω cannot be maintained. Assuming the contrary, if a mean queue length of greater than ω exists, then less than ω packets arrive during each token cycle, but ω packets are served, resulting in more packets being served than are arriving. This, however, implies that the queue length is constantly decreasing, which is inconsistent with the assumption of a mean queue length. Thus the mean queue length at token arrival must be less than ω. □

COROLLARY 3. *Given nondeterministic network operation, the mean number of packets present at token departure is less than ω if the mean exists.*

PROOF. The number of packets present at token arrival is equal to the sum of the number present at token departure plus those which arrive during the token vacation. Therefore, the number present at token departure must be less than or equal to the number present at token arrival. □

8. *Token Vacation Time Second Moment and Variance*

Using the definition of the variance of a random variable, the variance $Var[\Gamma_v]$ of the token vacation time is given by

$$Var[\Gamma_v] = (N-1)^2 \, Var[\Gamma_u] \tag{29}$$

Since the variance is a function of the first and second moments, we have

$$E[\Gamma_v^2] = E[((N-1)\,\Gamma_u + \gamma)^2] \tag{30}$$

where Γ_u is the token hold time random variable. Equation (30) resolves to

$$E[\Gamma_v^2] = (N-1)\left[2\gamma E[\Gamma_u] + (N-1)E[\Gamma_u^2]\right] + \gamma^2 \tag{31}$$

Because of Theorem 7, we may use the proof in [5] that

$$E[\Gamma_u] = \frac{\gamma\rho}{1-N\rho} \tag{32}$$

The only unknown is $E[\Gamma_u^2]$. It is easily shown that

$$E[\Gamma_u^2] = E[\Omega^2]\,Var[U] + E^2[U]\,Var[\Omega] + E^2[\Gamma_u] \tag{33}$$

where Ω is the random variable of the number of packets served, and U is the service time random variable for a single packet.

To find $E[\Omega^2]$, we need $P[\Omega=k \mid R]$ for $k = 1, 2,..., \omega$. Using $\pi_{i \mid j}$ as defined in Equation (9), we have

LEMMA 2.

$$P[\Omega \mid R] = \pi_{\Omega-R \mid R}, \quad \Omega - R > 0$$

$$P[\Omega \mid R] = 1, \quad \Omega - R \le 0 \tag{34}$$

PROOF. If Ω packets are to be served (where $\Omega \le \omega$), and R packets were left behind at the previous token departure, then exactly $\Omega - R$ packets must arrive. However, the probability of $\Omega - R$ packets arriving is $\pi_{\Omega-R \mid R}$. If $R \ge \omega$, then exactly ω packets will be served with probability 1. □

We may now remove the conditioning on the expectation, since the state probabilities are known. Thus $E[\Omega^2]$ is given by

$$E[\Omega^2] = \sum_{k=0}^{\omega} k^2 P[\Omega=k] \tag{35}$$

To compute $Var[\Omega]$ we need $E^2[\Omega]$, but this is equal to the square of the mean number of packets which arrive during a token cycle time, or $\left[\lambda E[\Gamma]\right]^2$. Thus

$$Var[\Omega] = E[\Omega^2] - \left[\lambda E[\Gamma]\right]^2 \tag{36}$$

Thus, after substitution and algebra, we have

$$Var[\Gamma_v] = (N-1)^2\, Var[\Gamma_u] \tag{37}$$

COROLLARY 4. *As* $E[\Omega] \to \omega$, $Var[\Gamma_v] \to (N-1)^2\, \omega^2\, Var[U]$, *and furthermore, if packet lengths are constant, as* $E[\Omega] \to \omega$, $Var[\Gamma_v] \to 0$.

PROOF. The proof follows from substitution into Equation (37). □

9. *State of the System at Token Arrival*

We now have a description of the system at token departure times. However, a description of the system at token *arrival* is also of interest. Since we have the z-transform and state probabilities of the system at token departure, it is possible to derive from this the z-transform of the state of the system at token arrival instants.

Given that $G(z)$ is the unconditional transform of the number of packets left at a station at token departure, the unconditional transform of the number of packets present at token arrival $G_A(z)$ is given by

COROLLARY 5.

$$G_A(z) = G(z)\, V^*(\lambda - \lambda z) \tag{38}$$

PROOF. The unconditional transform of the token vacation time is $V^*(\lambda - \lambda z)$. The number of packets present at token arrival cannot be less than the number present at token departure, since it is equal to the number left behind plus those which arrive during the token vacation, and the number which arrive during a mean token vacation is certainly independent of the number left behind. Since the addition of independent random variables is equivalent to the multiplication of their transforms, we arrive at Equation (44). □

The unconditional state probabilities at token arrival time may be found using Corollary 5. Defining P'_k to be the unconditional probability of k packets present at token arrival, and $P[A_v = k \mid j]$ to be the probability of k arrivals occurring during the token vacation time given token departure state j, we have

$$P'_k = \sum_{j=0}^{k} P_k P[A_v = k-j \mid j] \tag{39}$$

PROOF. Since the number of packets present at token arrival must be greater than or equal to the number present at token departure, we may state the following relationship:

$$P'_0 = P_0 P[A_v = 0 \mid 0]$$

$$P'_1 = P_0 P[A_v = 1 \mid 0] + P_1 P[A_v = 0 \mid 1]$$

$$P'_0 = P_0 P[A_v = 2 \mid 0] + P_1 P[A_v = 1 \mid 1] + P_2 P[A_v = 0 \mid 2] \tag{40}$$

etc. The general form of this relationship is given by Equation (39). □

10. *Mean Waiting Time*

For our purposes, we define the queueing delay of a packet to be the time between its arrival to the queue and the time it begins to receive service, that is, its total time in the queue minus its own service time. It has been proven in [6] that the mean waiting time of a packet $E[W]$ is given by

$$E[W] = E[Q] + \frac{E[\Gamma_v^2]}{2E[\Gamma_v]} \tag{41}$$

where $E[Q]$ is queueing delay, and $E[\Gamma_v^2] \,/\, 2E[\Gamma_v]$ is the time spent waiting for the token to arrive. The time spent waiting for the token can be found by simple substitution. It has been proven in [2] that queueing delay cannot be found using the Pollaczek-Khinchin mean value equation. We must therefore derive our own mean value equation.

We are interested in the mean queueing delay for an arbitrary arriving packet. Defining ξ as the number of packets which arrive ahead of our arbitrary packet, and β as the number of times its transmission is blocked (that is, it cannot be transmitted at the current token reception because at least ω packets precede it), we may state the following theorem:

THEOREM 8.

$$E[Q] = \frac{E[R] + E[\xi]}{\mu} + E[\beta]\, E[\Gamma_v] \tag{42}$$

PROOF. It is evident that an arbitrary arriving packet will experience two sources of queueing delay: the time needed to serve packets left behind at the previous token departure, and the time needed to serve packets which have arrived prior to its arrival. That is,

$$Q = \frac{R + \xi}{\mu} \tag{43}$$

where Q is the total queueing delay, ξ is the number of packet arrivals prior to the arrival of our arbitrary arrival, and R is the number of packets left behind at token departure.

Since $E[R]$ is derived above, the unknown in Equation (45) is ξ. A well-known result from renewal theory states that the expected value of residual life $E[\psi_X]$ of the random variable X is

$$E[\psi_X] = \frac{E[X^2]}{2E[X]} \tag{44}$$

where X is the random variable whose duration is in question. Since the age ϕ_X of a random variable X is merely the total life of X minus its residual life, we have

$$E[\phi_X] = E[X] - E[\psi_X] \tag{45}$$

Using Little's law, this allows us to state that the expected number of arrivals prior to the arrival of our arbitrary arrival is

$$E[\xi] = \lambda\rho\,(E[\phi_u] + E[\Gamma_v]) + \lambda(1-\rho)\,E[\phi_v] \tag{46}$$

where ϕ_u is the age of the token hold time, and ϕ_v is the age of the token vacation time.

LEMMA 3.

$$E[\beta] = \frac{E[R]}{\omega} \tag{47}$$

PROOF. Since it is guaranteed that $E[R] < \omega$, the mean number of packets blocked at any token departure can be transmitted at the next token arrival. Therefore, the mean number of times a packet is blocked is identical to the mean number of packets blocked divided by ω. □

Thus by Lemma 3, the average packet will be further delayed by an amount of time equal to $E[\beta]$ token vacation times, resulting in a total mean queuing delay of

$$E[Q] = \frac{E[R] + E[\xi]}{\mu} + E[\beta]\,E[\Gamma_v] \tag{48}$$

□

We may now find the total expected waiting time from Equation (41). The residual life of the token vacation time can easily be found from Equation (44) and Equation (33), and the residual life of the token hold time can be derived from Equation (44) through Equation (46). The derivation of mean queue length is implicit in the above proof.

11. *Numerical Results*

In this section we show the results of our analysis. This is done by comparing the results from our simulation system to the numerical values produced by the analysis. We show that the analysis has good agreement with the results produced by the simulation system.

The simulation system used to produce the results shown here is a discrete event, continuous time token ring simulator. It assumes error-free operation. Each value on each graph is the result of 10 completely independent experiments (more than 10 experiments is generally infeasible because of time restraints, however, 10 independent experiments generally produce statistically valid results).

The basic operation of the simulator is as follows. All events are prequeued in a timewise fashion using a random number generator to produce interarrival times. Each token arrival constitutes an event at a station. At each token arrival, queue lengths, token cycle times and packet delays are measured. Packets are "transmitted" by resetting pointers and updating time values. A packet is considered to be enqueued at token arrival if its arrival time at that queue is less than or equal to the current time as seen by the token. From these measurements, it is a simple matter to calculate mean values.

The protocol used in our simulations is identical to the IEEE 802.5 token ring protocol [4] running at 16 Mbps, with the exception that a packet counter was substituted for the token hold timer. That is, instead of a station being allowed to hold the token for a certain length of time, it is allowed to hold the token for a certain maximum number of packet transmissions. Whenever this number has been reached or the transmit queue becomes empty, the token departs.

The results shown are taken from equilibrium operation of the simulator (see [7] for the methods used to determine when the simulation has achieved equilibrium). The results from the mathematical analysis are within 95% confidence intervals (see [7]), except for the token departure queue length measurements for both Figures 1 and 2 at the highest loads measured. Token arrival queue lengths are all within 95% confidence intervals.

Figure 2 shows the token arrival and token departure queue lengths as produced both by the simulation and the mathematical analysis. Figure 3 shows identical graphs for a different configuration. Note in both cases the behavior of the predictive model: the residual packets are slightly overestimated at low loads, and slightly underestimated at high loads. Figure 4 shows mean waiting time (queueing delay plus the time spent waiting for the token to arrive).

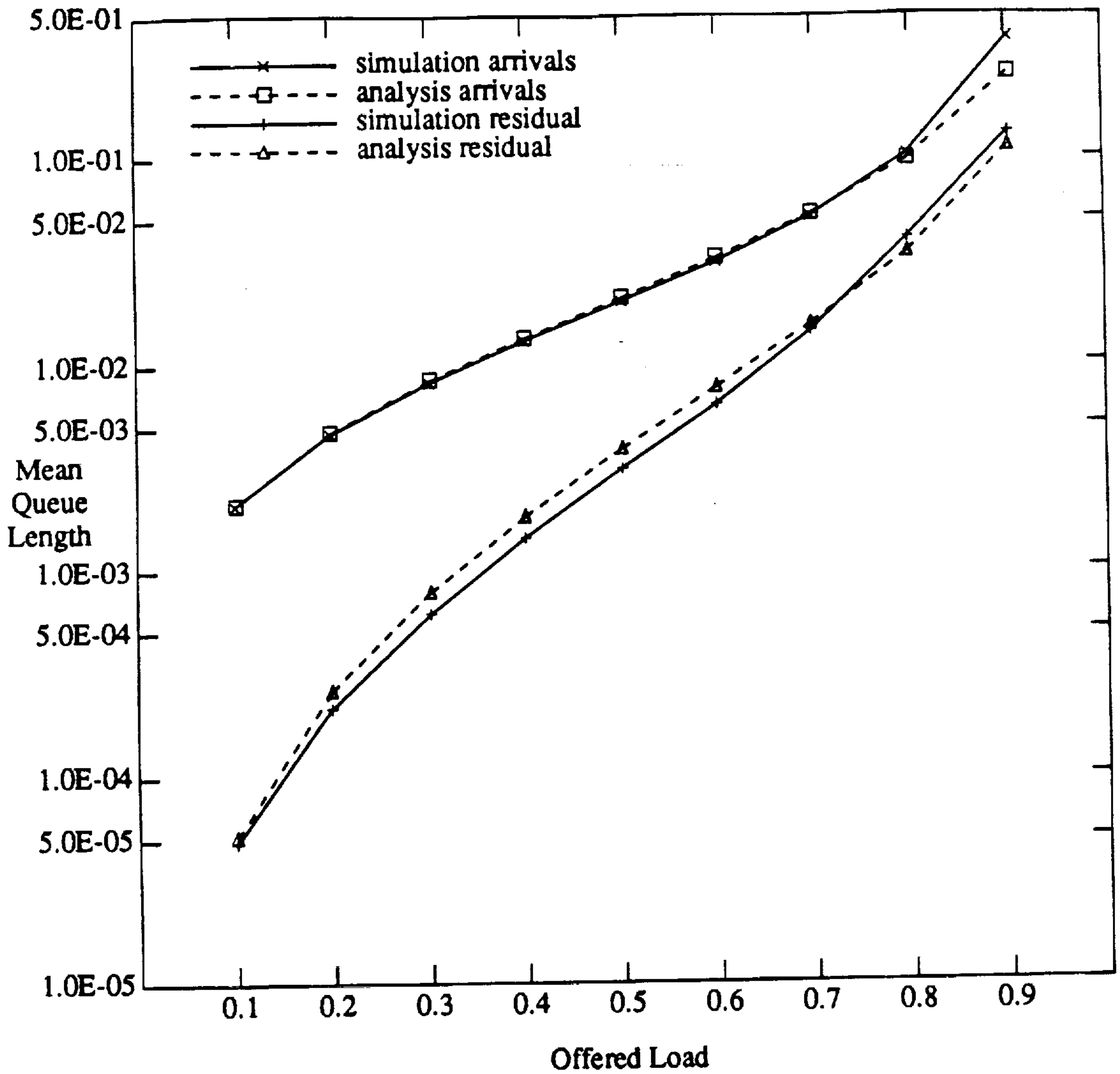

FIG. 2. Mean queue lengths vs. offered load for token arrival and departure instants for 5 stations, 32 byte packets, and $\omega = 1$.

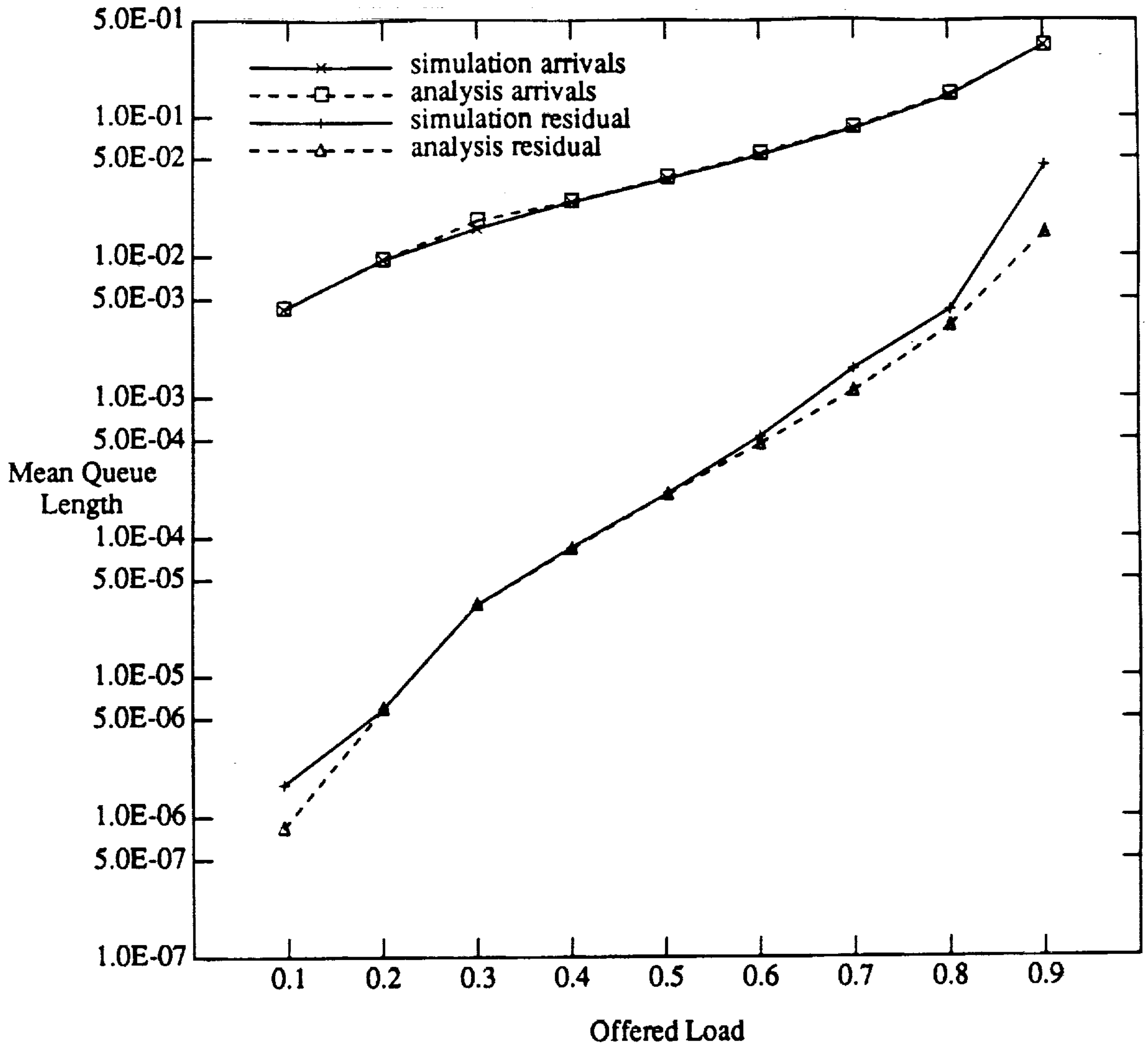

FIG. 3. Mean queue lengths vs. offered load at token arrival and departure instants for 10 stations, 32 byte packets, and $\omega = 2$.

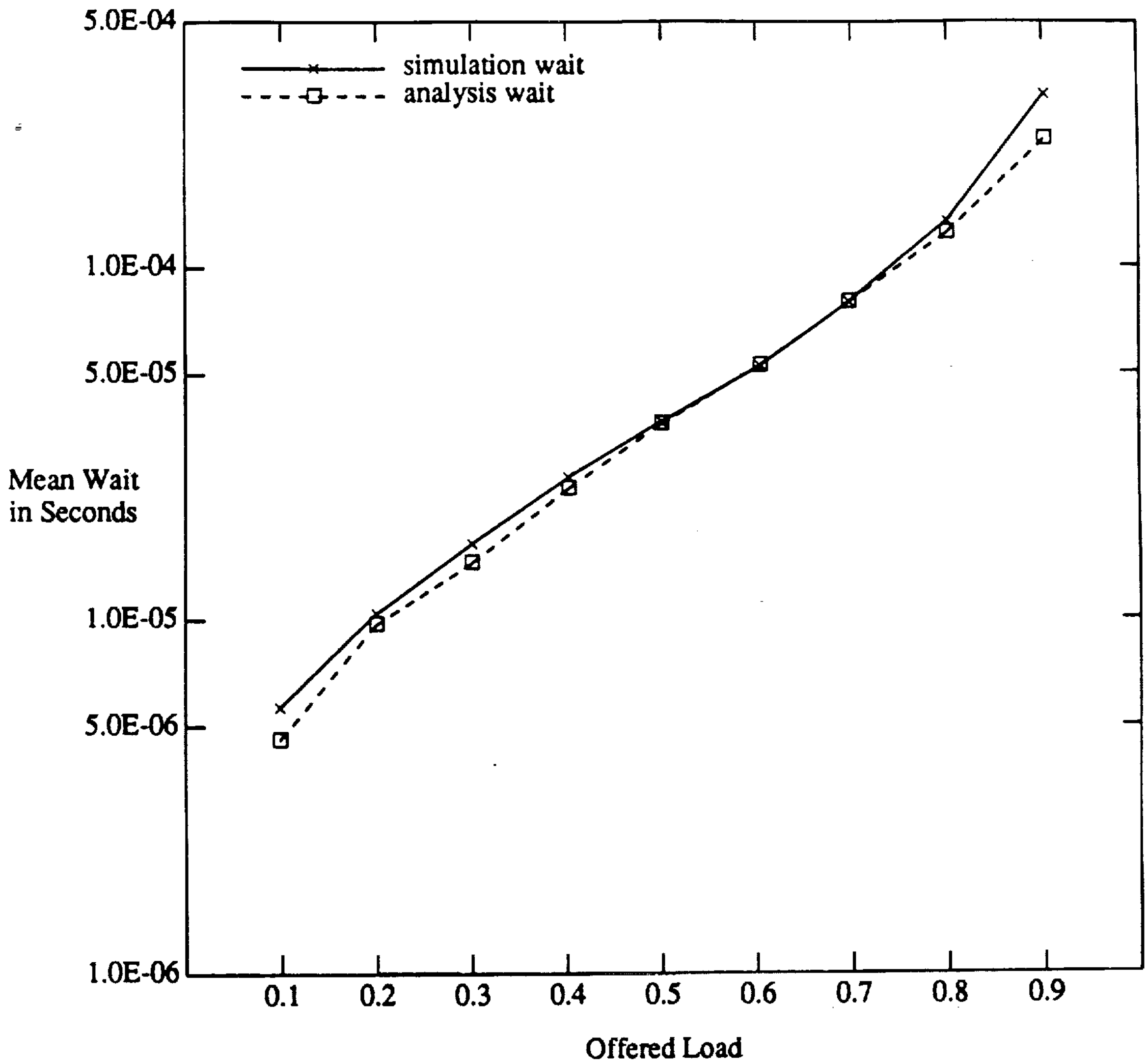

FIG. 4. Mean queueing delay vs. offered load for 10 stations, 128 byte packets, and $\omega = 2$.

The network configuration for all the simulations (with the exception of the number of stations and packet lengths) are identical. We assumed an internal station latency of 6 bit times, a uniform distance between each station of 3.0 meters, and a signal propagation delay of 5.085 nanoseconds per meter. In all cases we assumed Poisson arrivals and constant packet lengths. Packet lengths include all framing.

12. *Conclusions*

In this paper we have derived mathematical models describing the behavior of an exhaustive limited service discipline for token ring networks. The models we have derived are the transforms of the distribution of the number of packets present at token departure and token arrival instants. Our models also predict queue lengths at both token arrival and token departure. We have compared the mathematical results to simulation results, and agreement between the two is close. We have also shown how to find any state probability desired for both token arrival and token departure instants.

The complexity of our model is of order $\omega!$, where ω is the maximum number of packets per token allowed to be transmitted at token reception. The model requires the solution of ω independent linear equations. The complexity of our model, however, is small when compared to the time and memory needed to run a computer simulation of an identical configuration.

The importance of our model is that this service discipline closely approximates protocols which use a token hold timer to limit the token visitation time at a station. Also included in our model is an implicit solution to single packet per token, or ordinary cyclic service — this solution results when ω is set to 1.

ACKNOWLEDGEMENTS

I wish to acknowledge the assistance of my Ph.D. dissertation advisor, Professor Alfred C. Weaver of the Department of Computer Science at the University of Virginia. This work was supported by the Flight Data Systems Branch of NASA-Johnson Space Center, with technical supervision by Mr. William R. Reed and Mr. Frank W. Miller.

REFERENCES

1. FDDI, FDDI Token Ring Media Access Control Standard, American National Standards Institute, draft proposed Standard X3T9.5/84-49, rev. 10, Feb. 1986.

2. W. L. Genter, *Performance Analysis of the Token Passing Bus Network with Priority Classes*, Doctoral Dissertation, Rensselaer Polytechnic Institute, Troy, New York, 1988.

3. IEEE, *IEEE Standards for Local Area Networks: Token Passing Bus Access Method and Physical Layer Specifications*, The Institute of Electrical and Electronics Engineers, Inc., New York, 1985.

4. IEEE, *IEEE Standards for Local Area Networks: Token Ring Access Method and Physical Layer Specifications*, The Institute of Electrical and Electronics Engineers, Inc., New York, 1985.

5. A. G. Konheim and B. W. Meister, Waiting Lines and Times in a System with Polling, *J. ACM 21*, (July 1974), 470-490.

6. Y. Levy and U. Yechiali, Utilization of Idle Time in an M/G/1 Queueing System, *Management Science 22*, 2 (Oct. 1975), 202-211.

7. S. S. Lavenberg, ed., *Computer Performance Modeling Handbook*, Academic Press, New York, New York, 1983.

8. H. Takagi, *Analysis of Polling Systems*, The MIT Press, Cambridge, Massachusetts, 1986.

BIBLIOGRAPHY

J. M. Appleton and M. M. Peterson, Traffic Analysis of a Token Ring PBX, *IEEE Transactions on Communications COM-34*, 5 (May 1986), 417-422.

L. N. Bhuyan, D Ghosal and Q. Yang, Approximate Analysis of Single and Multiple Ring Networks, *IEEE Transactions on Computers 38*, 7 (July 1989), 1027-1040.

D. Everitt, A Note on the Pseudoconservation Laws for Cyclic Service Systems with Limited Service Disciplines, *IEEE Transactions on Communications COM-37*, 7 (July 1989), 781-783.

D. P. Heyman, Data-Transport Analysis of Fasnet, *Bell System Tech. J. 62*, 8 (Oct. 1983), 2547-2560.

L. K. Kleinrock, *Queueing Systems, Volume I: Theory*, John Wiley & Sons, New York, 1975.

P. J. Kuehn, Multiqueue Systems with Nonexhaustive Cyclic Service, *Bell System Tech. J. 58*, 3 (Mar. 1979), 671-698.

H. J. Larson and B. O. Shubert, *Probabilistic Models in Engineering Sciences, Volume I: Random Variables and Stochastic Processes*, John Wiley & Sons, New York, 1979.

V. J. Rego and L. M. Ni, *Performance Modeling of Token-Passing Protocols*, Michigan State University Technical Report, 1984.

N. Schult, *Analytic Models for Token-Ring Networks*, Ph.D. Dissertation, University of Virginia, Charlottesville, Virginia, May, 1989.

H. Takagi and M. Murata, Queueing Analysis of Nonpreemptive Reservation Priority Discipline, *Proceedings of Performance '86 and ACM Sigmetrics (Joint Conference)*, Raleigh, NC, May 27-30, 1986, 237-241.