PSEUDO-CONSERVATION LAWS IN CYCLIC-SERVER, MULTIQUEUE SYSTEMS WITH A CLASS OF LIMITED SERVICE POLICIES

Kow Chang and Deniz Sandhu

Electrical, Computer and Systems Engineering Department
Rensselaer Polytechnic Institute, Troy, NY 12180.

ABSTRACT

This paper considers a multiqueue system with a cyclic-server and a class of limited service policies. In particular, exhaustive limited (EL), gated limited (GL) and general decrementing (GD) service policies are investigated. The major results in this paper are the derivations of the expected amount of work left in the queue at the server departures for these three policies. These results are all in terms of unknown boundary probabilities. Using corresponding server vacation models, these unknown probabilities are estimated. Pseudo-conservation laws for these policies are subsequently obtained. Numerical results obtained for the EL policy are noted to be very accurate compared with simulation results. Finally, conservation laws with mixed service policies, and exact expressions of mean waiting times in symmetric systems are given.

1. INTRODUCTION

Token-passing protocols in local-area networks, using bus or ring topologies, are often modeled as Cyclic-Server, Multi-Queue systems (CSMQs) for performance evaluation. Bux uses a discrete-time model of CSMQ to compare the performances of single-token and multiple-token operations in a symmetric token ring [4]. Colvin [7] and Sach [28] use two different analytic models of CSMQs to obtain approximate mean waiting times for a single access class in the IEEE 802.4 token bus. Brooks and Yue use an approximate model of a symmetric CSMQ to investigate the impact of the token holding time constraint on the performance of the Automation Manufacturing Protocol (MAP) [3,35]. Karvelas applies a CSMQ model to obtain the average data delay in the Fiber Distributed Data Interface (FDDI) network [22].

1.1 Cyclic-Server, Multiqueue System

A description of CSMQ, with one-to-one correspondence to the parameters in token-passing protocols, is given as follows. A CSMQ consists of N queues: $Q_1,Q_2,\ldots,Q_N$ (network stations) with infinite capacities (buffers) and a single server (token). Customers (messages) arrive at the queues according to independent Poisson processes with rate λ_i for Q_i. The single server visits the queues in a predetermined, cyclic order. When the server visits Q_i, the customers are served on a first-come-first-served basis and the service times (transmission times) are assumed to be generally distributed with first moment b_i and second moment $b_i^{(2)}$. The server utilization of Q_i is denoted as ρ_i $(\lambda_i b_i)$ and the total server utilization ρ $(\sum_{i=1}^{N}\rho_i)$. According to a certain service policy (medium access control protocol), the server switches from Q_i to the next queue with a non-zero *walk time* (propagation delay) which is assumed to be generally distributed with first moment s_i and second moment $s_i^{(2)}$. Let the first moment of the total walk time be s $(\sum_{i=1}^{N}s_i)$ and its second moment $s^{(2)}$. We assume that the arrival, service and switch-over processes are independent.

With the same system parameters, the CSMQs differ from one another in the service policies. The service policy determines *when* the server will switch from one queue to another. The earlier analyses of CSMQs concentrated on the *exhaustive* (E), *gated* (G), *decrementing* (D), and *non-exhaustive* (NE) service policies [8,25]. In the E policy, the server switches only when the queue is empty. In the G policy, the server switches after serving all the messages in the queue upon his arrival. Those messages which arrive after the server arrival will be served in the next server visit. In the D policy, the server switches when the queue length decreases to one less than the queue length at the server arrivals. In the NE policy, the server switches after one message is served. However, the recent standards for token-passing protocols impose limits on the token holding times [14,21,26]. In order to provide more accurate performance models for the protocols, more flexible limited service policies need to be investigated.

1.2 A Class of Limited Service Policies

Here we consider a class of limited service policies for the CSMQs. These policies impose limits on *the number of messages* to be served consecutively in any queue, by varying parameter K_i for Q_i. If the server finds Q_i empty upon his arrival, he will immediately switch to the next queue. Otherwise the server will act as follows, depending on the service policy:

1. Exhaustive limited (EL): the server switches when either K_i messages are served consecutively in Q_i or the queue is empty, whichever occurs first.
2. Gated limited (GL): let L_i be the queue length of Q_i upon the server arrival. The server switches after serving $\min(K_i,L_i)$ messages in Q_i.
3. General decrementing (GD): the server switches when the queue length of Q_i decreases to either (L_i-K_i) or zero, whichever occurs first.

Note that these three policies cover all other service policies described previously. Furthermore we also allow mixed policy in the CSMQs. For instance, Q_1 may employ the EL policy while other queues employ the G policy. Nevertheless, we assume in this paper that the same service policy is applied to all queues, unless otherwise stated. We adopt the following notation in the rest of this paper: $E(N_i)$: mean queue length of Q_i (including the one in service); $E(W_i)$: mean message waiting time for Q_i (not including the service time); $E(M_i)$: mean queue length of Q_i *when* the server leaves; and $E(U_i)$: mean amount of work left in Q_i *when* the server leaves.

1.3 Pseudo-Conservation Laws

In general, queueing analysis of CSMQs is a difficult problem. For the E and G policies, numerical algorithms are available to compute the exact mean waiting times by solving $O(N^2)$ linear equations [15]. At the same time, others also suggest to use simpler approximate models [5,10,27]. The queueing analyses for other service policies are even more difficult. Exact mean waiting times are not available even for symmetric systems, i.e. system parameters are independent of i, except for the NE and D policies [19,31,32]. For asymmetric systems, a number of approximate models are recently proposed for the NE policy [1,20,30], and there are few heuristic approximations suggested for the EL and GL policies [11,12,17]. On the other hand, Watson has derived simple analytic expressions for a *weighed sum* of mean waiting times at individual queues in the CSMQs for the E, G, and NE policies [34]. Later, Boxma and Groenendijk have generalized these results into the following form and they are referred to as pseudo-conservation laws [2]:

$$\sum_{i=1}^{N}\rho_i E(W_i) = A+\sum_{i=1}^{N}E(U_i) \tag{1.1}$$

where

$$A = \rho\frac{\sum_{i=1}^{N}\lambda_i b_i^{(2)}}{2(1-\rho)}+\rho\frac{s^{(2)}}{2s}+\frac{s}{2(1-\rho)}\left(\rho^2-\sum_{i=1}^{N}\rho_i^2\right)$$

Note that A is independent of the service policies, and only $E(U_i)$ reflects the differences in the service policies. Clearly, $E(U_i)=0$ for the E policy. Using probabilistic arguments and Little's law, Boxma and Groenendijk have derived $E(U_i)$ for the G, NE, and D policies [2]:

$$E(U_i) = \frac{s}{1-\rho}\rho_i^2 \quad \text{G} \tag{1.2}$$

$$E(U_i) = \rho_i\frac{\lambda_i s}{1-\rho}E(W_i)+\frac{s}{1-\rho}\rho_i^2 \quad \text{NE} \tag{1.3}$$

$$E(U_i) = \rho_i\frac{\lambda_i s(1-\rho_i)}{1-\rho}E(W_i)-\rho_i\frac{s\lambda_i^2 b_i^{(2)}}{2(1-\rho)} \quad \text{D} \tag{1.4}$$

The corresponding pseudo-conservation laws can then be obtained by substituting $E(U_i)$s into (1.1), and the results for the E, G, NE policies agree with the results in [15,34]. When the walk times are zero, the first term in A is the only non-zero term, thus reducing to the *Kleinrock's law of conservation for priority queues* [23].

These pseudo-conservation laws are proved to be very useful in understanding the CSMQs. First, the conservation laws give exact expressions of mean waiting times in symmetric systems. Second, they are the bases for a number of simple and yet accurate approximate models of the asymmetric CSMQs: E and G policies [10]; NE policy [1,30]; EL and GL policies [11,12,17]. Third, they provide validity check for the accuracy of simulation models [30]. Fourth, they can be used for asymptotic analyses when queues are heavily loaded or the number of queues in the system becomes very large [34]. Nevertheless, the derivations of $E(U_i)$ for the EL, GL, and GD policies do not seem to be as straightforward as others. Fuhrmann derives upper bounds on $E(U_i)$ for both EL and GL policies [16]. By modifying the derivation in [16], Servi and Yao obtain an lower bound on $E(U_i)$ for the EL policy [29]. In this paper, we derive exact expressions of $E(U_i)$ for the EL, GL, and GD policies. It turns out that the expressions of $E(U_i)$ for these three policies are all in terms of K_i unknown boundary probabilities for Q_i. These probabilities can then be estimated by using the results in the corresponding *server vacation models*. We recently discover that Everitt has also derived pseudo-conservation laws for the EL and GL policies. The expressions are in terms of the second moments of the number of messages served during a server visit. He then estimates the second moment from a truncated negative binomial distribution [9,13].

The rest of this paper is organized as follows. In section 2, we present some basic results of CSMQs and stability conditions for these policies. Section 3 presents the derivation of $E(U_i)$ for the EL policy. The pseudo-conservation laws using the estimated probabilities are shown to be very accurate when compared with simulation results. An assessment of the performance of the upper and lower bounds on the weighed sum of mean waiting times is also given. Section 4 presents the derivation of $E(U_i)$ for the GL and GD policies. In section 5, we give the pseudo-conservation law for mixed service policies, and the exact mean waiting times in symmetric systems. Finally we conclude this paper in section 6, with a discussion on future research.

2. STABILITY CRITERIA

In this paper, we consider stable CSMQs in which all queue lengths are finite. We first define some important parameters for the CSMQs. *Cycle time* of Q_i is the time between two successive visits at Q_i by the server. It is well known that the mean cycle time, E(C), is independent of i and is given by [24]

$$E(C) = \frac{s}{1-\rho} \tag{2.1}$$

Visit period of Q_i, T_i, is the total time that the server spends at Q_i during each visit. Using a balancing argument, the mean visit period of Q_i is also equal to the total mean amount of work arriving during a cycle. Therefore

$$E(T_i) = b_i\frac{\lambda_i s}{1-\rho} = \rho_i\frac{s}{1-\rho} \tag{2.2}$$

Intervisit time of Q_i, V_i, is the period of time between the server arrival at Q_i and his last departure from Q_i. Since a cycle time is the sum of a visit period and an intervisit time, $E(V_i)$ can be obtained as

$$E(C) = E(V_i)+\rho_i\frac{s}{1-\rho} \Rightarrow E(V_i) = \frac{1-\rho_i}{1-\rho}s \tag{2.3}$$

For the E and G policies, $\rho<1$ is the only condition to ensure stability. However, $\rho<1$ is only a necessary condition for the limited service policies, and additional conditions are required for stability. For the EL and GL policies, we need the following additional conditions:

$$\lambda_i\frac{s}{1-\rho}<K_i \Rightarrow 0<1-\frac{\lambda_i s}{K_i(1-\rho)}<1 \qquad i=1,\dots,N \tag{2.4}$$

This reflects the fact that the mean number of message arrivals at Q_i during a cycle should be less than K_i, since the number of messages is reduced at most by K_i during a cycle. For the GD policy, the additional conditions are:

$$\lambda_i \frac{1-\rho_i}{1-\rho} s < K_i \Rightarrow 0 < 1 - \frac{\lambda_i s(1-\rho_i)}{K_i(1-\rho)} < 1 \quad i=1,\ldots,N \tag{2.5}$$

Here the system is stable if the mean number of message arrivals at Q_i during an intervisit time is less than K_i, since the number of messages at the server arrivals is at most reduced by K_i at the end of a visit period.

3. EXHAUSTIVE LIMITED SERVICE POLICY

Lemma 1:

Consider a stable CSMQ system where Q_i employs exhaustive limited service policy with parameter K_i, then

$$E(U_i) = \rho_i \frac{\lambda_i s}{K_i(1-\rho)} E(W_i) + \rho_i \frac{s}{2K_i(1-\rho)} \times \left\{ (1-\rho_i) \sum_{j=1}^{K_i-1} j(K_i-j) p_{i,0,j} - [K_i - 1 - (K_i+1)\rho_i] \right\}$$

where $p_{i,0,j}$, $j=1,\ldots,K_i-1$ are the joint probabilities that at the instant of a message departure, the message is the jth message served during a visit period and Q_i is empty.

Proof:

We observe Q_i at the service completion epoches. Denote a message at Q_i as a j-message if it is the jth message served by the server during a visit. Let $p_{i,k,j}$ be the joint probability that at the instant of a message departure, the message is a j-message and Q_i is empty. Therefore

$$\sum_{k=0}^{\infty} \sum_{j=1}^{K_i} p_{i,k,j} = 1 \tag{3.1}$$

The mean queue length of Q_i is hence given by

$$E(N_i) = \sum_{j=1}^{K_i} \sum_{k=1}^{\infty} k p_{i,k,j} \equiv \sum_{j=1}^{K_i} \Phi_{i,j} \tag{3.2}$$

where $\Phi_{i,j}$ can be interpreted as the mean queue length of Q_i when a j-message leaves the queue. From the structure of the EL policy, we note that the server will perform a jth service if Q_i is non-empty at the departure of a (j-1)-message, for $j=2,\ldots,K_i$. Therefore

$$\sum_{k=1}^{\infty} p_{i,k,j-1} = \sum_{k=0}^{\infty} p_{i,k,j} \quad j=2,\ldots,K_i \tag{3.3a}$$

Let $\phi_{i,j} \equiv \sum_{k=0}^{\infty} p_{i,k,j}$ for $j=1,\ldots,K_i$ and rewrite (3.3a) as

$$\phi_{i,j-1} - p_{i,0,j-1} = \phi_{i,j} \quad j=2,\ldots,K_i \tag{3.3b}$$

From (3.3b), we can therefore express $\phi_{i,j}$ in terms of ϕ_{i,K_i}, for $j=1,\ldots,K_i-1$:

$$\phi_{i,j} = \phi_{i,K_i} + \sum_{r=j}^{K_i-1} p_{i,0,r} \quad j=1,\ldots,K_i-1 \tag{3.4}$$

Also by adding ϕ_{i,K_i} with $\phi_{i,j}$ in (3.4), for $j=1,\ldots,K_i-1$, we obtain an equality for this system:

$$\sum_{j=1}^{K_i-1} j p_{i,0,j} + \sum_{k=0}^{\infty} K_i p_{i,k,K_i} = 1 \tag{3.5}$$

Multiply both sides in (3.5) by $\lambda_i s/(1-\rho)$:

$$\sum_{j=1}^{K_i-1} j \frac{\lambda_i s}{1-\rho} p_{i,0,j} + \sum_{k=0}^{\infty} K_i \frac{\lambda_i s}{1-\rho} p_{i,k,K_i} = \frac{\lambda_i s}{1-\rho} \tag{3.6}$$

The right hand side of (3.6) is the mean number of message arrivals at Q_i during a cycle. Therefore, for a stable CSMQ, the left hand side of (3.6) should be equal to the mean number of messages served during a server cycle time. This implies that $[\lambda_i s/(1-\rho)]p_{i,0,j}$ is the conditional probability that Q_i is empty and j messages have been served consecutively *when* the server departs, for $j=1,2,\ldots,K_i-1$. Similarly, $[\lambda_i s/(1-\rho)]p_{i,k,K_i}$ is the conditional probability that the queue length of Q_i is k and K_i messages have been served consecutively *when* the server departs, for $k=0,1,2,\ldots$. Thus the mean queue length of Q_i when the server leaves is given by

$$E(M_i) = \sum_{k=1}^{\infty} k \frac{\lambda_i s}{1-\rho} p_{i,k,K_i} \equiv \frac{\lambda_i s}{1-\rho} \Phi_{i,K_i} \tag{3.7}$$

and the mean amount of unfinished work in Q_i when the server leaves is

$$E(U_i) = b_i E(M_i) = \frac{\rho_i s}{1-\rho} \Phi_{i,K_i} \tag{3.8}$$

The remaining step is therefore to obtain an exact expression of Φ_{i,K_i}. Again from the structure of the exhaustive limited policy. If the queue length of Q_i at the departure of a (j-1)-message is non-zero, then the queue length at the departure of a j-message is equal to the queue length at the departure of a (j-1)-message, minus one, and plus the total number of message arrivals during a service time. Thus

$$\Phi_{i,j} = \Phi_{i,j-1} - (1-\rho_i) \sum_{k=1}^{\infty} p_{i,k,j-1} \quad j=2,\ldots,K_i \tag{3.9}$$

From (3.9), we can therefore express $\Phi_{i,j}$ in terms of Φ_{i,K_i}, for $j=1,\ldots,K_i-1$:

$$\Phi_{i,j} = \Phi_{i,K_i} + (1-\rho_i) \sum_{k=1}^{\infty} \sum_{r=j}^{K_i-1} p_{i,k,r} \quad j=1,\ldots,K_i-1 \tag{3.10}$$

Substituting (3.10) into (3.2), for $j=1,\ldots,K_i-1$, we obtain

$$E(N_i) = K_i \Phi_{i,K_i} + (1-\rho_i) \sum_{k=1}^{\infty} \sum_{j=1}^{K_i-1} j p_{i,k,j} \tag{3.11}$$

Using (3.8) in (3.11) yields

$$\frac{\rho_i s}{1-\rho} E(N_i) = K_i E(U_i) + \frac{\rho_i (1-\rho_i) s}{1-\rho} \sum_{k=1}^{\infty} \sum_{j=1}^{K_i-1} j p_{i,k,j} \tag{3.12}$$

Using Little's law: $E(N_i)=\lambda_i[E(W_i)+b_i]$ in (3.12) and rearrange the terms, we obtain

$$E(U_i) = \frac{\rho_i s}{K_i(1-\rho)} \left[\lambda_i E(W_i) + \rho_i - (1-\rho_i) \sum_{k=1}^{\infty} \sum_{j=1}^{K_i-1} j p_{i,k,j} \right] \tag{3.13}$$

By interchanging the summations in $\sum_{k=1}^{\infty} \sum_{j=1}^{K_i-1} j p_{i,k,j}$, we get

$$\sum_{k=1}^{\infty} \sum_{j=1}^{K_i-1} j p_{i,k,j} = \phi_{i,1} + \ldots + (K_i-1)\phi_{i,K_i-1} - \sum_{j=1}^{K_i-1} j p_{i,0,j} \tag{3.14}$$

Using (3.4) to express $\phi_{i,j}$ in terms of ϕ_{i,K_i}, for $j=1,\ldots,K_i-1$

and (3.5) to express ϕ_{i,K_i} in terms of $p_{i,0,j}$, $j=1,...,K_i-1$:

$$\sum_{k=1}^{\infty}\sum_{j=1}^{K_i-1} j p_{i,k,j} = \frac{K_i-1}{2} - \frac{1}{2}\sum_{j=1}^{K_i-1} j(K_i-j)p_{i,0,j} \quad (3.15)$$

Lemma 1 is thus proved by substituting (3.15) into (3.13).§

Using Lemma 1 and (1.1), the pseudo-conservation law for the EL policy is

$$\sum_{i=1}^{N}\rho_i\left(1-\frac{\lambda_i s}{K_i(1-\rho)}\right)E(W_i) = A + \frac{s}{1-\rho}\sum_{i=1}^{N}\frac{\rho_i}{2K_i}\times$$
$$\left\{(1-\rho_i)\sum_{j=1}^{K_i-1} j(K_i-j)p_{i,0,j} - [K_i-1-(K_i+1)\rho_i]\right\} \quad (3.16)$$

Note that the last term in (3.13) is positive. The upper bound in [16] can therefore be obtained by ignoring the last term. It can also be shown that $[\lambda_i s/(1-\rho)]\sum_{k=1}^{\infty})p_{i,k,j}<1$, for $j=1,...,K_i-1$. Therefore, the lower bound in [29] can be obtained by setting that term to 1 in (3.14). Therefore

$$\sum_{i=1}^{N}\rho_i\left(1-\frac{\lambda_i s}{K_i(1-\rho)}\right)E(W_i) \le A + \frac{s}{1-\rho}\sum_{i=1}^{N}\frac{\rho_i^2}{K_i} \equiv UB \quad (3.17)$$

$$\sum_{i=1}^{N}\rho_i\left(1-\frac{\lambda_i s}{K_i(1-\rho)}\right)E(W_i) \ge UB - \sum_{i=1}^{N}\frac{(K_i-1)(1-\rho_i)b_i}{2} \quad (3.18)$$

If $K_i=\infty$ for all i, (3.16) reduces to the result for the E policy. If $K_i=1$ for all i, (3.16) reduces to the result for the NE policy. For other cases, there are still K_i-1 unknown boundary probabilities for Q_i, $i=1,...,N$. Here we propose to use the *M/G/1 queue with vacations and EL policy* to estimate these probabilities. In the setting of the vacation model, the server intervisit time of Q_i can be interpreted as the server's vacation from the queue. However, the successive vacation periods in the vacation model are assumed to be independent, but this is not true for the server intervisit times in the CSMQs. By adding subscripts i to the parameters in the vacation model for Q_i and assuming exponentially distributed vacation periods, then the resulting vacation model can readily be used to estimate the boundary probabilities. The estimation of the probabilities for Q_i involves the solutions of a K_ith order polynomial equation and a set of K_i linear, independent equations [6], [33]. Now we give numerical examples of the conservation laws for the EL policy. We first define the percentage error of an approximation as

$$\%\ \text{Error} = \frac{\text{Simulation result} - \text{Approximation}}{\text{Simulation result}} \times 100\% \quad (3.19)$$

Tables 1-3 present Simulation (results), Estimate: (3.16) with estimated probabilities, Upper Bound: (3.17) and Lower Bound: (3.18). The percentage errors are given in parentheses. We note that the errors in the column of Estimate are less than 2% for most of the cases. Both bounds are exact for $K_i=1$, and the upper bound is also exact for $K_i=\infty$. For other values of K_i, the upper bound is reasonably tight (within 10%). However, it is reported that the performance of the upper bound begins to deteriorate when the walk times increase [16]. This result is confirmed in Table 2, and the percentage errors can be as high as 40%. On the other hand, the performance of the lower bound is quite poor, and it sometimes give negative values in which cases we assume zero for the bound.

4. GATED LIMITED AND G-DECREMENTING SERVICE POLICIES

Lemma 2:

Consider a stable CSMQ system where Q_i employs gated limited service policy with parameter K_i, then

$$E(U_i) = \rho_i\frac{\lambda_i s}{K_i(1-\rho)}E(W_i) + \frac{\rho_i^2 s}{1-\rho}$$
$$-\frac{b_i(1+\rho_i)\left\{\sum_{k=2}^{K_i-1}k(k-1)q_{i,k} + K_i(K_i-1)[1-\sum_{k=0}^{K_i-1}q_{i,k}]\right\}}{2K_i}$$

where $q_{i,k}$ is the steady-state probability that the queue length of Q_i is k at the server arrivals, for $k=0,1,...,K_i-1$.

Proof:

Let $q_{i,k}$ be the steady-state probability that the queue length of Q_i is k *when* the server arrives. Therefore

$$\sum_{k=0}^{\infty}q_{i,k} = 1 \quad (4.1)$$

According to the policy, the probabilities that k messages are served in a visit period is also equal to $q_{i,k}$ for $k=0, 1,...,K_i-1$. Thus, the probability that K_i messages are served in a visit period is $(1-\sum_{k=0}^{K_i-1}q_{i,k})$. As a result, we obtain an expression for the mean number of messages served in a visit period which should also equal to the mean number of message arrivals during a cycle:

$$\sum_{k=1}^{K_i-1}kq_{i,k} + K_i(1-\sum_{k=0}^{K_i-1}q_{i,k}) = \frac{\lambda_i s}{1-\rho} \quad (4.2)$$

The mean queue length of Q_i at the server arrivals is equal to the sum of the mean queue length at the server departures and the mean number of message arrivals during an intervisit time. Thus we have

$$E(M_i) = \sum_{k=1}^{\infty}kq_{i,k} - \lambda_i\frac{(1-\rho_i)s}{1-\rho} \quad (4.3)$$

and the mean amount of work left in Q_i when the server leaves:

$$E(U_i) = b_i\sum_{k=1}^{\infty}kq_{i,k} - \rho_i\frac{(1-\rho_i)s}{1-\rho} \quad (4.4)$$

Recall that a j-message is the jth message served by the server during a visit. Likewise, we denote a visit period of k services as *k-visit period*. Next we let $A_{i,k}$ be the events that a random message in Q_i is served in a k-visit period, $k=1,...,K_i$; $B_{i,j}$ the events that a random message in Q_i is a j-message, $j=1,...,K_i$; and $F_{i,k}$ the events that the queue length of Q_i at the server arrival is k, $k=0,1,2....$ Conditioned on the events $A_{i,k}$, the queue length of Q_i when a random message departs is:

$$E(N_i) = \sum_{k=1}^{K_i-1}E(N_i|A_{i,k})Pr(A_{i,k})$$
$$+\sum_{k=K_i}^{\infty}E(N_i|A_{i,K_i}\cap F_{i,k})Pr(A_{i,K_i}\cap F_{i,k}) \quad (4.5)$$

Given the probabilities $q_{i,k}$, the lengths of visit periods are independent, identically distributed, and we can therefore use the standard length biasing argument in renewal theory for $Pr(A_{i,k})$, $k=1,...,K_i-1$:

$$\Pr(A_{i,k}) = \frac{kq_{i,k}}{\sum_{k=1}^{K_i-1} kq_{i,k} + K_i \sum_{k=K_i}^{\infty} q_{i,k}} = \frac{kq_{i,k}}{\lambda_i s/(1-\rho)} \quad (4.6)$$

Using the same argument for $k \geq K_i$:

$$\Pr(A_{i,K_i} \cap F_{i,k}) = \frac{K_i q_{i,k}}{\lambda_i s/(1-\rho)} \quad (4.7)$$

For $E(N_i|A_{i,k})$, $k=1,\ldots,K_i-1$, we change the conditioning to $A_{i,k} \cap B_{i,j}$. Further, given a visit period of k services, it is equal likely that a random message comes from any one of the k services. Therefore

$$E(N_i|A_{i,k}) = \sum_{j=1}^{k} \frac{1}{k} E(N_i|A_{i,k} \cap B_{i,j}) \quad (4.8)$$

Similarly for $E(N_i|A_{i,K_i} \cap F_{i,k})$, $k \geq K_i$:

$$E(N_i|A_{i,K_i} \cap F_{i,k}) = \sum_{j=1}^{K_i} \frac{1}{K_i} E(N_i|A_{i,K_i} \cap B_{i,j} \cap F_{i,k}) \quad (4.9)$$

Combining (4.5) to (4.9):

$$\frac{\lambda_i s}{1-\rho} E(N_i) = \sum_{k=1}^{K_i-1} q_{i,k} \sum_{j=1}^{k} (k-j+\rho_i j) + \sum_{k=K_i}^{\infty} q_{i,k} \sum_{j=1}^{K_i} (k-j+\rho_i j) \quad (4.10)$$

From (4.10), we can express the mean queue length at the server arrivals in terms of other parameters:

$$\sum_{k=1}^{\infty} kq_{i,k} = \frac{\lambda_i s}{K_i(1-\rho)} E(N_i) - \frac{1+\rho_i}{2K_i} \sum_{k=2}^{K_i-1} k(k-1) q_{i,k} + \left(1-\frac{\rho_i}{K_i}\right) \sum_{k=1}^{K_i-1} kq_{i,k} + \frac{(1-\rho_i)(1+K_i)}{2} \left(1-\sum_{k=0}^{K_i-1} q_{i,k}\right) \quad (4.11)$$

Using Little's law, and (4.2) in (4.11):

$$\sum_{k=1}^{\infty} kq_{i,k} = \frac{\lambda_i s}{K_i(1-\rho)} [\lambda_i E(W_i) + \rho_i] + \frac{\lambda_i s}{1-\rho}\left(1-\frac{\rho_i}{K_i}\right) - \frac{(1+\rho_i)\left[\sum_{k=2}^{K_i-1} k(k-1) q_{i,k} + K_i(K_i-1)\left(1-\sum_{k=0}^{K_i-1} q_{i,k}\right)\right]}{2K_i} \quad (4.12)$$

Lemma 2 is thus proved by substituting (4.12) into (4.4).§

From Lemma 2 and (1.1), the pseudo-conservation law for the GL policy is

$$\sum_{i=1}^{N} \rho_i \left(1-\frac{\lambda_i s}{K_i(1-\rho)}\right) E(W_i) = A + \sum_{i=1}^{N} \left\{ \frac{\rho_i^2 s}{1-\rho} - \frac{b_i(1+\rho_i)}{2K_i} \mathrm{x} \left[\sum_{k=2}^{K_i-1} k(k-1) q_{i,k} + K_i(K_i-1)\left(1-\sum_{k=0}^{K_i-1} q_{i,k}\right)\right]\right\} \quad (4.13)$$

If $K_i=\infty$ for all i, (4.13) reduces to the result for the G policy. If $K_i=1$ for all i, (4.13) corresponds to the result for the NE policy. Similar to the EL policy, the K_i unknown boundary probabilities can be estimated from the corresponding vacation model. This again involves solving of a K_ith order polynomial equation and a set of K_i linear, independent equations [18], [33]. Note that the last term of $E(U_i)$ in Lemma 2 is non-negative, and the upper bound in [16] can therefore be obtained by ignoring the last term:

$$\sum_{i=1}^{N} \rho_i \left(1-\frac{\lambda_i s}{K_i(1-\rho)}\right) \leq A + \frac{s}{1-\rho} \sum_{i=1}^{N} \rho_i^2 \quad (4.14)$$

This bound in (4.14) is exact for both $K_i=1$ and $K_i=\infty$. According to the results in [16], the upper bound on the weighed sum of mean waiting times for the GL policy is less accurate than the upper bound for the EL policy in (3.17).

Lemma 3:

Consider a stable CSMQ system where Q_i employs general decrementing service policy with parameter K_i, then

$$E(U_i) = \rho_i \frac{\lambda_i s(1-\rho_i)}{K_i(1-\rho)} E(W_i) - \rho_i \frac{s\lambda_i^2 b_i^{(2)}}{2K_i(1-\rho)} - \frac{b_i\left\{\sum_{k=2}^{K_i-1} k(k-1) q_{i,k} + K_i(K_i-1)\left[1-\sum_{k=0}^{K_i-1} q_{i,k}\right]\right\}}{2K_i}$$

where $q_{i,k}$ is the steady-state probability that the queue length of Q_i is k at the server arrival, for $k=0,1,\ldots,K_i-1$.

Proof:

$q_{i,k}$ is again defined as the steady-state probability that the queue length of Q_i at the server arrivals is k, $k=0,1,2\ldots.$ According to this policy and by changing the order of service, the length of the visit period will consist of k busy periods of a standard M/G/1 queue (where each of them is initiated by a single service) if the server finds $k<K_i$ messages upon his arrival. For the case of $k \geq K_i$, the length of the visit period will consist of K_i busy periods. Therefore, we can obtain an expression of the mean number of messages served in a visit period which is also equal to the mean number of message arrivals during a cycle:

$$\sum_{k=1}^{K_i-1} k \frac{1}{1-\rho_i} q_{i,k} + \sum_{k=K_i}^{\infty} K_i \frac{1}{1-\rho_i} q_{i,k} = \frac{\lambda_i s}{1-\rho} \quad (4.15)$$

where $1/(1-\rho_i)$ is the mean number of messages served during a busy period, initiated by a message. And the mean queue length of Q_i when the server leaves is

$$E(M_i) = \sum_{k=K_i+1}^{\infty} (k-K_i) q_{i,k} \quad (4.16)$$

By using (4.15) in (4.16), we again obtain the same expression for $E(U_i)$ as in (4.4):

$$E(U_i) = b_i \sum_{k=1}^{\infty} kq_{i,k} - \rho_i \frac{(1-\rho_i)s}{1-\rho} \quad (4.17)$$

Here we follow similar procedures as in the proof of Lemma 2. But we first modify the definitions of the k-visit period and j-message. We redefine k-visit period to be a visit period that consists of k *busy periods* (instead of k services) and a j-message to be the one served in a jth *busy period* (instead of a jth service). The events $A_{i,k}$ and $B_{i,j}$ are also redefined according to the these new definitions. We start with (4.5) and note that the expressions of $E(N_i \mid A_{i,k})$ for $k=1,\ldots,K_i$ in (4.8) and (4.9) also apply here. That is, given k busy periods in a visit period, it is equally likely for a random message to come from any one of these busy periods. Using the standard length biasing argument for $\Pr(A_{i,k})$, $k=1,\ldots,K_i-1$, we have

$$\Pr(A_{i,k}) = \frac{kq_{i,k}}{\lambda_i(1-\rho_i)s/(1-\rho)} \quad (4.18)$$

and for $k \geq K_i$:

$$\Pr(A_{i,K_i} \cap F_{i,k}) = \frac{K_i q_{i,k}}{\lambda_i(1-\rho_i)s/(1-\rho)} \quad (4.19)$$

Substituting (4.8), (4.9), (4.18), and (4.19) into (4.5), we have

$$\frac{\lambda_i(1-\rho_i)s}{1-\rho}E(N_i) = \sum_{k=1}^{K_i-1} q_{i,k} \sum_{j=1}^{k} E(N_i|A_{i,k} \cap B_{i,j}) + \sum_{k=K_i}^{\infty} q_{i,k} \sum_{j=1}^{K_i} E(N_i|A_{i,k} \cap B_{i,j} \cap F_{i,k}) \quad (4.20)$$

Notice that both $E(N_i|A_{i,k} \cap B_{i,j})$ for $k=1,\ldots,K_i-1$ and $E(N_i|A_{i,k} \cap B_{i,j} \cap F_{i,k})$ for $k \geq K_i$ consist of two independent components: $(k-j)$, and the mean queue length in a standard M/G/1 queue with parameters λ_i and b_i. Therefore

$$\frac{\lambda_i(1-\rho_i)s}{1-\rho}E(N_i) = \sum_{k=1}^{K_i-1} q_{i,k} \sum_{j=1}^{k} \left(k-j+\frac{\lambda_i^2 b_i^{(2)}}{2(1-\rho_i)}+\rho_i \right) + \sum_{k=K_i}^{\infty} q_{i,k} \sum_{j=1}^{K_i} \left(k-j+\frac{\lambda_i^2 b_i^{(2)}}{2(1-\rho_i)}+\rho_i \right) \quad (4.21)$$

Using (4.21) and Little's law, we can express the mean queue length of Q_i at the server arrivals as

$$\sum_{k=1}^{\infty} kq_{i,k} = \frac{\lambda_i(1-\rho_i)s}{K_i(1-\rho)} \left(\lambda_i E(W_i) - \frac{\lambda_i^2 b_i^2}{2(1-\rho_i)} \right) + \frac{\lambda_i(1-\rho_i)s}{1-\rho} - \frac{\sum_{k=2}^{K_i-1} k(k-1)q_{i,k} + K_i(K_i-1)[1-\sum_{k=0}^{K_i-1} q_{i,k}]}{2K_i} \quad (4.22)$$

Lemma 3 is thus proved by substituting (4.22) into (4.17). §

From Lemma 3 and (1.1), the pseudo-conservation law for the GD policy is:

$$\sum_{i=1}^{N} \rho_i \left(1-\frac{\lambda_i s(1-\rho_i)}{K_i(1-\rho)} \right) E(W_i) = A - \sum_{i=1}^{N} \left\{ \rho_i \frac{s\lambda_i^2 b_i^{(2)}}{2K_i(1-\rho)} + \frac{b_i \left\{ \sum_{k=2}^{K_i-1} k(k-1)q_{i,k} + K_i(K_i-1)[1-\sum_{k=0}^{K_i-1} q_{i,k}] \right\}}{2K_i} \right\} \quad (4.23)$$

If $K_i=1$ for all i, then (4.23) reduces to the result for the D policy. If $K_i=\infty$ for all i, then (4.23) reduces to the result for the E policy. The K_i unknown boundary probabilities can be approximated by solving the corresponding vacation model. As in the cases with the EL and GL policies, the equations here also involve a K_ith order polynomial equation and a set of K_i linear, independent equations for Q_i [33]. Similarly, an upper bound on $E(U_i)$ for the GD policy can be obtained by ignoring the last term in Lemma 3, and it is exact for $K_i=1$ and $K_i=\infty$:

$$\sum_{i=1}^{N} \rho_i \left(1-\frac{\lambda_i s(1-\rho_i)}{K_i(1-\rho)} \right) E(W_i) \leq A - \frac{s}{1-\rho} \sum_{i=1}^{N} \rho_i \frac{\lambda_i^2 b_i^{(2)}}{2K_i} \quad (4.24)$$

5. OTHER RELATED RESULTS

5.1 CSMQ with Mixed Service Policies

As mentioned before, the set of EL, GL and GD policies cover all other service policies mentioned in this paper, by adjusting the parameter K_i. Therefore, a new result for the mixed services, which is more general than the result in [2], is stated in the following theorem:

Theorem 1:
Consider a stable CSMQ in which the queues can employ EL, GL, or GD policies, then

$$\sum_{i \in EL,GL} \rho_i \left(1-\frac{\lambda_i s}{K_i(1-\rho)} \right) E(W_i) + \sum_{i \in GD} \rho_i \left(1-\frac{\lambda_i s(1-\rho_i)}{K_i(1-\rho)} \right) E(W_i) = A + \frac{s}{1-\rho} \sum_{i \in EL} \frac{\rho_i}{2K_i} \left\{ (1-\rho_i) \sum_{j=1}^{K_i-1} j(K_i-j) p_{i,0,j} - [K_i-1-(K_i+1)\rho_i] \right\} + \sum_{i \in GL} \left\{ \frac{\rho_i^2 s}{1-\rho} - \frac{b_i(1+\rho_i)}{2K_i} E(G_i^2) \right\} - \sum_{i \in GD} \left\{ \frac{\rho_i s \lambda_i^2 b_i^{(2)}}{2K_i(1-\rho)} + \frac{b_i E(G_i^2)}{2K_i} \right\}$$

where

$$E(G_i^2) = \sum_{k=2}^{K_i-1} k(k-1)q_{i,k} + K_i(K_i-1)(1-\sum_{k=0}^{K_i-1} q_{i,k})$$

and the probabilities are defined in the Lemmas 1, 2 and 3.

5.2 Symmetric CSMQ

The pseudo-conservation laws readily give exact expressions of mean waiting times in symmetric CSMQ for the EL, GL, and GD policies. Let $\lambda_i=\lambda$, $b_i=b$, $K_i=K$, $E(W_i)=E(W)$ for all i, but the definitions of ρ, s, $s^{(2)}$ and are unchanged. In addition, $p_{i,0,j}$ is set to $p_{0,j}$ for all i in the EL policy, and $q_{i,k}$ is set to q_k for all i in GL and GD policies. The indices i in $E(W_i)$ denote the service policy. From (3.16), (4.13) and (4.23), the exact mean waiting times are

$$E(W_{EL}) = \frac{1}{1-\rho-\lambda s/K} \left\{ \rho\frac{b^{(2)}}{2b} + (1-\rho)\frac{s^{(2)}}{2s} + \frac{\rho s}{2}(1-1/N) + \frac{s}{2K}\left[(1-\rho/N) \sum_{j=1}^{K-1} j(K-j)p_{0,j} - [K-1-(K+1)\rho/N] \right] \right\} \quad (5.1)$$

$$E(W_{GL}) = \frac{1}{1-\rho-\lambda s/K} \left\{ \rho\frac{b^{(2)}}{2b} + (1-\rho)\frac{s^{(2)}}{2s} + \frac{\rho s}{2}(1+1/N) - \frac{(1+\rho/N)(1-\rho)}{2\lambda K}\left[\sum_{k=2}^{K-1} k(k-1)q_k + K(K-1)(1-\sum_{k=0}^{K-1} q_k) \right] \right\} \quad (5.2)$$

$$E(W_{GD}) = \frac{1}{1-\rho-\lambda s(1-\rho/N)/K} \left\{ \rho\frac{b^{(2)}}{2b} + (1-\rho)\frac{s^{(2)}}{2s} + \rho s(1-1/N) - \frac{s\lambda^2 b^{(2)}}{2K} - \frac{1-\rho}{2\lambda K}\left[\sum_{k=2}^{K-1} k(k-1)q_k + K(K-1)(1-\sum_{k=0}^{K-1} q_k) \right] \right\} \quad (5.3)$$

6. CONCLUSION AND FUTURE RESEARCH

In this paper, we have derived exact expressions of $E(U_i)$s for the exhaustive limited, gated limited, and general decrementing service policies in the CSMQs. Substitutions of $E(U_i)$s into (1.1) give the pseudo-conservation laws for these policies which are in terms of K_i unknown boundary probabilities (K_i-1 for EL). Then we propose to use the corresponding server vacation models to estimate these probabilities for those queues with $1<K_i<\infty$. The pseudo-conservation law for the EL policy using these estimated probabilities are shown to be very accurate, when compared with the simulation results. Subsequently, we also give a pseudo-conservation law for mixed services and exact mean waiting times in symmetric CSMQs.

Given the results in this paper, there are two potential areas for the future research. The first area is the results of pseudo-conservation laws for *time-limited* service policies. Similar to the EL and GL policies, exhaustive-timed limited (ETL) and gated-time limited (GTL) are two viable time-limited service policies. In these policies, there is a maximum allowed visit time for each queue: $S_{i,max}$ for Q_i, i=1,...,N. In the ETL, the server continues to serve Q_i until either $S_{i,max}$ is reached or the queue is empty, whichever comes first. In the GTL policy, the server will switch to the next queue when either $S_{i,max}$ is reached or the server has served all those messages in the queue upon his arrival, whichever comes first. If the message service times are deterministic and $S_{i,max}$ is an integral number of the service times, then Lemma 1 (Lemma 2) for the EL (GL) policy can be directly applied to the ETL (GTL) policy. Furthermore, conservation laws for other more complicated time-limited policies, such as the timed token rotation protocols in FDDI and MAP, need to be investigated. The second area is the need for accurate conservation law-based approximations of the mean waiting times. For the E, G and NE policies, approximate mean waiting time can be expressed as a linear function of the mean residual cycle time. By ignoring the small differences in the second moments of the cycle times, the mean residual cycle time can be obtained from the pseudo-conservation law. Hence, approximate mean waiting times at individual queues can be obtained. For the EL, GL, and GD policies, similar linear relationships also exist, however, these expressions are much more complex.

ACKNOWLEDGEMENT

The authors would like to thank the three reviewers for the correction, and other suggestions to improve this paper.

REFERENCES

[1] O. Boxma and B. Meister, "Waiting Time Approximations for Cyclic-Service Systems with Switch-Over Times," *Performance Evaluation Review*, vol.14, no.1, pp.254-262, May 1986.

[2] O. Boxma and W. Groenendijk, "Pseudo-Conservation Laws in Cyclic-Service Systems," *J. Appl. Prob.*, vol.24, pp.949-964, Dec. 1987.

[3] C. Brooks and O. Yue, "Effect of the Token Holding Timer on MAP Performance," *INFOCOM*, pp.342-347, 1989.

[4] W. Bux, "Local-Area Subnetworks: A Performance Comparison," *IEEE Trans. Commun.*, vol. COM-29, no.10, pp.1465-1473, Oct. 1981.

[5] W. Bux and H. Truong, "Mean-Delay Approximation for Cyclic-Service Queueing Systems," *Performance Evaluation*, vol.3, pp.187-196, 1983.

[6] K. Chang and S. Sandhu,"Performance Evaluation of a Single-Server Queue with Vacations and Exhaustive Limited Service Policy." *Report 0489*, Information and Decision System Laboratory, Rensselaer Polytechnic Institute, Troy, NY, July 1989.

[7] M. Colvin and A. Weaver, "Performance of Single Access Class on the IEEE 802.4 Token Bus," *IEEE Trans. Commun.*, vol. COM-34, no.12, pp.1253-1256, Dec. 1986.

[8] R. Cooper, "Queues Served in Cyclic Order: Waiting Times," *Bell Syst. Tech. J.*, vol.49, no.3, pp.399-413, Mar. 1970.

[9] D. Everitt, "A Conservation-Type Law for the Token Ring with Limited Service." *British Telecom. Tech. J.*, vol.4, no.2, pp.51-61, Apr. 1986.

[10] D. Everitt, "Simple Approximations for Token Rings," *IEEE Trans. Commun.*, vol. COM-34, no.7, pp.719-721 July 1986.

[11] D. Everitt, "Approximations for Asymmetric Token Rings with a Limited Service Discipline," *British Telecom. Tech. J.*, vol.6, no.3, pp.46-51, July 1988.

[12] D. Everitt, "An Approximation Procedure for Cyclic Service Queues with Limited Service," Presented at *International Seminar on Performance of Distributed and Parallel Systems*, Kyoto, Japan, Dec. 1988.

[13] D. Everitt, "A Note on the Pseudo-Conservation Laws for Cyclic Service Systems with Limited Service Disciplines," *IEEE Trans. Commun.*, vol. COM-37, no.7, pp.781-783, July 1989.

[14] *FDDI Token Ring Media Access Control*, Draft Proposed—ANSI Standard X3T9.5/83-16, rev.10, Feb. 28. 1986.

[15] M. Ferguson and Y. Aminetzah, "Exact Results for Nonsymmetric Token Ring Systems," *IEEE Trans. Commun.*, vol. COM-33, no.3, pp.223-231, Mar. 1985.

[16] S. Fuhrmann, "Inequalities for Cyclic Service Systems with Limited Service Disciplines," *GLOBECOM*, pp.182-186, 1987.

[17] S. Fuhrmann and Y. Wang, "Analysis of Cyclic Service Systems with Limited Service: Bounds and Approximations," *Performance Evaluation*, vol.9, no.1, 1988.

[18] W. Genter, *Performance Analysis of the Token Passing Bus Network with Priority Classes*. Doctoral Dissertation, Rensselaer Polytechnic Institute, Troy, New York, 1989.

[19] O. Hashida, *A Study of Multiqueues in Communication Control*. Doctoral Dissertation, University of Tokyo, 1981.

[20] O. Ibe and X. Cheng, "Approximate Analysis of Asymmetric Single-Service Token-Passing Systems," *IEEE Trans. Commun.*, vol.COM-37, no.6, pp.572-577, June 1989.

[21] IEEE Standard 802.5-1985. *Token-Passing Ring Access Method and Physical Layer Specifications*. 1985.

[22] D. Karvelas and A. Leon-Garcia, "Performance Analysis of the Medium Access Control Protocol of the FDDI Token Ring Network," *GLOBECOM*, pp.1119-1123, 1988.

[23] L. Kleinrock, "A Conservation Law for a Wide Class of Queueing Disciplines," *Naval Res. Log. Quart.*, vol.12, pp.181-192, 1965.

[24] P. Kuehn, "Multiqueue Systems with Non-exhaustive Cyclic Service," *Bell Syst. Tech. J.*, vol.58, pp.671-698, 1979.

[25] M. Leibowitz, "An Approximate Method for Treating a Class of Multiqueue Problems," *IBM J. of Res. and Develop.*, vol.5, no.3, pp.204-209, July 1961.

[26] F. Meier et al., *MAP in the Process Industries*, White paper, First Draft. The MAP Process Industries Focus Group, Instrument Society of America, 1987.

[27] J. Pang and R. Donaldson, "Approximate Delay Analysis and Results for Asymmetric Token-Passing and Polling Networks," *IEEE J. Sel. Commun.*, vol. SAC-4, no.6, pp.783-793, Sept. 1986.

[28] S. Sach, K. Kan and J. Silvester, "Performance Analysis of a Token-Bus Protocol and Comparison with Other LAN Protocols," *Proc. 10th Conf. on Local Computer Networks*, pp. 1492-1498, 1985.

[29] L. Servi and D. Yao, "Stochastic Bounds for Queueing Systems with Limited Service Schedules,"*Performance Evaluation*, vol.9, no.4, pp.247-260, Aug. 1989.

[30] M. Srinivasan, "An Approximation for Mean Waiting Times in Cyclic Server Systems with Non-Exhaustive Service," *Performance Evaluation*, vol.9, no.1, pp.17-33 Nov. 1988.

[31] H. Takagi, "Mean Message Waiting Times in Symmetric Multi-Queue Systems with Cyclic Service," *Performance Evaluation*, vol.5, no.4, pp.271-277, Nov. 1985.

[32] H. Takagi, *Analysis of Polling Systems*. The MIT Press, Cambridge, Mass. 1986.

[33] H. Takagi, *Queueing Analysis of Vacation Models*, IBM Tokyo Research Laboratory Report: TR87-0032, Tokyo, Japan, 1987.

[34] K. Watson, "Performance Evaluation of Cyclic Service Strategies—A Survey," *Proceedings of the Tenth International Symposium on Computer Performance*, North-Holland, Amsterdam, 1984.

[35] O. Yue, "Performance of the Timed Token Scheme in MAP," *GLOBECOM*, pp.187-192, 1987.

Table1: Numerical Results of Pseudo-Conservation Law for Exhaustive Limited Policy: Case 1

Three queues with asymmetric arrival rates: $K_1=3$, $K_2=K_3=1$; $\lambda_1=0.6$, $\lambda_2=\lambda_3=0.2$; exponentially distributed service times with equal means; deterministic walk times, s_i.

ρ	Simulation	Estimate	Upper bound	Lower bound
		$s_i=0.05\ \forall i$		
0.3	0.0674	0.0672 (0.34)	0.0703 (-4.30)	0.000 (100.0)
0.5	0.3140	0.3144 (-0.11)	0.3235 (-3.00)	0.000 (100.0)
0.8	2.8031	2.8094 (-0.22)	2.8500 (-1.69)	2.434 (13.15)
		$s_i=0.10\ \forall i$		
0.3	0.0954	0.0942 (1.26)	0.1021 (-7.06)	0.000 (100.0)
0.5	0.3760	0.3754 (0.20)	0.3970 (-5.53)	0.047 (87.51)
0.8	3.0401	3.0468 (-0.21)	3.1408 (-3.30)	2.725 (10.38)

Table 2: Numerical Results of Pseudo-Conservation Law for Exhaustive Limited Policy: Case 2

Three queues with asymmetric arrival rates: $K_1=3$, $K_2=K_3=1$; exponentially distributed service times with $b_i=0.2\ \forall i$; deterministic walk times, s_i which are greater than the mean service times.

Cases	Simulation	Estimate	Upper bound	Lower bound
$\lambda_1=1.5$, $\lambda_2=\lambda_3=0.5$; $s_i=0.3\ \forall i$	0.4209	0.4187 (0.53)	0.5410 (-28.54)	0.4010 (4.72)
$\lambda_1=1.2$, $\lambda_2=\lambda_3=0.4$; $s_i=0.4\ \forall i$	0.3381	0.3332 (1.44)	0.4469 (-32.20)	0.2949 (12.76)
$\lambda_1=1.0$, $\lambda_2=\lambda_3=0.3$; $s_i=0.6\ \forall i$	0.3133	0.3096 (1.18)	0.4455 (-42.19)	0.2855 (8.87)

Table 3: Numerical Results of Pseudo-Conservation Law for Exhaustive Limited Policy: Cases 3

Sixteen queues with asymmetric arrival rates: $K_1=\ldots K_4=3$, $K_5=\ldots=K_{16}=1$; $\lambda_1=\ldots=\lambda_4=0.16$, $\lambda_5=\ldots=\lambda_{16}=0.03$; exponentially distributed service times with equal means; deterministic walk times, $s_i=0.05\ \forall i$.

ρ	Simulation	Estimate	Upper bound	Lower bound
0.3	0.1981	0.1939 (2.13)	0.2088 (-5.41)	0.000 (100.0)
0.5	0.6052	0.5989 (1.04)	0.6453 (-6.63)	0.000 (100.0)
0.8	3.7922	3.8150 (-0.60)	4.1301 (-8.91)	1.340 (64.67)