Optimal Routing to Parallel Heterogeneous Servers-Small Arrival Rates

ZVI ROSBERG AND ARMAND M. MAKOWSKI, MEMBER, IEEE

Abstract-Consider a set of $k(\geq 2)$ heterogeneous and exponential servers which operate in parallel. Customers arrive into a single infinite capacity buffer according to a Poisson process, and are routed to available servers in accordance with some routing policy. We show that for arrival rates in some positive interval (0, λ_0], every routing policy which minimizes the long-run expected holding cost is contained in the set of routing policies that minimize the expected flow time for a system with fixed initial population and no new arrivals.

I. INTRODUCTION

TN this paper we consider a queueing system which is composed of an infinite capacity buffer (or queue) attended by k exponential servers operating at rates $\mu_1 > \mu_2 > \cdots > \mu_k$. Customers arrive into the system according to a Poisson process with rate λ , and are served on a first-come first-served basis in accordance with a routing policy (to be defined below). Customers in service cannot be preempted. Throughout, we assume the stability

condition $\lambda < \mu := \sum_{i=1}^{k} \mu_i$. To fix the notation, for all $t \ge 0$, let N(t) denote the number of customers in the queue at time t, and let e(t) = $(e_1(t), e_2(t), \dots, e_k(t))$ denote the state of the k servers at time t, with the understanding that $e_i(t) = 1$ if server i is busy and $e_i(t) = 0$ otherwise. Clearly, X(t) = (N(t), e(t)) is a natural state variable, and we use the notation $\mathbf{X} = \{X(t), t \ge 0\}$ for the stochastic process which describes the evolution of the buffer content and of the activity level of the servers. The state space of **X** is $S = \{0, 1, \dots\} \times \{0, 1\}^k$ and, for every $x = (n, \hat{e})$ in S, we set $|x| = n + e_1 + \cdots + e_k$.

A routing policy π is any rule which at every time $t \ge 0$ stipulates which idle servers to activate; this decision is made on the basis of past states $\{X(s), 0 \le s \le t\}$ and past decisions up to time t; the set of all such routing policies is denoted by Π . With a holding cost accrued at a fixed rate of c > 0 per unit time, the long-run average cost associated with any policy π in Π is then defined by

$$J_{\pi}(x) := \limsup_{T \to \infty} E_{x}^{\pi} \left[\frac{1}{T} \int_{0}^{T} c \cdot |X(t)| dt \right], \qquad x \in S \quad (1)$$

where $E_{x}^{\pi}[\cdot]$ denotes the expectation with respect to the probability measure induced by the policy π on the process X starting in state x. Note that $|X(t)| = N(t) + e_1(t) + \cdots + e_k(t)$ is the total number of customers in the system at time t. A routing policy π^* is said to be *average cost optimal* if it minimizes (1), i.e., if

$$J_{\pi^*}(x) < J_{\pi}(x), \qquad x \in S \tag{2}$$

for any other policy π .

Manuscript received October 21, 1988; revised October 10, 1989 and March 16, 1990. Paper recommended by Associate Editor, S. Verdu. This work was performed while the authors were visiting the IBM T. J. Watson Research Center, Yorktown Heights, NY 10958. Z. Rosberg is with the Department of Computer Science, Technion—Israel

Institute of Technology, Haifa, Israel.

A. M. Makowski is with the Department of Electrical Engineering and the Systems Research Center, University of Maryland, College Park, MD 20742. IEEE Log Number 9035880.

For the exponential system considered here, the optimization problem associated with (1) falls within the purview of continuous-time Markov decision processes which are uniformizable, i.e., which are equivalent to uniformized discrete-time Markov decision processes [7]. The reader is referred to [5] for details where the same problem with k = 2 is treated. To define the discrete-time decision process, consider that at any given instant, each server is working either on a real customer, if activated, or on a dummy customer otherwise. Dummy customers always return to the queue upon completing service and incur no contribution to the cost. Free transitions are associated either with arrivals or with service completions at one of the servers of a customer-either real or dummy. These free transitions occur according to a Poisson process of rate $\lambda + \mu$. A (free) transition due to an arrival occurs with probability $\lambda/(\lambda + \mu)$, whereas a transition due to a service completion at server *i* occurs with probability $\mu_i/(\lambda + \mu)$. If in state x before a transition, the process will jump after this transition to a state which depends on the current state x and on the action taken under the policy π in use. The cost function for using policy π which corresponds to (1) is then given by

$$V_{\pi}(x) := \limsup_{N \to \infty} E_x^{\pi} \left[\frac{1}{N+1} \sum_{m=0}^{N} c \cdot |X(m)| \right], \qquad x \in S \quad (3)$$

where X(m) now denotes the state sampled at the *m*th transition. We also need the total β -discounted cost ($0 < \beta < 1$) associated with the policy π , which is defined by

$$V_{\pi}^{\beta}(x) := E_{x}^{\pi} \left[\sum_{m=0}^{\infty} \beta^{m} c \left| X(m) \right| \right], \qquad x \in S.$$
(4)

A routing policy which is optimal for the β -discounted problem associated with (4) is called a β -optimal policy.

Since the cost function is linear in the state variable and the total number of customers in the system changes by at most one at every transition, it is well known that a β -optimal policy exists and that it can be taken in the class of Markov stationary policies [15]. One of the conclusions of Section II is that the same result also holds for the long-run average cost criterion (3). Furthermore, under the ergodicity condition $\lambda < \mu$, for every stationary policy π , (3) exists as a limit which is independent of the initial state x.

Here it is convenient to identify a stationary policy π with a function $\pi: S \to 2^{\{1, \cdots, k\}}$ in the following way. Assume that a free transition-either an arrival or a service completion-occurs that would make the state jump to x = (n, e) if no action were taken. The policy π activates the idle servers that make the state jump instantaneously from x to $\pi(x) = (\pi_0(n), \pi_1(e_1), \cdots, \pi_k(e_k)),$ where $|\pi(x)| = |x|$ and $\pi_i(e_i) \ge e_i$, $1 \le i \le k$. For the problem at hand it can be shown that the optimal policy satisfies $\pi(\pi(x)) =$ $\pi(x)$ for every state x in S. It therefore suffices to consider only policies with this property, as we do from now on.

The problem considered here has a fairly recent history. It was first studied by Larsen [4] who conjectured that the optimal policy would be of threshold type (as explained below). In [1], Agrawala et al. studied a version of the problem under the as-

0018-9286/90/0700-0789\$01.00 © 1990 IEEE

sumptions that the system has an initial load of *n* customers and no new customers enter the system, i.e., $\lambda = 0$. They determined a simple policy which minimizes the expected flow time, a cost structure which is the natural analog of the cost (1) for $\lambda = 0$. This optimal policy (denoted hereafter by π^*) has the following simple form [1]. Set

$$R_j := \frac{\mu_1 + \dots + \mu_{j-1}}{\mu_j} - (j-1), \qquad 1 < j \le k \qquad (5)$$

and define $R_1 = 0$. If there are *n* customers that remain unprocessed and server *j* is the fastest server available (i.e., with the largest μ_j), then the idle server *j* is activated—and a customer dispatched to it—if and only if $n > R_j$.

The conjecture of Larsen on the threshold form of the optimal policy was settled in the affirmative by Lin and Kumar [5] for k = 2. Using policy iteration, they showed that under the optimal policy, the fast server is always utilized and there is a number $R(\lambda)$ (which depends on λ) such that the slow server should be utilized if and only if the number of customers in the queue exceeds $R(\lambda)$. It was also conjectured there that as $\lambda \downarrow 0$, $R(\lambda)$ increases and converges to the quantity R_2 given by (5). In [16] Walrand provided simple stochastic coupling arguments to prove the optimality of the threshold policy.

As we review the conjecture made in [5] and the optimality results of [1], we are naturally led to entertain the idea that the threshold policy π^* defined above should be optimal for *small* enough values of the arrival rate λ . It is the very purpose of this paper to show in what precise sense this conjecture is indeed correct. To that end, we consider the following class of threshold policies. Let there be *n* customers waiting in the queue and let *j* be the fastest idle server. If $n > R_j$, then server *j* is activated and begins service on one of the customers in the queue. If $n = R_j$, then server *j* may either be activated or not. If $n < R_j$, then no server is activated. Note that these policies may not coincide with π^* if one of the threshold values (5) assumes an integer value.

Our main results, which were already announced in [12], are as follows. There exists $\lambda_0 > 0$ with the property that every optimal policy for the system with arrival rate λ in $(0, \lambda_0]$ is contained in this class of simple threshold policies. We also show that every such threshold policy is optimal when $\lambda = 0$, i.e., minimizes the expected flow time. As a result, if the thresholds R_i are all noninteger, then the optimal policy in the limiting case $\lambda = 0$ is unique and therefore also optimal for every λ in $[0, \lambda_0]$.

This light traffic result is established for an arbitrary k, and reduces in the case k = 2 to the conjecture of [5]. Independently of the authors, Reiman [9] has established the conjecture for k = 2 by a completely different method of proof. Reiman's approach is based on the light traffic theory developed by Reiman and Simon [10], and uses crucially the fact that only threshold policies need to be considered since the optimal policy is known to be of threshold type [5].

For arbitrary k > 2 and $\lambda > 0$, it is still open whether the optimal policy is of threshold type. Given that it is, we should not expect these thresholds to be the ones given by (5), for the optimal thresholds clearly should depend on the arrival rate λ and on the server states **e** [5]. However, from a practical viewpoint, there is value in trying to understand how knowledge of the optimal policy in the limiting case could be put to use in deriving good policies for an arbitrary value λ of the arrival rate. Although this problem is still very much open, an asymptotic analysis as in [2], [3], [13] might provide a possible starting point for dealing with this problem. Using this method, good policies may be obtained by minimizing the first few leading terms in the power series (in λ) of the value function. The limiting case with no arrivals is obtained by minimizing the first leading term.

In the process of deriving this optimality result for small arrival rates, we have gained information into the structural form of average cost optimal policies. For any arrival rate, a server is not activated unless all faster servers are busy, whereas for small

arrival rates, server i, $1 \le i \le k$ is activated if the queue size exceeds a threshold level that depends only on i. Of interest is also the method of proof used here. Unlike the approach in [2], [3] which is limited to problems with finite state-space and discounted cost, our method deals with an infinite state-space situation under the average cost criterion.

The paper is organized as follows. In Section II, we use pathwise comparisons as in [16] to show that the search for an optimal policy can be restricted to a finite set of Markov stationary policies with certain properties. In Section III, we establish various continuity properties in the arrival rate λ , from which the final result is derived.

II. REDUCTION TO A FINITE SET OF POLICIES

In this section we use arguments similar to those presented in [16] in order to show that there exists a *finite* set of states S_0 with the property that for every arrival rate $0 \le \lambda < \mu$, the average cost optimal policy always activates all servers whenever the state lies outside S_0 . In particular, this implies that we may restrict attention to a *finite* set of policies, a fact crucial for proving various continuity properties. We first show this reduction in the context of the β -discounted cost problem for $0 < \beta_0 \le \beta < 1$. The analog result for the average cost criterion, and the fact that the optimal policy exists and is stationary, is then an immediate consequence of [6].

The following two lemmas correspond to properties (2) and (1), respectively, of [16, Lemma 3.2], with similar proofs. We include the proofs for the sake of completeness, and elaborate on some details that were not explained in [16].

To fix the notation, all the proofs in this section are based on pathwise comparison arguments between an original state process \mathbf{X} under a given policy π , and the state process $\mathbf{\hat{X}}$ under another policy $\tilde{\pi}$ derived from π . The latter system is referred to as the *tilde system*, and we use a tilde to denote all relevant quantities in this tilde system.

Lemma 2.1: For every $0 < \beta < 1$, the β -optimal policy has the property that whenever it activates a server, it activates the fastest available one.

Proof: Let π be any given policy and let X(0) = x be an initial state in which π activates server i_2 and leaves the fastest available server i_1 idle. We shall show that π can be strictly improved.

Define a policy $\tilde{\pi}$ and a corresponding process $\tilde{\mathbf{X}}$ as follows. With initial state $\tilde{X}(0) = X(0) = x$, at time t = 0, $\tilde{\pi}$ takes the same action as π , except that it activates server i_1 instead of server i_2 . From then on, the realizations of \mathbf{X} and of $\tilde{\mathbf{X}}$ are coupled by feeding both systems with the same arrival process and by assuming that the first service time \tilde{T}_{i_1} at server i_1 in the tilde system is given by $\tilde{T}_{i_1} = (\mu_{i_1}/\mu_{i_1})T_{i_2}$ (where T_j is the service time of a customer at server j). This coupling is made possible by the fact that T_{i_2} being exponentially distributed with parameter μ_{i_1} , \tilde{T}_{i_1} is also exponential with parameter μ_{i_2} .

parameter μ_{i_2} , \tilde{T}_{i_1} is also exponential with parameter μ_{i_1} . After time t = 0, policy $\tilde{\pi}$ mimics the actions of policy π with one exception: with τ denoting the first time at which π activates server i_1 , if $\tau < T_{i_2}$, then $\tilde{\pi}$ activates server i_2 at time τ instead of server i_1 . (Observe that in the tilde system, server i_2 is available at time τ since $\tau < T_{i_2}$ which in turn implies that i_2 has not been activated under π until time τ , and therefore neither under $\tilde{\pi}$. Also, $\tilde{\pi}$ is feasible since the instant τ and the event { $\tau < T_{i_2}$ } can be emulated in the tilde system. This follows from the coupling between T_{i_2} and \tilde{T}_{i_1} and the inequality $T_{i_2} > \tilde{T}_{i_1}$. Indeed, as long as the service \tilde{T}_{i_1} is not over, neither is the service T_{i_2} . When \tilde{T}_{i_1} is over, then T_{i_2} is exactly known.)

For all realizations in **X** where the event $\{\tau < T_{i_2}\}$ occurs, we reach in both systems at time τ a state which is the same except for the following. In \tilde{X} the customer at server i_1 has been given some service while in **X** it has not; the converse is true for the customer at server i_2 . To continue the coupling, observe that at time τ , the residual service times of these two customers which have been given some service are still exponential given that the event $\{\tau < T_{i_2}\}$ occurs. Hence, $\hat{\mathbf{X}}$ and \mathbf{X} are in the same state at time τ (given that the event $\{\tau < T_{i_2}\}$ occurs), and therefore will evolve in the same manner from that time onward, i.e., $|\hat{\mathbf{X}}(t)| = |\mathbf{X}(t)|$ for all $t \ge 0$.

For all other realizations with $T_{i_2} \leq \tau$, we have $|\tilde{X}(t)| = |X(t)| - 1$ for $\tilde{T}_{i_1} < t < T_{i_2}$ and $|\tilde{X}(t)| = |X(t)|$ otherwise. Since $\beta < 1$ and the event that one customer leaves under $\tilde{\pi}$ before the corresponding one leaves under π occurs with positive probability, it follows that $\tilde{\pi}$ strictly improves π .

Hereafter, we may restrict attention to policies which always activate the fastest available servers. Another property of the β -optimal policies is given by the following lemma.

Lemma 2.2: For every $0 < \beta < 1$, the β -optimal policy always activates the fastest server 1.

Proof: Let X(0) = x be an initial state such that N(0) > 0and $e_1(0) = 0$, and let π be a policy that does not activate server 1 in that state. We show that π can be strictly improved.

From Lemma 2.1 we may assume that π does not activate any server at time t = 0 (otherwise π would not be optimal and we are done). Define a policy $\tilde{\pi}$ and a corresponding state process $\tilde{\mathbf{X}}$ in which $\tilde{X}(0) = X(0) = x$ and all arrivals and service requirements are coupled with those of the original system. At time t = 0, $\tilde{\pi}$ activates server 1 and continues as follows.

Let \tilde{T}_1 be the service time of the customer which is dispatched to server 1 in the tilde system, and let τ denote the first time at which π activates a server (which is necessarily server 1 by virtue of Lemma 2.1). For every realization where $\{T_1 \ge \tau\}$, we define $\tilde{\pi}$ to mimic all actions of π in **X**, except for its first activation of server 1 at time τ . For every realization where $\{T_1 < \tau\}$, we take $\tilde{\pi}$ to mimic all actions of π in **X** forever.

The policy $\tilde{\pi}$ is clearly feasible and we see that for every realization, $|\tilde{X}(t)| \leq |X(t)|$ for all $t \geq 0$. Furthermore, by coupling the realizations, we have $|\tilde{X}(t)| = |X(t)| - 1$ for a period of τ units of time. Therefore, every policy π that does not activate server 1 whenever it is available can be improved by a policy that does so.

Hereafter, we may further restrict attention to policies with the additional property that server 1 be kept active whenever possible.

Observe that multiple dispatching is not excluded. However, on the basis of Lemma 2.1, we may use the following convention. Whenever multiple dispatching takes place, we regard it as a sequence of single dispatchings which are performed instantaneously in increasing order. We also agree to consider as feasible states all instantaneous states that **X** undergoes. Thus, for every stationary policy, the set A_{π}^{i} of states in which server *i*, $1 \leq i \leq k$ is being activated under π is completely determined by

$$A_{\pi}^{i} = \{x = (n, \mathbf{e}) \in S | n > 0, \ e_{j} = 1$$

for $1 \le j \le i$, $e_{i} = 0$ and $\pi(x) = (n - 1, \mathbf{e} + 1_{i})\}$

where 1_i is a k-tuple whose elements are zero except for the *i*th element which equals 1.

The next lemma, roughly speaking, corresponds to properties (3) and (4) of [16, Lemma 3.2]. However, its proof requires a more delicate argument which is based on the uniform bounds we now discuss.

Let $0 < \beta_0 < 1$ be an arbitrary discount factor and let π_{β}^* be the β -optimal policy. For every server i < k and state $x = (0, \mathbf{e})$ with $e_i = e_k = 0$, define

$$\gamma_{i}^{\beta}((0, \mathbf{e})) = V_{\pi_{\beta}}^{\beta}((0, \mathbf{e} + 1_{k})) - V_{\pi_{\beta}}^{\beta}((0, \mathbf{e} + 1_{i})),$$

$$\beta_{0} \le \beta \le 1 \quad (6)$$

$$\phi^{\beta}((0, \mathbf{e})) = V^{\beta}_{\pi^{*}_{\beta}}((0, \mathbf{e} + 1_{k})) - V^{\beta}_{\pi^{*}_{\beta}}((0, \mathbf{e})),$$

$$\beta_{0} < \beta < 1.$$

For $\beta = 1$, these differences are understood as the limit (when $N \to \infty$) of the differences between the costs until step N.

The function $\gamma_i^{\beta}((0, \mathbf{e}))$ expresses the difference in the optimal discounted cost between a system that starts in state (1, \mathbf{e}) and activates server k at time t = 0, and a system that activates server i. The function $\phi^{\beta}((0, \mathbf{e}))$ expresses the difference in the optimal discounted cost between a system that starts in state (1, \mathbf{e}) and activates server k at time t = 0, and a system that starts in state (0, \mathbf{e}) and follows the optimal policy. The latter is used for its simpler structure and for the fact that it bounds the former.

By a simple pathwise comparison it is easy to verify that $\mathcal{V}^{\beta}_{\pi^{*}_{\theta}}((0, \mathbf{e})) \leq \mathcal{V}^{\beta}_{\pi^{*}_{\theta}}((0, \mathbf{e} + \mathbf{1}_{i}))$ which immediately implies

$$\gamma_i^{\beta}((0, \mathbf{e})) \le \phi^{\beta}((0, \mathbf{e})), \qquad 1 \le i \le k.$$
(8)

Next, we shall derive a uniform upper bound to $\phi^{\beta}((0, \mathbf{e}))$ that does not depend on \mathbf{e} , λ , or β . From (8), this will also be a uniform bound to $\gamma_i^{\beta}((0, \mathbf{e}))$ for every *i*.

For every initial state $X(0) = (0, \mathbf{e})$ as above, and every given β , $0 \le \beta_0 \le \beta < 1$, define a policy $\tilde{\pi}^*_{\beta}$ and a corresponding process \mathbf{X} that mimics π^*_{β} as follows.

Again, arrivals and service requirements are coupled in both systems and the initial state in the tilde system is taken to be $\hat{X}(0) = (0, \mathbf{e} + \mathbf{1}_k)$. Let $\tilde{\sigma}_k$ be the service requirement of the first customer served by server k in the tilde system. The policy $\tilde{\pi}_{\beta}^*$ mimics all actions of π_{β}^* in X as long as possible. For a realization where $\tilde{\sigma}_k$ is longer than the first time that π_{β}^* activates server k, $\tilde{\pi}^*_{\beta}$ mimics the actions of π^*_{β} except for activating server k. The customer dispatched by π_{β}^{*} to server k is "marked" by $\tilde{\pi}_{\beta}^{*}$ and is not served until $\tilde{X}(t) = (\tilde{1}, \mathbf{e}')$ for some \mathbf{e}' with $|\mathbf{e}'| < k$. Such an event thereafter is referred to as a boundary hitting event. This enables $\tilde{\pi}^*_{\beta}$ to mimic the actions of π^*_{β} until the boundary hitting instant. Observe that at a boundary hitting instant, $\tilde{N}(t) = \tilde{0}$ and the servers state in **X** is the same as in **X**. At that instant $\tilde{\pi}_{\beta}^{*}$ dispatches the marked customer to the fastest available server and causes the state in the tilde system to jump to $(0, e'+1_j)$ for some j. The state of the original system at that instant is (0, e'). For a realization where $\tilde{\sigma}_k$ is not longer than the first time at which π_B^* activates server k, $\tilde{\pi}^*_{\beta}$ mimics the actions of π^*_{β} forever.

At a boundary hitting epoch, both systems are in similar states as at time t = 0, with the difference that server j in the tilde system is playing the role of server k. From then on (under a boundary hitting event), $\tilde{\pi}^*_{\beta}$ mimics π^*_{β} in a similar manner as before (replace the index k by j), etc. Since π^*_{β} is assumed to be known, it is easy to see that $\tilde{\pi}^*_{\beta}$ is feasible in the tilde system.

Since $\tilde{\pi}^*_{\beta}$ is not necessarily optimal, we conclude from (7) that

$$\phi^{\beta}((0, \mathbf{e})) \leq V^{\beta}_{\bar{\pi}^{*}_{\beta}}((0, \mathbf{e} + \mathbf{1}_{k})) - V^{\beta}_{\pi^{*}_{\beta}}((0, \mathbf{e})), \qquad \beta_{0} \leq \beta \leq 1.$$

For every initial state $X(0) = (0, \mathbf{e})$, let $\tau(\mathbf{e})$ be the first time $\tilde{X}(t)$ hits a boundary state $(1, \mathbf{e}')$, with $|\mathbf{e}'| < k$, and define

$$\bar{\tau} = \max E[\tau(\mathbf{e})] < \infty. \tag{10}$$

(9)

The finiteness of $E[\tau(\mathbf{e})]$ will become apparent in Section III. Noting

$$0 \le |\tilde{X}(t)| - |X(t)| \le 1, \qquad t \ge 0$$
(11)

we are now in a position to bound $\phi^{\beta}((0, e))$.

(7)

Let $\{\tau_n, n = 0, 1\cdots\}$ (with $\tau_0 = 0$) be the successive boundary hitting epochs in the process \tilde{X} ; if for some $n = 0, 1, \cdots, \tilde{X}(\tau_n) = X(\tau_n)$, then $\tau_m = \infty$ for m > n. For every boundary hitting epoch τ_n , if $\tilde{X}(\tau_n) \neq X(\tau_n)$, then from (10) and (11), we get

$$E\left[\sum_{t=\tau_n}^{\tau_{n+1}-1} (|\tilde{X}(t)| - |X(t)|)\beta^t\right]$$
$$\leq E\left[\sum_{t=\tau_n}^{\tau_{n+1}-1} (|\tilde{X}(t)| - |X(t)|)\right] \leq \bar{\tau}.$$
 (12)

Otherwise, the left-hand side of (12) equals zero. From (6)-(12),

$$\gamma_{i}^{\beta}((0, \mathbf{e})) \leq V_{\bar{\pi}_{\beta}^{*}}^{\beta}((0, \mathbf{e} + \mathbf{1}_{k})) - V_{\pi_{\beta}^{*}}^{\beta}((0, \mathbf{e}))$$

$$= \sum_{n=0}^{\infty} c \cdot E \left[\sum_{l=\tau_{n}}^{\tau_{n+1}-1} (|\tilde{X}(l)| - |X(l)|)\beta^{l} \right]$$

$$\leq \sum_{n=0}^{\infty} c \cdot \left(1 - \frac{\mu_{k}}{\lambda + \mu} \right)^{n} \cdot \bar{\tau}$$

$$= c \cdot \frac{\lambda + \mu}{\mu_{k}} \cdot \bar{\tau} \leq \frac{2c\mu\bar{\tau}}{\mu_{k}} := \gamma < \infty.$$
(13)

The second inequality follows from two facts. The first is that $(1 - \mu_k / (\lambda + \mu))^n$ is an upper-bound on the probability that after *n* consecutive boundary hittings, **X** and $\tilde{\mathbf{X}}$ still differ by 1. The second is that once $\mathbf{X}(t)$ and $\tilde{\mathbf{X}}(t)$ agree at time *t*, they will agree forever.

Note that the bound in (13) is independent of the server index *i*, the server state **e**, the arrival rate λ , and the discount factor β .

Lemma 2.3: For every $0 < \beta_0 \le \beta < 1$ and $0 \le \lambda < \mu$, there exists an integer n_0 and a β -optimal policy that activates server k whenever $N(t) \ge n_0$.

Proof: Let X(0) = (n, e) be an initial state such that n > kand $e_k = 0$. Arguing by contradiction, suppose that the β -optimal policy π_{β}^* does not activate server k at that state. We show below that if n is large enough, then π_{β}^* can be improved irrespectively of $0 < \beta_0 \le \beta < 1$.

Let σ_k be an exponential r.v. with parameter μ_k which is independent of everything else in the system. Also, let τ_k be the first time that server k is activated under π_{β}^* ; and τ_0 be the first time that N(t) under π_{β}^* equals zero. The r.v.'s τ_k and τ_0 are integer-valued and are measured from time 0 in steps of the process **X**. Note that τ_k could be infinite. Define $\tau = \min{\{\tau_0, \tau_k\}}$

Define a policy $\tilde{\pi}^*_{\beta}$ and a corresponding process $\tilde{\mathbf{X}}$ that mimics π^*_{β} as follows. The arrivals and service requirements are coupled in both systems and the tilde system starts in state $\tilde{X}(0) = X(0)$. At time 0, $\tilde{\pi}^*_{\beta}$ activates server k in addition to all the servers that π^*_{β} does activate. One customer in the queue of \mathbf{X} that is not dispatched at time t = 0 is "marked" and without loss of generality is treated differently. That is, if $\{\tau_k < \tau_0\}$, then t is the marked customer who is dispatched only if it is the last one in the queue. There is no loss of generality in assuming that the marked customer is dispatched to server k in the tilde system at time t = 0. Observe that the service time of the marked customer in the tilde system has the same distribution as σ_k . After time t = 0, $\tilde{\pi}^*_{\beta}$ mimics the actions of π^*_{β} with the following differences.

For every realization where $\{\sigma_k \leq \tau\}$, $\tilde{\pi}^*_{\beta}$ mimics all the actions of π^*_{β} forever. For every realization where $\{\sigma_k > \tau_k, \tau_k \leq \tau_0\}$, $\tilde{\pi}^*_{\beta}$ mimics all the actions of π^*_{β} except for activating of server k at time τ_k . For every realization where $\{\sigma_k > \tau_0, \tau_0 < \tau_k\}$, $\tilde{\pi}^*_{\beta}$ mimics all the actions of π^*_{β} until time τ_0 , and then acts as the optimal policy π^*_{β} would have acted in $\tilde{\mathbf{X}}$. (Note that since at time τ_0 , $\tilde{X}(\tau_0) \neq X(\tau_0)$, $\tilde{\pi}^*_{\beta}$ does not mimic the actions of π^*_{β} after τ_0 .) Note also that under such realizations, $X(\tau_0) = (0, \mathbf{e} + \mathbf{1}_i)$ and $\tilde{X}(\tau_0) = (0, \mathbf{e} + \mathbf{1}_k)$ for some \mathbf{e}

IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 35, NO. 7, JULY 1990

with $e_i = e_k = 0$. (Here comes in the requirement for a uniform bound on $\gamma_i^{\beta}((0, \mathbf{e}))$.)

Under all realizations with $\{\sigma_k \leq \tau\}$, $\tilde{\pi}^*_{\beta}$ gains at least *c* times the waiting time of the marked customer during $[\sigma_k, \tau]$. Under all realizations with $\{\sigma_k > \tau_0, \tau_0 < \tau_k\}$, the difference between $\tilde{\mathbf{X}}$ and \mathbf{X} is at most γ on the average [as follows from (13)]. Otherwise, we have realizations with $\{\sigma_k > \tau_k, \tau_k < \tau_0\}$. Hence, the difference between \mathbf{X} and $\tilde{\mathbf{X}}$ is only due to the fact that the marked customer in \mathbf{X} was dispatched at time t = 0, while in $\tilde{\mathbf{X}}$ it was dispatched at time τ_k . At that time, the residual service time of the marked customer in $\tilde{\mathbf{X}}$ is still exponential with rate μ_k , and therefore both systems will evolve in the same way later on. Moreover, we have $|\tilde{\mathbf{X}}(t)| = |\mathbf{X}(t)|$ for $0 \le t \le \tau_k$.

Thus, for every $\beta_0 \leq \beta \leq 1$ and $0 \leq \lambda < \mu_k$, we have

$$V_{\pi_{\beta}^{*}}^{\beta}((n, \mathbf{e})) - V_{\bar{\pi}_{\beta}^{*}}^{\beta}((n, \mathbf{e})) \geq E\left[c \cdot I\{\sigma_{k} \leq \tau\} \sum_{i=\sigma_{k}}^{\tau} \beta^{i} - I\{\sigma_{k} > \tau_{0}, \tau_{0} < \tau_{k}\} \cdot \gamma\right] := \psi(n, \beta). \quad (14)$$

Next we show that there is an integer n_0 such that $\psi(n, \beta) > \eta > 0$ for every $n \ge n_0$, $0 \le \lambda < \mu$ and $0 < \beta_0 \le \beta \le 1$. Since $\tau \ge 1$ a.s. and σ_k is independent of τ , we conclude from (14) that

$$\psi(n,\beta) \ge \psi(n,\beta_0) \ge c \cdot \frac{\mu_k}{\lambda+\mu} \cdot \beta_0 - \gamma \cdot P_n\{\sigma_k > \tau_0\} \quad (15)$$

where $P_n\{\cdot\}$ is the probability induced by the process **X** under π_{β}^* with initial state $X(0) = (n, \mathbf{e})$.

⁶ Since the queue length is reduced by at most one at every step, we get

$$P_n\{\sigma_k > \tau_0\} \le \left(1 - \frac{\mu_k}{\lambda + \mu}\right)^n \cdot \frac{\mu_k}{\lambda + \mu} \to 0.$$
 (16)

Since $\beta_0 > 0$ and $0 < \mu_k < \lambda + \mu$, it follows from (15) and (16) that there exists some $\eta > 0$ and an integer n_0 (independent of β and λ) such that,

$$\psi(n, \beta) > \eta$$
 for every $n \ge n_0$, $0 \le \lambda < \mu$ and $\beta_0 \le \beta \le 1$.

Therefore, it follows from (14) that for every $0 \le \lambda < \mu_k$ and $\beta_0 \le \beta < 1$, if the β -optimal policy π_{β}^* does not activate server k whenever $N(t) \ge n_0$, then it can be strictly improved, thus contradicting the optimality of π_{β}^* .

From Lemmas 2.1–2.3, it follows that for every $\lambda < \mu$ and every β -discounted problem with $\beta_0 \leq \beta < 1$, we may restrict attention to policies that activate all available servers whenever the queue length is larger than n_0+k (irrespective of the state **e** of the servers). Since **e** assumes only a finite number of values, we have shown that there is a stationary β -optimal policy which activates all servers at states outside a finite set of states. Hence, for all discount factors $\beta_0 \leq \beta < 1$, only a finite number of policies need to be considered for optimality. It immediately follows from [6] that an average cost optimal policy exists and is one of the policies from the finite set above.

III. THE OPTIMAL DISPATCHING FOR SMALL ARRIVAL RATES

In the previous section we have concluded that for every $0 \le \lambda < \mu$, attention may be restricted to a *finite* set of stationary policies when solving the problem under the long-run average cost criterion (1). This finite set of policies is denoted by Π_0 , and consists of the Markov stationary policies under which: i) server 1 is always kept busy whenever possible, i.e., customers (if available) are always dispatched to server 1; and ii) all the servers are

ROSBERG AND MAKOWSKI: OPTIMAL ROUTING TO PARALLEL HETEROGENEOUS SERVERS

activated at states outside a finite set S_0 defined by

$$S_0 = \{ (n, \mathbf{e}) \in S : n < n_0 + k \}$$
(17)

with n_0 as given in Lemma 2.3.

To facilitate the discussion, we need some additional notation as well as several facts regarding the value function of a Markov decision process with a long-run average cost criterion.

Lemma 3.1: Assume $0 < \lambda < \mu$. Under every policy π in II₀, the process **X** is an ergodic Markov chain and the equalities

$$g_{\pi} := J_{\pi}(x) = V_{\pi}(x)$$
$$= \lim_{N \to \infty} E_{x}^{\pi} \left[\frac{1}{N+1} \sum_{m=0}^{N} c \cdot |X(m)| \right], \qquad x \in S \quad (18)$$

hold true, with the interpretation that g_{π} is the long-run average cost incurred by using policy π .

Proof: Since every policy in Π_0 activates all the servers outside a finite set of states S_0 , it is easy to see that for every x = (n, e) not in S_0 , the inequality

$$E^{\pi}[|X(m+1)| - |X(m)| |X(m) = x] \le \frac{\lambda - \mu}{\lambda + \mu} < 0$$
 (19)

holds true for each $m = 0, 1, \cdots$. The first part of the lemma follows from (19) by a generalization of Foster's criterion [8], whereas (18) is now immediate by the mean Ergodic Theorem.

To proceed, we need several additional quantities which we now introduce. Let τ denote the first return time (after time 0) to the empty state $\theta = (0, 0)$, i.e.,

$$\tau := \inf \{ m > 0 : X(m) = (0, 0) \}.$$
(20)

For every policy π in Π , set

$$T_{\lambda}(\pi, x) = E^{\pi}_{\lambda, x}[\tau], \qquad x \in S$$
(21)

and

$$C_{\lambda}(\pi, x) = E_{\lambda, x}^{\pi} \left[\sum_{m=0}^{\tau-1} c \cdot |\mathbf{X}(m)| \right], \qquad x \in S \qquad (22)$$

for all arrival rate $\lambda \geq 0$, with $E_{\lambda,x}^{\pi}[\cdot]$ denoting the expectation operator under policy π when starting in state x given that the arrival rate is λ .

Since the instantaneous cost is linear in the state variable, it follows from [11] that whenever $0 \le \lambda < \mu$, the quantities defined by (21) and (22) are finite, i.e., for every policy π in Π_0 ,

$$T_{\lambda}(\pi, x) < \infty \text{ and } C_{\lambda}(\pi, x) < \infty, \qquad x \in S.$$
 (23)

If, in analogy with (18), $g_{\pi}(\lambda)$ denotes the long-run average cost incurred by using policy π given that the arrival rate is λ , then it is well known that

$$g_{\pi}(\lambda) = \frac{E_{\lambda,0}^{\pi}\left[\sum_{m=0}^{\tau-1} c \cdot |X(m)|\right]}{E_{\lambda,0}^{\pi}[\tau]}.$$
 (24)

Moreover, for every policy π in Π_0 , set

$$h_{\pi}(\lambda, x) := E_{\lambda, x}^{\pi} \left[\sum_{m=0}^{\tau-1} c \cdot |X(m)| - g_{\pi}(\lambda) \cdot \tau \right], \qquad x \in S \ (25)$$

whenever $0 \leq \lambda < \mu$, so that

$$g_{\pi}(\lambda) = \frac{C_{\lambda}(\pi, 0)}{T_{\lambda}(\pi, 0)}$$
(26a)

$$h_{\pi}(\lambda, x) = C_{\lambda}(\pi, x) - g_{\pi}(\lambda) \cdot T_{\lambda}(\pi, x), \qquad x \in S.$$
 (26b)

For every x in S, let Y_x denote the state to which the process will jump next, given that it is at state x and that no dispatching action is taken. Since the distribution of Y_x is independent of the policy in use, we denote by $\mathcal{E}_{\lambda,x}[\cdot]$ the expectation operator of this distribution given that the system is in state x just before this "free" transition and that customers arrive at rate λ . It follows from (24) that for every policy π in Π_0 ,

$$h_{\pi}(\lambda, x) + g_{\pi}(\lambda) = c \cdot |x| + \mathcal{E}_{\lambda, x}[h_{\pi}(\lambda, \pi(Y_{x}))], \qquad x \in \mathbb{S}.$$
(27)

Finally, for every $0 \le \lambda < \mu$, we set

$$g^*(\lambda) := \inf_{\pi \in \Pi} g_{\pi}(\lambda) = \inf_{\pi \in \Pi_0} g_{\pi}(\lambda),$$

i.e., $g^*(\lambda)$ is the optimal value of the long-run average cost. Moreover, we denote by π^*_{λ} any policy in Π_0 which is optimal for the long-run average cost (of course, interpreted as the expected flow time when $\lambda = 0$), so that $g^*(\lambda) := g_{\pi^*}(\lambda)$. The next proposition shows in what sense the quantities defined by (24) and (25) characterize optimality.

Lemma 3.2: Fix λ in $(0, \mu)$. Every average cost optimal policy π_{λ}^{*} in Π_{0} satisfies the relations

$$h_{\pi_{\lambda}^{*}}(\lambda, x) + g^{*}(\lambda) = c \cdot |x| + \mathcal{E}_{\lambda, x}[h_{\pi_{\lambda}^{*}}(\lambda, \pi_{\lambda}^{*}(Y_{x}))]$$
$$= c \cdot |x| + \min_{\pi \in \Pi_{\lambda}} \mathcal{E}_{\lambda, x}[h_{\pi_{\lambda}^{*}}(\lambda, \pi(Y_{x}))]$$

for all x in S.

Proof: The first equality follows from (27) with $\pi = \pi_{\lambda}^*$. To prove the second equality, define the value function for the β -discounted problem by

$$V^{\beta}(\lambda, x) := \inf_{\pi \in \Pi} V^{\beta}_{\pi}(\lambda, x), \qquad x \in S.$$

Observe that for $\beta_0 < \beta < 1$, this definition becomes

$$V^{\beta}(\lambda, x) := \inf_{\pi \in \Pi_0} V^{\beta}_{\pi}(\lambda, x), \qquad x \in S$$
(28)

so that the corresponding optimality equation [14] now takes the form

$$V^{\beta}(\lambda, x) = c \cdot |x| + \beta \min_{\pi \in \Pi_0} \mathcal{E}_{\lambda, x}[V^{\beta}(\lambda, \pi(Y_x))], \qquad x \in S.$$
(29)

This relation can be rearranged to read

$$\begin{split} V^{\beta}(\lambda, x) &- V^{\beta}(\lambda, \theta) + (1 - \beta) V^{\beta}(\lambda, \theta) \\ &= c \cdot |x| + \beta \min_{\pi \in \Pi_0} \mathcal{E}_{\lambda, x} [V^{\beta}(\lambda, \pi(Y_x)) - V^{\beta}(\lambda, \theta)], \qquad x \in S \end{split}$$

(30)

Since $0 < \lambda < \mu$, standard arguments [14] now imply that for every policy π in Π_0 ,

$$\lim_{\beta \uparrow 1} \left[V_{\pi}^{\beta}(\lambda, x) - V_{\pi}^{\beta}(\lambda, \theta) \right] = h_{\pi}(\lambda, x), \qquad x \in S$$
(31)

and

$$\lim_{\beta \uparrow 1} (1-\beta) V^{\beta}(\lambda, \theta) = \lim_{\beta \uparrow 1} \min_{\pi \in \Pi_0} (1-\beta) V^{\beta}_{\pi}(\lambda, \theta)$$
$$= \min_{\pi \in \Pi_0} g_{\lambda}(\pi) = g^*(\lambda).$$

Since Π_0 is finite, it follows from [6] that there exists a sequence $\beta_n \uparrow 1$ with the property that all β_n -optimal policies are the same. Furthermore, this policy is also average cost optimal. Let π_{λ}^* be such a β_n -optimal policy, i.e., $V^{\beta_n}(\lambda, x) = V^{\beta_n}_{\pi_{\lambda}^*}(\lambda, x)$ for all x in S. It is plain from (31) that

$$\lim_{n\to\infty} \left[V^{\beta_n}(\lambda, x) - V^{\beta_n}(\lambda, \theta) \right] = h_{\pi^*_{\lambda}}(\lambda, x), \qquad x \in S.$$
(32)

To conclude, we let β go to 1 along the subsequence $\{\beta_n, n = 0, 1, \dots\}$ in (30). The result now follows upon using (31) and (32) together with the fact that Π_0 is finite. \Box

We now present several continuity properties which are useful in studying the small arrival rate situation.

Theorem 3.1: For every policy π in Π_0 and every state x in S, the mappings $\lambda \to g_{\pi}(\lambda)$ and $\lambda \to h_{\pi}(\lambda, x)$ are all continuous on the interval $[0, \mu_k]$.

Before giving a proof of Theorem 3.1, we obtain a simple and useful consequence of it, that derives from the observation $T_{\lambda}(\pi, \theta) = 1$ and $C_{\lambda}(\pi, \theta) = 0$ for $\lambda = 0$.

Corollary 3.1: For every policy π in Π_0 ,

$$\lim_{\lambda\to 0}g_{\pi}(\lambda)=0 \text{ and } \lim_{\lambda\to 0}h_{\pi}(\lambda, x)=h_{\pi}(0, x), \qquad x\in S.$$

Proof of Theorem 3.1: It is plain from (26) that the conclusion of Theorem 3.1 will be obtained if the mappings $\lambda \to T_{\lambda}(\pi, x)$ and $\lambda \to C_{\lambda}(\pi, x)$ can be shown to be continuous for each x in S.

Fix x in S. For all arrival rate $\lambda \ge 0$ and $n = 0, 1, \dots$, set

$$T_{\lambda,n}(\pi, x) = E_{\lambda,x}^{\pi}[1_{\{\tau \leq n\}}\tau \wedge n]$$

and

$$C_{\lambda,n}(\pi, x) = E_{\lambda,x}^{\pi} \left[\mathbb{1}_{\{\tau \leq n\}} \sum_{m=0}^{\tau \wedge n-1} c \cdot |X(m)| \right].$$

It is plain from the monotone convergence theorem that

$$\lim_{n \to \infty} T_{\lambda, n}(\pi, x) = T_{\lambda}(\pi, x) \text{ and } \lim_{n \to \infty} C_{\lambda, n}(\pi, x) = C_{\lambda}(\pi, x).$$
(33)

Observe that the one-step transition probability matrix $P_{\lambda}(\pi)$ that governs **X** under the policy π given that the arrival rate is λ has continuous entries (in λ). Moreover, for every $n = 0, 1, \cdots$, the expressions $C_{\lambda,n}(\pi, x)$ and $T_{\lambda,n}(\pi, x)$ are both sums with a *finite* number of terms, each of which is continuous in λ on $[0, \infty)$. Consequently, the mappings $\lambda \to C_{\lambda,n}(\pi, x)$ and $\lambda \to T_{\lambda,n}(\pi, x)$ are both continuous on $[0, \infty)$ for every n = $0, 1, \cdots$. It follows from (33) that the mappings $\lambda \to C_{\lambda}(\pi, x)$ and $\lambda \to T_{\lambda}(\pi, x)$ will be *continuous* on the positive interval $[0, \mu_k)$ if the convergence in (33) is *uniform* on every closed subinterval contained in that interval.

To obtain this uniform convergence, note from (23) that whenever $0 \le \lambda < \mu$,

$$0 \leq T_{\lambda}(\pi, x) - T_{\lambda, n}(\pi, x) = E_{\lambda, x}^{\pi}[1_{\{\tau > n\}}\tau], \qquad n = 0, 1, \cdots$$
(34)

and

$$0 \leq C_{\lambda}(\pi, x) - C_{\lambda, n}(\pi, x)$$

= $E_{\lambda, x}^{\pi} \left[1_{\{\tau > n\}} \sum_{m=0}^{\tau-1} c \cdot |X(m)| \right], \qquad n = 0, 1, \cdots.$ (35)

Now consider a new system where all k servers are serving at rate μ_k under a routing policy π' derived from π as follows. The arrivals to the new system are coupled with the arrivals to the

original one. Furthermore, at every instant the states and service requirements of the original system are emulated in the new one. Every customer that is dispatched by π to server *i*, $1 \le i \le k$ is marked accordingly and his service duration T_i is converted to $T'_i = (\mu_i/\mu_k)T_i$. In the new system, π' dispatches to server *i* the same customers as π dispatches to it in the original system, while preserving the system service order within each server. However, the service requirement is now taken to be T'_i . Since arrivals and service requirements are coupled, the policy π is known and since $T' > T_i$, it is easy to verify that π' is indeed a faction.

 $T'_i > T_i$, it is easy to verify that π' is indeed a feasible policy. Since the new system has the same stochastic evolution as the original one, we see by a simple pathwise comparison that $|X(m)| \le |X'(m)|$ for all $m = 0, 1, \cdots$. Therefore, if τ' denotes the first return time (after time 0) to the empty state $\theta = (0, 0)$, i.e.,

$$\tau' := \inf \{m > 0 : X'(m) = (0, 0)\}$$

then $\tau \leq \tau'$, and the inequalities

L

$$0 \le T_{\lambda}(\pi, x) - T_{\lambda, n}(\pi, x)$$

$$\le E_{\lambda, x}^{\pi} [1_{\{\tau' > n\}} \tau'], \qquad n = 0, 1, \cdots$$
(36)

(37)

$$0 \leq C_{\lambda}(\pi, x) - C_{\lambda, n}(\pi, x)$$

$$\leq E_{\lambda, x}^{\pi} \left[1_{\{\tau' > n\}} \sum^{\tau' - 1} c \cdot |X'(m)| \right], \qquad n = 0, 1, \cdots$$

m=0

now follow from (34) and (35). The finiteness of these bounds is guaranteed whenever $\lambda \leq k\mu_k$ [in analogy with (23)].

Next we bound the right-hand side of (36) and (37) by quantities which do not depend on π . To do this, observe that since π is a policy in Π_0 , in the new system π' always dispatches customers to server 1 (which now operates at rate μ_k), and occasionally dispatches customers to other servers (which now also operate at rate μ_k). Therefore, by a similar pathwise comparison as above, we see that $|X'(m)| \leq |X''(m)|$ for all $m = 0, 1, \cdots$ where $\{X''(m), m = 0, 1, \cdots\}$ is the state process obtained by operating the auxiliary system as an M/M/1 queue with arrival rate λ and service rate μ_k . This amounts to modifying π' into a new policy π'' that shuts off all servers except server 1 (which is always kept active since π is a policy in Π_0). In that case, $\{|X''(m)|, m = 0, 1, \cdots\}$ can be interpreted as the queue size process of an M/M/1 queue with arrival rate λ and service rate μ_k . Again, with τ'' denoting the first return time (after time 0) to the empty state θ , i.e., $\tau'' := \inf\{m > 0: X''(m) = (0, 0)\}$, we have $\tau' \leq \tau''$ and the inequalities

$$0 \leq T_{\lambda}(\pi, x) - T_{\lambda, n}(\pi, x)$$
$$\leq E_{\lambda, x}^{\pi}[1_{\{\tau'' > n\}}\tau''], \qquad n = 0, 1, \cdots$$
(38)

and

and

$$0 \leq C_{\lambda}(\pi, x) - C_{\lambda, n}(\pi, x)$$

$$\leq E_{\lambda, x}^{\pi} \left[1_{\{\tau'' > n\}} \sum_{m=0}^{\tau''-1} c \cdot |X''(m)| \right], \qquad n = 0, 1, \cdots$$
(39)

hold true. The right-hand sides of (38) and (39) are finite if $\lambda < \mu_k$, for in that case

$$E^{\pi}_{\lambda,x}[au''] < \infty ext{ and } E^{\pi}_{\lambda,x}\left[\sum_{m=0}^{ au''-1} c \cdot |X''(m)|
ight] < \infty$$

794

in analogy with (23). Moreover, these right-hand sides do not depend on π and are increasing in λ . This last fact can be seen by splitting a Poisson arrival process whose rate is $\lambda + \delta$ into two independent Poissonian arrival streams with rates λ and δ , respectively. The customers that arrive in the δ -stream are served with a lower preemptive priority. By a similar pathwise comparison as above, the monotonicity in λ is easily verified.

From the monotone convergence theorem and from the monotonicity in λ , we now conclude that for every $0 < \lambda_0 < \mu_k$,

$$\lim_{n\to\infty}\sup_{0\leq\lambda\leq\lambda_0}E^{\pi}_{\lambda,x}[1_{\{\tau''>n\}}\tau'']=0$$

and

$$\lim_{n\to\infty}\sup_{0\leq\lambda\leq\lambda_0}E^{\pi}_{\lambda,x}\left[\mathbf{1}_{\{\tau''>n\}}\sum_{m=0}^{\tau''-1}c\cdot|X''(m)|\right]=0.$$

This shows via (36) and (37) that the convergence in (33) is uniform over $[0, \lambda_0]$ and the proof is now complete.

Now, recall that π_{λ}^{*} is any average cost optimal policy given that the arrival rate is λ . Consider the accumulation points of the family of average cost optimal policies $\{\pi_{\lambda}^{*}, 0 < \lambda < \mu\}$ as λ approaches zero. A policy π is such an accumulation point if there exists a sequence $\lambda_{n} \downarrow 0$ such that $\pi_{\lambda_{n}}^{*} = \pi$ for all n = $0, 1, \cdots$. Since Π_{0} is finite, there are only finitely many such accumulation points, say π_{1}, \cdots, π_{J} , each one of them of course in Π_{0} . Consequently, there exists λ_{0} such that $0 < \lambda_{0} < \mu_{k}$, with the property that whenever $0 < \lambda < \lambda_{0}$, then $\pi_{\lambda}^{*} = \pi_{j}$ for some *j*.

The following theorem shows in what sense optimal policies for $\lambda = 0$ are also average cost optimal for small arrival rates.

Theorem 3.2: There exists an interval $(0, \lambda_0]$ with $0 < \lambda_0 < \mu_k$ such that for all λ in $(0, \lambda_0]$, every average cost optimal policy π_{λ}^* is also an optimal policy for $\lambda = 0$. Moreover, if the optimal policy π_0^* for $\lambda = 0$ is unique, then π_{λ}^* is also unique and coincides with π_0^* .

Proof: Clearly, only the first part needs to be established for the second part immediately follows from it.

Take λ_0 as defined in the remarks preceding the statement of the theorem. Fix j, $1 \le j \le J$, and denote by I_j the set of points in $(0, \lambda_0)$ with the property that π_j is average cost optimal when the arrival rate λ is in I_j . From the very definition of I_j we conclude that I_j is nonempty and at least countable. By Lemma 3.2, whenever λ lies in I_j , we have

$$h_{\pi_j}(\lambda, x) + g^*(\lambda) = c \cdot |x| + \mathcal{E}_{\lambda, x}[h_{\pi_j}(\lambda, \pi_j(Y_x))], \quad x \in S$$

while for every other policy π in Π_0 ,

$$h_{\pi_j}(\lambda, x) + g^*(\lambda) \leq c \cdot |x| + \mathcal{E}_{\lambda, x}[h_{\pi_j}(\lambda, \pi(Y_x))], \qquad x \in S.$$

Letting $\lambda \downarrow 0$ in I_i , we conclude from Corollary 3.1 that

$$h_{\pi_j}(0, x) = c \cdot |x| + \mathcal{E}_{0,x}[h_{\pi_j}(0, \pi_j(Y_x))], \qquad x \in S \quad (40)$$

while for every other policy π in Π_0 ,

$$h_{\pi_j}(0, x) \leq c \cdot |x| + \mathcal{E}_{0,x}[h_{\pi_j}(0, \pi(Y_x))], \qquad x \in S.$$
 (41)

Now applying standard arguments we show that (40) and (41) imply that π_j is optimal for $\lambda = 0$. From (40) and (41) we readily conclude for all $m = 0, 1, \dots$, that

$$E_{0,x}^{\pi_j}[h_{\pi_j}(0, X(m))] = E_{0,x}^{\pi_j}[c \cdot |X(m)|] + E_{0,x}^{\pi_j}[h_{\pi_j}(0, X(m+1))], \quad x \in S \quad (42)$$

while for every other policy π in Π_0 ,

$$E_{0,x}^{\pi}[h_{\pi_j}(0, X(m))] \le E_{0,x}^{\pi}[c \cdot |X(m)|] + E_{0,x}^{\pi}[h_{\pi_j}(0, X(m+1))], \qquad x \in S.$$
(43)

Telescoping (42) and (43) in the usual way we find that

$$h_{\pi_j}(0, x) = E_{0, x}^{\pi_j} \left[\sum_{n=0}^m c \cdot |X(n)| \right] + E_{0, x}^{\pi_j} [h_{\pi_j}(0, X(m+1))], \quad x \in S \quad (44)$$

while for every other policy π in Π_0 ,

$$h_{\pi_{j}}(0, x) \leq E_{0, x}^{\pi} \left[\sum_{n=0}^{m} c \cdot |X(n)| \right] + E_{0, x}^{\pi} [h_{\pi_{j}}(0, x(m+1))], \quad x \in S \quad (45)$$

for all $m = 0, 1, \cdots$

It is plain that for every policy π in Π_0 , the growth estimate

$$|h_{\pi}(0, x)| \le K \cdot (1 + |x|^2), \qquad x \in \mathbb{R}$$

holds for some positive constant K. This can be seen by comparing the system under the policy π to an M/M/1 queue with no arrival where the exponential server operates at rate μ_k as was done in the proof of Theorem 3.1.

Recall that server 1 is always active if possible under the policy π in Π_0 . Consequently, when $\lambda = 0$, i.e., when there are no arrivals into the system, then |X(m)| can only decrease until it becomes zero. As a result of these remarks, we conclude that for all $m = 0, 1, \dots$,

$$|h_{\pi_j}(0, X(m))| \le K \cdot (1 + |X(m)|^2) \le K \cdot (1 + |x|^2), \qquad x \in S$$

a.s. under $P_{0,x}^{\pi}$, and invoking the bounded convergence theorem, we obtain

$$\lim_{m \to \infty} E^{\pi}_{0,x}[|h_{\pi_j}(0, X(m))|] = 0, \qquad x \in S$$
(46)

since $X(m) = \theta$ whenever $\tau \leq m$ and therefore $h_{\pi_j}(0, \mathbf{0}) = 0$. Letting *m* go to infinity in (44) and (45), and using (46), we readily conclude that

$$h_{\pi_j}(0, x) = E_{0, x}^{\pi_j} \left[\sum_{m=0}^{\tau-1} c \cdot |X(m)| \right], \qquad x \in S$$

in obvious agreement with one's intuition (see (47) below), while

$$h_{\pi_j}(0,x) \leq E_{0,x}^{\pi}\left[\sum_{m=0}^{\tau-1} c \cdot |X(m)|\right] = h_{\pi}(0,x), \qquad x \in S$$

for every other policy π in Π_0 . In other words, the policy π_j is indeed optimal for $\lambda = 0$.

Our final result complements Theorem 3.2 by focusing on the structure of the optimal policies π_0^* for $\lambda = 0$. These optimal policies are not necessarily unique as we now discuss.

For any policy π in Π_0 , we readily see that for $\lambda = 0$, $h_{\pi}(0, x)$ coincides with the flow time for using π when the system starts in state x [1], i.e.,

$$h_{\pi}(0, x) = E_{0, x}^{\pi} \left[\sum_{m=0}^{\tau-1} c \cdot |X(m)| \right], \qquad x \in S.$$
 (47)

Therefore, the value function for the case $\lambda = 0$ is defined by

$$h^*(x)=\min_{\pi\in\Pi_0}h_{\pi}(0,x), \qquad x\in S.$$

It follows from Theorem 3.1 and Lemma 3.2 that every optimal policy π_0^* (for $\lambda = 0$) satisfies the relations

$$h^{*}(x) = c \cdot |x| + \mathcal{E}_{0,x}[h^{*}(\pi_{0}^{*}(Y_{x}))]$$

= $c \cdot |x| + \min_{x \in \Pi_{0}} \mathcal{E}_{0,x}[h^{*}(\pi(Y_{x}))]$ (48)

as x ranges over S.

The optimal actions at every state x are either to dispatch a customer to the fastest available server, say $\pi_1(x)$, or not to dispatch a customer to any server, say $\pi_0(x)$. These choices are determined by comparing $\mathcal{E}_{0,x}[h^*(\pi_1(Y_x))]$ and $\mathcal{E}_{0,x}[h^*(\pi_0(Y_x))]$. When these quantities are equal, both actions are optimal; otherwise, only one of them is optimal. From [1], it can be verified that if R_i is an integer, then at every state x where the fastest available server is j and the number of customers in the queue equals R_i , these expected values are equal. In that case, to dispatch a customer to server j or not to activate any server are both optimal in state x. Therefore, when at least one of the thresholds is an integer, π_0^* is not unique. Otherwise, if the thresholds R_i are all noninteger, then only one action minimizes the right-hand side of (48) and π_0^* is therefore unique.

In other words, the policy π_0^* is optimal for $\lambda = 0$ if and only if it is of the following form. Let j be the fastest idle server when n customers are waiting in the queue. If $n > R_i$, then one customer from the queue is dispatched to server j. If $n = R_j$, then a customer may or may not be dispatched to server j. If $n < R_j$, then no customer is dispatched to any of the idle servers.

With this information in hand, we now conclude with the following corollary to Theorem 3.2.

Corollary 3.2: If the thresholds R_j , $1 \le j \le k$, are all noninteger, then π_0^* is unique and is also the unique optimal policy for small arrival rates λ . Otherwise, if the thresholds R_j , $1 \le j \le k$, assume an integer value, then π^*_{λ} is given by any π^*_0 except for states x in which j is the fastest available server and the number of customers in the queue equals R_i .

Note Added in Proof: The authors wish to thank Dr. R. Righter for bringing their attention to her recent paper which generalizes the result of Agrawala et al. [1].

References

- A. K. Agrawala, E. G. Coffman, Jr., M. R. Garey, and S. K. Tri-pathi, "A stochastic optimization algorithm minimizing expected flow times on uniform processors," *IEEE Trans. Comput.*, vol. C-33, pp. 351-356, Apr. 1984. M. N. Katehakis and A. Levine, "A dynamic routing problem—
- [2]
- [3]
- M. N. Katehakis and A. Levine, "A dynamic routing problem— Numerical procedures for light traffic conditions," Appl. Math. Com-put., vol. 17, pp. 267-276, 1985.
 , "Allocation of distinguishable servers," Comput. Operat. Res., vol. 13, pp. 85-93, 1986.
 R. L. Larsen, "Control of multiple exponential servers with application to computer systems," Ph.D. dissertation, Univ. Maryland, College Park, MD, Tech. Rep. 1041, 1981.
 W. Lin and P. R. Kumar, "Optimal control of queueing systems with two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-29, pn. 696-703. Aug. 1984. [4]
- [5]
- two heterogeneous servers," *IEEE Trans. Automat. Contr.*, vol. AC-29, pp. 696-703, Aug. 1984.
 S. A. Lippman, "Semi-Markov decision processes with unbounded rewards," *Management Sci.*, vol. 19, pp. 717-731, 1973.
 —, "Applying a new device in the optimization of exponential queueing systems," *Management Sci.*, vol. 23, pp. 687-710, 1975.
 A. G. Pakes, "Some conditions on the ergodic theory of Markov chains," *Oper. Res.*, vol. 17, pp. 1058-1061, 1969.
 M. I. Reiman, "Optimal control of a heterogeneous two server queue in light traffic," AT&T Bell Labs., Murray Hill, NJ, 1989. [6]
- [7]
- [8]
- [9]

- [10] f111
- [12]
- M. I. Reiman and B. Simon, "Open queueing systems in light traffic," Math. Oper. Res., vol. 14, pp. 26-59, 1989.
 Z. Rosberg, "Semi-Markov decision processes with polynomial reward," J. Appl. Prob., vol. 19, pp. 301-309, 1982.
 Z. Rosberg and A. M. Makowski, "Optimal dispatching to parallel heterogeneous servers in light traffic," in Proc. 28th IEEE Conf. Decision Contr., Tampa, FL, Dec. 1989, pp. 1539-1543.
 Z. Rosberg and P. Kermani, "Customer routing to different servers with complete information," Adv. Appl. Prob., vol. 21, pp. 861-882, 1990.
- [13] 1990.
- [14]
- S. Ross, Introduction to Stochastic Dynamic Programming. New York: Academic, 1983.
 M. Schäl, "Conditions for optimality in dynamic programming and for the limit of n-stage optimal policies to be optimal," Z. Warscheinlichkeitstheorie Verw. Gebiete, vol. 32, pp. 179–196, 1975.
 J. Walrand, "A note on the optimal control of a queueing system with two heterogeneous servers," Syst. Contr. Lett., vol. 4, pp. 131–134, 1984 [15]
- [16] 1984
- R. Righter, "Job scheduling to minimize expected weighted flow time on uniform processors," *Syst. Contr. Lett.*, vol. 10, pp. 211-216, [17] 1988.



Zvi Rosberg received the B.Sc., M.A., and Ph.D. degrees from the Hebrew University, Jerusalem, in 1971, 1974, and 1978, respectively.

From 1972 to 1978 he was a Senior Systems Analyst in the General Computer Bureau of the Israeli Government. From 1978 to 1979 he had a research fellowship at the Center of Operation Research and Econometrics (CORE), University of Louvain, Belgium. From 1979 to 1980 he was a Visiting Assistant Professor at the University of Illinois. Since 1980 he has been with the Department of Com-

puter Science, Technion-Israel Institute of Technology, Haifa. During 1985-1987 he was on leave at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. His main research interests include probabilistic models and architecture of communication networks and computer systems, performance evaluation, control of queueing systems, and applied probability.



Armand M. Makowski (M'83) was born in Watermael-Boitsfort, Belgium, on October 9, 1953. He received the Licence en Sciences Mathématiques from the Université Libre de Bruxelles in 1975, the M.S. degree in engineering-systems science from the University of California, Los Angeles, in 1976, and the Ph.D. degree in applied mathematics from the University of Kentucky, Lexington, in 1981. In August 1981, he joined the Faculty of the De-

partment of Electrical Engineering, University of Maryland, College Park, where he is presently an

Associate Professor. He is also affiliated with the Systems Research Center, one of the NSF Engineering Research Centers, since its establishment in 1985 on the College Park campus. Over the past few years, he has held visiting positions at the Technion (Israel), INRIA (France), the IBM T. J. Watson Research Center (Hawthorne), and AT&T Bell Laboratories (Murray Hill). His research interests lie mainly in the applications of the theory of stochastic processes to the modeling, (optimal) design and performance evaluation of discrete-event dynamical systems, with applications to computer and communication systems, and manufacturing. More recently, his efforts have focused on the adaptive control of queueing systems, and on developing computational techniques and bounding methodologies for nonstandard queueing systems with synchronization constraints.

Dr. Makowski was a C.R.B. Graduate Fellow of the Belgian-American Educational Foundation for the academic year 1975-1976; he was also a 1984 recipient of the NSF Presidential Young Investigator Award.