**STABILITY OF TOKEN PASSING RINGS**

February 3, 1992

Leonidas Georgiadis
High Perf. Comp. and Comm.
IBM T. J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598
U.S.A.

Wojciech Szpankowski*
Department of Computer Science
Purdue University
W. Lafayette, IN 47907
U.S.A.

## Abstract

Sufficient stability condition for the standard token passing ring is "known" since the seminal paper of Kuehn in 1979. However, this condition was derived without formal proof, and the proof seems to be of considerable interest to research community. In fact, Watson observed that in the performance evaluation of token passing rings "*it is convenient to derive stability conditions ... (without proof)*". Our intention is to fill this gap, and provide a formal proof of the *sufficient and necessary* stability condition for the token passing ring. In this paper we present the case when the arrival process to each queue is Poisson but service times and switchover times are generally distributed. We consider in depth gated $\ell$-limited ($\ell \leq \infty$) service discipline for each station. We also indicate that the basic steps of our technique can be used to study the stability of some other multiqueue systems.

### Keywords and Phrases

Token passing rings, stability, substability, ergodicity, Markov chains, Loynes' scheme, stochastically dominant, Little's formula, regenerative processes.

*This research was supported by NSF Grant CCR-8900305, and in part by AFOSR Grant 90-0107, and by Grant R01 LM05118 from the National Library of Medicine.

## 1. INTRODUCTION

Distributed multiqueue systems which share a single scarce resource (i.e., server) such as a communication channel or a processor, have received a considerable amount of attention in the recent literature. Important examples of such distributed multiqueue systems are local area networks (e.g., ALOHA system, Ethernet, token passing ring, FDDI ring, etc.), multiprocessor systems, distributed computations, distributed data bases, and so forth. Of special interest is the token passing ring (cf. Kuehn [12], Takagi [26], Takagi [27]) due to a number of reasons. In particular, it appears that determination of sound measures of performance for such a system, under realistic assumptions such as asymmetric traffic, finite or infinite buffers, non-exhaustive service and general input are fairly difficult to obtain, as can be witnessed from the literature (Boxma [2], Coffman and Gilbert [5], Kleinrock and Levy [11], Levy *et al.* [13], Takagi [27]). For example, it is known that obtaining the distribution of the number of messages queued in each station is a formidable open problem, as is the problem of obtaining the waiting time distribution. Surprisingly enough, the stability condition for the token passing ring was *heuristically* predicted by Kuehn [12] in 1979, and then reproduced with some minor changes in many other papers (e.g., Ibe and Cheng [9]). But ... , Watson [30] observed that in the performance evaluation of token passing rings "*it is convenient to derive stability conditions ... (without proof)*". In fact, no formal proof of the stability condition for the token passing ring was published (for some preliminary results see Szpankowski [24]). Our intention is to fill this gap, and provide a formal proof of the *sufficient and necessary* stability condition for the token passing ring.

There is a version of the token passing ring, one on which the original token passing LAN was defined, which is particularly formidable for the analysis, and therefore it will be of our prime interest. This is the problem of *nonexhaustive* service on an *asymmetric* system with $M$ stations, where at most $\ell_i$ messages are transmitted by station $i \in \mathcal{M} = \{1, \ldots, M\}$ each time the $i$th station acquires the free token. Following the literature we call such a system $\ell$-limited token passing ring. It must be stressed that up-to-date no exact analysis of such a system exists except for two stations system with $\ell_i = 1$ for $i = 1, 2$ (cf. Boxma [2]). Nevertheless, even without such an explicit analysis we present in this paper a rigorous proof for the stability conditions of such a system with Poisson arrivals, and general service and switchover times.

Stability is of considerable importance to the engineering and scientific communities. It is a fundamental issue in the design of any distributed system since only stable systems can work in practice. Hereafter, by stability we understand the existence of the limiting distribution of a quantity of interest. This will imply that the queue lengths process stays in a bounded region with high probability.

Despite a vigorous research in the area of stability over last twenty years (cf. Tweedie [28], Szpankowski [24], Walrand [29]), very few computable stability criteria are known for multidimensional processes, in particular multidimensional Markov chains. The most popular approach through the *Lyapunov test function* (cf. Tweedie [28]) did not succeed in the past to provide general computable criteria for multidimensional Markov chains. However, due to pioneering work of Malyshev [15], continued by Mensikov [17], and Malyshev and Mensikov [16] some progress has been made in obtaining stability condition for a class of two-dimensional and three-dimensional Markov chains. Recently, stronger stability criteria for two dimensional chains have been presented by Fayolle [7] and Rozenkrantz [21]. Unfortunately, these conditions are still difficult to apply in practice for higher dimensional processes (see Karatzoglu and Ephremides [10] for an application of this to a multidimensional ALOHA system). A more practical approach to stability of multidimensional Markov chains arising in queueing applications was discussed in Szpankowski [23] (for more details see the survey in Szpankowski [24]).

Our approach to the stability of token passing rings follows the idea suggested in Szpankowski [25], and differs significantly from the standard methodology of the test function (cf. Tweedie [28]). Our approach is based on a simple idea of stochastic dominance technique, and application of Loynes [14] stability criteria for an isolated queue. We use the stochastic dominance to verify technical stationarity requirements in Loynes' criteria. We shall indicate that this approach is *not* restricted to $\ell$-limited token passing rings, and stability of several other distributed systems can be assessed by this methodology (cf. [8] and [25]).

In the rest of this paper, we will consider the gated version of the $\ell$-limited policy, i.e., the customers that are allowed to be served at queue $i$ are only those that are present at the instant of token arrival at that queue. We shall analyze the token passing ring with Poisson arrivals with parameter $\lambda_i$ for the $i$th station, general distribution of service times $\{S_i^k\}_{k=1}^{\infty}$ and switchover times $\{U_i^k\}_{k=1}^{\infty}$. Our main result can be formulated as follows (see also Theorem 7 and Theorem 10).

**PROPOSITION.** *Consider a token passing ring consisting of $M$ stations with $\ell_i$-limited service schedule for the $i$th station, and Poisson arrivals. Then the system is stable if and only if $\rho_0 = \sum_{j=1}^{M} \rho_j < 1$ and*

$$\lambda_j < \frac{\ell_j}{u_0}(1 - \rho_0) \quad \text{for all} \quad j \in \mathcal{M} = \{1, \ldots M\}, \quad \text{with} \quad \ell_j < \infty,$$

*where* $u_0 = \sum_{j=1}^{M} EU_j$ *is the average total switchover time, and* $\rho_j = \lambda_j s_j$ *with* $s_i = ES_i$ *being the average service time at the ith station.* ■

Note that the above stability criteria are represented in terms of a set of *linear* inequalities with respect to input rates $\lambda_i$ for $i \in \mathcal{M}$. Figure 1 shows the stability region for $M = 3$.

The paper is organized as follows. In the next section we present our preliminary results that are of their own interests for the performance evaluation of the token passing ring. In particular, we find Markovian representations of the system (cf. Theorem 1), establish some Wald's type formulas (cf. Theorem 3), and prove a crucial stochastic dominance relationship (cf. Theorem 4). Finally, in Section 3 we present our main construction that leads to the proof of the above Proposition.

## 2. PRELIMINARY RESULTS

In this section we present several results that are required to establish our main finding regarding the stability of the token passing ring. These results are of their own interests, and can be used to obtain some estimates for the performance evaluation of the system. In the sequel, we list our main assumptions, prove Markovian character of an imbedded queueing process, show two simple Wald's type identities, and finally establish a stochastic dominance relationship.

We start with a precise definition of our stochastic model. We shall adopt the following assumptions.

(A1) There are $M$ stations (queues) on a loop, each having infinite capacity buffer.

(A2) Maximum number of customers served during the token visit at the $i$th queue is limited to $\ell_i < \infty$. Only customers that are present at the instant of token arrival can be served. This assumption will be relaxed later, to include the case $\ell_i = \infty$ (see Corollary 9).

(A3) Arrival process $A_i^t$, $t \in [0, \infty)$, to the $i$th queue is a Poisson process with parameter $\lambda_i > 0$. Here, $A_i^t$ is the number of arrivals at queue $i$ up to time $t$. The arrival process at a queue is independent of the arrival processes to other queues.

(A4) Service time process $\{S_i^k\}_{k=1}^{\infty}$ at queue $i$ is i.i.d. with $s_i = ES_i^1 > 0$. The service time process at a queue is independent of the arrival processes at all queues and independent of the service time processes at other queues.

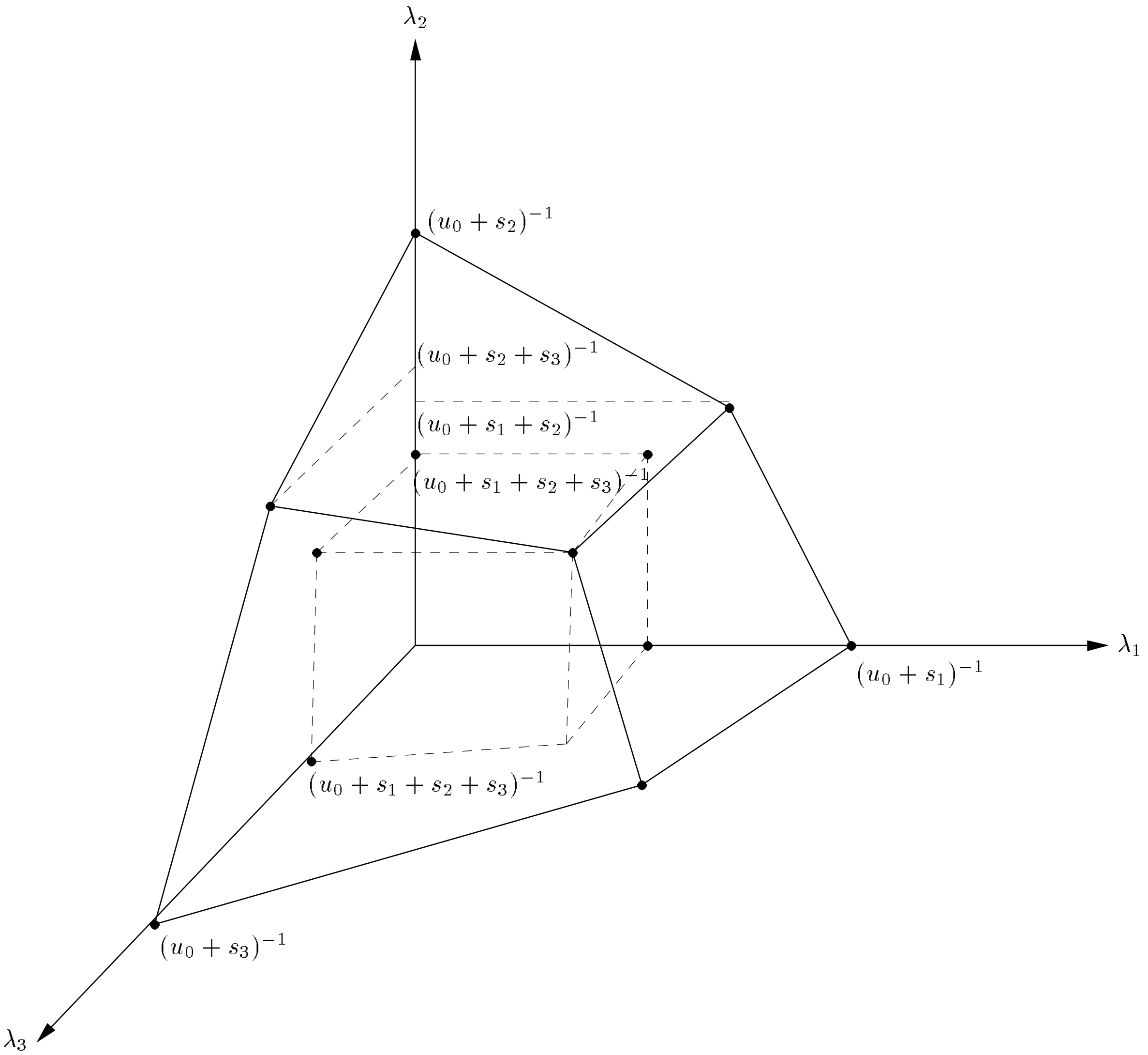


Figure 1: Stability region for the token passing ring with $M = 3$ users

(A5) The switchover time between $i$ and $i+1 \bmod M$ queue $\{U_i^k\}_{k=1}^{\infty}$ is i.i.d. with the average total switchover time equal to $u_0 = \sum_{i=1}^{M} EU_i^1$, and the process is independent of the arrival and the service time processes. To avoid unnecessary complications we assume that $\Pr\{U_i^1 > 0\} = 1$ for $i = 1, \ldots, M$.

Now we are ready to present a Markovian description of the system. We need a little bit of notation. By (A1), the token visits stations in a cyclic order. Let $n$ denote the $n$th visit of the token to any queue. Then, $k_n = \lfloor (n-1)/M \rfloor + 1$ denotes the cycle number in which the $n$th visits occurs (we start counting cycles from one and assume that the token starts from queue 1). Note that the queue visited at the $n$th visit is just $J_n = n - M(k_n - 1)$. Let also $T_n$ be the time instant of the $n$th visit of the token to any queue. Define an $M$-dimensional process $\widetilde{\mathbf{N}}^n = (\widetilde{N}_1^n, \ldots, \widetilde{N}_M^n)$, where $\widetilde{N}_i^n$ is the number of customers in queue $i$ at time $T_n$. In addition, by $\mathcal{N}_i^n$ we mean the total number of customers served from queue $i$ up to time $T_n$. Theorem 1 below proves that $\widetilde{\mathbf{N}}^n$ is a Markov chain.

**Theorem 1.** *The process $\widetilde{\mathbf{N}}^n$ is a ( nonhomogeneous in general) Markov chain.*

**Proof.** Let $L_i^n$ be the number of customers served from queue $J_n$ at the $n$th visit of the token. According to (A2) $L_{J_n}^n = \min\{\widetilde{N}_{J_n}^n \ , \ \ell_{J_n}\}$ and $L_i^n = 0$ for $i \neq J_n$. The time $B_n$ that elapses between $n$th and $n+1$st visit of the token to *any* queue is

$$B_n = \sum_{j=1}^{L_{J_n}^n} S_{J_n}^{\mathcal{N}_{J_n}^n + j} + U_{J_n}^{k_n} \ , \tag{1}$$

and the number of arrivals $X_i^n$ to queue $i$ between the $n$th and $n+1$st visits are

$$X_i^n = A_i^{T_n + B_n} - A_i^{T_n} \ . \tag{2}$$

Finally, the following recursions hold for the queue size in the $i$th station

$$\begin{aligned} \widetilde{N}_i^{n+1} &= \widetilde{N}_i^n + X_i^n & & if \quad i \neq J_n \\ \widetilde{N}_i^{n+1} &= [\widetilde{N}_{J_n}^n - \ell_{J_n}]^+ + X_i^n & & if \quad i = J_n \ , \end{aligned} \tag{3}$$

where $[x]^+ = \max\{x, 0\}$. Since the transmission policy is nonpreemptive and does not depend on the service times of the customers, no information is obtained from the history up to time $T_n$ about the service times of the customers that are in the queues at time $T_n$. Taking also into account assumptions (A3)-(A5), we conclude that the processes $\{S_i^{\mathcal{N}_i^n + j}\}_{j=1}^{\infty}$ $A_i^{T_n+t} - A_i^{T_n}$, $t \in [0, \infty)$, and the random variable $U_{J_n}^{k_n}$, are independent of $\widetilde{\mathbf{N}}^m$, $1 \leq m \leq n$. From the above discussion we conclude that $\widetilde{\mathbf{N}}^{n+1}$ is of the form $\widetilde{\mathbf{N}}^{n+1} = f(\widetilde{\mathbf{N}}^n, \mathbf{Y}^n)$ for some (measurable) function $f(\cdot)$ where $\mathbf{Y}^n$ is composed of the processes $S_i^{\mathcal{N}_i^n+j}$, $A_i^{T_n+t} - A_i^{T_n}$ and $U_{J_n}^{k_n}$. Therefore, $\tilde{\mathbf{N}}^n$ is a Markov chain (see, for example, page 34 of Nevelson and Hasminskii [20]). ■

There are other Markovian descriptions of the system. For example, define $N_j^n(i)$ to be the number of customers at queue $j$ when the token visits queue $i$ for the $n$th time. Then, the process $\mathbf{N}^n(i) = (N_1^n(i), \ldots, N_M^n(i)\,)$ can be deduced from $\tilde{\mathbf{N}}^n$ since $\mathbf{N}^n(i) = \tilde{\mathbf{N}}^{(n-1)M+i}$. This implies that for a fixed $i$, $\mathbf{N}^n(i)$ is a Markov chain too. In fact, repeating the arguments of the proof of Theorem 1 it can be seen that $\mathbf{N}^n(i)$ is a homogeneous Markov chain. It is easily verified also that the chain is irreducible and aperiodic. Hence, we have the following.

**Corollary 2**. *The process $\mathbf{N}^n(i)$ of the queue lengths registered by the token when it visits (reference) queue $i$, is a homogeneous, irreducible and aperiodic Markov chain.* ■

We will need some Wald's type relationships between the average number of customers served per token visit and the average cycle time. Let $L_i^n$ be the number of customers served by queue $i$ during the $n$th visit of the token to this queue. Let also $C_i^n$ be the cycle length, that is, the length of time between the $n$th and $n+1$st visits of the token to the reference queue $i$. By $EL_i$ and $EC$ we denote the long run averages of $L_i^n$ and $C_i^n$, if they exist (it will be seen that the limiting average of the cycle length $C_i^n$ does not depend on the reference queue $i$). The following result is known since Kuehn [12] (cf. Takagi [26]). We provide here a proof based on regenerative arguments since we will need some of the steps of the proof in later sections.

**Theorem 3**. *Let the Markov chain $\mathbf{N}^n(i)$ be positive recurrent (ergodic) for some $i \in \mathcal{M}$. Then*

1. *$\mathbf{N}^n(j)$ is ergodic for all $j \in \mathcal{M}$.*
2. *Under any initial condition on $\mathbf{N}^1(i)$, the long run averages of $\{L_i^n\}_{n=1}^{\infty}$ and $\{C_i^n\}_{n=1}^{\infty}$ exist and are unique. Moreover, $\rho_0 = \sum_{j=1}^M \rho_j < 1$,*

$$EL_j = \lambda_j EC, \qquad j \in \mathcal{M} \tag{4}$$

*and*

$$EC = \frac{u_0}{1 - \sum_{j=1}^M \rho_j}\ , \tag{5}$$

*where $u_0$ is the total average switchover time (cf. assumption (A5)) and $\rho_i = \lambda_i s_i$ is the utilization coefficient for the $i$th queue.*

**Proof.** Without loss of generality, let $i = 1$. By the assumption, $\mathbf{N}^n(1)$ is an ergodic Markov chain. Note that $\mathbf{N}^n(1)$ has a natural regeneration structure, namely when all queues are empty, that is, when the process returns to zero state $\mathbf{0} = (0, 0, \ldots, 0)$. Assume $\mathbf{N}^1(1) = \mathbf{0}$, and $K^1 = 1$. Define

$$K^{n+1} = \min\{m > K^n : \ \mathbf{N}^m(1) = \mathbf{0}\} \ ,$$

and $R^n = K^{n+1} - K^n$. We shall also denote $R = R^1$. It is well known that $\{R^n\}_{n=1}^{\infty}$ are i.i.d. random variables. Due to the ergodicity of $\mathbf{N}^n(1)$ we have $ER < \infty$. Observe that for $j \in \mathcal{M}$, $\mathbf{N}^n(j)$ is regenerative with respect to $R^n$. Since it is easily seen that $R^n$ is aperiodic, it follows (see Asmussen [1, Chapter V]) that the process $\mathbf{N}^n(j)$ has a steady state distribution and therefore, is ergodic.

The sequences $\{C_i^n\}_{n=1}^{\infty}$ and $\{L_i^n\}_{n=1}^{\infty}$ are regenerative with respect to $R^n$. Therefore (cf. Asmussen [1] Corollary 1.5 and Theorem 3.1 in Chapter V),

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} L_j^k}{n} = \frac{E\left(\sum_{k=1}^{R} L_j^k\right)}{ER}, \qquad \lim_{n\to\infty} \frac{\sum_{k=1}^{n} C_1^k}{n} = \frac{E\left(\sum_{k=1}^{R} C_1^k\right)}{ER} \qquad a.s. \ . \tag{6}$$

Moreover, $L_j^n$ and $C_1^n$ converge in distribution to $L_j$ and $C_1$ such that

$$EL_j = \frac{E\left(\sum_{k=1}^{R} L_j^k\right)}{ER} \quad , \quad EC_1 = \frac{E\left(\sum_{k=1}^{R} C_1^k\right)}{ER} \ . \tag{7}$$

If $\mathbf{N}^1(1) \neq \mathbf{0}$, then the involved sequences constitute delayed regenerative processes for which $R^1 = \min\{m > 1 : \ \mathbf{N}^m(1) = \mathbf{0}\} - 1$, has different distribution than $\{R^n\}_{n=2}^{\infty}$. Since $\mathbf{N}^n(1)$ is ergodic, $R^1$ is an honest random variable ($\Pr(R^1 = \infty) = 0$), and formulas (6), (7) still hold, with the provision that the averages involve now the corresponding quantities in the regeneration cycle $R^2$. Now we are in position to prove (4) and (5). Note first, that $\sum_{k=1}^{R} L_1^k \leq R\ell_1$ and since $ER < \infty$, we also have that $E\left(\sum_{k=1}^{R} L_1^k\right) < \infty$. Observe next that in the interval $[0, \sum_{k=1}^{R} C_1^k)$ *all the arriving customers from all queues must be served.* If $\mathcal{A}_j$ is the number of arrivals to queue $j$ in the interval $[0, \sum_{k=1}^{R} C_1^k)$, then $E\mathcal{A}_j = E\left(\sum_{k=1}^{R} L_j^k\right)$, and due to the Poisson assumption (A3) we also have $E\mathcal{A}_j = \lambda_j E\left(\sum_{k=1}^{R} C_1^k\right)$. The last formula follows from the fact that $\sum_{k=1}^{R} C_1^k$ is a *stopping time* for the Poisson arrival process to the $j$th station. Therefore,

$$E\left(\sum_{k=1}^{R} C_1^k\right) = E\left(\sum_{k=1}^{R} L_1^k\right) / \lambda_1 < \infty, \tag{8}$$

and

$$E\left(\sum_{k=1}^{R} L_j^k\right) = \lambda_j E\left(\sum_{k=1}^{R} C_1^k\right) \quad j \in \mathcal{M}. \tag{9}$$

The above and (7) lead to $EL_j = \lambda_j EC_1$, which completes the proof of (4).

To prove (5) we note that the cycle length $C_1^n$ is

$$C_1^n = U^n + \sum_{j=1}^{M} \sum_{m=1}^{L_j^n} S_j^{m+\sum_{k=1}^{n-1} L_j^k} , \tag{10}$$

where $U^n = \sum_{j=1}^{M} U_j^n$. Summing the above over first $R$ visits of the token, taking the expectation of it, and using (9) one obtains the following

$$E\left(\sum_{n=1}^{R} C_1^n\right) = u_0 \cdot ER + \left(\sum_{j=1}^{M} \lambda_j s_j\right) E\left(\sum_{n=1}^{R} C_1^n\right) . \tag{11}$$

Since $ER > 1$ and by (8) $E\left(\sum_{n=1}^{R} C_1^n\right) < \infty$, using (7) we obtain from the above that $\sum_{j=1}^{M} \rho_j < 1$ and $EC_1 = EC = u_0/(1 - \sum_{i=1}^{M} \rho_i)$ as needed for (5). ■

**Remark.** As can be seen from the proof, for the first assertion of the theorem, the finiteness of $\ell_i$ is not needed. Also, for the second assertion, only the finiteness of $\ell_i$ for some $i \in \mathcal{M}$ is needed.

The next result is our main finding in this section. Before we plunge into technical details, we first give a brief overview of our approach. In the process of estimating stability we need to build several dominant systems of the original token passing ring. For example, we partition the set of users into a class $\mathcal{S}$ of *nonpersistent* queues and a class $\mathcal{U}$ of *persistent* queues. A nonpersistent queue serves customers in the normal way as in the original token passing ring. A persistent queue, however, always sends the *maximum* allowable number of customers, that is, $\ell_i$ for $i \in \mathcal{U}$, by sending if necessary 'dummy' customers. In other words, the token spends $\ell_i$ i.i.d. service times, identically distributed to $S_i^1$, in the $i$th queue before it starts walking to the next queue. A question is whether such a new system dominates the original token passing ring in some sense. If the answer is *yes*, then by proving stability of the dominant system we establish stability of the original token passing ring.

Let $(\mathcal{U}, \mathcal{S})$ be a partitioning the set of $M$ queues. The system that results from this partitioning can be viewed as a token ring that operates under the same policy as the original system, but in which the vacation times have been increased. In Theorem 4 we show that under the Poisson assumption of the arrival processes, an increase in the vacation times implies, under certain statistical assumptions, an increase (in a stochastic sense) of the queue sizes seen by the token at the instants it visits the queues. The Poisson assumption is crucial to this result. To see that, in general, this may not be true even if the vacation times are i.i.d. consider the following example.

EXAMPLE. *Counter-example.*

Consider a single queue with gated service. Assume that the service time is 1 and that the i.i.d. interarrival times $A_k$ have distribution

$$\Pr\{A_k = 1.9\} = \alpha, \quad \Pr\{A_k = 5\} = 1 - \alpha.$$

Consider two versions of the system: the first with vacations 1.5 and the second with vacations 2. Let $N_k^i$, $k = 1, 2, \cdots$; $i = 1, 2$ be the queue size in system $i$ at the instant of the $k$th visit of the token to the queue. We also assume that $N_1^i = 0$, $i = 1, 2$, and at time zero the first vacation begins. Let us consider in detail the first system. The first vacation ends at time 1.5, and since there is no arrival the token resumes the second vacation that ends at time 3, which is also the time of the third token arrival to the queue. Consider the queue length at this time, that is, $N_3^1$. Clearly, $N_3^1 \leq 1$, and $N_3^1 = 0$ if and only if $A_1 = 5$, that is, $\Pr\{N_3^1 = 0\} = 1 - \alpha$, $\Pr\{N_3^1 = 1\} = \alpha$. For the second system, we also have $N_3^2 \leq 1$, however, $N_3^2 = 0$ if and only if one of the following mutually exclusive events happen: (i) $A_1 = 5$ or (ii) $A_1 = 2$ and $A_2 = 5$. Therefore, $\Pr\{N_3^2 = 0\} = 1 - \alpha + \alpha(1 - \alpha)$. Since $\Pr\{N_3^1 \leq 0\} \leq \Pr\{N_3^2 \leq 0\}$, we have that $N_3^1 \geq_{st} N_3^2$. □

The next result holds for general service disciplines of the type introduced in Levy *et al.* [13]. Specifically, in the terminology of Levy *et al.* [13], we consider the class of 'monotonic', 'contractive' policies. This amounts to replacing assumption (A2) with the following more general one.

(A2') Let $f_i(n)$ be the number of customers served from queue $i$ when there are $n$ queued messages at the instant of token arrival at queue $i$. We assume that $f_i(n)$ is a *non-decreasing* function of the number of customers in the $i$th queue. In addition, the following relation holds.

$$f_i(n_1) - f_i(n_2) \leq n_1 - n_2 \qquad if \qquad n_1 > n_2 \ . \tag{12}$$

Now we are ready to formulate our result. Consider two token passing rings, say $\theta$ and $\Theta$. Both satisfy assumptions (A1)-(A4) with (A2) replaced by the weaker assumption (A2'). System $\theta$ satisfies also assumption (A5) and represents our original token passing ring. System $\Theta$ differs only in the switchover times, namely, we assume that the switchover times for $\Theta$ are replaced by $\{\Delta_i^k + U_i^k\}_{i=1}^M$ for $k = 0, 1, \ldots$, where $U_i^k$ are the corresponding switchover times in system $\theta$. We assume that for every $i \in \mathcal{M}$ and every $k \geq 0$ we have $\Delta_i^k \geq 0$. The processes $\{\Delta_i^k\}_{k=1}^{\infty}$, $i \in \mathcal{M}$, may depend on the rest of the processes, however, we make the following "independence of future" assumption.

(A6) The random variable $\Delta_i^k$ is independent of the service times, switchover times and the Poisson increments of the arrival processes to all stations *after* time $T_{M(k-1)+(i+1)} - U_i^k$ (see Fig.2).

**Theorem 4.** *Let* $\tilde{\mathbf{N}}^n(\theta)$ *and* $\tilde{\mathbf{N}}^n(\Theta)$ *denote the queue lengths in both systems. Then, under the above assumptions, and under the condition that the token starts from the same queue, say queue number one, and with the same number of initial customers in both systems, the following holds*

$$\tilde{\mathbf{N}}^n(\theta) \leq_{st} \tilde{\mathbf{N}}^n(\Theta), \tag{13}$$

*where* $\leq_{st}$ *means stochastically smaller.*

**Proof.** To avoid cumbersome notation we present the proof only for $M = 2$ users. The proof can be easily extended to any number of users.

We define some new variables. For system $\theta$ let $T_n^\theta$ and $D_n^\theta$ denote the instances of the $n$th visit and the $n$th token departure from any queue respectively. As before, $J_n^\theta$ denotes the queue number visited at the $n$th visit of the token. Finally, $L_i^n(\theta)$ denotes the number of customers served from queue $i$ at the $n$th visit of the token. Clearly, for our two station system $L_1^n(\theta) = 0$ for $n$ even, and $L_2^n(\theta) = 0$ for $n$ odd. In a similar manner we define respective quantities in the $\Theta$ system.

We assume that $S_i^k$ are assigned upon the beginning of the service. Since the service policies we are considering do not depend on the knowledge of the service requirements of the customers, this assumption does not change the distributions of the processes involved, hence also stochastic dominance property of the systems. Under this modification, we show how to construct from the system $\Theta$ a token passing ring $\bar{\theta}$, which is stochastically equivalent to the system $\theta$ and for which we have that

$$\tilde{\mathbf{N}}^n(\bar{\theta}) \leq \tilde{\mathbf{N}}^n(\Theta) \; . \tag{14}$$

Figure 2 should help to understand our construction. Assume $\tilde{N}_i^1(\bar{\theta}) = \tilde{N}_i^1(\Theta)$ for $i = 1, 2$. Now, we assign to the $\tilde{N}_1^1(\bar{\theta})$ customers of $\bar{\theta}$, the service times $S_i^k$ exactly as in $\Theta$. The same functions $f_i(n)$, $i = 1, 2$ are used in both systems. Therefore, the decision to switch to queue 2 will occur at the same time, namely $D_1^{\bar{\theta}} = D_1^\Theta$.The switchover time for $\bar{\theta}$ becomes now $U_1^1$, and of course $T_2^{\bar{\theta}} \leq T_2^\Theta$ since $\Delta_1^1 \geq 0$ (see Fig. 2).

The interarrival times in the interval $[T_1^{\bar{\theta}}, D_1^{\bar{\theta}}) = [T_1^\Theta, D_1^\Theta)$, are identical for both systems. The interarrival times in system $\bar{\theta}$ in $[D_1^{\bar{\theta}}, T_2^{\bar{\theta}})$ *are constructed identical* to the interarrival times in $[D_1^\Theta + \Delta_1^1, T_2^\Theta)$ in system $\Theta$. Therefore clearly $\tilde{N}_i^2(\bar{\theta}) \leq \tilde{N}_i^2(\Theta)$ for $i = 1, 2$. We

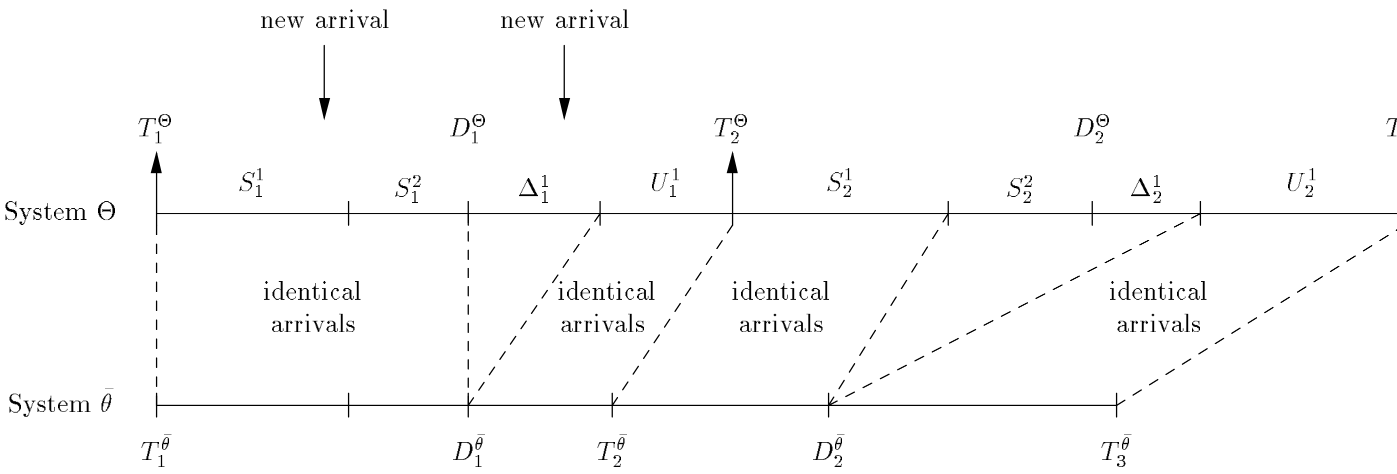


Figure 2: Illustration to the proof of Theorem 4

.

continue our construction, and at time $T_2^{\bar\theta}$ service times are assigned from $S_2^k$ in the same order as in $\Theta$. Also, the interarrival times to system $\bar\theta$ in $[T_2^{\bar\theta}\ ,\ T_2^{\bar\theta} + S_2^1 + \cdots + S_2^{L_2^1(\bar\theta)})$ are taken to be identical to the interarrival times in $[T_2^\Theta\ ,\ T_2^\Theta + S_2^1 + \cdots + S_2^{L_2^1(\bar\theta)})$. Note that by (A2') we have $L_2^1(\bar\theta) \le L_2^1(\Theta)$ and therefore, $D_2^{\bar\theta} = T_2^{\bar\theta} + S_2^1 + \cdots + S_2^{L_2^1(\bar\theta)} \le D_2^\Theta$ (cf. Fig. 2).

To complete the description of the system $\bar\theta$ we have to specify the interarrival times in $[D_2^{\bar\theta}\ ,\ D_2^{\bar\theta} + U_2^1)$. These are taken to be the same as the interarrival times in $[D_2^\Theta + \Delta_2^1\ ,\ D_2^\Theta + \Delta_2^1 + U_2^1)$ in system $\Theta$ (see Fig. 2). Note from the construction that

$$\tilde N_1^3(\bar\theta) = \tilde N_1^2(\bar\theta) + A_1^{[T_2^{\bar\theta},T_3^{\bar\theta})} \le \tilde N_1^2(\Theta) + A_1^{[T_2^\Theta,T_3^\Theta)} = \tilde N_1^3(\Theta)$$

and also by (12),

$$\tilde N_2^3(\bar\theta) = \tilde N_2^2(\bar\theta) - L_2^2(\bar\theta) + A_2^{[T_2^{\bar\theta},T_3^{\bar\theta})} \le \tilde N_2^2(\Theta) - L_2^2(\Theta) + A_2^{[T_2^\Theta,T_3^\Theta)} = \tilde N_2^3(\Theta)\ .$$

We can now repeat exactly the same procedure to construct $\bar\theta$ in the interval $[T_n^{\bar\theta}, T_{n+1}^{\bar\theta})$, $n \ge 3$, in the same manner as it was constructed in the interval $[T_2, T_3)$. By construction the service times and switchover times of system $\bar\theta$ are identically distributed to the corresponding variables of system $\theta$ and are independent of the interarrival processes. In addition, assumption (A6) and the independence of the increments of the Poisson process implies that the constructed interarrival process in system $\bar\theta$ is Poisson with rate $\lambda_i$ for queue $i$. Moreover, by construction (14) holds. Since $\bar\theta$ is stochastically equivalent to $\theta$, we have that the distribution of $\tilde N^n(\theta)$ is identical to the distribution of $\tilde N^n(\bar\theta)$. This completes the proof of Theorem 4. ∎

## 3. MAIN RESULTS

In this section we present a proof of our proposition in the Introduction. However, before we plunge into technical details an overview of our stability approach is discussed. We shall argue that our idea is novel, and can be successfully used to establish stability of some other distributed systems (see Szpankowski [25], [24] for applications to the ALOHA system and coupled-processors system).

Let us introduce some notations. If $\mathbf{N}^n = (N_1^n, \ldots, N_M^n)$ is a stochastic process – not necessary a Markov chain – then we say the process is stable if the distribution of $\mathbf{N}^n$ as $n \to \infty$ *exists* and the distribution is *honest*. In other words, $\mathbf{N}^n$ is stable if for $\mathbf{x} \in I\!R^M$, where $I\!R$ is the set of real numbers, the following holds for all points of continuity of $F(\mathbf{x})$ below.

$$\lim_{n\to\infty} \Pr\{\mathbf{N}^n \leq \mathbf{x}\} = F(\mathbf{x}) \qquad \text{and} \qquad \lim_{\mathbf{x}\to\infty} F(\mathbf{x}) = 1 \tag{15}$$

where $F(\mathbf{x})$ is the limiting distribution function, and by $\mathbf{x} \to \infty$ we understand that $x_j \to \infty$ for all $j \in \mathcal{M} = \{1, \ldots, M\}$. If a weaker condition holds, namely,

$$\lim_{\mathbf{x}\to\infty} \lim_{n\to\infty} \inf \ \Pr\{\mathbf{N}^n \leq \mathbf{x}\} = 1 \ , \tag{16}$$

then the process is called *substable* or *tight* or *bounded in probability* sense. Otherwise, the system is unstable (for more details see Loynes [14]). The isolation property mentioned above can be formally presented as follows.

**Lemma 5.** (i) *If for all $j \in \mathcal{M}$ the one dimensional processes $N_j^n$ are stable (substable), then the $M$-dimensional process $\mathbf{N}^n = (N_1^n, N_2^n, \ldots, N_M^n)$ is substable.*

(ii) *If for some $j$, say $j^*$, $N_{j*}^n$ is unstable, then $\mathbf{N}^n$ is also unstable.*

**Proof.** The proof is simple and can be found in Szpankowski [24] and [25]. For example, part (i) follows from the following inequality

$$\begin{aligned} 1 \geq \lim_{x\to\infty} \liminf_{n\to\infty} \ \Pr\{N_j^n \leq x_j, \ \text{for} \quad j &= 1, 2, \ldots, M\} \\ &\geq \ 1 - \sum_{j=1}^{M} \lim_{x_j\to\infty} \limsup_{n\to\infty} \ \Pr\{N_j^n > x_j\} = 1 \ , \end{aligned}$$

and the last equality is a simple consequence of the substability of $\mathbf{N}_j^t$ for every $j \in \mathcal{M}$. ∎

Our approach is based on the following observations. If the Markov chain defined on a countable state space is irreducible and aperiodic, then substability implies ergodicity of the process since every such a Markov chain converges to a distribution (not necessary honest).

This is a well known fact, and the reader is referred to any book treating Markov chains, for example Chung [4] (cf. Meyn and Tweedie [19]).[1]

By Lemma 5 and the above, we need only to establish substability of every isolated queue. To obtain such stability conditions, we apply the technique of Loynes [14] who proved that a single $G|G|1$ queue is stable if the input rate is smaller than the average service time provided that service times and interarrival times are jointly stationary and ergodic. To verify a technical stationarity condition in Loynes' criteria we apply the stochastic dominance result of Theorem 4. More precisely, we partition the set of queues $\mathcal{M}$ into a set $\mathcal{S}$ of nonpersistent queues and into a set $\mathcal{U}$ of persistent queues as was described in Section 2. By Theorem 4 the new system stochastically dominates the original one, and by proving stability of it, we clearly establish stability conditions for the original token passing ring. We use the mathematical induction to establish stability conditions for the nonpersistent queues in the new system, while the stability condition for a persistent queue is shown by using Loynes' criteria.

To fulfill the above plan, we start by considering the stability condition of a queue that is related to the operation of a persistent queue in the dominant system. More formally, we consider a single queue that always services $\ell$ (dummy if necessary) customers, even if there are less than $\ell$ customers in the queue when the token (server) arrives. The server is of walking type, and after servicing $\ell$ customers it goes for a vacation. It is assumed that the cycle time $C^n$ represents a *stationary and ergodic sequence* with mean $EC$ (no independence is required). The arrival process $A^t$ to this queue is a Poisson process with parameter $\lambda$, independent of the process of cycle times. Let $N^n$ represent the queue length at the beginning of the $n$th cycle. By $X^n$ we denote the number of customers arrived during the $n$th cycle. Note that since the processes $C^n$ and $A^t$ are independent and $A^t$ is Poisson, the process $X^n$ is a stationary and ergodic sequence with mean $EX = \lambda EC$. Then, the queue length satisfies the following recurrence

$$N^{n+1} = \max\{N^n + X^n - \ell \ , \ X^n\} \ . \tag{17}$$

Lemma 6 below provides the stability condition for the system governed by (17). The proof is standard and it is based on Loynes' technique [14]. Therefore, it is omitted.

**Lemma 6**. *Consider the queueing system just described. If $\lambda EC < \ell$, then the queue is stable in the sense of definition (15) .*

[1] For a general Markov chains defined on a general space this is not necessarily true, however, as proved by Meyn and Tweedie [19] very weak conditions regarding compact sets are sufficient for this assertion to be true.

Now we are ready to prove our main result described already in the proposition of the Introduction. In the next theorem we show that the conditions of the proposition are sufficient. The proof uses the idea presented in the overview above, however due to technical reasons we carry it out formally through the mathematical induction.

**Theorem 7**. *The Markov chain $\mathbf{N}^n(i)$ representing the queue lengths in the token passing ring when it visits queue $i \in \mathcal{M}$ is ergodic if*

$$\lambda_j < \frac{\ell_j}{u_0}(1-\rho_0) \qquad \text{for all} \quad j \in \mathcal{M} \tag{18}$$

*where* $\rho_0 = \sum_{j=1}^{M} \rho_j$.

**Proof.** We use mathematical induction. For $M = 1$ the proof is simple. Applying the Lyapunov test function method (cf. Szpankowski [24], Tweedie [28]) to the Markov chain $\{\mathbf{N}^n(1)\}_{n=1}^{\infty}$ we have $E\{N^{n+1} - N^n \mid N^n = k \geq \ell\} = \lambda(\ell s_1 + u_0) - \ell = \lambda u_0 - \ell(1-\rho_0)$, where $\rho_0 = \lambda s_1$. Note that this drift is negative when $\lambda < \ell(1-\rho_0)/u_0$, as needed.

Now we assume that the theorem is true for $M-1$ and prove that it can be extended to the $M \geq 2$ queue case. We will show that if there is a partition $(\mathcal{U},\ \mathcal{S})$, with $\mathcal{U} \neq \emptyset$ and $\mathcal{S} \neq \emptyset$, of the the set $\mathcal{M}$ of $M$ queues such that

$$\lambda_i < \frac{\ell_i}{u_0 + \sum_{k\in\mathcal{U}} \ell_k s_k}\left(1 - \sum_{k\in\mathcal{S}} \rho_k\right) \quad \text{for all} \quad i \in \mathcal{M}\ , \tag{19}$$

then the system is stable. As will be shown below, this will imply that the system is stable when (18) holds. Assuming (19), we construct first a token ring system that dominates stochastically our original token ring system and has stationary cycles. Next we show that the dominating system is substable, which implies the substability of the original token ring system.

Along these lines we consider the system in which the queues in $\mathcal{U}$ are *persistent* and the queues in $\mathcal{S}$ are *nonpersistent*. Recall that a persistent queue $i$ always sends $\ell_i$ i.i.d. messages distributed according to $S_i^k$ (if necessary dummy messages distributed in the same manner are transmitted). Note that the cardinality $|\mathcal{S}|$ of $\mathcal{S}$ is not larger than $M-1$. Let $\overline{\mathbf{N}}^n(i) = \{\overline{N}_1^n(i), \cdots, \overline{N}_M^n(i)\}$ be the queue lengths when the token visits the $i$th queue for the $n$th time in the $(\mathcal{U},\ \mathcal{S})$ system. Observe that the modified system differs from the original token ring system only in the switchover time from a persistent queue to the successor of that queue in the ring. Specifically, if $i \in \mathcal{U}$, then the switchover times become,

$$\overline{U}_i^k = \Delta_i^k + U_i^k,$$

where $\Delta_i^k$ is the time needed to service the dummy messages at node $i$ (if any). Note that the number of dummy messages serviced by node $i$ at the $k$th visit of the token to queue $i$ is equal to $\max\{\ell_i - N_i^k(i), 0\}$. Since $N_i^k(i)$ is independent of future arrivals or future service times, and the service times of the dummy messages are independent of the rest of the processes in the system, it is clear that $\Delta_i^k$ satisfies condition (A6) of Section 2. Therefore, according to Theorem 4, if $\mathbf{N}^1(1) = \overline{\mathbf{N}}^1(1)$, then

$$\mathbf{N}^n(j) \leq_{st} \overline{\mathbf{N}}^n(j), \quad \text{for all } n, \quad j \in \mathcal{M}. \tag{20}$$

Note now that in the $(\mathcal{U},\ \mathcal{S})$ system the queues in $\mathcal{S}$ constitute a token passing ring with $|\mathcal{S}|$ stations satisfying conditions (A1)-(A5) of Section 2, whose operation is independent of the interarrival processes in the persistent queues. Observe that our assumption (A5) holds for the token ring composed of the queues in $\mathcal{S}$, with new switchover times that are i.i.d., however, with a new distribution. This is due to the fact that every persistent queue $i$ sends *exactly* $\ell_i$ i.i.d. messages. Clearly, the total average switchover time $\hat{u}_0$ in such a ring (composed of the queues in $\mathcal{S}$) is equal to

$$\hat{u}_0 = u_0 + \sum_{i \in \mathcal{U}} \ell_i s_i.$$

Let the queue lengths at the $n$th visit of the token to queue $i \in \mathcal{S}$ in such a system be denoted as $\overline{\mathbf{N}}_{\mathcal{S}}^n(i) = \{\overline{N}_j^n(i)\}_{j \in \mathcal{S}}$. Clearly, $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(i)\}_{n=1}^{\infty}$ is a Markov chain, and since $|\mathcal{S}| \leq M - 1$ we can apply the induction hypothesis. Hence, for $i \in \mathcal{S}$, $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(i)\}_{n=1}^{\infty}$ is ergodic if

$$\lambda_i < \frac{\ell_i}{u_0 + \sum_{k \in \mathcal{U}} \ell_k s_k} \left(1 - \sum_{k \in \mathcal{S}} \rho_k\right) \quad \text{for all } \ i \in \mathcal{S}\ . \tag{21}$$

Consider now a queue in $\mathcal{S}$, say queue 1 and let $C_{\mathcal{S}}^n(1)$ be the process of cycle lengths (successive visits to queue 1). Under our assumption (19), (21) holds and the ergodicity of $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(1)\}_{n=1}^{\infty}$, which is true by the induction hypothesis, implies the existence of a unique honest stationary distribution $\pi$ for this process. Let us assume that the process $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(1)\}_{n=1}^{\infty}$ starts with the initial distribution $\pi$, that is, $\overline{\mathbf{N}}_{\mathcal{S}}^1(1)$ has distribution $\pi$. Then, it is well known that the resulting process $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(1)\}_{n=1}^{\infty}$ is stationary and ergodic. Observe now that the process $\{\overline{\mathbf{N}}_{\mathcal{S}}^n(1), C_{\mathcal{S}}^{n-1}(1)\}_{n=2}^{\infty}$ is also a Markov chain with uncountable (in general) state space, whose transition probabilities have the following properties.

$$\begin{aligned}
\Pr\{(\overline{\mathbf{N}}_{\mathcal{S}}^{n+1}(1), C_{\mathcal{S}}^n(1)) \in B \mid \overline{\mathbf{N}}_{\mathcal{S}}^n(1), C_{\mathcal{S}}^{n-1}(1)\} &= \Pr\{(\overline{\mathbf{N}}_{\mathcal{S}}^{n+1}(1), C_{\mathcal{S}}^n(1)) \in B \mid \overline{\mathbf{N}}_{\mathcal{S}}^n(1)\} \\
&= q(\overline{\mathbf{N}}_{\mathcal{S}}^n(1), B), \quad n \geq 2,
\end{aligned} \tag{22}$$

and

$$\Pr\{(\overline{\mathbf{N}}^2_{\mathcal{S}}(1), C^1_{\mathcal{S}}(1)) \in B \mid \overline{\mathbf{N}}^1_{\mathcal{S}}(1)\} = q(\overline{\mathbf{N}}^1_{\mathcal{S}}(1), B), \tag{23}$$

where $B$ is a Borel set in $I\!N_0^{|\mathcal{S}|} \times I\!R$ ($I\!N_0$ is the set of nonnegative integers) and $q(\cdot, B)$ is a stationary transition probability function with domain $I\!N_0^{|\mathcal{S}|}$. From (22) it easily follows that the process $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1), C^{n-1}_{\mathcal{S}}(1)\}_{n=2}^{\infty}$ has a stationary distribution

$$\bar{\pi}(B) = \sum_{\mathbf{x} \in I\!N_0^{|\mathcal{S}|}} q(\mathbf{x}, B)\pi(\mathbf{x}) \ .$$

Since $\overline{\mathbf{N}}^1_{\mathcal{S}}(1)$ has distribution $\pi$, we see from (23) that $(\overline{\mathbf{N}}^2_{\mathcal{S}}(1), C^1_{\mathcal{S}}(1))$ has distribution $\bar{\pi}$ and therefore, the process $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1), C^{n-1}_{\mathcal{S}}(1)\}_{n=2}^{\infty}$ is stationary [3, Proposition 7.11]. Using (22) and the fact that $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1)\}_{n=1}^{\infty}$ is irreducible and therefore indecomposable, we show in Appendix A that $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1), C^{n-1}_{\mathcal{S}}(1)\}_{n=2}^{\infty}$ is indecomposable as well. It follows from [3, Theorem 7.16] that $\bar{\pi}$ is unique and $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1), C^{n-1}_{\mathcal{S}}(1)\}_{n=2}^{\infty}$ is ergodic. To complete the definition of the initial conditions we set $\overline{N}^1_i(1) = 0$ for $i \in \mathcal{U}$.

Having completed the construction of $\{\overline{\mathbf{N}}^n(1)\}_{n=1}^{\infty}$ it remains to show that this process is substable when (19) holds. It will follow that under the same initial conditions the irreducible and aperiodic Markov chain $\{\mathbf{N}^n(1)\}_{n=1}^{\infty}$ is substable and therefore ergodic. The fact that $\mathbf{N}^n(j)$ is ergodic for all $j \in \mathcal{M}$ will follow from Theorem 3.

To show the substability of $\{\overline{\mathbf{N}}^n(1)\}_{n=1}^{\infty}$ in the presence of (19), recall first from Lemma 5 (i), that it suffices to show that for all $i \in \mathcal{M}$, the process $\{\overline{N}^n_i(1)\}_{n=1}^{\infty}$ is substable. For $i \in \mathcal{S}$ the stability of $\{\overline{N}^n_i(1)\}_{n=1}^{\infty}$ follows by the construction of the initial conditions. Indeed, for all $n \geq 1$, $\{\overline{N}^n_i(1)\}_{i \in \mathcal{S}}$ has distribution $\pi$. It remains to show the substability of a persistent queue, that is, of the process $\{\overline{N}^n_i(1)\}_{n=1}^{\infty}$ for $i \in \mathcal{U}$. The idea is to upper bound the process $\{\overline{N}^n_i(1)\}_{n=1}^{\infty}$, $i \in \mathcal{U}$, by a process which satisfies (17) and then to show stability of the last process using Lemma 6.

Let $T^n(i)$ be the time of the $n$th visit of the token to queue $i$. Let also $X^n_1(i)$, $X^n_2(i)$ be the number of arrivals to queue $i$ in the intervals $[T^n(1),\ T^n(i))$ and $[T^n(i),\ T^{n+1}(1))$ respectively. Define also $X^n(i) = X^n_1(i) + X^n_2(i)$. The process $\overline{N}^n_i(1)$, $i \in \mathcal{U}$ satisfies the relations,

$$\overline{N}^{n+1}_i(1) = \max\left(\overline{N}^n_i(1) + X^n_1(i) - l,\ 0\right) + X^n_2(i) \leq \max\left(\overline{N}^n_i(1) + X^n(i) - l,\ X^n(i)\right). \tag{24}$$

Define now the process $\{M^n_i\}_{n=1}^{\infty}$ as follows. Let $M^1_i = 0$, and

$$M^{n+1}_i = \max(M^n_i + X^n(i) - l,\ X^n(i)), \quad n = 1, 2, \cdots \tag{25}$$

By definition, $X^n(i)$ is the number of Poisson arrivals at queue $i$ in the cycle $C^n_{\mathcal{S}}(1)$. Since the sequence $C^n_{\mathcal{S}}(1)$ is stationary by the above construction, and independent of the arrival process at queue $i$, using Lemma 6 we have that the process $\{M_i^n\}_{n=1}^{\infty}$ is stable if

$$\lambda_i < \frac{\ell_i}{EC^1_{\mathcal{S}}(1)} = \frac{\ell_i}{u_0 + \sum_{k\in\mathcal{U}} \ell_k s_k}\left(1 - \sum_{k\in\mathcal{S}} \rho_k\right) \qquad i \in \mathcal{U} \ , \tag{26}$$

where the equality in (26) follows from the fact that by Theorem 3,

$$EC^1_{S}(1) = \frac{\hat{u}_0}{1 - \sum_{j\in\mathcal{S}} \rho_j} = \frac{u_0 + \sum_{j\in\mathcal{U}} \ell_j s_j}{1 - \sum_{j\in\mathcal{S}} \rho_j} \ . \tag{27}$$

Note that (19) implies (26). Since $M_i^1 = \overline{N}_1^1(1) = 0$, using (24), (25) it follows that

$$\overline{N}_i^n(1) \le M_i^n, \quad n = 1, 2, \cdots, \quad i \in \mathcal{U}.$$

Therefore, the process $\{\overline{N}_i^n(1)\}_{n=1}^{\infty}, \quad i \in \mathcal{U}$, is substable under our hypothesis (19).

Putting everything together, from (21) and (26) we finally conclude that the Markov chain $\mathbf{N}^n(j)$ is ergodic for every $j \in \mathcal{M}$ if *for some* partition $\mathcal{P} = (\mathcal{U}, \mathcal{S})$ the inequality (19) holds. Therefore, we conclude that the stability region $\mathcal{R}$ of the whole system becomes

$$\mathcal{R} = \bigcup_{S\subset\mathcal{M}} \mathcal{R}_{\mathcal{S}} \ , \tag{28}$$

where

$$\mathcal{R}_{\mathcal{S}} = \{\lambda = (\lambda_1, \ldots, \lambda_M) : \text{condition (19) holds}\} \ . \tag{29}$$

The union in (28) ranges over all nonempty strict subsets of $\mathcal{M}$. Finally, to complete the proof we need to show that

$$\bigcup_{\mathcal{S}\subset\mathcal{M}} \mathcal{R}_{\mathcal{S}} = \{\lambda = (\lambda_1, \ldots, \lambda_M) : \lambda_i < \frac{\ell_i}{u_0}(1 - \sum_{k=1}^{M} \rho_k) \ \ i \in \mathcal{M}\} \ . \tag{30}$$

This requires only algebraic manipulations, and it is delayed till Appendix B (to assist the reader to see graphically how (30) arises, we construct in the example below the stability region $\mathcal{R}$ from $\mathcal{R}_{\mathcal{S}}$ for $M = 2$ users). This completes the proof of Theorem 7. ∎

To illustrate the construction of the stability region $\mathcal{R}$ in the simplest possible case, we discuss $M = 2$ users token passing ring.

EXAMPLE. *Stability Region for $M = 2$*

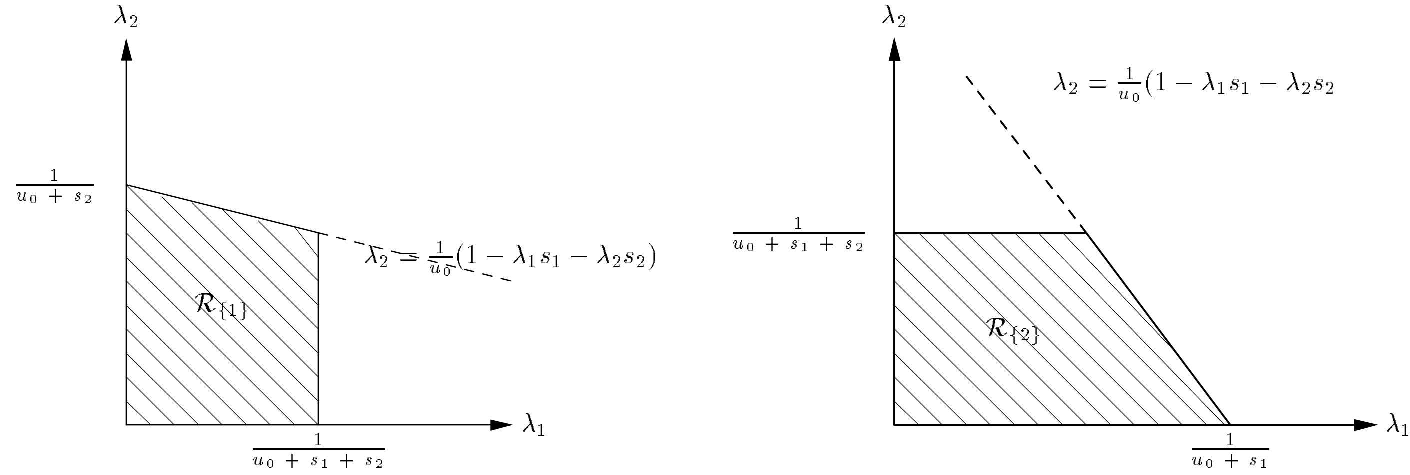


Figure 3: Stability subregions $\mathcal{R}_{\{1\}}$ and $\mathcal{R}_{\{2\}}$ for $M = 2$ users system. Stability region $\mathcal{R}$ is $\mathcal{R} = \mathcal{R}_{\{1\}} \cup \mathcal{R}_{\{2\}}$.

Let us assume $M = 2$ and $\ell_1 = \ell_2 = 1$. Consider first $\mathcal{S} = \{1\}$. In this case, (19) leads to

$$\begin{aligned} \lambda_1 &< \frac{1}{u_0 + s_1 + s_2} \\ \lambda_2 &< \frac{1 - \lambda_1 s_1}{u_0 + s_2}\ , \end{aligned}$$

which defines region $\mathcal{R}_{\{1\}}$ shown in Figure 3. In a similar manner, for $\mathcal{S} = \{2\}$ we have

$$\begin{aligned} \lambda_1 &< \frac{1 - \lambda_2 s_2}{u_0 + s_1} \\ \lambda_2 &< \frac{1}{u_0 + s_1 + s_2}\ , \end{aligned}$$

which leads to the stability region $\mathcal{R}_{\{2\}}$ shown also in Figure 3. The total stability region $\mathcal{R}$ is the union of both regions $\mathcal{R}_{\{1\}}$ and $\mathcal{R}_{\{2\}}$, that is, $\mathcal{R} = \mathcal{R}_{\{1\}} \cup \mathcal{R}_{\{2\}}$. From Figure 3, it is easy to see that $\mathcal{R}$ can be equivalently represented as

$$\begin{aligned} \lambda_1 &< \frac{1}{u_0}(1 - \lambda_1 s_1 - \lambda_2 s_2) \\ \lambda_2 &< \frac{1}{u_0}(1 - \lambda_1 s_1 - \lambda_2 s_2)\ . \end{aligned}$$

This is in an agreement with (30).

In a similar manner, one can identify six different sets $\mathcal{S}$ in the case $M = 3$ and construct the stability regions $\mathcal{R}_{\mathcal{S}}$. The whole stability region in this case is presented in Figure 1. $\Box$

We can use Theorem 7 to establish some other stability results. We concentrate here on two problems. First, Theorem 7 can be extended to the process of queue lengths at arbitrary time instants, that is the process $\tilde{\mathbf{N}}(t) = (\tilde{N}_1(t), \cdots, \tilde{N}_M(t))$, where $\tilde{N}_i(t)$ is the queue length at queue $i$ at time $t$. The second extension deals with the *gated-unlimited service* discipline in which we set $\ell_i = \infty$ in assumption (A2).

Let us first consider the stability of $\tilde{\mathbf{N}}(t)$. Assume that $\mathbf{N}^n(1)$ is ergodic. Using the notation of the proof of Theorem 3, we have from (9) that $E\left(\sum_{k=1}^{R} C_1^k\right) < \infty$, i.e., the renewal process $\tilde{C}'^n$ of the *length of time* between two successive returns to state $\mathbf{0}$ of the process $\mathbf{N}^n(1)$ has finite expectation. Since the interarrival times are exponential, this renewal process is non-lattice. Since $\tilde{\mathbf{N}}(t)$ is regenerative with respect to $\tilde{C}^n$, we conclude that

**Corollary 8**. *The process $\tilde{\mathbf{N}}(t)$ is stable if (18) holds.* ■

The next result extends assumption (A2) to gated-unlimited service disciplines. In fact, the basic steps of the methodology presented here can be useful in establishing rigorously stability conditions for some other service disciplines such as *Bernoulli*, *geometric*, *time limited* and so forth (cf. Levy *et al.* [13], Takagi [27]). However, additional work may be needed to fill the various steps in each case. Recently, in [8] we rigorously proved stability condition for the time-limited token passing ring with non-preemptive discipline. We plan to extend this analysis to preemptive time-limited token passing rings. This latter case is particularly interesting from the theoretical point of view.

Here, we only concentrate on the extension to the gated-unlimited service discipline. Assume that a subset $\mathcal{M}_\infty$ of the queues employs the gated-unlimited service discipline and let $\mathbf{N}^n_\infty(i)$ be the vector of queue sizes at *all* stations when the token visits the $i$th queue for the $n$th time. Then we have the following

**Corollary 9**. *For $i \in \mathcal{M}$, the Markov chain $\mathbf{N}^n_\infty(i)$ is ergodic if $\rho_0 = \sum_{j=1}^{M} \rho_j < 1$ and*

$$\lambda_j < \frac{\ell_j}{u_0}(1-\rho_0), \quad j \in \mathcal{M} - \mathcal{M}_\infty. \tag{31}$$

**Proof.** For arbitrary $l > 0$, let $\mathbf{N}^n_l(i)$ be the queue sizes when the queues in $\mathcal{M}_\infty$ are employing the $\ell$-limited policy with threshold $l$. As in Theorem 3, define for $1 \le l \le \infty$, $\mathbf{N}^1_l(1) = \mathbf{0}$, $K^1_l = 0$,

$$K^{n+1}_l = \min\{m > K^n_l \ : \ \mathbf{N}^m_l(1) = \mathbf{0}\}$$

and $R^n_l = K^{n+1}_l - K^n_l$. In Appendix C we show that $R^1_\infty \le R^1_l$, $l = 1, 2, \cdots$. Therefore, for any $l$, if $\mathbf{N}^n_l(1)$ is ergodic, so is $\mathbf{N}^n_\infty(1)$. Since $\rho_0 < 1$ we can pick $l$ large enough so that

$$\lambda_j < \frac{l}{u_0}(1-\rho_0), \quad j \in \mathcal{M}_\infty \ .$$

This, together with (31) implies that the Markov chain $\mathbf{N}_l^n(1)$ is ergodic, which in turn implies that $\mathbf{N}_\infty^n(1)$ is ergodic. ■

Finally, we show in the next theorem that the conditions of Corollary 9 are also necessary for the ergodicity of the Markov chain $\mathbf{N}_\infty^n(i),\quad i \in \mathcal{M}$. In particular, this will establish the necessary condition for stability of the $\ell$-limited token passing ring, and therefore it completes the proof of our Proposition from the Introduction.

**Theorem 10.** *If for some $i \in \mathcal{M}$ the Markov chain $\mathbf{N}_\infty^n(i)$ is ergodic, then $\mathbf{N}_\infty^n(j)$ is ergodic for every $j \in \mathcal{M}$. Moreover, $\sum_{j=1}^M \rho_j < 1$, and*

$$\lambda_j < \frac{\ell_j}{u_0}(1-\rho_0), \quad j \in \mathcal{M} - \mathcal{M}_\infty$$

**Proof.** The first assertion follows from Theorem 3 and the remark following that theorem. The second assertion is well known when $\mathcal{M} = \mathcal{M}_\infty$, Eisenberg [6]. Assume now that $\mathcal{M} \neq \mathcal{M}_\infty$. Without loss of generality, let $\ell_1 < \infty$. The fact that $\sum_{j=1}^M \rho_j < 1$ follows from Theorem 3 (see also the remark following Theorem 3). All cycles in the following will refer to queue 1. For simplicity of notation we omit the queue index from the various variables. Let us define,

$C^n$**:** length of the $n$th cycle.

$C^n(r)$**:** length of the $n$th cycle during which $r$ customers from queue 1 were served.

$M^n(r)$ **:** number of cycles in regeneration cycle $R^n$, (see proof of Theorem 3 for the definition of $R^n$) during which $r$ customers from queue 1 were served. Clearly,

$$R = \sum_{r=0}^{\ell_1} M(r), \tag{32}$$

where $M(r) = M^1(r)$ and $R = R^1$.

Since (by the ergodicity of the chain $\mathbf{N}_\infty^n(1)$) $ER < \infty$, using regenerative arguments again we have the following formulas for the long run averages:

- average length of a cycle during which $r$ customers from queue 1 were served,

$$EC(r) = \lim_{n\to\infty} \frac{\sum_{k=1}^n C^k(r)}{n} = \frac{E(\sum_{k=1}^{M(r)} C^k(r))}{EM(r)}\ ; \tag{33}$$

- probability (proportion) of cycles during which $r$ customers from queue 1 were served,

$$P(r) = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} M^k(r)}{n} = \frac{EM(r)}{ER}. \tag{34}$$

Consider now the following system.

**System S.** Upon arrival of the token to queue 1, the number of customers (from queue 1) that will be served in the next cycle enters system **S**. These customers stay in **S** until the token visits queue 1 for the next time, at which time all customers depart.

Clearly, the number of customers that enter system **S** in the $n$th cycle is $L^n$. Let $A_S^t$ be the number of customers that arrived in system **S** by time $t$. Recall the definition of the renewal process $\tilde{C}^n$ in the paragraph before Corollary 8. $A_S^t$ is regenerative with respect to $\tilde{C}^n$, and the ergodicity of $\mathbf{N}_\infty^n(1)$ implies by Theorem 3 that $E\tilde{C}^n < \infty$. Hence we have that

$$\lambda_S = \lim_{t\to\infty} \frac{A_S^t}{t} = \frac{E\left(\sum_{k=1}^{R} L^k\right)}{E\left(\sum_{k=1}^{R} C^k\right)} = \lambda_1, \tag{35}$$

where the last equality follows from (9). Similarly, we have the following formulas for the long-run average queue size $EN_S$, and the long-run average waiting time $EW_S$, in system **S**.

$$EN_S = \frac{E\left(\sum_{r=1}^{\ell_1} r \sum_{k=1}^{M(r)} C^k(r)\right)}{E\left(\sum_{k=1}^{R} C^k\right)} = \frac{\sum_{r=1}^{\ell_1} rE\left(\sum_{k=1}^{M(r)} C^k(r)\right)}{ECER} \tag{36}$$

$$EW_S = \frac{E\left(\sum_{r=1}^{\ell_1} r \sum_{k=1}^{M(r)} C^k(r)\right)}{E\left(\sum_{r=1}^{\ell_1} rM(r)\right)} = \frac{\sum_{r=1}^{\ell_1} rE\left(\sum_{k=1}^{M(r)} C^k(r)\right)}{\sum_{r=1}^{\ell_1} rE\left(M(r)\right)}. \tag{37}$$

We explain the middle term in (36). Let $N_S(t)$ be the queue size in system $S$ at time $t$. Observe that $r$ customers were served during $C^k(r)$ and the waiting time in the system **S** of each of these customers is $C^k(r)$. Since there are $M(r)$ cycles of length $r$ in a regeneration cycle, the sum in the nominator is the sum of the waiting times of the customers that were served in system $S$ during a regeneration cycle. Since there are no customers in $S$ when a regeneration cycle starts and ends, this sum is equal to the integral of $N_S(t)$ during a regeneration cycle. The denominator is simply $E\tilde{C}^1$. Therefore, the ratio in the middle term of (36) is exactly the ratio required by the regenerative theorem. The middle term in (37) is similarly explained. Note that since $\tilde{C}^1 = \sum_{r=1}^{\ell_1}\sum_{k=1}^{M(r)} C^k(r)$ and $E\tilde{C}^1 < \infty$, we have that $E\left(\sum_{k=1}^{M(r)} C^k(r)\right) < \infty, \quad r = 1, \cdots, \ell_1$. Also, since $\lambda_1 > 0$, it follows from

(35) that $0 < E\left(\sum_{k=1}^{R} L^k\right) = E\left(\sum_{r=1}^{\ell_1} M(r)\right)$. From the previous discussion we see that $EN_S < \infty$ and $EW_S < \infty$. Using (33), (34), we derive from (36), (37),

$$EN_S = \frac{\sum_{r=1}^{\ell_1} rP(r)EC(r)}{EC}, \tag{38}$$

$$EW_S = \frac{\sum_{r=1}^{\ell_1} rP(r)EC(r)}{\sum_{r=1}^{\ell_1} rP(r)}. \tag{39}$$

Since $0 < \sum_{r=1}^{\ell_1} rP(r) \leq \ell_1(1-P(0))$ (the first inequality is implied by the fact that $\lambda_1 > 0$), we conclude from (39) that

$$EW_S \geq \frac{\sum_{r=1}^{\ell_1} rP(r)EC(r)}{\ell_1(1-P(0))}. \tag{40}$$

Using Little's law (cf. Stidham [22]), (35) (38) and (40) we have

$$EN_S = \lambda_S EW_S \geq \lambda_1 \frac{EN_S EC}{\ell_1(1-P(0))},$$

and therefore,

$$\lambda_1 EC \leq \ell_1(1-P(0)). \tag{41}$$

We claim that $P(0) > 0$. Otherwise, it follows from (34) that $EM(0) = 0$ which implies that $P(L^n \geq 1, \quad n = 1, 2\cdots) = 1$. Since $L^n \geq 1$ if and only if $N^n(1) \geq 1$, we have that $P(N^n(1) = 0, \quad n = 1, 2\cdots) = 0$ which contradict the ergodicity of the chain $\mathbf{N}^n(1)$. Using the last observation, (41) implies that $\lambda_1 < \ell_1/EC$, as desired. ■

## APPENDIX A

In this appendix we show that the if the Markov chain $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1)\}_{n=1}^{\infty}$ is indecomposable then so is $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1), C^{n-1}_{\mathcal{S}}(1)\}_{n=2}^{\infty}$. Recall that a Markov chain with values in the measurable space $(X, \mathcal{F})$ and the transition probability $p(x, B)$, is called indecomposable if there are no two disjoint nonempty measurable sets $B_1,\ B_2$ that are closed. A measurable set $B$ is closed if $p(x, B) = 1$ for all $x \in B$.

The state space of the Markov chain under consideration is $\mathbb{N}_0^{|\mathcal{S}|} \times \mathbb{R}$ and the transition probability has the special structure $p((\mathbf{x}, y), B) = q(\mathbf{x}, B),\ \mathbf{x} \in \mathbb{N}_0^{|\mathcal{S}|},\ y \in \mathbb{R}$. Let us assume that there are two disjoint, nonempty, closed measurable set $B_1,\ B_2$ in $\mathbb{N}_0^{|\mathcal{S}|} \times \mathbb{R}$. Then

$$q(\mathbf{x}, B_i) = 1 \quad \text{for all} \quad (\mathbf{x}, y) \in B_i, \quad i = 1, 2. \tag{A1}$$

Let $\Pi(B_i)$ be the projection of $B_i$ on $\mathbb{N}_0^{|\mathcal{S}|}$. If $\mathbf{x} \in \Pi(B_1)$ then by (A1), $q(\mathbf{x}, B_1) = 1$ and since $B_1 \cap B_2 = \emptyset$ we conclude that $q(\mathbf{x}, B_2) = 0$. Therefore, $\mathbf{x}$ does not belong to

$\Pi(B_2)$ (again by (A1)) and so, $\Pi(B_1) \cap \Pi(B_2) = \emptyset$. Also, since $B_i \neq \emptyset$, we have $\Pi(B_i) \neq \emptyset$. Finally, since $B_i \subset \Pi(B_i) \times I\!R$, we have that for all $\mathbf{x} \in \Pi(B_i)$, $q(\mathbf{x}, \Pi(B_i) \times I\!R) = 1$. But $q(\mathbf{x}, \Pi(B_i) \times I\!R) = \Pr\{\overline{\mathbf{N}}^2_{\mathcal{S}}(1) \in \Pi(B_i) | \overline{\mathbf{N}}^1_{\mathcal{S}}(1) = \mathbf{x}\}$ and the discussion in this paragraph shows that with respect to the Markov chain $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1)\}_{n=1}^{\infty}$, the sets $\Pi(B_1)$, $\Pi(B_2)$ are closed, nonempty and disjoint. The last statement, however, contradicts the fact that $\{\overline{\mathbf{N}}^n_{\mathcal{S}}(1)\}_{n=1}^{\infty}$ is indecomposable.

## APPENDIX B

We prove (30). Let $\mathcal{M}_i = \mathcal{M} - \{i\}$, and denote the RHS of (30) as $\tilde{\mathcal{R}}$, that is,

$$\tilde{\mathcal{R}} = \{\lambda = (\lambda_1, \ldots, \lambda_M) : \lambda_i < \frac{\ell_i}{u_0}(1 - \sum_{k=1}^{M} \rho_k) \;\; i \in \mathcal{M}\} \; . \tag{B1}$$

We will prove that

$$\bigcup_{i=1}^{M} \mathcal{R}_{\mathcal{M}_i} = \tilde{\mathcal{R}} \; . \tag{B2}$$

At first, we show that for every $i \in \mathcal{M}$ we have $\mathcal{R}_{\mathcal{M}_i} \subset \tilde{\mathcal{R}}$. Let $\lambda \in \mathcal{R}_{\mathcal{M}_i}$. Then, by (19)

$$\lambda_i < \frac{\ell_i}{u_0 + \ell_i s_i} \left(1 - \sum_{k \neq i} \rho_k\right)$$

and this is equivalent to

$$\lambda_i < \frac{\ell_i}{u_0}(1 - \rho_0) \; . \tag{B3}$$

But (B3) implies that $u_0 \rho_i \leq \ell_i s_i (1 - \sum_{k=1}^{M} \rho_k)$. Using this, after some algebra, we conclude that the following also holds

$$\frac{\left(1 - \sum_{k \neq i} \rho_k\right)}{u_0 + \ell_i s_i} \leq \frac{\left(1 - \sum_{k=1}^{M} \rho_k\right)}{u_0} \; . \tag{B4}$$

Therefore, for every $j \in \mathcal{M}_i$ we have

$$\lambda_j < \frac{\ell_j \left(1 - \sum_{k \neq i} \rho_k\right)}{u_0 + \ell_i s_i} \leq \frac{\ell_j \left(1 - \sum_{k=1}^{M} \rho_k\right)}{u_0}, \tag{B5}$$

and (B3), (B5) imply $\mathcal{R}_{\mathcal{M}_i} \subset \tilde{\mathcal{R}}$, as needed.

Now we prove that $\bigcup_{i=1}^{M} \mathcal{R}_{\mathcal{M}_i} \supset \tilde{\mathcal{R}}$. Note that

$$\mathcal{R}_{\mathcal{M}_i} = \{\lambda \in I\!R^M : \; \rho_j < \frac{\ell_j s_j (1 - \sum_{k \neq j} \rho_k)}{u_0 + \ell_j s_j}, \;\; j = 1, \cdots, M\} \; . \tag{B6}$$

Let $\lambda \in \widetilde{\mathcal{R}}$, and let $k$ be such that $\rho_k/(\ell_k s_k) \geq \rho_j/(\ell_j s_j)$ for all $j \in \mathcal{M}$. Then, $\lambda \in \widetilde{\mathcal{R}}$ implies $\rho_j u_0 < \ell_j s_j(1-\rho_0) + \rho_k \ell_j s_j - \rho_j \ell_k s_k, \ \ j = 1, \cdots, M$, and this leads to $\lambda \in \mathcal{R}_{\mathcal{M}_k}$, as required. This proves (B2).

Observe that as far as the proof of Theorem 7 is concerned, (B2) is sufficient since we know that the system is stable when the vector of arrival rates is in $\bigcup_{i=1}^{M} \mathcal{R}_{\mathcal{M}_i}$. Since the stability condition is also necessary by Theorem 10, the stability region cannot be larger and this shows indirectly that $\bigcup_{\mathcal{S} \subset \mathcal{M}} \mathcal{R}_{\mathcal{S}} = \bigcup_{i=1}^{M} \mathcal{R}_{\mathcal{M}_i}$. This equality can also be proved directly by applying similar algebraic manipulations as above. Details are left to the reader.

## APPENDIX C

Following the notation in Levy et.al. [13] we denote by $\mathbf{f} = \{f_n(\cdot)\}, \ n = 1, 2, \cdots$, a gated policy that during the $n$th visit of the token to any queue serves $f_n(x)$ customers, where $x$ is the number of customers in that queue at the instant of token arrival. Let $\widetilde{\mathbf{N}}_n^{\mathbf{f}}$ be the vector of queue sizes at all queues at the instant of the $n$th token arrival to any queue. Let also let $T_n^{\mathbf{f}}, \ D_n^{\mathbf{f}}$ be the successive instants of token arrival and departure to any queue under a policy $\mathbf{f}$. The following Lemma is derived by a simple modification of the proof of Theorem 1 in Levy et.al. [13]

**Lemma C.** *Let* $\mathbf{f} = \{f_n(\cdot)\}, \ n = 1, 2, \cdots$ *and* $\mathbf{g} = \{g_n(\cdot)\}, \ n = 1, 2, \cdots$ *be two gated-type policies that serve customers in a queue in FCFS order. Assume that the system is empty at* $n = 0$ *and that the two policies operate with the same realizations of arrivals, service times, switchover times and and the same polling order. If* $f_n(x) \geq g_n(x), \ n = 1, 2, \cdots$ *and* $\mathbf{g}$ *is monotonic and contractive policy then*

$$\widetilde{\mathbf{N}}_k^{\mathbf{g}} = (0, \ldots, 0) \quad \textit{implies that} \quad T_k^{\mathbf{f}} = T_k^{\mathbf{g}} \quad \textit{and} \quad \widetilde{\mathbf{N}}_k^{\mathbf{f}} = (0, \ldots, 0).$$

**Proof.** Let $W(t)$ be the sum of the service times (total work) of the customers that arrive in the system (that is, in all queues) by time $t$ (notice that $W(t)$ is independent of the policy.) We need only consider $k \geq 2$ since the case $k = 1$ is obvious by the initial conditions. If $\widetilde{\mathbf{N}}_k^{\mathbf{g}} = (0, \ldots, 0)$ for some $k \geq 2$, that is, if all queues are empty at the instant of the $k$th arrival of the token to a queue, then

$$W(T_k^{\mathbf{g}}) = T_k^{\mathbf{g}} - \sum_{j=1}^{k-1} \widetilde{U}_j, \tag{C1}$$

where $\widetilde{U}_n, \ n = 1, 2, \cdots$, are the successive switchover times, that is, with $k_n, \ J_n$ as defined in the paragraph before Theorem 1, $\widetilde{U}_n = U_{J_n}^{k_n}$. We claim now that $T_k^{\mathbf{f}} \leq T_k^{\mathbf{g}}$. If this holds, then taking into account the fact that $D^{\mathbf{f}}_{k-1} \geq D^{\mathbf{g}}_{k-1}$ (see Lemma 1 in Levy et.al [13]) and that the switchover times are identical we will have $T^{\mathbf{f}}_k \geq T^{\mathbf{g}}_k$ and therefore, $T^{\mathbf{f}}_k = T^{\mathbf{g}}_k$. But then, (C1) implies $\tilde{\mathbf{N}}^{\mathbf{f}}_k = (0, \cdots, 0)$ as desired.

To see that $T^{\mathbf{f}}_k \leq T^{\mathbf{g}}_k$, assume the contrary, namely $T^{\mathbf{f}}_k > T^{\mathbf{g}}_k$. Then, if $I^{\mathbf{f}}(t)$ is the amount of time that the token is idle (switching from one queue to another) up to time $t$ under policy $\mathbf{f}$, we would have that $\sum_{j=1}^{k-1} \tilde{U}_j > I^{\mathbf{f}}(T^{\mathbf{g}}_k)$ (recall that $P(\tilde{U}_j > 0) = 1$). Since we also have $W(t) \geq t - I^{\mathbf{f}}(t),\ t \geq 0$, we would have that

$$W(T^{\mathbf{g}}_k) > T^{\mathbf{g}}_k - \sum_{j=1}^{k-1} \tilde{U}_j,$$

which contradicts (C1). ∎

If we now identify $f_n(x) = x$ and $g_n(x) = \min(x, l_n)$, where $l_n = \ell_{J_n}$, we immediately have from Lemma C that $R^1_\infty \leq R^1_l$, which is what we need in Corollary 9.

**ACKNOWLEDGMENT**

We are grateful to two referees and Professor Boxma for many suggestions and comments that improved substantially the clarity of the presentation.

# References

[1] S. Asmussen, *Applied Probability and Queues*, John Wiley&Sons, Chichester 1987.

[2] O.J. Boxma, W.P. Groenendijk, Two queues with alternating service and switching times, In: O.J. Boxma, R. Syski (eds.) *Queueing Theory and its Applications - Liber Amicorum for J.W. Cohen*, North-Holland, Amsterdam, 1988, p. 261-282.

[3] L. Breiman, *Probability*, Addison-Wesley Pub. Company, Reading, Mass., 1968.

[4] K.L. Chung, *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag 1967.

[5] E. Coffman, Jr., and E. Gilbert, A Continuous Polling Systems with Constant Service Time, *IEEE Trans. Information Theory*, 32 (1986) 584-591.

[6] M. Eisenberg, Queues with Periodic Service and Changeover Time, *Operations Research*, 20, (1972) pp. 440-451.

[7] G. Fayolle, On random walks arising in queueing systems: Ergodicity and transience via quadratic forms as Lyapunov functions. Part I., *Queueing Systems*, 5, (1989) pp. 167-184.

[8] L. Georgiadis and W. Szpankowski, Stability Criteria for Yet Another Multidimensional Distributed System, Purdue University, CSD-TR-91-071, 1991.

[9] O. Ibe and X. Cheng, Stability Conditions for Multiqueue Systems with Cycle Service, *IEEE Trans. Automatic Control*, 33 (1988) 102-103.

[10] M. Karatzoglu and A. Ephremides, Ergodicity of $M$-dimensional Random Walks and Random Access Systems, preprint, 1989.

[11] L. Kleinrock and H. Levy, The Analysis of Random Polling Systems, *Oper. Res.* 36 (1988) 716-732.

[12] P. Kuehn, Multiqueue Systems with Nonexhaustive Cycle Service, *The Bell System Technical Journal*, 58 (1979) 671-698.

[13] H. Levy, M. Sidi and O. Boxma, Dominance Relations in Polling Systems, *Queueing Systems*, 6 (1990) 155-172.

[14] R. Loynes, The stability of a queue with non-independent inter-arrival and service times, *Proc. Camb. Philos. Soc.*, 58 (1962), 497–520.

[15] V.A. Malyshev, Classification of two-dimensional positive random walks and almost linear semimartingals, *Dokl. Akad., Nauk SSSR*, 22.3 (1972), 136–138.

[16] V.A. Malyshev and M.V Mensikov, Ergodicity, continuity and analyticity of countable Markov chains, *Trans. Moscov Math. Soc.* (1981), 1-18.

[17] Mensikov, M.V., Ergodicity and transience conditions for random walks in the positive octant of space, *Soviet. Math. Dokl.*, 15 (1974), 1118–1121.

[18] D. R. Miller, Existence of Limits in Regenerative Processes *The Annals of Mathematical Statistics*, 43, (1972), 1275-1282.

[19] S. Meyn and R.L. Tweedie, Criteria for Stability of Markovian Processes I: Discrete Time Chains, *Adv. Appl. Prob.*, 1992.

[20] Nevelson, M.B. and Hasminskii, R.Z., *Stochastic Approximation and Recursive Estimation*, American Mathematical Society, vol. 47, Providence. Rhode Island, 1973

[21] W. Rosenkrantz, Ergodicity conditions for two dimensional Markov chains on the positive quadrant, *Probab. Theory and Related Fields*, 83, (1989), 309-319.

[22] S. Stidham, Jr., A last word on $L = \lambda W$. *Oper. Res.,* 22 (1974), pp. 417-421.

[23] W. Szpankowski, Stability conditions for multi-dimensional queueing systems with computer applications, *Oper. Res*, 36, (1988), 944–957.

[24] W. Szpankowski, Towards computable stability criteria for some multidimensional stochastic processes, in *Stochastic Analysis of Computer and Communication Systems* (ed. H. Takagi), Elsevier Science Publications B. V. (North-Holland), (1990), pp. 131-172.

[25] W. Szpankowski, Towards Stability Criteria for Multidimensional Distributed Systems: The Buffered ALOHA Case, Purdue University, CSD TR-983 (1990); revised (1991).

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

[27] H. Takagi, Queueing Analysis of Polling Models, *ACM Computing Surveys*, 20 (1988) 5-28.

[28] R. L. Tweedie, Criteria for classifying general Markov chains, *Adv. Appl. Prob.*, 8 (1976), 737–771.

[29] J. Walrand, *An Introduction to Queueing Networks*, Prentice Hall, New Jersey 1988.

[30] K. Watson, Performance Evaluation of Cyclic Service Strategies – A Survey. *Proc. PERFORMANCE'84*, E. Gelenbe Ed. (Paris 1984).