Acta Cybernetica, Tom. 10, Fasc. 1-2, Szeged, 1991

Modelling of heterogeneous multiprocessor systems with randomly changing parameters

J. Sztrik*†

Abstract

A queueing theoretic approach is developed to analyse the performance of heterogeneous multiprocessor computer systems evolving in random environments. The time intervals from the completion of the previous bus usage to the generation of a new request as well as the holding times of the common bus are assumed to be exponentially distributed random variables with parameter depending on the state of the corresponding random environment. Each processor is characterised by its own acces and service rate. The bus arbiter selects the processor to use the common bus according to a First-Come, First-Served (FCFS) discipline. Supposing that the acces rates of the processors are much greater than the corresponding service rates ("fast" arrival), it is shown that the busy period length of the bus converges weakly, under appropriate norming, to an exponentially distributed random variable. As a consequence the main steady-state performance measures, such as utilizations, throughput, mean delay time, expected waiting time, the average number of requests served during a busy period, and mean number of active processors can be calculated. Moreover, exact and approximate validation results are presented to illustrate the credibility of the proposed method.

Keywords: queueing, multiprocessor system, performance measures, weak convergence, random environments, utilization.

1 Introduction

In multiprocessor systems the contention for a common bus is one of the major factors affecting the computer performance. Several papers have been devoted to the analysis of such systems under different conditions on acces rates, the distribution function of holding times, and bus arbitration protocols (c.f., Ajmone Marsan *et al.* (1986), Bodnar and Liu (1989), Gelenbe (1989), Noyami and Sumita (1989)). More recently Ishigaki *et al.* (1990) suggested a queueing theoretic approach to analyse the system and a numerical technique was used for the evaluation of the basic

*Department of Mathematics University of Debrecen Debrecen, P.O. Box 12, Hungary, 4010

†Research done while J. Sztrik was visiting the Department of Mathematics, University of Bradford, England. The work is partially supported by István Széchenyi Foundation, Hungary, Hungarian National Science Foundation under grant OTKA-1648.

performance measures. In this work an asymptotic queueing theoretic approach is proposed to study the performance of a First-Come, First-Served (FCFS) heterogeneous single bus multiprocessor system evolving in random environments. All random times in the system are considered to be exponentially distributed, while each processor is characterised by its own acces and service rates depending on the state of the corresponding random environment. Under a heavy traffic assumption (i.e., "fast" arrivals), it is shown that the busy period length of the bus converges weakly, under appropriate norming, to an exponentially distributed rendom variable. This result facilitates the calculation of several steady-state performance measures of interest.

Note that the asymptotic technique has a widespread applicability in the field of reliability theory (c.f., Anisimov *et al* (1987), Anisimov and Sztrik (1989), Gertsbakh (1984, 1989)). Refinements in the model are often needed when the the system environment is subject to randomly occuring fluctuations which appear as changes in the parameters of the model. These fluctuations may be due to changes in the physical environment, personnel changes, alteration of computer system usage intensity, etc., (c.f., Baccelli and Makowski (1986), Gaver *et al.* (1984), Gelenbe and Rosenberg, Neuts (1978), Rosenberg *et al.* (1990), Sengupta (1990)).

2 Preliminary results

This section presents a brief survey of results (c.f., Anisimov *et al.* (1987)) to be applied in the next section.

Let $(X_\varepsilon(k), k \geq 0)$ be a Markov chain with state space

$$\bigcup_{q=0}^{m+1} X_q, \quad X_i \cap X_j = 0, \; i \neq j,$$

with $m+2$ levels of states, $i, j = 0, 1, \ldots, m+1$, defined by the transition matrix $(p_\varepsilon(i^{(q)}, j^{(z)}))$, $i^{(q)} \in X_q$, $j^{(z)} \in X_z, q, z = 0, 1, \ldots, m+1$ satisfying the following conditions:

1. $p_\varepsilon(i^{(0)}, j^{(0)}) \longrightarrow p_0(i^{(0)}, j^{(0)})$, as $\varepsilon \longrightarrow 0$, $i^{(0)}, j^{(0)} \in X_0$, and matrix $P_0 = (p_0(i^{(0)}, j^{(0)}))$ is irreducible;

2. $p_\varepsilon(i^{(q)}, j^{(q+1)}) = \varepsilon\alpha^{(q)}, j^{(q+1)}) + o(\varepsilon)$, $i^{(q)} \in X_q$, $j^{(q+1)} \in X_{q+1}$, where $\alpha^{(q)}(i^{(q)}, j^{(q+1)})$ is an appropriate transition matrix;

3. $p_\varepsilon(i^{(q)}, f^{(q)}) \longrightarrow 0$, as $\varepsilon \longrightarrow 0$, $i^{(q)}, f^{(q)} \in X_q, q \geq 1$;

4. $p_\varepsilon(i^(q), f^{(z)}) \equiv 0 \; i^{(q)} \in X_q$, $f^{(z)} \in X_z$, $\quad z - q \geq 2$.

In the sequeel the set of states X_q is called the q-th level of the chain, $q = 0, \ldots, m+1$. Let us single out the subset of states

$$\langle \alpha_m \rangle = \bigcup_{q=0}^{m} X_q.$$

Denote by $\{\pi_\varepsilon(i^{(q)}), \; i^{(q)} \in X_q\}$, $\quad q = 1, \ldots, m$ the stationary distribution of a chain with transition matrix

$$\Big(\frac{p_\varepsilon(i^{(q)}, j^{(z)})}{1 - \sum\limits_{k^{(m+1)} \in X_{m+1}} p_\varepsilon(i^{(q)}, k^{(m+1)})}\Big), i^{(q)} \in X_q, j^{(z)} \in X_z, \; q, z \leq m.$$

Furthermore denote by $g_\varepsilon(\langle\alpha_m\rangle)$ the steady state probability of exit from $\langle\alpha_m\rangle$, that is

$$g_\varepsilon(\langle\alpha_m\rangle) = \sum_{i^{(m)}\in X_m} \pi_\varepsilon(i^{(m)}) \sum_{j^{(m+1)}\in X_{m+1}} p_\varepsilon(i^{(m)}, j^{(m+1)}).$$

Denote by $\{\pi_o(i^{(0)}),\quad i^{(0)}\in X_0\}$ the stationary distribution corresponding to P_o and let

$$\overline{\pi}_o = \{\pi_o(i^{(0)}),\ i^{(0)}\in X_0\},\qquad \overline{\pi}_\varepsilon^{(q)} = \{\pi_\varepsilon(i^{(q)}), i^{(q)}\in X_q\},$$

be row vectors. Finally, let the matrix

$$A^{(q)} = (\alpha^{(q)}(i^{(q)}, j^{(q+1)})),\quad i^{(q)}\in X_q,\ j^{(q+1)}\in X_{q+1},\ q=0,\ldots,m$$

defined by condition 2.

Conditions (1)-(4) enable us to compute the main terms of the asymptotic expression for $\overline{\pi}_\varepsilon^{(q)}$ and $g_\varepsilon(\langle\alpha_m\rangle)$. Namely, we obtain

$$\overline{\pi}_\varepsilon^{(q)} = \varepsilon^q\overline{\pi}_o A^{(0)}A^{(1)}\ldots A^{(q-1)} + o(\varepsilon^q),\quad q=1,\ldots,m,$$

$$g_\varepsilon(\langle\alpha_m\rangle) = \varepsilon^{m+1}\overline{\pi}_o A^{(0)}A^{(1)}\ldots A^{(m)}\underline{1} + o(\varepsilon^{m+1}), \tag{1}$$

where $\underline{1} = (1,\ldots,1)^*$ is a column vector, (c.f., Anisimov *et al.* (1987), pp. 141-153).

Let $(\eta_\varepsilon(t),\quad t\geq 0)$ be a Semi Markov Process (SMP) given by the embedded Markov chain $(X_\varepsilon(k),\quad k\geq 0)$ satisfying conditions (1)-(4). Let the times $\tau_\varepsilon(j^{(s)}, k^{(z)})$ - transition times from state $j^{(s)}$ to state $k^{(z)}$ - fulfill the condition

$$E\exp\{i\Theta\beta_\varepsilon\tau_\varepsilon(j^{(s)}, k^{(z)})\} = 1 + a_{jk}(s,z,\Theta)\varepsilon^{m+1} + o(\varepsilon^{m+1}),\quad (i^2=-1)$$

where β_ε is some normalizing factor. Denote by $\Omega_\varepsilon(m)$ the instant at which the SMP reaches the $(m+1)$-th level for the first time, exit time from $\langle\alpha_m\rangle$, provided $\eta_\varepsilon(0)\in\langle\alpha_m\rangle$. Then we have:

Theorem 1 *(c.f., Anisimov et al. (1987), pp. 153) If the above (1)-(4) conditions are satisfied then*

$$\lim_{\varepsilon\longrightarrow 0} E\exp\{i\Theta\beta_\varepsilon\Omega_\varepsilon(m)\} = (1-A(\Theta))^{-1},$$

where

$$A(\Theta) = \frac{\sum\limits_{j^{(0)},k^{(0)}\in X_o} \pi_o(j^{(0)})p_o(j^{(0)},k^{(0)})a_{jk}(0,0,\Theta)}{\overline{\pi}_o A^{(0)}A^{(1)}\ldots A^{(m)}\underline{1}}.$$

Corollary 1 *In particular, if $a_{jk}(s,z,\Theta) = i\Theta m_{jk}(s,z)$ then the limit is an exponentially distributed random variable with mean*

$$\frac{\sum\limits_{j^{(0)},k^{(0)}\in X_o} \pi_o(j^{(0)})p_o(j^{(0)},k^{(0)})m_{jk}(0,0)}{\overline{\pi}_o A^{(0)}A^{(1)}\ldots A^{(m)}\underline{1}}.$$

3 The Queueing Model

Consider a multiprocessor computer system in which N different processors with a common memory are connected by a single bus. A processor that generates a request to use the bus is said to be active, otherwise it is called inactive or idle. The bus arbitration protocol (selection rule) is assumed to be FCFS, that is, the arbiter selects the next processor to use the bus amongst the active ones in order of requests' arrivals. The time intervals from the completion of the previous bus usage to the generation of a new request as well as the holding times of the common bus are exponentially distributed random variables with parameter depending on the state of the corresponding random environment. Each processor is characterised by its own acces and service rate. The processors operate in a random environment governed by an ergodic Markov chain $(\xi_1(t), \quad t \geq 0)$ with state space $(1, \ldots, r_1)$ and with transition rate matrix $a^{(1)}_{i_1 j_1}$, $i_1, j_1 = 1, \ldots, r_1$, $\quad a^{(1)}_{i_1 i_1} = \sum_{j \neq i_1} a^{(1)}_{i_1 j}$). Moreover, it is assumed that each processor can have at most one outstanding request at any time, i.e., each processor can generate a new request only after the bus usage of the previous request has been completed. Whenever the environmental process is in state i_1, let $\lambda_p(i_1, \varepsilon)$ be the access rate for processor p, $p = 1, \ldots, N$, respectively. Similarly, the shared bus is supposed to operate in a random environment governed by an ergodic Markov chain $(\xi_2(t), \quad t \geq 0)$ with state space $(1, \ldots, r_2)$ and with transition rate matrix $(a^{(2)}_{i_2 j_2}$, $i_2, j_2 = 1, \ldots, r_2$, $\quad a^{(2)}_{i_2 i_2} = \sum_{j \neq i_2} a^{(2)}_{i_2 j}$). Whenever the environmental process is in state i_2, let $\mu_p(i_2)$ be the service rate for processor p, $\quad p = 1, \ldots, N$, respectively. To this end the probability that processor p generates a request in the time interval $(t, t+h)$ is $\lambda_p(i_1, \varepsilon)h + o(h)$, where $\varepsilon > 0$, $\quad i_1 = 1, \ldots r_1$, and the probability that processor p completes the bus usage in time interval $(t, t+h)$ is $\mu_p(i_2)h + o(h)$, $i_2 = 1, \ldots, r_2$, $p = 1, \ldots, N$.

All random variables and the random environment are assumeed to be independent of each other.

Let us consider the system under the heavy traffic assumption, i.e., $\lambda_p(i_1, \varepsilon) \longrightarrow \infty$ as $\varepsilon \longrightarrow 0$. For simplicity let $\lambda_p(i_1, \varepsilon) = \lambda_p(i_1)/\varepsilon$, $\quad p = 1, \ldots, N$, $i_1 = 1, \ldots, r_1$.

Denote by $Y_\varepsilon(t)$ the number of inactive processors at time t, and let

$$\Omega_\varepsilon(m) = \inf\{t : t > 0, \ Y_\varepsilon(t) = m + 1 \ / \ Y_\varepsilon(0) \leq m\},$$

i.e., the instant at which the number of inactive processors reaches the $(m+1)$-th level for the first time, provided that at the beginning their number is not greater than m, $m = 1, \ldots, N-1$. In particular, if $m = N-1$ then the bus becomes idle since there is no active processor and, hence $\Omega_\varepsilon(N-1)$ can be referred to as the busy period length of the bus.

Denote by $\pi_o(i_1, i_2 : 0; k_1, \ldots, k_N)$ the steady- state probability that $\xi_1(t)$ is in state i_1, $\xi_2(t)$ is in state i_2, there is no idle processor and the order of requests' arrival to the bus is $(k_1, \ldots, k_N)$. Similarly, denote by $\pi_o(i_1, i_2 : 1; k_2, \ldots, k_N)$ the steady-state probability that the first random environment is in state i_1, the second one is in state i_2, processor k_1 is inactive and the other processors sent their requests in order $(k_2, \ldots, k_N)$. Clearly $(k_s, \ldots, k_N) \in V_N^{N-s+1}$, $s = 1, 2$, where V_N^{N-s+1} denotes the set of all variations of order $N - s + 1$ of integets $1, \ldots, N$. Now we have:

Theorem 2 *For the system in question under the above assumptions, independently of the initial state, the distribution of the normalized random variable*

$\varepsilon^m \Omega_\varepsilon(m)$ *converges weakly to an exponentially distributed random variable with parameter*

$$\wedge = \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \sum_{(k_1,\ldots,k_N)\in V_N^N} \pi_o(i_1, i_2 : 1; k_2, \ldots, k_N)$$

$$\times \frac{\mu_{k_2}(i_2)}{\lambda_{k_1}(i_1)} \frac{\mu_{k_3}(i_2)}{\lambda_{k_1}(i_1) + \lambda_{k_2}(i_1)} \times \ldots \times \frac{\mu_{k_{m+1}}(i_2)}{\lambda_{k_1}(i_1) + \ldots + \lambda_{k_m}(i_1)} \frac{1}{D},$$

where

$$D = \sum_{\substack{i_1, j_1 = 1 \\ j_1 \neq i_1}}^{r_1} \sum_{\substack{i_2, j_2 = 1 \\ j_2 \neq i_2}}^{r_2} \sum_{(k_1,\ldots,k_N)\in V_N^N} \pi_o(i_1, i_2, 0; k_1, \ldots, k_N)$$

$$\times \frac{a^{(1)}_{i_1 j_1} + a^{(2)}_{i_2 j_2}}{(a^{(1)}_{i_1 i_1} + a^{(2)}_{i_2 i_2} + \mu_{k_1}(i_2))^2}.$$

Proof. Let us introduce the following stochastic process

$$Z_\varepsilon(t) = (\xi_1(t), \xi_2(t) : Y_\varepsilon(t); \beta_1(t), \ldots, \beta_{N-Y_\varepsilon(t)}(t))$$

where $\beta_1(t), \ldots, \beta_{N-Y_\varepsilon(t)}(t)$ denotes the indices of the active processors in the order of their request arrival to the bus. It is easy to see that $(Z_\varepsilon(t), t \geq 0)$ is a multidimensional Markov chain with state space

$$E = ((i_1, i_2 : s; k_1, \ldots, k_{N-s}), \quad i_1 = 1, \ldots, r_1, \quad i_2 = 1, \ldots, r_2,$$

$$(k_1, \ldots, k_{N-s}) \in V_N^{N-s}, s = 0, \ldots, N)$$

where $k_o = \{0\}$ by definition.
Furthermore, let

$$\langle \alpha_m \rangle = ((i_1, i_2 : s; k_1, \ldots, k_{N-s}), \quad i_1 = 1, \ldots, r_1, \quad i_2 = 1, \ldots, r_2,$$

$$(k_1, \ldots, k_{N-s}) \in V_N^{N-s}, s = 0, \ldots, m).$$

Hence our aim is to determine the distribution of the first exit time of $Z_\varepsilon(t)$ from $\langle \alpha_m \rangle$, provided that $Z_\varepsilon(o) \in \langle \alpha_m \rangle$.

It can easily be verified that the transition probabilities for the embedded Markov chain are

$$p_\varepsilon[(i_1, i_2 : s; k_1, \ldots, k_{N-s}), (j_1, i_2 : s; k_1, \ldots, k_{N-s})]$$

$$= \frac{a^{(1)}_{i_1 j_1}}{a^{(1)}_{i_1 i_1} + a^{(2)}_{i_2 i_2} + \sum_{p \neq k_1, \ldots, k_{N-s}} \lambda_p(i_1)/\varepsilon + \mu_{k_1}(i_2)}, \quad s = 0, \ldots, N-1,$$

$$p_\varepsilon[i_1,i_2:N;0),(j_1,i_2:N;0)=\frac{a^{(1)}_{i_1j_1}}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p=1}^{N}\lambda_p(i_1)/\varepsilon},\quad s=N,$$

$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(i_1,j_2:s;k_1,\ldots,k_{N-s})]$$
$$=\frac{a^{(2)}_{i_2j_2}}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p\neq k_1,\ldots,k_{N-s}}\lambda_p(i_1)/\varepsilon+\mu_{k_1}(i_2)},\quad s=0,\ldots,N-1,$$

$$p_\varepsilon[(i_1,i_2:N;0),=\frac{a^{(2)}_{i_2j_2}}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p=1}^{N}\lambda_p(i_1)/\varepsilon},\quad s=N,$$

$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(i_1,i_2:s+1;k_2,\ldots,k_{N-s})]$$
$$=\frac{\mu_{k_1}(i_2)}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p\neq k_1,\ldots,k_{N-s}}\lambda_p(i_1)/\varepsilon+\mu_{k_1}(i_2)},\quad s=0,\ldots,N-1,$$

$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(i_1,i_2:s-1;;k_1,\ldots,k_{N-s+1})]$$
$$=\frac{\lambda_{k_{N-s+1}}(i_1)/\varepsilon}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p\neq k_1,\ldots,k_{N-s}}\lambda_p(i_1)/\varepsilon+\mu_{k_1}(i_2)},\quad s=1,\ldots,N-1,$$

$$p_\varepsilon[(i_1,i_2:N;0),(i_1,i_2:N-1;k)]=\frac{\lambda_k(i_1)}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\sum\limits_{p=1}^{N}\lambda_p(i_1)/\varepsilon},\quad s=N.$$

As $\varepsilon\longrightarrow 0$ this implies

$$p_\varepsilon[(i_1,i_2:0;k_1,\ldots,k_N),(j_1,i_2:0;k_1,\ldots,k_N)]=\frac{a^{(1)}_{i_1j_1}}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\mu_{k_1}(i_2)},\quad s=0,$$

$$p_\varepsilon[(i_1,i_2:0;k_1,\ldots,k_N),(i_1,j_2:0;k_1,\ldots,k_N)]=\frac{a^{(2)}_{i_2j_2}}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\mu_{k_1}(i_2)},\quad s=0,$$

$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(j_1,i_2:s;k_1,\ldots,k_{N-s})]=o(1),\quad s=1,\ldots,N,$$
$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(i_1,j_2:s;k_1,\ldots,k_{N-s})]=o(1),\quad s=1,\ldots,N,$$

$$p_\varepsilon[(i_1,i_2:0;k_1,\ldots,k_N),(i_1,i_2:1;k_2,\ldots,k_N)]=\frac{\mu_{k_1}(i_2)}{a^{(1)}_{i_1i_1}+a^{(2)}_{i_2i_2}+\mu_{k_1}(i_2)},\quad s=0,$$

$$p_\varepsilon[(i_1,i_2:s;k_1,\ldots,k_{N-s}),(i_1,i_2:s+1;;k_2,\ldots,k_{N-s})]$$
$$=\frac{\mu_{k_1}(i_2)\varepsilon}{\sum\limits_{p\neq k_1,\ldots,k_{N-s}}\lambda_p(i_1)}(1+o(1)),\quad s=1,\ldots,N-1.$$

This agrees with the conditions (1)-(4), but here the zero level is the set

$$((i_1,i_2:0,k_1,\ldots,k_N),(i_1,i_2:1;k_1,\ldots,k_{N-1},\ i_1=1,\ldots,r_1,\ i_2=1,\ldots,r_2,$$
$$(k_1,\ldots,k_{N-s})\in V_N^{N-s},\ s=0,1),$$

while the q-th level is the set

$$((i_1, i_2 : q+1; k_1, \dots, k_{N-q-1}), \quad i_1 = 1, \dots, r_1, \quad i_2 = 1, \dots, r_2,$$
$$(k_1, \dots, k_{N-q-1}) \in V_N^{N-q-1}).$$

Since the level 0 in the limit forms an essential class, the probabilities

$\pi_o(i_1, i_2 : 0 k_1, \dots, k_N)$, $\pi_o(i_1, i_2 : 1; k_1, \dots, k_{N-1})$, $\quad i_1 = 1, \dots, r_1$, $i_2 = 1, \dots, r_2$, $(k_1, \dots, k_{N-s}) \in V_N^{N-s}$, $\quad s = 0, 1$, satisfy the following system of equations

$$\begin{aligned}\pi_o(j_1, j_2 : 0; k_1, \dots, k_N) &= \sum_{i_1 \neq j_1} \pi_o(i_1, j_2 : 0; k_1, \dots, k_N) a^{(1)}_{i_1 j_1} / [a^{(1)}_{i_1 i_1} + a^{(2)}_{j_2 j_2} + \mu_{k_1}(j_2)] \\ &+ \sum_{i_2 \neq j_2} \pi_o(j_1, i_2 : 0; k_1, \dots, k_N) a^{(2)}_{i_2 j_2} / [a^{(2)}_{j_1 j_1} + a^{(2)}_{i_2 i_2} + \mu_{k_1}(i_2)] \\ &+ \pi_o(j_1, j_2 : 1; k_1, \dots, k_{N-1}), \qquad (2)\end{aligned}$$

$$\begin{aligned}&\pi_o(j_1, j_2 : 1; k_1, \dots, k_{N-1}) \\ &\quad = \pi_o(j_1, j_2 : 0; k_N, k_1, \dots, k_{+N-1}) \mu_{k_N}(j_2) / [a^{(1)}_{j_1 j_1} + a^{(2)}_{j_2 j_2} + \mu_{k_N}(j_2)]. \qquad (3)\end{aligned}$$

To apply the asymptotic expressions (1), it is necessary to solve system (2), (3), subject to normalizing condition

$$\sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \sum_{(k_1, \dots, k_N)} \{\pi_o(i_1, i_2 : 0; k_1, \dots, k_N) + \pi_o(i_1, i_2 : 1; k_1, \dots, k_{N-1})\} = 1.$$

Suppose this solution is known. Then by substituting it into (1) it follows that

$$\begin{aligned}g(\langle \alpha_m \rangle) = \varepsilon^m \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \sum_{(k_1, \dots, k_N) \in V_N^N} & \pi_o(i_1, i_2 : 1; k_2, \dots, k_N) \frac{\mu_{k_2}(i_2)}{\lambda_{k_1}(i_1)} \frac{\mu_{k_3}(i_2)}{\lambda_{k_1}(i_1) + \lambda_{k_2}(i_1)} \\ & \times \dots \times \frac{\mu_{k_{m+1}}(i_2)}{\lambda_{k_1}(i_1) + \dots + \lambda_{k_m}(i_1)} (1 + o(1)). \qquad (4)\end{aligned}$$

Taking into account the exponentiality of $\tau_\varepsilon(j_1, j_2 : s; k_1, \dots, k_{N-s})$ for fixed Θ it is implied that

$$E \exp\{i\varepsilon^m \Theta \tau_\varepsilon(j_1, j_2 : 0; k_1, \dots, k_N)\} = 1 + \varepsilon^m \frac{i\Theta}{a^{(1)}_{j_1 j_1} + a^{(2)}_{j_2 j_2} + \mu_{k_1}(j_2)} (1 + o(1)),$$

$$E \exp\{i\varepsilon^m \Theta \tau_\varepsilon(j_1, j_2 : s; k_1, \dots, k_{N-s})\} = 1 + o(\varepsilon^m), \quad s > 0.$$

Notice that $\beta_\varepsilon = \varepsilon^m$ and therefore from Corollary 1 our statement immediately follows.

However, if $\mu_p(i_2) = \mu(i_2)$, $p = 1, \ldots, N$, $i_2 = 1, \ldots, r_2$, then by substituting (3) into (2) then we get

$$\begin{aligned}\pi_o(j_1, j_2 : 0; k_1, \ldots, k_N) &= \sum_{i_1 \neq j_1} \pi_o(i_1, j_2 : 0; k_1, \ldots, k_N) a^{(1)}_{i_1 j_1} / [a^{(1)}_{i_1 i_1} + a^{(2)}_{j_2 j_2} + \mu(j_2)] \\ &+ \sum_{i_2 \neq j_2} \pi_o(j_1, i_2 : 0; k_1, \ldots, k_N) a^{(2)}_{i_2 j_2} / [a^{(1)}_{j_1 j_1} + a^{(2)}_{i_2 j_2} + \mu(i_2)] \\ + \pi(j_1, j_2 : 0; k_N, k_1, \ldots, k_{N-1}) \mu(j_2) / [a^{(1)}_{j_1 j_1} &+ a^{(2)}_{j_2 j_2} + \mu(j_2)]. \end{aligned} \tag{5}$$

Since the steady-state distributions of the governing Markov chains satisfy

$$\pi^{(1)}_{j_1} a^{(1)}_{j_1 j_1} = \sum_{i_1 \neq j_1} \pi^{(1)}_{i_1} a^{(1)}_{i_1 j_1}, \qquad \pi^{(2)}_{j_2} a^{(2)}_{j_2 j_2} = \sum_{i_2 \neq j_2} \pi^{(2)}_{i_2} a^{(2)}_{i_2 j_2}, \tag{6}$$

it can easily be verified, that the solution of (5) together with (6) is

$$\pi_o(i_1, i_2 : 0; k_1, \ldots, k_N) = B \pi^{(1)}_{i_1} \pi^{(2)}_{i_2} (a^{(1)}_{i_1 i_1} + a^{(2)}_{i_2 i_2} + \mu(i_2)),$$

$$\pi_o(i_1, i_2 : 1; k_1, \ldots, k_{N-1}) = B \pi^{(1)}_{i_i} \pi^{(2)}_{i_2} \mu(i_2),$$

where B is the normalising constant, i.e.

$$1/B = N! \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \pi^{(1)}_{i_1} \pi^{(2)}_{i_2} (a^{(1)}_{i_1 i_1} + a^{(2)}_{i_2 i_2} + 2\mu(i_2)).$$

Thus, from Th. 2 follows that $\varepsilon^m \Omega_\varepsilon(m)$ converges weakly to an exponentially distributed random variable with parameter

$$\wedge = \frac{\mu(i_2)^{m+1}}{N!} \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \sum_{(k_1, \ldots, k_N) \in V_N^N} \pi^{(1)}_{i_1} \pi^{(2)}_{i_2} \frac{1}{\lambda_{k_1}(i_1)} \frac{1}{\lambda_{k_1}(i_1) + \lambda_{k_2}(i_1)}$$

$$\times \ldots \times \frac{1}{\lambda_{k_1}(i_1) + \ldots + \lambda_{k_m}(i_1)}.$$

Consequently, the distribution of the time while the number of idle processors reaches the $(m+1)$-th level for the first time is approximated by

$$P(\Omega_\varepsilon(m) > t) = P(\varepsilon^m \Omega_\varepsilon(m) > \varepsilon^m t) \approx \exp(-\varepsilon^m \wedge t).$$

In particular, when $m = N - 1$, we get that the busy period length of the bus is asymptotically an exponentially distributed random variable with parameter

$$\varepsilon^{N-1} \wedge = \varepsilon^{N-1} \frac{\mu(i_2)^N}{N!} \sum_{i_1=1}^{r_1} \sum_{i_2=1}^{r_2} \sum_{(k_1, \ldots, k_N) \in V_N^N} \pi^{(1)}_{i_1} \pi^{(2)}_{i_2} \frac{1}{\lambda_{k_1}(i_1)} \frac{1}{\lambda_{k_1}(i_1) + \lambda_{k_2}(i_1)}$$

$$\times \ldots \times \frac{1}{\lambda_{k_1}(i_1) + \ldots + \lambda_{k_N}(i_1)}. \tag{7}$$

In the case when there are no random environments, i.e., $\mu(i_2) = \mu$, and $\lambda_p(i_1) = \lambda_p$, $i_1 = 1,\ldots,r_1$, $i_2 = 1,\ldots,r_2$, $p = 1,\ldots,N$, from (7) it follows that

$$\varepsilon^{N-1}\Lambda = \frac{\mu^N}{N!} \sum_{(k_1,\ldots,k_N)\in V_N^N} \frac{1}{\lambda_{k_1}/\varepsilon} \frac{1}{\lambda_{k_1}/\varepsilon + \lambda_{k_2}/\varepsilon} \times \ldots \times \frac{1}{\lambda_{k_1}/\varepsilon + \ldots + \lambda_{k_{N-1}}/\varepsilon}. \tag{8}$$

Finally, for the special case of totally homogeneous processors (i.e., $\lambda_p = \lambda$, $p = 1,\ldots,N$) expression (8) reduces to

$$\varepsilon^{N-1}\Lambda = \frac{1}{(N-1)!} \frac{\mu^N}{(\lambda/\varepsilon)^{N-1}}. \tag{9}$$

4 Performance Measures

This section deals with the derivation of the main steady-state performance measures relating to the heterogeneous multiprocessor model treated in the previous section.

4.1 Utilizations

The utilization U of the bus is defined as the fraction of time during which it is busy. The idle period of the bus starts when each processor is idle at the end of a service completion, and terminates when a processor generates a request. It is clear that the mean idle period length is

$$\sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\sum_{p=1}^{N} \lambda_p(i_1)/\varepsilon}.$$

Hence for U the following expression is obtained

$$U = \frac{\frac{1}{\varepsilon^{N-1}\Lambda}}{\frac{1}{\varepsilon^{N-1}\Lambda} + \sum_{i_1=1}^{r_1} \pi_{i_1} \frac{1}{\sum_{p=1}^{N} \lambda_p(i_1)/\varepsilon}}. \tag{10}$$

The bus utilization U_p of processor p is defined as the fraction of time that processor p uses the bus. Since the processors have identically distributed holding times we get

$$U_p = U \sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \left(\lambda_p(i_1) \Big/ \sum_{k=1}^{N} \lambda_k(i_1) \right). \tag{11}$$

4.2 Throughput

The throughput γ_p of processor p is defined as the mean number of requests of processor p served per unit time. It is well-known that

$$U_p = \gamma_p b_p$$

where b_p is the mean bus usage (service) time of a request by processor p.

In this case

$$U_p = \gamma_p \sum_{i_2=1}^{r_2} \pi_{i_2}^{(2)} \frac{1}{\mu(i_2)}$$

and thus

$$\gamma_p = U_p \Big/ \sum_{i=1}^{r_2} \pi_{i_2}^{(2)} \frac{1}{\mu(i_2)}.$$

4.3 Mean delay and waiting times

The mean delay T_p of processor p is the average time from the instant at which a request is generated at processor p to the instant at which the bus usage of that request has been completed. In other words, T_p is the mean duration of an active state at processor p. Since the state of processor p alternates betweenn the active state of average duration T_p and the inactive state of mean duration

$$\sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\lambda_p(i_1)/\varepsilon}$$

the following relationship clearly holds

$$\gamma_p = \frac{1}{T_p + \sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\lambda_p(i_1)/\varepsilon}}.$$

Thus,

$$T_p = \frac{1}{\gamma_p} - \sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\lambda_p(i_1)/\varepsilon}.$$

Furthermore, for the mean waiting time W_p of processor p it follows that

$$W_p = T_p - \sum_{i_2=1}^{r_2} \pi_{i_2}^{(2)} \frac{1}{\mu(i_2)}.$$

4.4 Average number of requests served during a busy period

A pair of an idle period followed by an busy period is called a cycle, whose mean length is dennoted by C. Clearly,

$$C = \frac{1}{\varepsilon^{N-1}\Lambda} + \sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\sum_{p=1}^{N} \lambda_p(i_1)/\varepsilon}.$$

Denote by N_p the mean number of requests of processor p served during a cycle. The throughput γ_p of processor p is then given by $\gamma_p = N_p/C$, which yields that the total number of requests served during an busy peirod is

$$\sum_{p=1}^{N} N_p = \sum_{p=1}^{N} \gamma_p C.$$

4.5 Mean number of active processors

Let us denote by $Q^{(p)}$ the steady-state probability that processor p is idle. Clearly, we have

$$Q^{(p)} = \gamma_p \sum_{i_1=1}^{r_1} \pi_{i_1}^{(1)} \frac{1}{\lambda_p(i_1)/\varepsilon}.$$

Hence, the mean number of active processors is

$$\sum_{p=1}^{N}(1 - Q^{p)}) = N - \sum_{p=1}^{N} Q^{(p)}.$$

5 Numerical Results

This section presents a number of validation experiments (c.f., Tables 1-8) examining the credibility of the proposed approximation against exact results for the performance measure of processor utilization at equilibrium. Note that an exact formula for the utilization is known only when the system is not effected by random environment and it is given (via Palm-formula) by

$$U_p^* = \frac{1}{N} \frac{\sum_{k=1}^{N} \binom{N}{k} k! \rho^k}{1 + \sum_{k=1}^{N} \binom{N}{k} k! \rho^k},$$

where $\rho = \frac{\lambda/\varepsilon}{\mu}$.
In this case relations (9-11)) reduce to the following approximation

$$U_p = \frac{1}{N} \frac{N!}{N! + \left(\frac{\mu}{\lambda/\epsilon}\right)^N}.$$

The following results are derived:

Table 1

$N = 3$

ρ	U_p^*	U_p
1	0.3125	0.285714286
2	0.329113924	0.326530612
2^2	0.332657201	0.332467532
2^3	0.333237575	0.333224862
2^4	0.333320592	0.333319771
2^5	0.333333169	0.333331638
2^6	0.333333125	0.333333121
2^7	0.333333307	0.333333307
2^8	0.333333333	0.333333333

Table 2

$N = 4$

ρ	U_p^*	U_p
1	0.246153846	0.24
2	0.249605055	0.249350649
2^2	0.249968310	0.249959317
2^3	0.249997756	0.249997457
2^4	0.249999999	0.249999999
2^5	0.25	0.25

Table 3

$N = 5$

ρ	U_p^*	U_p
1	0.199386503	0.198347107
2	0.199968409	0.199947930
2^2	0.199998732	0.199998372
2^3	0.199999955	0.199999949
2^4	0.199999998	0.199999998
2^5	0.2	0.2

Table 4

$N = 6$

ρ	U_p^*	U_p
1	0.166581502	0.166435506
2	0.166664473	0.166666305
2^2	0.166666623	0.166666661
2^3	0.166666666	0.166666666

Table 5

$N = 7$

ρ	U_p^*	U_p
1	0.142846715	0.1428828804
2	0.142857009	0.142856921
2^2	0.142857142	0.142857141
2^3	0.142857143	0.142857141

Table 6

$N = 8$

ρ	U_p^*	U_p
1	0.124998860	0.1249969
2	0.124999993	0.124999988
2^2	0.125	0.125

Table 7

$N = 9$

ρ	U_p^*	U_p
1	0.111110998	0.111110805
2	0.111111111	0.111111111

Table 8

$N = 10$

ρ	U_p^*	U_p
1	0.099999999	0.99999999
2	0.1	0.1

It can be observed from Tables 1-8 that the approximate values for $\{U_p\}$ are very much comparable in accuracy to those provided by the exact results for $\{U_p^*\}$. However, the computational complexity, due to the proposed approximation, has been considerably reduced. As λ/ε becomes greater that μ, the $\{U_p\}$ approximations, as expected, approach the exact values of $\{U_p^*\}$. Clearly, the greater the number of processors the less number of steps are needed to reach the exact results.

Acknowledgement. The author is very grateful to Prof. Hideaki Takagi for providing him with the paper Ishigaki *et al.* (1990). His thanks are also due to Prof. M. Arató for his helpful remarks which greatly improved the presentation.

References

[1] Ajmone Marsan M., Balbao G., Conte G. (1986) *Performance Modeling of Multiprocessor Systems,* The MIT Press, Cambridge, Massachusetts.

[2] Anisimov V. V., Zakusilo O.K., Donchenko V.S. (1987) *Elements of queueing theory and asymptotic analysis of systems,* Visha Skola, Kiev (in Russian).

[3] Anisimov V.V., Sztrik J. (1989) Asymptotic analysis of some controlled finite-source queueing systems, *Acta Cybernetica* **9** 27-38.

[4] Baccelli F., Makowski A. (1986) Stability and bounds for single server queues in random environment, *Communications in Statistics: Stochastic Models* **2** 281-292.

[5] Bodnar B.D., Liu A.C. (1989) Modeling and Performance Analysis of Single-Bus Tightly-Coupled Multiprocessors, *IEEE Transactions on Computers* **38** 464-470.

[6] Gaver D.P., Jacobs P.A., Latouche G. (1984) Finite birth-and-death models in randomly changing environments, *Advances in Applied Probability* **16** 715-731.

[7] Gelenbe E. (1988) *Multiprocessor Performance,* John Wiley and Sons, New York.

[8] Gelenbe E. Rosenberg C. (1990) Queues with slowly varying arrival and service processes, *Management Science* **36** 928-937.

[9] Gertsbakh I.B. (1984) Asymptotic methods in reliability theory; A review, *Advances in Applied Probability* **16** 147-175.

[10] Gertsbakh I.B. (1989) *Statistical Reliability Theory,* Marcel Dekker, Inc. New York.

[11] Ishigaki M., Takagi H., Takahashi Y., Hasewaga T. (1990) Throughput and Fairness Analysis of Prioritized Multiprocessor Bus Arbitration Protocols, *Research Report* RT 0051 08/17/90, IBM Research, Tokyo Research Laboratory.

[12] Neuts M. (1978) The M/M/1 Queue with Randomly Varying Arrival and Service Rates, *Opsearch* **15** 139-157.

[13] Nogami S., Sumita S., (1989) Approximate Analysis of Memory Contetion in a Multiprocessor System with a Single Bus, *Electronics and Communications in Japan,* Part 1, Vol. 27, No. 8 65-74.

[14] Rosenberg C., Mazumdar R., Kleinrock L. (1990) On the Analysis of Exponential Queueing Systems with Randomly Changing Arrival Rates: Stability Conditions and Finite Buffer Scheme with Resume Level, *Performance Evaluation* **11** 283-292.

[15] Sengupta B. (1990) A queue with service interruptions in an alternating random environment, *Operations Research* **38** 308-318.

(Received April 4, 1991)