Stability of Multi-server Polling Models

Christine Fricker, M.R. Jaïbi

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Stability of multi-server polling models

Christine Fricker et M.R. Jaïbi

N ° 3347

Janvier 1998

THÈME 1

Rapport de recherche

ISSN 0249-6399

Stability of multi-server polling models

Christine Fricker * et M.R. Jaïbi †

Thème 1 — Réseaux et systèmes
Projet Méval

Rapport de recherche n° 3347 — Janvier 1998 — 17 pages

Abstract: The stability of polling models with several servers and Markovian routing of the servers is analyzed. Customers arrive according to independent renewal processes. The service times and the switch-over times are independent with general distributions, independent of the servers. Service policies are attached to the queues. We established the necessary and sufficient condition for stability for general service policies covering the classical ones. The sufficiency of the condition is proved using the associated fluid model. For the necessity of the condition, a straightforward argument is given. Moreover, local stability is given.

Key-words: polling models, Markovian scheme, multi-dimensional Markov chains, ergodicity of Markov processes, fluid models.

(Résumé : tsvp)

* Christine.Fricker@inria.fr
† Département de Mathématiques, Faculté des Sciences, Campus Universitaire, 1060 Tunis, Tunisie

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Stabilité des systèmes à polling à plusieurs serveurs

Résumé : On étudie la stabilité d' un modèle général de systèmes à polling à plusieurs serveurs, ces derniers ayant des routages Markoviens. Les clients arrivent aux différentes files suivant des processus de renouvellement indépendants. Les durées de services et de déplacements des serveurs entre les files sont indépendantes avec des lois générales, indépendantes des serveurs. Les politiques de service sont définies par file. Nous considérons des politiques de service générales recouvrant les politiques de service classiques. Nous établissons la condition nécessaire et suffisante de stabilité d'un tel modèle. On prouve que la condition est suffisante en utilisant le modèle fluide associé. Pour prouver que la condition est nécessaire, un argument direct est donné. De plus, on donne les résultats de stabilité locale.

1 Introduction

A polling system is a set of queues separately served by one or several servers. When a server moves from a queue to another queue, it takes a time called switch-over time. The main specifications of a polling model are the way in which the servers attend to the queues, the service policies that are applied (the policy indicates the customers that are served in a queue during one visit of a server), and the statistical assumptions on the arrivals to the queues, the required service times and the switch-over times between the queues.

An important issue is the derivation of the stability condition. In the case of single server models, this question has been studied by many authors. Under the usual independence assumptions and Poisson arrivals, Georgiadis and Szpankovsky [14] studied the l-limited token ring, Fricker and Jaibi [12, 13] the periodic and the Markovian routing with monotonic service policies by monotonicity arguments, and Borovkov and Schassberger [1] the Markovian routing with the limited gated policy by Lyapounov functions (see also Kotler [15]). For a general stationary input, Massoulie [17] gives a sufficient but not necessary condition of stability, Foss and Last studied a state-dependent routing (for example [11]) and, for i.i.d inter-arrivals times, Dai and Meyn [5] establish that the stability condition of the l-limited one token ring is sufficient via the associated fluid model.

We consider a polling model with several servers and Markovian routing with general service policies verifying monotonicity and "work conserving" assumptions. We establish the stability condition for this model. The proof of sufficiency here involves the fluid model associated to Markov processes, developed by many authors: Malyshev [16], Dupuis and Williams [8], and for queueing networks Chen and Mandelbaum [3], Rybko and Stolyar [18], Dai [4], Bramson [2] and Dumas [7] and others. Our proof of the stability generalizes the proof of Dai and Meyn [5]. The fluid approach has been applied to other polling models by Foss and Kovalevskii [10] and by Down [6] and Foss, Kovalevskii and Chernova [9] who investigates the stability of multiple-server polling models. For the necessity of the condition, the straightforward argument developed in [12] applies.

The model is composed of c infinite-buffered queues, labeled from 1 to c, attended to by s servers. The servers attend to the queues according to independent Markov chains $(X_n^i)_n$ $(1 \leq i \leq s)$ on $\{1, \ldots, c\}$ with the same transition matrix $(r_{k,l})$ and unique invariant measure (p_k). Each server stays serving at a queue a period of time called a visit and determined by the service policy which is attached to the queue; we consider general service policies as defined in the next section. At the completion of a visit to a queue, the transition of a server to (another) queue takes a time called switch-over time. Customers arrive to queue k at rate λ_k, and are served at rate μ_k. The sequences of inter-arrival times $(\tau_k(n))_n$, service times $(\sigma_k^i(n))_n$ and switch-over times $(\sigma_{k,l}^{0,i}(n))_n$ required in queue k (respectively from queue k to l) for server i are independent i.i.d. sequences with general distribution independent of i, with mean $\sigma_k = \frac{1}{\mu_k}$ (respectively $\sigma_{k,l}^0 = \frac{1}{\mu_{k,l}^0}$), and independent of the Markov chains $(X_n^i)_n$. Moreover inter-arrival times are assumed to be unbounded and spread out. For each k we assume $\rho_k = \lambda_k \sigma_k < s$ to ensure the stability of queue k when it operates as a standard M/G/s queue in isolation.

Our main result is that the polling model is stable if and only if

$$\mathcal{C}: \qquad \hat{\rho} + \max_{1 \leq k \leq c} \frac{\lambda_k}{p_k l_k} S < s$$

where $\hat{\rho} = \sum_{k=1}^{c} \rho_k$ is the total traffic load of the system, $S = \sum_{1 \leq k,l \leq c} p_k r_{k,l} \sigma^0_{k,l}$ is the expected switch-over time per visit and $l_k \leq \infty$ is the maximum expected number of customers served by a server in a visit to queue k (when there are infinitely many customers waiting at the queue) which may be infinite. The quantity l_k is a characteristic of the service policy attached to queue k and is defined in the next section. Then local stability results are stated: when the system is not stable, the subset of stable queues is given.

The paper is organized as follows. In Section 2, the service policies are precised and the state process which is Markov is defined. In Section 3, the associated fluid model is introduced and the main result is proved. In section 4, the local stability is analyzed.

2 Service policies and state process

2.1 Service policies

A service policy determines the (number of) customers that are served without interruption in a service period by the servers which attend to the queue.

For a single server, it is is defined through the (random) function f where $f(x, a)$ is *the (random) number of customers that are served without interruption when the server arrives to a queue and finds x customers waiting given the elapsed inter-arrival time a* (for more details see [12, 13]). Let $v(x, a)$ be the duration of the service period by a single server for initial condition (x, a). We consider general service policies satisfying the following properties.

i) There is work-conservation so that $v(x, a) = \sum_{l=1}^{f(x,a)} \sigma(l)$ with $f(0, a) = v(0, a) = 0$, where $(\sigma(l))_l$ is the i.i.d. sequence of requested service times.

ii) The selection of a customer for service is done independently of its particular service time and of the past up to the start of the service period. Thus the distribution (f, v) does not depend on the order in which the customers are served.

iii) The service policy is monotonic in the sense that for each $a \geq 0$ the numbers $f(x, a)$ are monotonic in distribution in x and their limit in distribution as $x \to \infty$ is a (possibly degenerate) random variable F^* which does not depend on a. In fact, F^* would be the number of customers served in a service period if there were infinitely many customers waiting at the start of the service period with $l = E(F^*) \leq \infty$. The integrability of this stochastic bound plays an important role in the analysis and leads to the following classification of the service policies:

Policy f is said to be of limited type $E(F^*) < \infty$ and of unlimited type otherwise.

When several servers are available at a queue, they provide service such that the first two properties (work-conservation and selection of the customers) and the bound F^* apply to each server. Moreover, eventually the number of customers served by each server is bounded in distribution from below by (random) $f^{min}(x)$ of the queue length x which also monotonically converges in distribution to F^* as $x \to \infty$. This last requirement makes that in a heavily charged queue, all available servers are "heavily charged".

The four properties are satisfied by the classical policies, for example by the exhaustive and the gated policies, some limited versions of these policies and their Binomial versions.

Service policy f_k with bound F_k^* and $l_k = E(F_k^*) \leq \infty$ is attached to queue k ($k = 1, \dots, c$).

2.2 State process

We describe the polling system by the Markov process $(X(t))_{t \in \mathbf{R}}$ defined below. The state of the system at time $t \geq 0$ is given by

$$X(t) = (Q(t), P(t), A(t), R(t), C(t))$$

where
$Q(t) = (Q_k(t), 1 \leq k \leq c)$,
$P(t) = (P^i(t), 1 \leq i \leq s)$,
$A(t) = (A_k(t), 1 \leq k \leq c)$,
$R(t) = (R_k^i(t), R_{k,l}^{0,i}(t), 1 \leq k, l \leq c, 1 \leq i \leq s)$ and
$C(t) = (C_k^i(t), 1 \leq k \leq c, 1 \leq i \leq s)$
are defined as follows:
$Q_k(t)$ is the number of customers in queue k at time t,
$P^i(t)$ is the position of server i: $P^i(t) = k$ or (k, l) depending on whether server i is serving queue k or switching from queue k to queue l at time t,
$A_k(t)$ is the residual arrival time in queue k at time t,
$R_k^i(t)$ is the residual service time of the customer served by server i in queue k at time t, with $R_k^i(t) = 0$ if $P^i(t) \neq k$,
$R_{k,l}^{0,i}(t)$ is the residual switch-over time of server i from queue k to queue l at time t, with $R_{k,l}^{0,i}(t) = 0$ if $P^i(t) \neq (k, l)$, and
$C_k^i(t)$ is a complementary component relative to queue k and server i which is determined by the service policy in queue k and which makes the state process $X(t)$ a Markov process. Because the service policies only depend on numbers of customers, $C_k^i(t)$ must have a finite number of components, some of them being bounded by the number $Q_k(t)$ of customers in queue k and the others by random variables that are independent of the state of the system.

All the processes above are taken to be right-continuous.

The components C_k^i can be defined for general service policies. Let us define $C_k^i(t)$ for two "classical" policies in the context of several servers, assuming for simplicity FIFO service.
a) The exhaustive policy. Here each server serves until no more customers are waiting (there may be still customers in service if there are other servers at the queue). The components $C_k^i(t)$ ($i = 1, \dots, s$) are superfluous and we can take them equal to 0. Another possible definition is $C_k^i(t) = Q_k^w(t)$ where $Q_k^w(t) = Q_k(t) - \sum_{i=1}^s 1_{\{P^i(t)=k\}}$ is the number of waiting

customers and $C_k^i(t) = 0$ otherwise. Any server leaves the queue as soon as he completes a service and $C_k^i = 0$.

Consider now the L_k-limited exhaustive policy, where L_k is a non-negative integrable random variable, and where the limit L_k is relative to each server (no more than L_k customers may be served by a server in a visit). Let $(L_k^i(n))_n$ be independent sequences of i.i.d. non-negative random variables having the distribution of L_k. Here we define the process $C_k^i(t)$ by $C_k^i(t) = 0$ when server i is not in queue k at time t, C_k^i takes the value $(L_k^i(n)-1)^+$ at the n-th arrival of server i to queue k, decreases by one at each service completion by server i in queue k. No more customers of queue k are admitted for service by server i in the n-th visit of server i as soon as C_k^i reaches 0; if positive, C_k^i is reset to 0 at the end of the visit.

b) The gated policy. It can have different versions when there are several servers. Let us consider the first-server-gated policy which is such that the only customers that are served during a service session are those that were waiting at the start of the service session (when a server arrives to the queue and is the only server there); if other servers arrive to the queue during the session, they contribute to the service of those customers. Here $C_k^i(t)$ does not depend on i, $C_k^i(t) = 0$ if no server is in queue k at time t, $C_k^i(t) = (Q_k(t) - 1)^+$ if the n-th service session starts at time t and then decreases by one at each service completion in the queue; no more customers are admitted for service in the session as soon as $C_k^i(t) = 0$ (even in the case of an arrival of a server) and the session ends when all services under progress are resumed.

Another version is the last-server-gated policy: waiting customers are admitted for service during a service session at soon as they arrive before the last server who participate to the service session (all available servers contribute to the service of those customers). Here also $C_k^i(t)$ does not depend on i, $C_k^i(t) = 0$ if no server is in queue k, $C_k^i(t) = Q_k^w(t)$ if a server arrives to queue k at time t and C_k^i decreases by one at each service completion in the queue; no more customers are admitted for service in the n-th session as soon as $C_k^i(t) = 0$ and $Q_k(t) = Q_k^w(t)$.

It is also possible to define the $C_k^i(t)$ corresponding to the L_k-limited gated policies similarly to the limited exhaustive policy.

In the examples above, $C(t) \in \mathcal{C} = \mathbb{N}^{cs}$.

Process X has the strong Markov property and its state space is the product space

$$\mathcal{X} = \mathbb{N}^c \times E^s \times (\mathbb{R}^+)^c \times (\mathbb{R}^+)^{cs} \times (\mathbb{R}^+)^{c^2 s} \times \mathcal{C}$$

where E is the finite set $\{1, \dots c\}^2 \cup \{1, \dots c\}$ and $\mathcal{C}$ depends on the service policies.

Remark The component $R(t)$ of $X(t)$ may be simplified and replaced by a vector in $(\mathbb{R}^+)^s$ whose component i is either the residual service time or the residual switch-over time of server i depending on the value of $P^i(t)$. But our definition is convenient in the sequel.

3 Stability

For any $x \in \mathbf{X}$, we denote by X^x the Markov process X with initial condition $X^x(0) = x$; in the sequel we refer to initial condition $X^x(0) = x$ by adding the superscript ^{x}. On the product space $\mathbf{X} = \Pi_{i\in I}\mathbf{X_i}$, let $|x| = \sum_{i\in I}|x_i|$ where x_i $(i \in I)$ are the coordinates of x.
Let
$B_k^{i,x}(t)$ be the cumulative amount of time spent by server i serving queue k up to t,
$B_{k,l}^{0,i,x}(t)$ the cumulative amount of time server i is switching from queue k to queue l up to t.
It defines
$B^{i,x}(t) = (B_k^{i,x}(t), 1 \le k \le c)$,
$B^x(t) = (B_k^{i,x}(t), 1 \le k \le c, 1 \le i \le s)$,
$B^{0,i,x}(t) = (B_{k,l}^{0,i,x}(t), 1 \le k, l \le c)$ and
$B^{0,x}(t) = (B_{k,l}^{0,i,x}(t), 1 \le k, l \le c, 1 \le i \le s)$.
Let
$N_k^{i,x}(t)$ be the number of visit completions to queue k by server i between 0 and t,
$N_{k,l}^{0,i,x}(t)$ be the number of switch-over completions from queue k to queue l by server i between 0 and t,
$\mathcal{N}^{i,x}(t) = \sum_{k=1}^c N_k^{i,x}(t)$ be the number of visit completions by server i between 0 and t,
$\mathcal{N}^x(t) = \sum_{i=1}^s N^{i,x}(t)$ be the total number of visit completions between 0 and t and
$D_k^{i,x}(t)$ be the number of customers served by server i in queue k up to time t.

All the processes are taken right continuous with left limits. For any process H and any sequence $(x_n)_n \in \mathbf{X}$, let $\frac{1}{|x_n|}H^{x_n}(|x_n|t)$ be denoted by $\bar{H}^n(t)$. When it exists, the derivative of $\bar{H}(t)$ at t is denoted by $\dot{\bar{H}}\ (t)$ and point t is said to be regular. By $\xrightarrow{u.o.c.}$ we denote the uniform convergence on compact sets for right continuous with left limits real functions on $[0, \infty[$, and it means component by component when it concerns vectors.

Theorem and definition 1 *For any sequence $(x_n)_{n\in \mathbb{N}} \subset \mathbf{X}$ such that $|x_n| \to +\infty$, there is a subsequence $(x_{n_j})_j$ with $|x_{n_j}| \to +\infty$ such that:*
$(\bar{Q}^{n_j}(0), \bar{A}^{n_j}(0), \bar{R}^{n_j}(0))$ converges to some limit $(\bar{Q}(0), \bar{A}, \bar{R})$
and that with probability one

$$\left(\bar{Q}^{n_j}, \bar{B}^{n_j}, \bar{B}^{0,n_j}, \bar{N}^{n_j}, \bar{N}^{0,n_j}\right) \xrightarrow[j\to\infty]{u.o.c.} (\bar{Q}, \bar{B}, \bar{B}^0, \bar{N}, \bar{N}^0)$$

for some limit process $(\bar{Q}, \bar{B}, \bar{B}^0, \bar{N}, \bar{N}^0)$.

Any limit process $(\bar{Q}, \bar{B}, \bar{B}^0, \bar{N}, \bar{N}^0)$ is called a fluid limit of the system and satisfies the following:
(i) the components of $\bar{B}$ and $\bar{B}^0$ are non-decreasing with $\bar{B}(0) = \bar{B}^0(0) = 0$.
(ii) For all $t \ge 0$, $t = \sum_{k=1}^c \bar{B}_k^i(t) + \sum_{1\le k,l\le c} \bar{B}_{k,l}^{0,i}(t) \quad (1 \le i \le s)$.
(iii) $\bar{Q}_k(t) = \bar{Q}_k(0) + \lambda_k(t - \bar{A}_k)^+ - \mu_k \sum_{i=1}^s (\bar{B}_k^i(t) - \bar{R}_k^i)^+ \quad (1 \le k \le c)$.

(iv) Let $J(t) = \{k, l_k = +\infty$ and $\bar{Q}_k(t) > 0\}$. There exists $T_0 > 0$ (independent of the fluid limit) such that, for $t \geq T_0$ and whenever the derivatives exist,

- *when $J(t) = \emptyset$, then*

$$\text{if } \bar{Q}_k(t) = 0 \text{ then } \dot{\bar{\mathcal{B}}}_k(t) = \rho_k, \text{ and}$$
$$\text{if } \bar{Q}_k(t) > 0 \text{ then } \mu_k \, \dot{\bar{B}}_k^i(t) = l_k \, \dot{\bar{N}}_k^i(t) \quad (1 \leq i \leq s)$$

- *when $J(t) \neq \emptyset$, then $\dot{\bar{\mathcal{B}}}^i(t) = 1 \quad (1 \leq i \leq s)$*

where $\bar{\mathcal{B}}^i(t) = \sum_{k=1}^{c} \bar{B}_k^i(t)$ and $\bar{\mathcal{B}}_k(t) = \sum_{i=1}^{s} \bar{B}_k^i(t)$.

Proof. By definition, $x_n = (Q^{x_n}(0), \ldots, A^{x_n}(0), R^{x_n}(0), \ldots)$. Thus the sequence

$$(\frac{1}{|x_n|}Q^{x_n}(0), \frac{1}{|x_n|}A^{x_n}(0), \frac{1}{|x_n|}R^{x_n}(0))$$

is bounded by 1 and admits a subsequence converging to some limit $(\bar{Q}(0), \bar{A}, \bar{R})$. Given $x \in \mathbf{X}$, $x \neq 0$ and $t \geq s \geq 0$, we have for each $i \in \{1, \ldots s\}$ and for each $k \in \{1, \ldots c\}$:

$$\frac{1}{|x|}B_k^{i,x}(|x|t) - \frac{1}{|x|}B_k^{i,x}(|x|s) \leq t - s$$

and the same inequality holds for $B_{k,l}^{0,i,x}$. Because the components of $\frac{1}{|x|}(B^x(|x|t), B^{0,x}(|x|t))$ are nondecreasing functions of t, there exists a subsequence $(x_{n_j})_j$ such that

$$\left(\bar{B}^{n_j}(t), \bar{B}^{0,n_j}(t)\right) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} (\bar{B}(t), \bar{B}^0(t)) \;\; a.s.$$

for some process $(\bar{B}(t), \bar{B}^0(t)) = (\bar{B}_k^i(t), \bar{B}_{k,l}^{0,i}(t),\ 1 \leq k, l \leq c,\ 1 \leq i \leq s)$.

Let
$E_k^x(t) = \sup\{r,\ A_k(0) + \sum_{j=1}^{r-1} \tau_k(j) \leq t\}$ be the number of arrivals in queue k at time t,
$S_k^{i,x}(t) = \sup\{r,\ R_k^i(0) + \sum_{j=1}^{r-1} \sigma_k^i(j) \leq t\}$ and
$S_{k,l}^{0,i,x}(t) = \sup\{r,\ R_{k,l}^{0,i}(0) + \sum_{j=1}^{r-1} \sigma_{k,l}^{0,i}(j) \leq t\}$.

Because $E_k^x(t)$, $S_k^{i,x}(t)$ and $S_{k,l}^{0,i,x}(t)$ are counting functions of renewal processes, the sequences $(\bar{E}_k^n(t))_n$, $(\bar{S}_k^{i,n}(t))_n$ and $(\bar{S}_{k,l}^{0,i,n}(t))_n$ are relatively compact and for a subsequence, say $(x_{n_j})_j$ for notational simplicity, we have that a.s.

$$\bar{E}_k^{n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} \bar{E}_k(t) = \lambda_k(t - \bar{A}_k)^+, \tag{1}$$
$$\bar{S}_k^{i,n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} \bar{S}_k^i(t) = \mu_k(t - \bar{R}_k^i)^+, \tag{2}$$
$$\bar{S}_{k,l}^{0,i,n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} \bar{S}_{k,l}^{0,i}(t) = \mu_{k,l}^0(t - \bar{R}_{k,l}^{0,i})^+. \tag{3}$$

Furthermore

$$D_k^{i,x}(t) = S_k^{i,x}(B_k^{i,x}(t)) \leq S_k^{i,x}(t)$$

and

$$N_{k,l}^{0,i,x}(t) = S_{k,l}^{0,i,x}(B_{k,l}^{0,i,x}(t)) \leq S_{k,l}^{0,i,x}(t).$$

Because $(\bar{S}_k^{i,n}(t))_n$ and $(\bar{S}_{k,l}^{0,i,n}(t))_n$ are relatively compact and by the law of large numbers, a.s.

$$\bar{D}_k^{i,n_j}(t) \quad \xrightarrow[j\to\infty]{u.o.c.} \quad \bar{D}_k^i(t) = \mu_k(\bar{B}_k^i(t) - \bar{R}_k^i)^+. \tag{4}$$

$$\bar{N}_{k,l}^{0,i,n_j}(t) \quad \xrightarrow[j\to\infty]{u.o.c.} \quad \bar{N}_{k,l}^{0,i}(t) = \mu_{k,l}^0(\bar{B}_{k,l}^{0,i}(t) - \bar{R}_{k,l}^{0,i})^+, \tag{5}$$

Note that $N_k^{i,x}(t)$ differs by at most 1 from $\sum_j N_{j,k}^{0,i,x}(t)$. Thus, a.s.

$$\bar{N}_k^{i,n_j}(t) \xrightarrow[j\to\infty]{u.o.c.} \bar{N}_k^i(t) = \sum_l \mu_{l,k}^0(\bar{B}_{l,k}^{0,i}(t) - \bar{R}_{l,k}^{0,i})^+. \tag{6}$$

Notice that

$$Q_k^x(t) = Q_k^x(0) + E_k^x(t) - \sum_{i=1}^{s} D_k^{i,x}(t). \tag{7}$$

Therefore, (7) implies that, for some subsequence say (x_{n_j}), with probability one, $(\bar{Q}^{n_j}(t))$ converges u.o.c. to some process $\bar{Q}(t)$ such that

$$\bar{Q}_k(t) = \bar{Q}_k(0) + \lambda_k(t - \bar{A}_k)^+ - \mu_k \sum_{i=1}^{s}(\bar{B}_k^i(t) - \bar{R}_k^i)^+. \tag{8}$$

The convergence statements of the theorem are proved and also point (*iii*). The points (*i*) and (*ii*) are obvious.

For the proof of (*iv*), first there exists $T_0 > 0$ (independent of the fluid limit) such that for $t \geq T_0$ we have $\bar{B}_k^i(t) \geq \bar{R}_k^i$, $\bar{B}_{k,l}^{0,i}(t) \geq \bar{R}_{k,l}^{0,i}$ and $t \geq \bar{A}_k$. Indeed, any fluid limit satisfies $|(\bar{Q}(0), \bar{A}, \bar{R})| \leq 1$ and we have $B_k^{i,x}(R_k^{i,x}(0)) = R_k^{i,x}(0)$ and $B_{k,l}^{0,i,x}(R_{k,l}^{0,i,x}(0)) = R_{k,l}^{0,i,x}(0)$.

Assume now that $J(t) = \emptyset$. If $\bar{Q}_k(t) = 0$, then $\dot{\bar{Q}}_k\,(t) = 0$ at each regular point $t > T_0$ and it follows from (8) that $\dot{\bar{B}}_k\,(t) = \rho_k$. Assume now that $J(t) = \emptyset$ and $\bar{Q}_k(t) > 0$, thus $l_k < +\infty$. Because $\bar{Q}_k(t)$ is continuous, there exists $\varepsilon > 0$ and $h > 0$ such that $\bar{Q}_k(u) > \varepsilon$ for $u \in [t, t+h]$. There is a sequence $\bar{Q}_k^n$ which converges u.o.c. to $\bar{Q}_k$ and thus satisfies for n large enough $Q_k^{x_n}(|x_n|u) \geq \varepsilon|x_n|$ for $u \in [t, t+h]$. Let $q_n = \lfloor \varepsilon|x_n| \rfloor$; at each visit of server i in the time interval $[|x_n|t, |x_n|(t+h)]$, for n large enough, the number of served customers lies (in distribution) between $f_k^{min}(q_n)$ and the bound F_k^* of policy f_k. More precisely, a coupling argument yields

$$\bar{D}_k^{i,n}(t+h) - \bar{D}_k^{i,n}(t) \quad \geq \quad \frac{1}{|x_n|} \sum_{j=N_k^{i,x_n}(|x_n|t)}^{N_k^{i,x_n}(|x_n|(t+h))} F_k^{min}(j, q_n) \tag{9}$$

$$\bar{D}_k^{i,n}(t+h) - \bar{D}_k^{i,n}(t) \quad \leq \quad \frac{1}{|x_n|} \sum_{j=N_k^{i,x_n}(|x_n|t)}^{N_k^{i,x_n}(|x_n|(t+h))} F_k^*(j) \tag{10}$$

where $(F_k^{min}(j,q_n))_j$ (respectively $(F_k^*(j))_j$) is a sequence of i.i.d. random variables having the distribution of $f_k^{min}(q_n)$ (respectively F_k^*). Because the mean of $f_k^{min}(q_n)$ converges monotonically to $l_k = E(F_k^*)$, the quantities on the right hand side in (9) and in (10) admit the same limit namely $l_k\left(\bar{N}_k^i(t+h) - \bar{N}_k^i(t)\right)$ and thus we have

$$\bar{D}_k^i(t+h) - \bar{D}_k^i(t) = l_k\left(\bar{N}_k^i(t+h) - \bar{N}_k^i(t)\right). \tag{11}$$

It follows that, for each regular t, we have

$$\dot{\bar{D}}_k^i(t) = l_k \, \dot{\bar{N}}_k^i(t).$$

On the other hand, from (4) and for $t > T_0$, $\dot{\bar{D}}_k^i(t) = \mu_k \, \dot{\bar{B}}_k^i(t)$. Thus when $J(t) = \emptyset$ and $\bar{Q}_k(t) > 0$, for each $i \in \{1,\ldots,s\}$:

$$\mu_k \, \dot{\bar{B}}_k^i(t) = l_k \, \dot{\bar{N}}_k^i(t).$$

It remains to prove (iv) when $J \neq \emptyset$. For each $k \in J$, $l_k = +\infty$ and $\bar{Q}_k(t) > 0$. Therefore, for a sequence $(x_n)_n$ and for n large enough, we have that $Q_k(|x_n|u) > q_n$ for $u \in [t, t+h]$. The inequality (9) is valid but here the mean of $f_k^{min}(q_n)$ converges monotonically to $\infty = l_k$. Because $\bar{D}_k^i(t+h) - \bar{D}_k^i(t)$ is finite, it implies that

$$\bar{N}_k^i(t+h) - \bar{N}_k^i(t) = 0.$$

The Markovian routing of the servers implies the relations

$$\bar{N}_{k,l}^{0,i}(t) = r_{k,l}\bar{N}_k^i(t) \quad \text{and} \quad \bar{N}_l^i(t) = \sum_{k=1}^{c} \bar{N}_{k,l}^{0,i}(t) \tag{12}$$

for each i and all (k,l). Therefore

$$\bar{N}_k^i(t) = p_k \bar{\mathcal{N}}^i(t). \tag{13}$$

From (5), for $t > T_0$, for each i and for all (k,l) we have $\bar{B}_{k,l}^{0,i}(t+h) - \bar{B}_{k,l}^{0,i}(t) = 0$. Thus for each i

$$\sum_{1 \le k,l \le c} \left(\bar{B}_{k,l}^{0,i}(t+h) - \bar{B}_{k,l}^{0,i}(t)\right) = 0.$$

Inserting in (ii) provides

$$h = \bar{\mathcal{B}}^i(t+h) - \bar{\mathcal{B}}^i(t)$$

which yields $\dot{\bar{\mathcal{B}}}^i(t) = 1$ for regular $t > T_0$ for each i. This ends the proof. □

Lemma 1 *If $\hat{\rho} + \max_{1 \le k \le c} \frac{\lambda_k}{p_k l_k} S < s$, then the fluid limit model is stable i.e. there exists T such that for any fluid limit with $|\bar{Q}(0)| + |\bar{A}| + |\bar{R}| = 1$, $\forall t \ge T,\ \bar{Q}(t) = 0$.*

Proof. Let $\bar{W}(t) = \sum_{k=1}^{c} \frac{1}{\mu_k} \bar{Q}_k(t)$ be the workload of the fluid model. Let $t \geq T_0$ and as previously $J(t) = \{k, l_k = +\infty \text{ and } \bar{Q}_k(t) > 0\}$. Define $K(t) = \{k, \bar{Q}_k(t) > 0\}$ and $U(t) = \{k, l_k = +\infty\}$ so that $J(t) = K(t) \cap U(t)$. From theorem 1-(iii) we get for regular $t > T_0$

$$\dot{\bar{W}}(t) = \hat{\rho} - \sum_{i=1}^{s} \dot{\bar{\mathcal{B}}}^{i}(t) = \hat{\rho} - \sum_{k=1}^{c} \dot{\bar{\mathcal{B}}}_k(t).$$

If $J(t) \neq \emptyset$, $\sum_{i=1}^{s} \dot{\bar{\mathcal{B}}}^{i}(t) = s$ and

$$\dot{\bar{W}}(t) = \hat{\rho} - s.$$

If $J(t) = \emptyset$, from theorem 1-(ii) and using (12) and (13), we get

$$\begin{aligned} \sum_{k=1}^{c} \dot{\bar{\mathcal{B}}}_k(t) &= s - \sum_{i=1}^{s} \sum_{1 \leq k,l \leq c} \dot{\bar{B}}_{k,l}^{0,i}(t) \\ &= s - \sum_{i=1}^{s} \sum_{1 \leq k,l \leq c} \frac{1}{\mu_{k,l}^{0}} \dot{\bar{N}}_{k,l}^{0,i}(t) \\ &= s - \sum_{i=1}^{s} \sum_{1 \leq k,l \leq c} \frac{p_k r_{k,l}}{\mu_{k,l}^{0}} \dot{\bar{\mathcal{N}}}^{i}(t) \\ &= s - S\, \dot{\bar{\mathcal{N}}}(t). \end{aligned} \tag{14}$$

where $S = \sum_{1 \leq k,l \leq c} \frac{p_k r_{k,l}}{\mu_{k,l}^{0}}$ and $\bar{\mathcal{N}}(t) = \sum_{i=1}^{s} \bar{\mathcal{N}}^{i}(t)$. Thus

$$\dot{\bar{W}}(t) = \hat{\rho} - s + S\, \dot{\bar{\mathcal{N}}}(t). \tag{15}$$

We can also calculate $\sum_{k=1}^{c} \dot{\bar{\mathcal{B}}}_k(t)$ by theorem 1-(iv). Simple algebra provides

$$\sum_{k=1}^{c} \dot{\bar{\mathcal{B}}}_k(t) = \sum_{k \notin K(t)} \rho_k + \dot{\bar{\mathcal{N}}}(t) \sum_{k \in K(t)} l_k p_k \sigma_k.$$

The previous equation and (14) yield

$$\dot{\bar{\mathcal{N}}}(t) = \frac{s - \hat{\rho} + \sum_{k \in K(t)} \rho_k}{S + \sum_{k \in K(t)} l_k p_k \sigma_k}$$

which inserted in (15) gives

$$\dot{\bar{W}}(t) = \frac{\sum_{k \in K(t)} ((\hat{\rho} - s) l_k p_k \sigma_k + S \rho_k)}{S + \sum_{k \in K(t)} l_k p_k \sigma_k}.$$

Thus $\dot{\bar{W}}\,(t)$ can only take a finite number of values, depending on the subset $K(t) \subset \{1,\ldots,c\}$ when $J(t) = \emptyset$ and the value $\hat{\rho} - s$ when $J(t) \neq \emptyset$. It is then easy to see that the maximum of this finite number of values is independent of t and is negative when the condition of the lemma holds. Thus there exists $\alpha > 0$ such that for regular $t > T_0$, $\dot{\bar{W}}\,(t) < -\alpha$ as soon as $\bar{W}(t) > 0$.

On the other hand, $\bar{W}$ is absolutely continuous, since Lipschitz, and consequently has a derivative almost everywhere. Therefore there exists T, such that for $t \geq T$, $\bar{W}(t) = 0$ and consequently $\bar{Q}(t) = 0$. Hence, the fluid model is stable. □

Theorem 2 *If* $\hat{\rho} + \max_{1\leq k\leq c} \frac{\lambda_k}{p_k l_k} S < s$ *where* $l_k = E(F_k^*)$*, then the Markov process* $(X(t))_{t\geq 0}$ *is ergodic.*

Proof. The proof is analogous to the proof of Theorem 4.2 of [4] p. 68 to which we mainly refer. To prove that X is Harris positive recurrent, it is sufficient to verify that there exists $T > 0$ such that $\lim_{|x|\to+\infty} \frac{1}{|x|} E(|X^x(|x|T)|) = 0$ under the assumption that the distributions of inter-arrival times are unbounded and spread out (see also Theorem 3.1 of [4] p. 60).

We have,

$$\begin{aligned}\frac{1}{|x|}\quad & E(|X^x(t)|) = \\ & \frac{1}{|x|}E(|Q^x(t)|) + \frac{1}{|x|}E(|P^x(t)|) + \frac{1}{|x|}E(|A^x(t)|) + \frac{1}{|x|}E(|R^x(t)|) + \frac{1}{|x|}E(|C^x(t)|).\end{aligned}$$

Let (x_n) be any sequence with $|x_n| \to +\infty$ and $t > T_0$.
By Theorem 1 and Lemma 1, there exists a subsequence $(x_{n_j})_j$ and $T_1 > T_0$ such that with probability one

$$\forall t > T_1 \quad \bar{Q}^{n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} \bar{Q}(t) = 0.$$

From the following expressions,

$$A_k^x(t) = A_k^x(0) - t + \sum_{j=1}^{E_k^x(t)} \tau_k(j)$$

and

$$R_k^{i,x}(t) = R_k^{i,x}(0) + \sum_{j=1}^{D_k^{i,x}(t)} \sigma_k^i(j) - B_k^{i,x}(t),$$

using (1) and (4), for $t \geq T_1$,

$$\begin{aligned}\bar{A}^{n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} & \quad \bar{A}_k - t + (t - \bar{A}_k)^+ = 0,\\ \bar{R}_k^{i,n_j}(t) \overset{u.o.c.}{\underset{j\to\infty}{\longrightarrow}} & \quad \bar{R}_k^i + \sigma_k \bar{D}_k^i(t) - \bar{B}_k^i(t)\\ = & \quad \bar{R}_k^i + \sigma_k \mu_k (\bar{B}_k^i(t) - \bar{R}_k^i)^+ - \bar{B}_k^i(t)\\ = & \quad 0.\end{aligned}$$

for $t \geq T_0$. By the same argument, for $t \geq T_0$,

$$R_k^{\bar{0},i^{n_j}}(t) \xrightarrow[j\to\infty]{u.o.c.} 0.$$

It suffices then to check that the sequences are uniformly integrable (see [4] Lemma 4.3 for details), to get

$$E(|\bar{Q}^{n_j}(t)|) + E(|\bar{A}^{n_j}(t)|) + E(|\bar{R}^{n_j}(t)|) \to 0.$$

On the other hand, the process P is valued in a finite set, thus, for any definition of the norm $|P(t)|$, for every t,

$$E(|\bar{P}^{n_j}(t)|) \to 0.$$

The component C is specific to the service policies that are involved. We consider the quite general case where $C(t) = (C_k(t), 1 \leq k \leq c)$ and $C_k(t) = (C'_k(t), C''_k(t))$ is relative to queue k. $C'_k(t)$ has $\gamma'_k \geq 0$ components, each of them being bounded by $Q_k(t)$ and $C''_k(t)$ has $\gamma''_k \geq 0$ components, each of them being bounded in distribution by an integrable random variable L_k. First,

$$E(|\bar{C'}^{n_j}_k(t)|) \leq \gamma'_k E(\bar{Q}^{n_j}_k(t)) \to 0.$$

Secund, because L_k is integrable,

$$E(|\bar{C''}^{n_j}_k(t)|) \to 0.$$

Therefore

$$E(|\bar{C}^{n_j}(t)|) \to 0.$$

Suppose now that, for some $T > T_1$, $\limsup_{|x|\to+\infty} \frac{1}{|x|} E|X^x(|x|T)| > 0$. This will be a contradiction. The proof is complete. □

The necessity of the condition is established in the following theorem.

Theorem 3 *Condition C is necessary for the ergodicity of $(X(t))$.*

Proof. Our proof for one server (see [12] p. 230) holds. Let us suppose (without loss of generality) that the c queues are numbered such that:
- the queues $1, \dots, b$ are served according to an unlimited policy ($l_k = \infty$ for $k = 1, \dots, b$, $b = 0$ if this set of queues is empty),
- the queues $b+1, \dots, c$ are served according to a limited policy ($l_k < \infty$ for $k = b+1, \dots, c$, $b = c$ if this set of queues is empty) and

$$\frac{\lambda_{b+1}}{l_{b+1}p_{b+1}} \leq \frac{\lambda_{b+2}}{l_{b+2}p_{b+2}} \leq \dots \leq \frac{\lambda_c}{l_c p_c}. \tag{16}$$

Suppose that X is ergodic. Consider the return times $(T_c^1(n))_n$ of server 1 to queue c ($p_c > 0$). We call the n-th cycle the time interval $[T_c^1(n), T_c^1(n+1)]$. The cycle durations $(T_c^1(n+1) - T_c^1(n))_n$ constitute an integrable stationary sequence. Let

$$G_j^i = E(D_j^i(T_c^1(n+1)) - D_j^i(T_c^1(n))$$

be the mean number of customers served by server i in queue j during the n-th cycle, which by stationarity does not depend on n. The number G_j^i do not either depend on i because, by the ergodic theorem,

$$\frac{G_j^i}{\tau_c^1} = \lim_{t\to+\infty} \frac{1}{t} \sum_{n,T_c^1(n)\le t} E(D_j^i(T_c^1(n+1)) - D_j^i(T_c^1(n))$$

where $\tau_c^1 = E(T_c^1(2) - T_c^1(1))$ and, on the other hand,

$$\lim_{t\to+\infty} \frac{1}{t} E(D_j^i(t) - D_j^k(t)) = 0.$$

Let G_j be the common value of the G_j^i $(1 \le i \le s)$. We have

$$\begin{aligned} Q_j(T_c^1(n+1)) - Q_j(T_c^1(n)) &= N_j(]T_c^1(n), T_c^1(n+1)]) \\ &\quad - \sum_{j=1}^{s} \left(D_j^i(T_c^1(n+1)) - D_j^i(T_c^1(n))\right) \end{aligned} \tag{17}$$

which gives

$$-\sum_{j=1}^{s} \left(D_j^i(T_c^1(2)) - D_j^i(T_c^1(1))\right) \le Q_j(T_c^1(2)) - Q_j(T_c^1(1)) \le N_j(]T_c^1(n), T_c^1(2)])$$

where both bounds are integrable. This implies that (see Lemma 7 in [12])

$$E(Q_j(T_c^1(2)) - Q_j(T_c^1(1))) = 0.$$

Thus (17) yields

$$\lambda_j(\frac{S}{p_c} + \sum_{l=1}^{c} \sigma_l G_l) = \sum_{i=1}^{s} G_j \quad (1 \le j \le c). \tag{18}$$

Multiplying by σ_j and summing from $j = 1$ to b,

$$s \sum_{j=1}^{b} \sigma_j G_j = \hat{\rho}_b(\frac{S}{p_c} + \sum_{l=1}^{c} \sigma_l G_l) \quad (1 \le j \le c)$$

where $\hat{\rho}_b = \sum_{k=1}^{b} \rho_k$ and

$$(s - \hat{\rho}_b) \sum_{j=1}^{b} \sigma_j G_j = \hat{\rho}_b(\frac{S}{p_c} + \sum_{l=b+1}^{c} \sigma_l G_l) \quad (1 \le j \le c)$$

which implies $\hat{\rho}_b < s$. This ends the proof for $b = c$. If $b < c$, the linear system (18) is easily solved and for $j = c$, $G_c = (s - \hat{\rho})^{-1}\lambda_c \frac{S}{p_c}$. But $G_c^i < l_c$. Indeed, by ergodicity, $P(Q_c(T_c^1(1)) = 0) > 0$ and

$$\begin{aligned} E(D_c^1(T_c^1(2)) - D_c^1(T_c^1(1))) &= E(f_c(Q_c(T_c^1(1)))) \\ &\leq E(F_c^*)P(Q_c(T_c^1(1)) > 0) + E(f(0))P(Q_c(T_c^1(1)) = 0) \\ &< l_c \end{aligned}$$

because $E(f(0)) = 0$. This yields that $(s - \hat{\rho})^{-1}\lambda_c \frac{S}{p_c} < l_c$ which is $(\mathcal{C})$. The proof is complete. □

4 Local stability

By local stability, we mean the stability of a subset of queues. Let the queues be numbered as in the proof of Theorem 3. Let $e \geq b$ and $\mathcal{S}_e$ be the polling system of queues $1, \ldots, e$ obtained by saturating queues $e + 1, \ldots, c$. It is clear that polling system $\mathcal{S}_e$, which is completely similar to $\mathcal{S}$, is stable if and only if

$$\mathcal{C}_e : \qquad \hat{\rho_e} + \frac{\lambda_e}{l_e p_e} S^e < s$$

where $\hat{\rho_e} = \sum_{k=1}^{e} \rho_k$ is the total traffic load and $S^e = S + \sum_{k=e+1}^{c} l_k p_k \sigma_k$ the expected switch-over time per visit in sub-system $\mathcal{S}_e$.

Here, we suppose that the polling system $\mathcal{S}$ is not stable, i.e. that condition $\mathcal{C}$ is violated and determine which queues are stable. A queue is unstable if the queue length converges to $+\infty$ with positive probability.

The previous analysis shows that the queues $1, \ldots, b$ are simultaneously stable or unstable, according to $\hat{\rho}_b < s$ or $\hat{\rho}_b \geq s$, respectively. In the second case, the mean cycle time converges to $+\infty$, excluding any stable behavior in the system and every queue is unstable in the preceding sense. Therefore, we assume that $\hat{\rho}_b < s$ to ensure the integrability of the cycle times, because the queues $b + 1, \ldots, c$ contribute for integrable visit durations and we focus on the behavior of the queues $b + 1, \ldots, c$ in a unstable polling system $\mathcal{S}$. Because of the numbering of the queues, (16) yields

$$\hat{\rho}_{b+1} + \frac{\lambda_{b+1}}{l_{b+1}p_{b+1}} S^{b+1} \leq \hat{\rho}_{b+2} + \frac{\lambda_{b+2}}{l_{b+2}p_{b+2}} S^{b+2} \leq \ldots \leq \hat{\rho}_c + \frac{\lambda_c}{l_c p_c} S^c. \tag{19}$$

Let

$$\kappa = \max\{b \leq j \leq c - 1, \hat{\rho}_j + \frac{\lambda_j}{l_j p_j} S^j < s\}.$$

The set is not empty, thus κ is finite. The local stability is determined by κ: Queues $1, \ldots, \kappa$ compose a stable polling subsystem while for $i > \kappa$, queue i is unstable. By definition of κ, polling subsystem $\mathcal{S}_\kappa$ is stable. The proof is then also based on the fact that (see [12]), for $k > \kappa$, because $(\frac{\lambda_j}{p_j l_j})_j$ is non-decreasing,

$$\begin{aligned}\rho_k + \frac{\lambda_k}{l_k p_k}(\sum_{j\leq\kappa} \frac{\sigma_j \lambda_j p_j S^\kappa}{s-\hat{\rho}_\kappa} + \sum_{j>\kappa, j\neq k} l_j\sigma_j p_j + S) \quad &\geq \quad \hat{\rho}_{\kappa+1} + \frac{\lambda_{\kappa+1}}{l_{\kappa+1}p_{\kappa+1}}(\sum_{\kappa+2\leq j\leq c} l_j \sigma_j p_j + S) \\ &> \quad s\end{aligned}$$

which is equivalent to

$$\rho_k - \frac{s l_k \sigma_k}{\sum_{j\leq\kappa} \frac{p_j\sigma_j\lambda_j S^\kappa}{p_k(s-\hat{\rho}_\kappa)} + \sum_{j>\kappa} \frac{l_j p_j \sigma_j}{p_k} + \frac{S}{p_k}} > 0$$

and gives that, for the z_k's large enough, on a set of positive probability,

$$Z_k(t) = z_k + M_k(t) - \mathcal{B}_k(t) \rightarrow_{t\rightarrow+\infty} +\infty,$$

where z_k is the load at time 0 and $M_k(t)$ is the load arrived between 0 and t at queue k.

References

[1] A. Borovkov and R. Schassberger, *Ergodicity of a polling network*, Stochastic Processes and their Applications, 50 (1994), pp. 253–262.

[2] M. Bramson, *Instability of FIFO queueing networks*, Ann. Appl. Probab., 4 (1994), pp. 414–431.

[3] H. Chen and A. Mandelbaum, *Discrete flow networks: bottleneck analysis and fluid approximations*, Math. Oper. Res., 16 (1991), pp. 408–446.

[4] J. G. Dai, *On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models*, Ann. Appl. Probab., 5 (1995), pp. 49–77.

[5] J. G. Dai and S. P. Meyn, *Stability and convergence of moments for multiclass queueing networks via fluid limit models*, IEEE Trans. Automat. Control, 40 (1995), pp. 1889–1904.

[6] G. Down, *On the stability of polling models with multiple servers*, Tech. Report BS-R9605, CWI, 1996.

[7] V. Dumas, *A multiclass network with non-linear, non-convex, non-monotonic stability conditions*, Queueing Systems, Theory and Applications, 25 (1997), pp. 1–44.

[8] P. Dupuis and R. J. Williams, *Lyapunov functions for semimartingale reflecting Brownian motions*, Ann. Probab., 22 (1994), pp. 680–702.

[9] S. Foss, N. Chernova, and A. Kovalevskii, *Stability of polling systems with state-independent routing*, in Proceedings of the 34-th Annual Allerton Conference in Communication, Control and Computing, Monticello, Illinois, oct 1996, pp. 220–227.

[10] S. Foss and A. Kovalevskii, *The fluid approximation approach: a generalized stability criterion and its application to a polling system.* preprint, 1997.

[11] S. Foss and G. Last, *Stability of polling systems with exhaustive service policies and state-dependent routing*, Ann. Appl. Probab., 6 (1996), pp. 116–137.

[12] C. Fricker and R. Jaibi, *Monotonicity and stability of periodic polling models*, Queueing Systems, Theory and Applications, 15 (1994), pp. 211–238.

[13] ——, *Stability of a polling model with a markovian scheme*, Tech. Report 2278, INRIA, May 1994.

[14] L. Georgiadis and W. Szpankowski, *Stability criteria for yet another multidimensional distributed system*, Tech. Report CSD-TR-91-071, Purdue University, Sept. 1991.

[15] M. L. Kotler, *Proof of stability conditions for token passing rings by Lyapunov functions*, IEEE Trans. Automat. Control, 41 (1996), pp. 908–912.

[16] V. A. Malyshev, *Networks and dynamical systems*, Adv. in Appl. Probab., 25 (1993), pp. 140–175.

[17] L. Massoulié, *Stability of non-Markovian polling systems*, Queueing Systems Theory Appl., 21 (1995), pp. 67–95.

[18] A. N. Rybko and A. L. Stolyar, *On the ergodicity of random processes that describe the functioning of open queueing networks*, Problemy Peredachi Informatsii, 28 (1992), pp. 3–26.

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399