Queueing Systems with Synergistic Servers

Sigrún Andradóttir and Hayriye Ayhan
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205, U.S.A.

Douglas G. Down
Department of Computing and Software
McMaster University
Hamilton, Ontario L8S 4L7, Canada

March 1, 2010

Abstract

We consider tandem lines with finite buffers and flexible, heterogeneous servers who are synergistic in that they work more effectively in teams than on their own. Our objective is to determine how the servers should be assigned dynamically to tasks in order to maximize the long-run average throughput. In particular, we investigate when it is better to take advantage of synergy among servers, rather than exploiting the servers’ special skills, to achieve the best possible system throughput. We show that when there is no tradeoff between server synergy and servers’ special skills (because the servers are generalists who are equally skilled at all tasks), the optimal policy has servers working in teams of two or more at all times. Moreover, for Markovian systems with two stations and two servers, we provide a complete characterization of the optimal policy and show that depending on how well the servers work together, the optimal policy either takes full advantage of servers’ special skills, or full advantage of server synergy (and hence there is no middle ground in this case). Finally, for a class of larger Markovian systems, we provide sufficient conditions that guarantee that the optimal policy should take full advantage of server synergy at all times.

1 Introduction

In the last decade, there has been a growing interest in queueing systems with flexible servers. There is now a significant body of research addressing the question of how servers should be assigned dynamically to tasks as the state of the system evolves. The objective is to utilize each server’s training and abilities to achieve optimal system performance (e.g., to maximize throughput or minimize holding costs).

Previous work on the optimal assignment of servers to tasks has assumed that when multiple servers are assigned to the same task, their combined service rate is additive. This is appropriate
when servers work in parallel (as long as the servers do not interfere with one another, e.g., via sharing of tools). However, when several servers are collaborating on a single customer, the assumption that their combined rate is additive does not take into account the fact that their collaboration may be synergistic (e.g., due to factors such as complementarity of skills, motivation, etc.) or not (e.g., due to bad team dynamics, lack of space or tools, etc.). This is obviously restrictive.

This paper is concerned with the optimal assignment of servers to tasks when server collaboration is synergistic. Not surprisingly, we will show that when servers are homogeneous (in that all servers have the same service rate for each task), then synergistic servers should collaborate at all times. However, when servers are heterogeneous with respect to the tasks they are trained for and their service rates at those tasks, we will see that there is a tradeoff between taking advantage of server synergy on the one hand and of each server’s training and abilities on the other hand. Thus, it may be more effective for servers to work on their own at tasks they are good at, rather than collaborate synergistically on tasks where they are less effective. Our objective is to understand this tradeoff and determine when it is important to take advantage of server synergy, as opposed to exploiting their special skills, to improve the system throughput. Previous works do not address the tradeoff between individual and collaborative effectiveness, and hence this paper opens up a new, interesting, and important line of research.

More specifically, we focus on a queueing network with $N \geq 1$ stations and $M \geq 1$ flexible servers. There is an infinite amount of raw material in front of station 1, infinite room for departing customers after station $N$, and a finite buffer between stations $j$ and $j+1$, for $j \in \{1, \ldots, N-1\}$, whose size is denoted by $B_j$. The system operates under manufacturing blocking. At any given time, there can be at most one customer at each station and each server can work on at most one customer. We assume that server $i \in \{1, 2, \ldots, M\}$ works at a deterministic rate $\mu_{ij} \in [0, \infty)$ at station $j \in \{1, 2, \ldots, N\}$. Thus, server $i$ is trained to work at station $j$ if $\mu_{ij} > 0$ and the server’s skill at station $j$ is measured by the magnitude of $\mu_{ij}$. Several servers can work together on a single customer, in which case the combined rate of a server team is proportional to the sum of the rates of the individual servers. Thus, if servers $i_1, \ldots, i_k \in \{1, \ldots, M\}$ are simultaneously assigned to station $j \in \{1, \ldots, N\}$, the service rate is equal to $\alpha \sum_{r=1}^{k} \mu_{i_r,j}$. We assume that $1 \leq \alpha < \infty$, which implies that the servers are synergistic with $\alpha$ being a measure of the magnitude of the servers’ synergy. The service requirements of different customers at each station $j \in \{1, \ldots, N\}$ are independent and identically distributed (i.i.d.) random variables whose rate we take to be equal to 1 without loss of generality, and the service requirements at different stations are independent of each other. We assume that travel and set-up times are negligible. Under these assumptions, our objective is to determine the dynamic server assignment policy that maximizes the long-run average throughput. We completely characterize the optimal server assignment policy for systems with exponential service requirements, $N = 2$ stations, and $M = 2$ servers, and also provide results about when synergistic servers should collaborate in more general systems.

There is a significant amount of literature on queues with flexible servers. In the interest of
space, we do not provide a complete literature review here, but refer the interested reader to Hopp and Van Oyen [11] for a comprehensive review of the literature in this area, and to Akşin, Armony, and Mehrotra [3], Akşin, Karaesmen, and Örmeç [4], and Gans, Koole, and Mandelbaum [10] for thorough reviews of the literature on flexible servers in call centers. This paper is most closely related to other works that employ Markov decision process techniques and sample path analysis in determining effective server allocation schemes, see for example Ahn, Duenyas, and Zhang [1], Ahn and Righter [2], Andradóttir and Ayhan [5], Andradóttir, Ayhan, and Down [6, 7], Kaufman, Ahn, and Lewis [12], Örmeç [14], Sennott, Van Oyen, and Iravani [16], Van Oyen, Gel, and Hopp [18], and Wu, Lewis, and Veatch [19]. However, these papers only consider cases where the combined rate of a set of collaborating servers is additive (i.e., $\alpha = 1$). We are only aware of two works that consider models of server collaboration other than the additive model. In particular, Argon and Andradóttir [8] provide sufficient conditions for partial pooling of multiple adjacent queueing stations to be beneficial in tandem lines, and Buzacott [9] considers team work involving task partitioning (with the team completing work when all servers have completed their assigned subtasks) in a single stage queue with identical servers. These works do not address the tradeoff between individual and collaborative effectiveness, or how system performance can be optimized by dynamically assigning flexible servers to tasks.

The outline of this paper is as follows. In Section 2, we show that for systems with generalist servers, synergistic servers should collaborate at all times. We then turn our attention to systems with specialist servers (so that servers who are particularly effective at some tasks may be ineffective at other tasks). In Section 3, we provide a complete characterization of the optimal policy for Markovian systems with two stations and two servers for all $1 \leq \alpha < \infty$. We will show that the optimal policy either ignores server synergy (for small $\alpha$) or takes full advantage of server synergy by having servers collaborate at all times (for large $\alpha$). In Section 4, we investigate whether server collaboration is beneficial in more general Markovian networks. Section 5 summarizes our findings and provides some directions for future research. Finally, the proofs of some of our results are provided in an online appendix to this paper.

2 Systems with Generalist Servers

In this section consider systems with generalist servers when the service requirements have arbitrary distributions. In systems with generalist servers, the service rate of each server at each station can be expressed as the product of two constants, one representing the server’s speed at every task and the other representing the intrinsic difficulty of the task at the station. Thus, $\mu_{ij} = \mu_i \gamma_j$ for all $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N\}$.

We use $\Pi$ to denote the set of all server assignment policies under consideration (which will be defined later) and $D_\pi(t)$ to denote the number of departures under policy $\pi$ by time $t \geq 0$. Define

$$T_\pi = \limsup_{t \to \infty} \frac{E[D_\pi(t)]}{t}$$
as the long-run average throughput corresponding to the server assignment policy $\pi \in \Pi$. Our objective is to solve the following optimization problem

$$\max_{\pi \in \Pi} T_\pi. \quad (1)$$

In this section, $\Pi$ may include all possible server assignment policies. We call a policy fully collaborative if all servers work in teams of two or more at all times. An example of a fully collaborative policy is the expedite policy which assigns all available servers to a single team that will follow each customer from the first to the last station and only starts work on a new customer once all work on the previous customer has been completed. The next result states that when the servers are generalists, the lower bound on $\alpha$ that guarantees that a fully collaborative policy is optimal is 1. This is not surprising because when the servers are generalists, there is no tradeoff between the servers’ special skills and server synergy. On the other hand, when $\alpha = 1$ (i.e., the service rates are additive), it has been shown in Andradóttir, Ayhan, and Down [7] that any non-idling policy is optimal. Thus, when the servers are synergistic and generalists, the optimal policy takes full advantage of this synergy, and a non-idling policy is no longer sufficient to achieve optimal throughput. The proof of the following theorem is provided in the online appendix to this paper.

**Theorem 2.1** Assume that for each $j = 1, \ldots, N$, the service requirements $S_{k,j}$ of customer $k \geq 1$ at station $j$ are i.i.d. with mean 1. Moreover, assume that for all $t \geq 0$, if there is a customer in service at station $j$ at time $t$, then the expected remaining service requirement at station $j$ of that customer is bounded above by a scalar $1 \leq S < \infty$. Finally, assume that service is either nonpreemptive or preemptive-resume. If $\mu_{ij} = \mu_i \gamma_j$ for all $i = 1, \ldots, M$ and $j = 1, \ldots, N$ and $\alpha > 1$, then for all $0 \leq B_1, B_2, \ldots, B_{N-1} < \infty$, any fully collaborative policy $\pi$ is optimal, with long-run average throughput

$$T_\pi = \frac{\alpha \sum_{j=1}^M \mu_j}{\sum_{j=1}^N \frac{1}{\gamma_j}}.$$  

### 3 Two Station Markovian Systems with Two Servers

In this section, we completely characterize the optimal policy for a Markovian tandem line with $N = 2$ stations and $M = 2$ synergistic servers. We start with the description of a continuous time Markov chain model for systems with $N$ stations, $M$ servers, and exponentially distributed service requirements. For all $\pi \in \Pi$ and $t \geq 0$, let $X_\pi(t) = \{X_{\pi,1}(t), \ldots, X_{\pi,N-1}(t)\}$, where $X_{\pi,j}(t) \in \{0, \ldots, B_j + 2\}$ denotes the number of customers that have been processed at station $j$ at time $t$ but are either waiting to be processed by station $j + 1$ or in process at station $j + 1$ at time $t$. Let $S$ denote the state space of $\{X_\pi(t)\}$. For the remainder of this paper, we assume that the class $\Pi$ of server assignment policies under consideration consists of Markovian stationary deterministic policies corresponding to the state space $S$, and that the set of possible actions is given by $A = \{a_{\sigma_1,\ldots,\sigma_M} : \sigma_i \in \{0,1,\ldots,N\}, \forall i = 1,\ldots,M\}$, where for all $i \in \{1,\ldots,M\}$, $\sigma_i = 0$
when server $i$ is idle and $\sigma_i = j \in \{1, \ldots, N\}$ when server $i$ is assigned to station $j$. Finally, we use $\Sigma_j$ to denote $\sum_{i=1}^{M} \mu_{ij}$, for all $j = 1, \ldots, N$, so that $\alpha \Sigma_j$ is the combined service rate of servers $1, \ldots, M$ when they collaborate at station $j$.

For the remainder of this section, we assume $N = 2$ and set $B_1 = B$ for notational convenience. Then it is clear that for $\pi \in \Pi$, $\{X_{\pi}(t)\}$ is a birth-death process with state space $S = \{0, \ldots, B + 2\}$ and that there exists a scalar $q_{\pi} \leq \alpha \sum_{i=1}^{2} \max_{1 \leq j \leq 2} \mu_{ij} < \infty$ such that the transition rates $\{q_{\pi}(s, s')\}$ of $\{X_{\pi}(t)\}$ satisfy $\sum_{s' \in S, s' \neq s} q_{\pi}(s, s') \leq q_{\pi}$ for all $s \in S$. Hence, $\{X_{\pi}(t)\}$ is uniformizable. Let $\{Y_{\pi}(k)\}$ be the corresponding discrete time Markov chain, so that $\{Y_{\pi}(k)\}$ has state space $S$ and transition probabilities $p_{\pi}(s, s') = q_{\pi}(s, s')/q_{\pi}$ if $s' \neq s$ and $p_{\pi}(s, s) = 1 - \sum_{s' \in S, s' \neq s} q_{\pi}(s, s')/q_{\pi}$ for all $s \in S$. Using the analysis in Section 3 of Andradóttir, Ayhan, and Down [6], one can show that the original optimization problem in (1) can be translated into an equivalent (discrete time) Markov decision problem. More specifically, for all $i \in S$, let

$$R_{\pi}(i) = \begin{cases} q_{\pi}(i, i - 1) & \text{if } i \in \{1, \ldots, B + 2\}, \\ 0 & \text{if } i = 0, \end{cases}$$

be the departure rate from state $i$ under policy $\pi$. Then the optimization problem (1) has the same solution as the Markov decision problem

$$\max_{\pi \in \Pi} \lim_{K \to \infty} \mathbb{E} \left[ \frac{1}{K} \sum_{k=1}^{K} R_{\pi}(Y_{\pi}(k - 1)) \right].$$

In other words, maximizing the steady-state throughput of the original queueing system is equivalent to maximizing the steady-state departure rate for the associated embedded (discrete time) Markov chain.

The next theorem provides a complete characterization of the optimal server assignment policy for Markovian systems with $N = 2$ stations and $M = 2$ synergistic servers. We say that server $i \in \{1, \ldots, M\}$ has a primary assignment at station $j \in \{1, \ldots, N\}$ if server $i$ works at station $j$ unless he has no work at station $j$. For notational convenience, we set

$$C_{22} = \max \left\{ \frac{\mu_{11}}{\Sigma_1} + \frac{\mu_{22}}{\Sigma_2}, \frac{\mu_{21}}{\Sigma_1} + \frac{\mu_{12}}{\Sigma_2} \right\}.$$

Note that $C_{22} \geq 1$ since $\frac{\mu_{11}}{\Sigma_1} + \frac{\mu_{22}}{\Sigma_2} + \frac{\mu_{21}}{\Sigma_1} + \frac{\mu_{12}}{\Sigma_2} = 2$.

**Theorem 3.1** For a Markovian system of two stations and two servers, we have

(i) If $1 \leq \alpha \leq C_{22}$ and $\mu_{11}\mu_{22} \geq \mu_{21}\mu_{12}$ ($\mu_{21}\mu_{12} \geq \mu_{11}\mu_{22}$), then the policy where server 1 (2) has primary assignment at station 1 and server 2 (1) has primary assignment at station 2 and both servers work at station 1 (2) when station 2 (1) is starved (blocked) is optimal. Moreover, this is the unique optimal policy in the class of stationary deterministic policies if $\mu_{11}\mu_{22} > \mu_{21}\mu_{12}$ ($\mu_{21}\mu_{12} < \mu_{11}\mu_{22}$) and $\alpha < C_{22}$.

(ii) If $\alpha > C_{22}$, then a policy is optimal if and only if it is non-idling and fully collaborative.
Note that the uniqueness of the optimal policy in Theorem 3.1 is subject to the interpretation that assigning a server to a station where there is no work is equivalent to idling him. Also, note that $C_{22} = \frac{\mu_{11}}{\Sigma_1} + \frac{\mu_{22}}{\Sigma_2}$ when $\mu_{11} \mu_{22} \geq \mu_1 \mu_{12}$, and $C_{22} = \frac{\mu_{22}}{\Sigma_2} + \frac{\mu_{11}}{\Sigma_1}$ when $\mu_{21} \mu_{12} \geq \mu_{11} \mu_{22}$. One can observe that if $\alpha \leq C_{22}$, the optimal policy stated in Theorem 3.1 assigns servers to their primary stations (which are determined by the sign of $\mu_{11} \mu_{22} - \mu_1 \mu_{12}$) unless they have no work at these stations. Thus, the servers are allowed to collaborate only if their primary stations are blocked or starved. On the other hand if $\alpha > C_{22}$, the optimal policy has the servers work as a team at all times. Thus, the optimal policy switches from one that takes full advantage of servers’ skills in that the servers avoid collaboration unless they have no work at their primary assignments, to one that takes full advantage of server synergy (i.e., there is no middle ground). It is also important to point out that the expedite policy (which is one of the optimal policies when $\alpha > C_{22}$) minimizes the WIP (Work In Process) at all times, since the number of customers in the system is always one. Thus, in a Markovian system with two stations and two servers, one can simultaneously maximize the throughput and minimize the WIP if the synergy among the servers is sufficiently large.

The complete proof of Theorem 3.1 is given in an accompanying online Appendix. Here, we only provide a brief sketch of the proof of part (i) when $\mu_{11} \mu_{22} \geq \mu_1 \mu_{12}$ (because the remainder of the proof is similar). The policy $\pi_0$ described in part (i) of Theorem 3.1 when $\mu_{11} \mu_{22} \geq \mu_1 \mu_{12}$ yields an irreducible Markov chain $\{Y_{\pi_0}(k)\}$. Hence, the long-run average gain $g_0$ under this policy is a scalar, and it is given by

$$g_0 = \frac{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B+1} \mu_{11}^j \mu_{22}^{B+1-j}}{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B} \mu_{11}^j \mu_{22}^{B-j} + \mu_{22}^{B+1} \Sigma_2 + \mu_{11}^{B+1} \Sigma_1}.$$  \hspace{1cm} (2)

The main idea of the proof involves choosing the policy $\pi_0$ as the initial policy in the policy iteration algorithm for communicating Markov decision processes (see pages 479 and 480 of Puterman [15]) and showing that no improvement is possible.

Theorem 3.1 states that for Markovian systems with two stations and two synergistic servers, it suffices to consider two types of policies. The next proposition quantifies the loss in throughput performance if one employs the wrong policy in Theorem 3.1, implying that it is important to use the correct policy for any given value of $\alpha$.

**Proposition 3.2** Consider a Markovian tandem queue with two stations and two servers. Assume that $\Sigma_1 > 0$ and $\Sigma_2 > 0$. If $\alpha \leq C_{22}$, using a non-idling fully collaborative policy instead of the optimal policy described in part (i) of Theorem 3.1 can reduce the throughput performance by at most a factor of 2 (and this bound is tight). On the other hand, if $\alpha > C_{22}$, using the policy described in part (i) of Theorem 3.1 rather than a non-idling fully collaborative policy can yield arbitrarily worse throughput performance.

**Proof:** We first obtain an expression for the ratio of the throughputs of the two policies described in parts (i) and (ii) of Theorem 3.1. As mentioned above, the quantity $g_0$ specified in equation
(2) is the long-run average throughput of the optimal policy when \( \alpha \leq C_{22} \) and \( \mu_{11}\mu_{22} \geq \mu_{21}\mu_{12} \). Similarly, one can compute the throughput of the expedite policy (which is one of the optimal policies when \( \alpha > C_{22} \)) as \( T_{\text{expedite}} = \frac{\mu_{11} \Sigma_2}{\Sigma_1 + \Sigma_2} \). Then

\[
\frac{g_0}{T_{\text{expedite}}} = \frac{(\Sigma_1 + \Sigma_2) \sum_{j=0}^{B+1} \mu_{11}^{j} \mu_{22}^{B+1-j}}{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B} \mu_{11}^{j} \mu_{22}^{B-j} + \mu_{22}^{B+1} \Sigma_2 + \mu_{11}^{B+1} \Sigma_1}.
\]  

(3)

Using the argument in Remark 4.1 of Andradóttir and Ayhan [5], one can verify that \( g_0 \) is nondecreasing in the buffer size \( B \). Thus, \( g_0/T_{\text{expedite}} \) is nondecreasing in \( B \) and nonincreasing in \( \alpha \). This implies that

\[
\lim_{B \to \infty} \frac{g_0}{T_{\text{expedite}}} = \frac{\max\{\mu_{11}, \mu_{22}\}(\Sigma_1 + \Sigma_2)}{\alpha \Sigma_1 \Sigma_2 + (\max\{\mu_{11}, \mu_{22}\} - \min\{\mu_{11}, \mu_{22}\})(\Sigma_1 + \Sigma_2)}
\]

\[
\leq \frac{\max\{\mu_{11}, \mu_{22}\}(\Sigma_1 + \Sigma_2)}{\Sigma_1 \Sigma_2 + (\max\{\mu_{11}, \mu_{22}\} - \min\{\mu_{11}, \mu_{22}\})(\Sigma_1 + \Sigma_2)},
\]  

(4)

where \( \mathbb{I}(E) = 1 \) if \( E \) occurs and \( \mathbb{I}(E) = 0 \) otherwise. The bound in (4) is obtained by replacing \( \alpha \) with 1. First assume that \( \mu_{11} \geq \mu_{22} \). Then

\[
\frac{g_0}{T_{\text{expedite}}} \leq \frac{\mu_{11}(\Sigma_1 + \Sigma_2)}{\Sigma_1 \Sigma_2 + (\mu_{11} - \mu_{22})\Sigma_1} \leq 2.
\]  

(5)

The second inequality in (5) follows immediately from the following argument. If

\[
\frac{\mu_{11}(\Sigma_1 + \Sigma_2)}{\Sigma_1 \Sigma_2 + (\mu_{11} - \mu_{22})\Sigma_1} > 2,
\]

then we need to have \( \mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12} < \mu_{11}\mu_{22} - 2\mu_{21}\mu_{12} \), which is not possible because \( \mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12} \geq \mu_{11}\mu_{22} \) and \( \mu_{21}\mu_{12} \geq 0 \). Note that since \( \mu_{11} \geq \mu_{22} \),

\[
\frac{\mu_{11}(\Sigma_1 + \Sigma_2)}{\Sigma_1 \Sigma_2 + (\mu_{11} - \mu_{22})\Sigma_1} \leq \frac{2\mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12}}{\mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12} + \mu_{21}\mu_{12}}.
\]

But

\[
\lim_{\mu_{11} \to \infty} \frac{2\mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12}}{\mu_{11}^2 + \mu_{11}\mu_{21} + \mu_{11}\mu_{12} + \mu_{21}\mu_{12}} = 2,
\]

which yields the desired result (i.e., the bound 2 is tight for \( \alpha = 1 \) as \( B \to \infty \) and \( \mu_{11} = \mu_{22} \to \infty \)). The same arguments can be used to obtain the tight upper bound of 2 when \( \mu_{11} \leq \mu_{22} \).

On the other hand, it immediately follows from the expression in (3) that employing the policy described in part (i) of Theorem 3.1 when \( \alpha > C_{22} \) yields arbitrarily worse throughput performance as \( \alpha \) gets larger. Replacing \( \mu_{11} \) (\( \mu_{21} \)) with \( \mu_{11} \) (\( \mu_{21} \)) and \( \mu_{22} \) (\( \mu_{12} \)) with \( \mu_{22} \) (\( \mu_{12} \)) in the above expressions provides the ratio of \( g_0 \) to \( T_{\text{expedite}} \) when \( \mu_{21}\mu_{12} \geq \mu_{11}\mu_{22} \), and using the arguments above, one can obtain the same results in this case. \( \square \)
4 Results for Larger Markovian Systems

In this section, we analyze more general Markovian tandem queues. We will need the following notation. Denote the number of servers at each station \( j \in \{1, \ldots, N\} \) under action \( a_{\sigma_1, \ldots, \sigma_M} \) as 
\[
n_j(a_{\sigma_1, \ldots, \sigma_M}) = \sum_{i=1}^{M} \mathbb{I}(\sigma_i = j).
\]
Similarly, for all \( j \in \{1, \ldots, N\} \), \( r_j(a_{\sigma_1, \ldots, \sigma_M}) = \sum_{i=1}^{M} \mathbb{I}(\sigma_i = j)\mu_{ij} \) is the cumulative service rate at station \( j \) under action \( a_{\sigma_1, \ldots, \sigma_M} \) (not taking into account the synergy among servers). For all \( a \in A \), define \( \mathcal{N}_1(a) = \{ j : n_j(a) = 1 \} \), \( \mathcal{N}_1(a) = \{ j : n_j(a) \geq 1 \} \), and \( \mathcal{N}_2(a) = \{ j : n_j(a) \geq 2 \} \). Thus, \( \mathcal{N}_1(a), \mathcal{N}_1(a), \) and \( \mathcal{N}_2(a) \) denote the sets of stations with exactly one server, at least one server, and at least two servers under action \( a \), respectively, and 
\[
\mathcal{N}_1(a) \cup \mathcal{N}_2(a) = \mathcal{N}_1(a). \] Finally, for notational convenience, we set \( R(a) = \sum_{l \in \mathcal{N}_1(a)} r_l(a) + \alpha \sum_{l \in \mathcal{N}_2(a)} r_l(a) \) for all \( a \in A \).

The next theorem states that a fully cross-trained server should never idle. Note that this generalizes Proposition 2.1 of Andradóttir and Ayhan [5] which implies the optimality of non-idling policies for Markovian systems with two stations in tandem, servers capable of working at both stations, and \( \alpha = 1 \).

**Theorem 4.1** Consider a Markovian tandem line with \( N \geq 1 \) stations and \( M \geq 1 \) servers. Suppose server \( i \) is such that \( \mu_{ij} > 0 \) for all \( j \in \{1, \ldots, N\} \). If 
\[
\alpha > \frac{\mu_{kj}}{\mu_{ij} + \mu_{kj}} \text{ for all } k \in \{1, \ldots, M\}, k \neq i, \text{ and } j \in \{1, \ldots, N\},
\] then the optimal policy should not allow server \( i \) to idle in any state that is recurrent under this policy.

Note that the condition in equation (6) guarantees that when server \( i \) works with any server \( k \) at any station \( j \), their rate as a team is faster than the individual service rate of server \( k \) at that station. Moreover, the lower bound on \( \alpha \) in equation (6) is strictly less than 1, implying that the servers need not be synergistic for the result of Theorem 4.1 to hold (clearly, the condition is satisfied if the servers are synergistic).

**Proof:** Let \( d^\infty \) denote a Markovian stationary deterministic policy that uses the decision rule \( d \) at each decision epoch with \( d(s) \in A \) for all \( s \in S \). Moreover, assume that under the decision rule \( d \), there exists a recurrent state \( s_0 \in S \) such that \( d(s_0) = a \) and \( \sigma_i = 0 \) under \( a \). Without loss of generality, we assume that \( \mathcal{N}_1(a) \neq \emptyset \) and \( R(a) \neq 0 \) (because otherwise \( s_0 \) is an absorbing state and the long-run average throughput equals zero).

For all \( j = 1, \ldots, N \), let the action \( a^j \) be the same as \( a \) except that \( \sigma_i = j \) under \( a^j \). Note that \( \alpha r_l(a^j) - r_l(a) > 0 \) for all \( l \in \mathcal{N}_1(a) \) and \( r_l(a^j) - r_l(a) > 0 \) for all \( l \in \mathcal{N}_2(a) \) under our assumptions (see (6)). Now consider the Markovian stationary randomized policy \( (d')^\infty \) such that \( d'(s) = d(s) \) for all \( s \in S \setminus \{s_0\} \) and 
\[
d'(s_0) = a^j \text{ with probability } p_j(a) \text{ for all } j \in \mathcal{N}_1(a),
\]
\[
p_j(a) = \frac{R(a^j)r_j(a) \prod_{l \in N_1(a)} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a))}{K(a)} \geq 0
\]

and
\[
K(a) = \sum_{m \in N_1(a)} R(a^m)r_m(a) \prod_{l \in N_1(a)} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a)) > 0.
\]

Thus, \(d'(s_0)\) is a randomized action that does not allow server \(i\) to idle in state \(s_0\). Moreover,
\[
R(a^j) = \begin{cases} R(a) + \alpha r_j(a^j) - r_j(a) & \text{if } j \in N_1(a), \\ R(a) + \alpha(r_j(a^j) - r_j(a)) & \text{if } j \in N_2(a), \end{cases}
\]

and hence
\[
K(a) = R(a) \left[ \sum_{m \in N_1(a)} r_m(a) \prod_{l \in N_1(a)} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a)) + \prod_{l \in N_1(a)} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a)) \right]. \tag{7}
\]

The transition probabilities out of state \(s_0\) for the underlying Markov chain are the same under decision rules \(d\) and \(d'\). In order to see this, consider the transition probability out of state \(s_0\) corresponding to a service completion at station \(j \in N_1(a)\). Clearly, this probability is equal to \(r_j(a)/R(a)\) under decision rule \(d\). Under decision rule \(d'\), this transition probability equals
\[
\frac{p_j(a) \frac{\alpha r_j(a^j)}{R(a^j)}}{R(a)} + \sum_{m \in N_1(a) \atop l \neq j} p_m(a) \frac{r_j(a)}{R(a^m)}
\]
\[
= p_j(a) \frac{\alpha r_j(a^j)}{R(a^j)} - p_j(a) \frac{r_j(a)}{R(a^j)} + \sum_{m \in N_1(a) \atop l \neq j} p_m(a) \frac{r_j(a)}{R(a^m)}
\]
\[
= r_j(a) \left[ \prod_{l \in N_1(a)} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a)) \right] K(a) + \sum_{m \in N_1(a)} r_m(a) \prod_{l \in N_1(a) \atop l \neq m} (\alpha r_l(a^l) - r_l(a)) \prod_{l \in N_2(a)} (r_l(a^l) - r_l(a)) \right] K(a)
\]
\[
= \frac{r_j(a)}{R(a)}.
\]

where the last equality follows from (7). Similarly, one can show that the transition probabilities are the same when \(j \in N_2(a)\) (both are equal to \(\alpha r_j(a)/R(a)\)). However, the total rate out of
state $s_0$ under decision rule $d$ is $R(a)$, and similarly, the total rate out of state $s_0$ under decision rule $d'$ is equal to

$$\left( \sum_{j \in N_1(a)} \frac{p_j(a)}{R(a^j)} \right)^{-1} = \frac{K(a)}{\sum_{j \in N_1(a)} r_j(a) \prod_{i \in N_i(a)} (\alpha r_i(a^i) - r_i(a)) \prod_{i \in N_2(a)} (\alpha r_i(a^i) - r_i(a))}$$

$$= R(a) + \frac{R(a) \prod_{i \in N_i(a)} (\alpha r_i(a^i) - r_i(a)) \prod_{i \in N_2(a)} (\alpha r_i(a^i) - r_i(a))}{\sum_{j \in N_1(a)} r_j(a) \prod_{i \in N_i(a)} (\alpha r_i(a^i) - r_i(a)) \prod_{i \in N_2(a)} (\alpha r_i(a^i) - r_i(a))}$$

$$> R(a),$$

where the last equality again follows from (7). Thus the expected time spent per visit to $s_0$ under $d'$ is shorter than under $d$. Since the transition probabilities are equal under both decision rules, the number of departures under $d'$ is larger. A simple sample path argument now shows that $(d')^\infty$ has a larger long-run average throughput than $d^\infty$ (because $s_0$ is a recurrent state visited infinitely often with probability 1). But Lemma 4.3.1 of Puterman [15] implies that there exists a policy in $\Pi$ whose throughput performance is at least as good as that of $(d')^\infty$. Repeating the same argument for all states where server $i$ idles under the decision rule $d$, it is clear that there is a Markovian stationary deterministic policy that does not allow server $i$ to idle with long-run average throughput strictly larger than the throughput of $d^\infty$, and the proof is complete. $\square$

Note that in a system with $\mu_{ij} = 0$ for some $j \in \{1, \ldots, N\}$, a policy that allows idling could yield a better throughput than a non-idling policy, because sometimes it might be better to idle a server rather than assigning him to a station where his efforts are counter-productive. In order to see this, consider a Markovian tandem line with $N = 2$ stations, $M = 3$ servers, and $B = 1$, so that $S = \{0, 1, 2, 3\}$. Suppose that $\alpha = 1$, $\mu_{11} = 3$, $\mu_{12} = \mu_{21} = \mu_{31} = 1$, $\mu_{22} = 8$, and $\mu_{31} = 0$. Consider the stationary policy $\delta^\infty$ such that

$$\delta(s) = \begin{cases} 
\text{all servers work at station 1} & \text{if } s = 0, \\
\text{server 2 works at station 1, servers 1 and 3 work at station 2} & \text{if } s = 1, \\
\text{server 1 works at station 1, server 2 works at station 2, and server 3 idles} & \text{if } s = 2, \\
\text{all servers work at station 2} & \text{if } s = 3.
\end{cases}$$

Let $(\delta')^\infty$ be the same as $\delta^\infty$ in all states except $s = 2$ where

$$\delta'(2) = \text{server 1 works at station 1, servers 2 and 3 work at station 2}.$$

Thus, $(\delta')^\infty$ does not allow server 3 to idle. But then we have $2.027 \simeq T_{(\delta')^\infty} < T_{\delta^\infty} \simeq 2.030$. Hence, it is counter-productive in this policy for server 3 to work at station 2 in state 2, and since $\mu_{31} = 0$, it is best to idle server 3 in this state. However, Theorem 4.1 shows that if $\alpha \geq 1$ and all service rates are positive, one can always find a server allocation policy that assigns all the servers to the stations where they are productive.

Next we study fully collaborative policies. Recall that a policy is fully collaborative if all servers work in teams of two or more at all times under this policy. Note that the team formations need not
be static. The next theorem shows that if $\alpha$ is large enough, the optimal policy is fully collaborative for a certain class of Markovian systems.

**Theorem 4.2** Consider a Markovian tandem line with $N = 2$ stations and $M \geq 2$ servers or $N > 2$ stations and $2 \leq M \leq 4$ servers. Suppose $\mu_{ij} > 0$ for all $i \in \{1, \ldots, M\}$ and all $j \in \{1, \ldots, N\}$. If

$$
\alpha > \max_{a \in A} \frac{\sum_{j \in N_1(a)} r_j(a) \frac{r_{ij}(a)}{\Sigma_j}}{1 - \sum_{j \in N_{2+}(a)} r_j(a) \frac{r_{ij}(a)}{\Sigma_j}}
$$

(with the convention that summation over an empty set is equal to 0), then the optimal policy is fully collaborative in all recurrent states under that policy.

**Proof:** Let $d^\infty$ denote a Markovian stationary deterministic policy that uses the decision rule $d$ at each decision epoch with $d(s) \in A$ for all $s \in S$. Assume that under the decision rule $d$, there exists a recurrent state $s_0 \in S$ such that $d(s_0) = a$ with $N_1(a) \neq \emptyset$. Without loss of generality, we assume that $R(a) \neq 0$ (because otherwise $s_0$ is absorbing and the throughput is zero). It suffices to show that there exists a policy which selects a fully collaborative action in state $s_0$ and yields a better throughput than $d^\infty$ if $\alpha$ is large enough. Consider the Markovian stationary randomized policy $(d')^\infty$ such that $d'(s) = d(s)$ for all $s \in S \setminus \{s_0\}$ and

$$
d'(s_0) = a_{j_1, \ldots, j_k} \text{ with probability } p_j(a) \text{ for all } j \in N_{1+}(a),
$$

where

$$
p_j(a) = \begin{cases} 
\frac{r_j(a)}{R(a)} & \text{if } j \in N_1(a), \\
\frac{\alpha r_j(a)}{R(a)} & \text{if } j \in N_{2+}(a). 
\end{cases}
$$

Thus, $d'(s_0)$ is a randomized, fully collaborative action.

We now show that the transition probabilities out of state $s_0$ for the underlying Markov chain are the same under decision rules $d$ and $d'$. Consider the transition probability out of state $s_0$ corresponding to a service completion at station $j \in N_1(a)$. Clearly, this probability is equal to $r_j(a)/R(a)$ under decision rule $d$, and it equals $p_j(a) = r_j(a)/R(a)$ under decision rule $d'$. Similarly, the transition probability out of state $s_0$ corresponding to a service completion at station $j \in N_{2+}(a)$ is equal to $\alpha r_j(a)/R(a)$ under both decision rules. However, the total rate out of state $s_0$ under decision rule $d$ is $R(a)$, and, similarly, the total rate out of state $s_0$ under decision rule $d'$ is equal to

$$
\frac{1}{\sum_{j \in N_{1+}(a)} p_j(a) \frac{r_j(a)}{\Sigma_j}} = \frac{\alpha R(a)}{\sum_{j \in N_1(a)} \frac{r_j(a)}{\Sigma_j} + \sum_{j \in N_{2+}(a)} \frac{\alpha r_j(a)}{\Sigma_j}}.
$$

Thus, we can conclude that if

$$
\alpha > \frac{\sum_{j \in N_1(a)} \frac{r_j(a)}{\Sigma_j}}{1 - \sum_{j \in N_{2+}(a)} \frac{r_j(a)}{\Sigma_j}},
$$

(8)
then the rate out of state $s_0$ under $(d')^\infty$ is larger than the rate under $d^\infty$. Note that the denominator of the expression in (8) is always positive because for all $a \in A$, $N_1(a)$ can have at most one element when $N_1(a) \neq \emptyset$ and either $N = 2$ or $M \leq 4$. Thus, if $\alpha$ is greater than the lower bound specified in (8), a simple sample path argument shows that $(d')^\infty$ has a larger long-run average throughput than $d^\infty$. But Lemma 4.3.1 of Puterman [15] implies that there exists a policy in $\Pi$ whose throughput performance is at least as good as that of $(d')^\infty$. Thus, there is a fully collaborative Markovian stationary deterministic policy whose long-run average throughput is strictly larger than the throughput of $d^\infty$ when $\alpha$ is greater than the lower bound specified in (8).

Taking the maximum over all possible server allocations in $A$ with $N_1(a) \neq \emptyset$ yields the desired result (note that the numerator of the right-hand side in (8) equals zero when $N_1(a) = \emptyset$).

We conclude this section by providing some of our observations on fully collaborative policies:

**Observation 1:** If $\mu_{ij} = 0$ for some $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N\}$, the optimal policy may not be fully collaborative even when $\alpha$ is large. For example, consider a Markovian tandem line with $N = 2$ stations, $M = 3$ servers, and $B = 0$, so that $S = \{0, 1, 2\}$. Suppose that $\mu_{32} = 0$ and all the other service rates are positive. Consider the stationary policy $\delta^\infty$ such that

$$
\delta(s) = \begin{cases} 
\text{all servers work at station 1} & \text{if } s = 0, \\
\text{servers 1 and 2 work at station 2, server 3 works at station 1} & \text{if } s = 1, \\
\text{servers 1 and 2 work at station 2, server 3 is idle} & \text{if } s = 2.
\end{cases}
$$

In this case, the only fully collaborative policy is the expedite policy. But

$$
T_{\delta^\infty} - T_{\text{expedite}} = \frac{\alpha \Sigma_1 \Sigma_2 \mu_{31}}{(\alpha \Sigma_2^2 + \alpha \Sigma_1 \Sigma_2 + \mu_{31} \Sigma_1)(\Sigma_1 + \Sigma_2)},
$$

which is strictly bigger than 0 for all $\alpha > 0$.

**Observation 2:** In general the lower bound in Theorem 4.2 is not tight. That is, one can easily construct systems where a fully collaborative policy outperforms all non-fully collaborative policies for values of $\alpha$ smaller than the bound specified in Theorem 4.2 (which one can also intuitively deduce from the proof of the theorem). For example, consider a system with $N = 2$ stations, $M = 4$ servers, and $B = 0$. Suppose that $\mu_{11} = \mu_{21} = \mu_{22} = \mu_{42} = 1$, $\mu_{32} = 2$, and $\mu_{12} = \mu_{31} = \mu_{41} = 4$. In this case, the lower bound in Theorem 4.2 is equal to 5, but the policy that assigns all servers to station 1 when station 2 is starved, all servers to station 2 when station 1 is blocked, and assigns servers 3 and 4 to station 1 and servers 1 and 2 to station 2 otherwise, has a better throughput than all non-idling, non-fully collaborative policies for all $\alpha \geq 1$. However, it follows from Theorem 3.1 in Section 3 that the bound is tight for systems with two stations and two servers.

**Observation 3:** The lower bound on $\alpha$ in Theorem 4.2 is always greater than or equal to 1. In order to see this, first consider systems with $M = 2$ servers and consider the allocations
\[ a_{12} \text{ and } a_{21}. \] Taking the maximum over these two allocations would yield \( \alpha > C_{22} \geq 1 \) (see the statement right before Theorem 3.1). For systems with \( M > 2 \), first consider the allocation where \( \sigma_i = 1 \) and \( \sigma_k = 2 \) for all \( k \neq i \) and then consider the allocation \( \sigma_i = 2 \) and \( \sigma_k = 1 \) for all \( k \neq i \). The first allocation yields the bound \( (\mu_{i1} \Sigma_2)/(\mu_{i2} \Sigma_1) \) and the second yields \( (\mu_{2i} \Sigma_1)/(\mu_{1i} \Sigma_2) \), which again implies that \( \alpha \geq 1 \). Moreover, the lower bound on \( \alpha \) in Theorem 4.2 is in general strictly bigger than one (for example when \( N = 2 \) and \( M \geq 2 \), it is clear from the above that one needs to have \( \mu_{11} \Sigma_2 = \mu_{12} \Sigma_2 = \ldots = \mu_{1M} \Sigma_2 \) for this bound to be equal to one). Thus, even when the servers are fully cross-trained and synergistic, the optimal policy need not be fully collaborative for all \( \alpha > 1 \) when the service rates are arbitrary (i.e., it may take advantage of the servers’ special skills for some \( \alpha > 1 \), see also Theorem 3.1). Hence, adding collaboration to a policy when \( \alpha > 1 \) is not necessarily desirable (see also the example following the proof of Theorem 4.1). However, if the servers are generalists, the lower bound is equal to one, as shown in Theorem 2.1.

**Observation 4:** It is not necessarily the case that increasing \( \alpha \) will increase the long-run average throughput of a particular server allocation policy. For example, consider a Markovian tandem line with \( N = 2 \) stations, \( M = 3 \) servers, and \( B = 1 \). Suppose that \( \mu_{11} = 100, \mu_{12} = 1, \mu_{21} = \mu_{22} = \mu_{31} = 0.1, \) and \( \mu_{32} = 200 \). Note that in this case \( S = \{0, 1, 2, 3\} \), and consider the stationary policy \( \delta^{\infty} \) such that

\[
\delta(s) = \begin{cases} 
\text{all servers work at station 1} & \text{if } s = 0 \\
\text{servers 1 and 2 work at station 1, server 3 works at station 2} & \text{if } s = 1, \\
\text{servers 2 and 3 work at station 1, server 1 works at station 2} & \text{if } s = 2, \\
\text{all servers work at station 2} & \text{if } s = 3.
\end{cases}
\]

It is easy to compute the long-run average throughput of \( \delta^{\infty} \) and to verify that it is decreasing in \( \alpha \) when \( \alpha \in [0.16, 3.11] \).

**Observation 5:** For a Markovian tandem line with \( N = 2 \) stations, \( M \geq 2 \) servers, and \( B = 0 \), the long-run average throughput of any non-idling policy is increasing in \( \alpha \). As is mentioned in Section 3, in this case \( \{X_\pi(t)\} \) is a birth-death process with state space \( S = \{0, 1, 2\} \). For a non-idling policy \( \pi \) and for all \( s \in S \), let \( \lambda_\pi(s) \) be the birth rate in state \( s \) and \( \nu_\pi(s) \) be the death rate in state \( s \). Note that \( \lambda_\pi(s) \) is non-increasing in \( s \) and \( \nu_\pi(s) \) is non-decreasing in \( s \). It follows from Shantikumar and Yao [17] (see page 437 of [17]) that this network can be equivalently viewed as a two-node cyclic network with 1 customer circulating among the nodes, where node 1 models the death process and node 2 models the birth process. The result then follows from Lemma 14.B.10 of Shantikumar and Yao [17].

**Observation 6:** For a Markovian system with \( N \geq 1 \) stations and \( M \geq 1 \) servers, the long-run average throughput \( T_\pi \) of a fully collaborative policy \( \pi \) is a strictly increasing function of \( \alpha \). This result follows immediately from the fact that since all the transition rates of
the corresponding continuous time Markov chain under a fully collaborative policy $\pi$ are multiplied by $\alpha$, increasing $\alpha$ does not affect the stationary distribution of the continuous time Markov chain but increases the departure rate. Note that this result also holds for general Markovian queueing networks with infinite buffers as long as the system is stable.

5 Concluding Remarks

We considered tandem lines with finite buffers and flexible synergistic servers where the combined rate of a server team is proportional to the sum of the rates of the individual servers. For systems with generalist servers, arbitrary number of stations and servers, and general service requirement distributions, we proved that any fully collaborative policy is optimal. Hence, when the servers have no special skills, the optimal policy always takes full advantage of the synergy among the servers. For Markovian systems with two stations in tandem and two servers, we have shown that the policy which is optimal for systems with additive service rates remains optimal if the synergy among servers is small. However, as the level of synergy between servers increases (so that collaboration gets more effective), the expedite policy becomes optimal. Thus, for Markovian systems with two stations and two servers, the optimal policy abruptly switches from one where the servers work together only if they have to (i.e., they have no work at their primary assignments) to one where they work together at all times. Moreover, employing the wrong one among these two policies can result in significant losses in throughput performance. Finally, we identify how much synergy is sufficient to guarantee that the optimal policy is non-idling or fully collaborative for certain classes of larger Markovian systems.

To the best of our knowledge, this paper is the first to analyze the tradeoff between exploiting synergies arising from the differences between workers and synergies arising from their collaboration, with collaboration becoming more desirable as the collaboration becomes more effective (not surprisingly). In the future, we will continue this new line of research by investigating whether similar structural results also hold for systems with more general models for the team service rates. In particular, we will allow the proportionality constant $\alpha$ to depend on both the team size and the station that the team is working at. Allowing $\alpha$ to depend on the team size is important because it takes into account the fact that there may be an optimal team size (e.g., it may be beneficial to assign two or three workers to a team, but assigning more than two or three workers to the team may yield diminishing returns) and allowing $\alpha$ to depend on the station captures the fact that collaboration may not be equally beneficial for all tasks. We are also interested in studying systems with different classes of servers with the synergy between servers in a team depending on the classes the servers belong to. Finally, we are currently studying systems where the servers are not synergistic (i.e., $\alpha < 1$). This case seems more complex than the synergistic case, but our observations so far indicate that for systems with two stations in tandem and two servers, the optimal policy depends on whether one server is better than the other one at both stations or not,
and may involve some server idling when \( \alpha \) is small.

**Acknowledgments**

This research was supported by the National Science Foundation under Grant CMMI–0856600. The research of the third author was also supported by the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the area editor, associate editor and the referees for their careful reviews and valuable comments.

**References**


Online Appendix

In this section, we prove Theorems 2.1 and 3.1.

**Proof of Theorem 2.1:** Define $A_\pi(t)$ as the number of customers that have entered the system by time $t$ under policy $\pi \in \Pi$. Then $A_\pi(t) = Q_\pi(t) + D_\pi(t)$, where $Q_\pi(t)$ denotes the number of customers in the system at time $t$ under policy $\pi \in \Pi$. Since $Q_\pi(t) \leq \sum_{j=1}^{N-1} B_j + N$ for all $t \geq 0$ and for all $\pi \in \Pi$, we have

$$T_\pi = \limsup_{t \to \infty} \frac{E[D_\pi(t)]}{t} = \limsup_{t \to \infty} \frac{E[A_\pi(t)]}{t},$$

(9)

Our model is equivalent to one where the service requirements of successive customers at station $j \in \{1, \ldots, N\}$ are i.i.d. with mean $1/\gamma_j$ and the service rates depend only on the server (i.e., $\mu_{ij} = \mu_i$ for all $i \in \{1, \ldots, M\}$). Let $\pi$ be a fully collaborative policy and define $W_{\pi,p}(t)$ as the total work performed by time $t$ for all servers under the policy $\pi$. Then $W_{\pi,p}(t) = \alpha \sum_{i=1}^{M} \mu_i t$. Let $S_k = \sum_{j=1}^{N} B_j/\gamma_j$ be the total service requirement (in the system) of customer $k$ for all $k \geq 1$. Let $W_\pi(t) = \sum_{k=1}^{A_\pi(t)} S_k$ and let $W_{\pi,r}(t) = W_\pi(t) - W_{\pi,p}(t)$ be the total remaining service requirement (work) at time $t$ for the customers that entered the system by time $t$. We have

$$E[W_{\pi,r}(t)] \leq (N + \sum_{j=1}^{N-1} B_j) \times S \times \sum_{j=1}^{N} \frac{1}{\gamma_j},$$

which implies that $\lim_{t \to \infty} E[W_{\pi,r}(t)]/t = 0$ and

$$\lim_{t \to \infty} \frac{E[W_\pi(t)]}{t} = \lim_{t \to \infty} \frac{E[W_{\pi,p}(t)]}{t} = \alpha \sum_{i=1}^{M} \mu_i.$$

(10)

For all $n \geq 0$, let $Z_n = (S_{n,1}, \ldots, S_{n,N})$. Since the event $\{A_\pi(t) = n\}$ is completely determined by the random vectors $Z_1, Z_2, \ldots, Z_{n-1}$ (and independent of $Z_n, Z_{n+1}, \ldots$), $A_\pi(t)$ is a stopping time for the sequence of random vectors $\{Z_n\}$. Moreover, for all $t \geq 0$, $A_\pi(t) \leq K(t) + 1$, where $K(t)$ is the number of customers departing station 1 by time $t$ if all servers work at station 1 at all times and there is unlimited room for completed customers after station 1. Since $\{K(t)\}$ is a nondecreasing process with $\lim_{t \to \infty} E[K(t)]/t = \alpha \sum_{i=1}^{M} \mu_i \gamma_i < \infty$ (which follows from the elementary renewal theorem), we have $E[A_\pi(t)] < \infty$ for each $t \geq 0$. Wald’s lemma yields

$$E[W_\pi(t)] = E\left[\sum_{k=1}^{A_\pi(t)} S_k\right] = E[A_\pi(t)] \times \sum_{j=1}^{N} \frac{1}{\gamma_j}.$$  

(11)

From (9), (10), and (11), we now have

$$\alpha \sum_{i=1}^{M} \mu_i = \lim_{t \to \infty} \frac{E[W_\pi(t)]}{t} = \lim_{t \to \infty} \frac{E[A_\pi(t)]}{t} \times \sum_{j=1}^{N} \frac{1}{\gamma_j} = T_\pi \times \sum_{j=1}^{N} \frac{1}{\gamma_j},$$

which yields the desired throughput. The optimality of this throughput follows from equations (9), (10), and (11) and the fact that $W_{\pi,p}(t) \leq \alpha \sum_{i=1}^{M} \mu_i t$ for all $t \geq 0$ and for all server assignment policies $\pi \in \Pi$. □
Proof of Theorem 3.1: We start with the proof of part (i). We only provide the proof for the case $\mu_{11}\mu_{22} \geq \mu_{21}\mu_{12}$ (so that $C_{22} = \frac{m_{11}}{\Sigma_1} + \frac{m_{22}}{\Sigma_2}$) because the proof for the case $\mu_{21}\mu_{12} \geq \mu_{11}\mu_{22}$ simply follows by relabeling the servers. First suppose that $\mu_{1j} = \mu_{2j} = 0$ for some $j \in \{1, 2\}$ (i.e., there is at least one station at which no server is capable of working). Then the long-run average throughput is zero under any policy and the policy described in part (i) of Theorem 3.1 is optimal. Thus, we can assume, without loss of generality, that there exist $i_1, i_2 \in \{1, 2\}$, such that $\mu_{i_11} > 0$ and $\mu_{i_22} > 0$, and hence $\Sigma_1 > 0$ and $\Sigma_2 > 0$. Similarly, if $\mu_{11}\mu_{22} = \mu_{12}\mu_{21} = 0$, $\Sigma_1 > 0$, and $\Sigma_2 > 0$, then we know that $C_{22} = 1$ and that there exists $i_1 \in \{1, 2\}$ such $\mu_{i_11} = \mu_{i_22} = 0$. But Theorem 2.1 of Andradóttir, Ayhan, and Down [7] implies that any policy, including the one stated in part (i) of Theorem 3.1, that does not allow server $i_2 \neq i_1$ to idle is optimal. Since $\mu_{11}\mu_{22} \geq \mu_{21}\mu_{12}$, we have shown that we can assume, without loss of generality, that $\mu_{11} > 0$ and $\mu_{22} > 0$.

The set $A_s$ of allowable actions in state $s$ is given as

$$A_s = \begin{cases} \{a_{11}\} & \text{for } s = 0, \\ \{a_{11}, a_{12}, a_{21}, a_{22}, a_{00}, a_{01}, a_{02}, a_{10}, a_{20}\} & \text{for } s \in \{1, \ldots, B + 1\}, \\ \{a_{22}\} & \text{for } s = B + 2, \end{cases}$$

where we use a sample path argument to eliminate actions that allow servers to idle in states $s = 0$ and $s = B + 2$ (see Lemma 2.1 of Kirkuzlar, Andradóttir, and Ayhan [13]). Since the number of possible states and actions are both finite, the existence of an optimal Markovian stationary deterministic policy follows from Theorem 9.1.8 of Puterman [15] which provides sufficient conditions under which such a policy exists.

Under our assumptions on the service rates neither $\mu_{11}$ nor $\mu_{22}$ can be equal to zero. Consequently, the policy described in part (i) of Theorem 3.1 corresponds to an irreducible Markov chain, and, hence, we have a communicating Markov decision process. Therefore, we use the policy iteration algorithm for communicating models (see pages 479 and 480 of Puterman [15]) to prove the optimality of the policy described in Theorem 3.1.

Let $p(s'|s, d(s))$ be the probability of going to state $s' \in S$ in one step when the action prescribed by decision rule $d$ is taken in state $s$ and $P_d$ be the corresponding $(B + 3) \times (B + 3)$-dimensional probability transition matrix. Similarly, $r(s, d(s))$ denotes the immediate reward obtained when the action prescribed by decision rule $d$ is taken in state $s$ and $r_d$ denotes the corresponding $(B + 3)$-dimensional reward vector.

As the initial policy of the policy iteration algorithm, we choose

$$d_0(s) = \begin{cases} a_{11} & \text{for } s = 0, \\ a_{12} & \text{for } 1 \leq s \leq B + 1, \\ a_{22} & \text{for } s = B + 2, \end{cases}$$

18
corresponding to the policy described in part (i) of Theorem 3.1. Then
\[
r(s, d_0(s)) = \begin{cases} 
0 & \text{for } s = 0, \\
\mu_{22} & \text{for } 1 \leq s \leq B + 1, \\
\alpha \Sigma_2 & \text{for } s = B + 2,
\end{cases}
\]
and
\[
p(s'|s, d_0(s)) = \begin{cases} 
\frac{\alpha \Sigma_1}{q} & \text{for } s = 0, s' = 1, \\
1 - \frac{\alpha \Sigma_1}{q} & \text{for } s = s' = 0, \\
\frac{\mu_{22}}{q} & \text{for } 1 \leq s \leq B + 1, s' = s - 1, \\
1 - \frac{\mu_{11} + \mu_{22}}{q} & \text{for } 1 \leq s \leq B + 1, s' = s, \\
\frac{\mu_{11}}{q} & \text{for } 1 \leq s \leq B + 1, s' = s + 1, \\
\frac{\alpha \Sigma_2}{q} & \text{for } s = B + 2, s' = B + 1, \\
1 - \frac{\alpha \Sigma_2}{q} & \text{for } s = s' = B + 2,
\end{cases}
\]
where \( q \) is the uniformization constant. Since the policy \((d_0)\infty\) (corresponding to the decision rule \(d_0\)) is irreducible, we find a scalar \( g_0 \) and a vector \( h_0 \) solving
\[
r_{d_0} - g_0 e + (P_{d_0} - I)h_0 = 0, 
\]
subject to \( h_0(0) = 0 \), where \( e \) is a column vector of ones and \( I \) is the identity matrix. Then \( g_0 \) is provided in equation (2), \( h_0(0) = 0 \), and
\[
h_0(s) = \frac{q g_0}{\alpha \Sigma_1} \left( \sum_{j=0}^{s-2} (j + 1) \mu_{22}^{s-j-2} \mu_{11}^{j-s+1} (\alpha \Sigma_1 - \mu_{11} + \mu_{22}) + s \right) - q \mu_{22} \sum_{j=0}^{s-2} (j + 1) \mu_{22}^{s-j-2} \mu_{11}^{j-s+1}
\]
for \( 1 \leq s \leq B + 2 \) constitute a solution to equation (12). Note that \( g_0 \) is the long-run average throughput of the policy \((d_0)\infty\) and \( h_0 \) denotes the bias vector under the policy \((d_0)\infty\) (see pages 338 and 339 of Puterman [15] for the interpretation of the bias vector).

For the next step of the policy iteration algorithm, we choose
\[
d_1(s) \in \arg\max_{a \in A_s} \left\{ r(s, a) + \sum_{j \in S} p(j|s, a)h_0(j) \right\}, \quad \forall s \in S,
\]
setting \( d_1(s) = d_0(s) \) if possible. We now show that \( d_1(s) = d_0(s) \) for all \( s \in S \). In particular, for all \( s \in S \setminus \{0, B + 2\} \) and \( a \in A_s \setminus \{d_0(s)\} \), we will compute the differences
\[
\Delta(s, a) = r(s, d_0(s)) + \sum_{j \in S} p(j|s, d_0(s))h_0(j) - \left( r(s, a) + \sum_{j \in S} p(j|s, a)h_0(j) \right)
\]
and show that the differences are non-negative. Note that for \( s = 0 \) and \( s = B + 2 \), there is nothing to prove because there is only one possible action in these states, namely \( d_0(0) = a_{11} \) and \( d_0(B + 2) = a_{22} \).
For $1 \leq s \leq B + 1$, we have that $d_0(s) = a_{12}$. Since the set $A_s$ of all possible actions is large, in the interest of space, we will specify $\Delta(s, \alpha)$ only for actions $a_{11}$, $a_{21}$, and $a_{22}$. With some algebra we obtain
\[
\Delta(s, a_{11}) = \frac{\alpha \Sigma_1 \mu_1^{B+1-s} \sum_{j=0}^{s-1} \mu_1^{s-1-j} (2\mu_1 \mu_2 - \alpha \Sigma_1 \Sigma_2 + \mu_2 \mu_2 + \mu_1 \mu_1 \mu_2)}{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B} \mu_{11}^{B+j} + \mu_{22}^{B+1} \Sigma_2 + \mu_{11}^{B+1} \Sigma_1} \geq 0
\]
since $\alpha \leq \frac{2\mu_1 \mu_2 + \mu_2 \mu_2 + \mu_1 \mu_1 \mu_2}{\Sigma_1 \Sigma_2} = \frac{\mu_1}{\Sigma_1} + \frac{\mu_2}{\Sigma_2}$, and the expression is equal to zero only when $\alpha$ attains the upper bound. Similarly,
\[
\Delta(s, a_{21}) = \frac{\Gamma(s, \alpha)}{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B} \mu_{11}^{B+j} + \mu_{22}^{B+1} \Sigma_2 + \mu_{11}^{B+1} \Sigma_1},
\]
where $\Gamma(s, \alpha) = \Gamma_1(s, \alpha) + \Gamma_2(s, \alpha)$ with
\[
\Gamma_1(s, \alpha) = \sum_{j=B-s+1}^{B} \mu_{11}^{B-j} \Sigma_1 (\alpha (\mu_1 - \mu_2) \Sigma_2 + \mu_2 \mu_2 - \mu_1 \mu_2)
\]
and
\[
\Gamma_2(s, \alpha) = \sum_{j=-1}^{B-s} \mu_{11}^{j+1} \mu_{22}^{B-j} \Sigma_2 (\alpha (\mu_2 - \mu_2) \Sigma_1 + \mu_1 \mu_2 - \mu_2 \mu_2).
\]
We next prove that $\Gamma(s, \alpha) \geq 0$ by showing that $\Gamma_1(s, \alpha) \geq 0$ and $\Gamma_2(s, \alpha) \geq 0$ when $\mu_1 \mu_2 \geq \mu_2 \mu_2$ and $1 \leq \alpha \leq C_{22}$. Note that if $\mu_1 \geq \mu_2$, then $\Gamma_1(s, \alpha)$ is a non-decreasing function of $\alpha$ and
\[
\Gamma_1(s, 1) = \sum_{j=B-s+1}^{B} \mu_{11}^{B-j} \Sigma_1 (\mu_1 \mu_2 - \mu_2 \mu_2) \geq 0
\]
(where the equality is attained if and only if $\mu_1 \mu_2 = \mu_2 \mu_2$) implying that $\Gamma_1(s, \alpha) \geq 0$ for all $\alpha \geq 1$. If $\mu_1 < \mu_2$, then $\Gamma_1(s, \alpha)$ is a decreasing function of $\alpha$. But
\[
\Gamma_1(s, \frac{\mu_1}{\Sigma_1} + \frac{\mu_2}{\Sigma_2}) = 2 \sum_{j=B-s+1}^{B} \mu_{11}^{j+1} \mu_{22}^{B-j} (\mu_1 \mu_2 - \mu_2 \mu_2) \geq 0
\]
(where the equality is attained if and only if $\mu_1 \mu_2 = \mu_2 \mu_2$) implying that $\Gamma_1(s, \alpha) \geq 0$ for all $\alpha \leq C_{22}$. Similarly, if $\mu_2 \geq \mu_2$, $\Gamma_2(s, \alpha)$ is a non-decreasing function of $\alpha$ and
\[
\Gamma_2(s, 1) = \sum_{j=-1}^{B-s} \mu_{11}^{j+1} \mu_{22}^{B-j} \Sigma_2 (\mu_1 \mu_2 - \mu_2 \mu_2) \geq 0
\]
(where the equality is attained if and only if $\mu_1 \mu_2 = \mu_2 \mu_2$) implying that $\Gamma_2(s, \alpha) \geq 0$ for all $\alpha \geq 1$. On the other hand, if $\mu_2 < \mu_2$, then $\Gamma_2(s, \alpha)$ is a decreasing function of $\alpha$. However,
\[
\Gamma_2(s, \frac{\mu_1}{\Sigma_1} + \frac{\mu_2}{\Sigma_2}) = 2 \sum_{j=-1}^{B-s} \mu_{11}^{j+1} \mu_{22}^{B-j} (\mu_1 \mu_2 - \mu_2 \mu_2) \geq 0
\]
(where the equality is attained if and only if $\mu_{11}\mu_{22} = \mu_{12}\mu_{21}$) implying that $\Gamma_2(s, \alpha) \geq 0$ for all $\alpha \leq C_{22}$.

Finally,

$$\Delta(s, a_{22}) = \frac{\alpha \Sigma_2 \sum_{j=0}^{B+1-s} \mu_{11}^{j} \mu_{22}^{B-j} (2\mu_{11}\mu_{22} - \alpha \Sigma_1 \Sigma_2 + \mu_{21}\mu_{22} + \mu_{11}\mu_{12})}{\alpha \Sigma_1 \Sigma_2 \sum_{j=0}^{B} \mu_{11}^{j} \mu_{22}^{B-j} + \mu_{22}^{B+1} + \mu_{11}^{B+1} \Sigma_1} \geq 0$$

because $\alpha \leq \frac{2\mu_{11}\mu_{22} + \mu_{21}\mu_{22} + \mu_{11}\mu_{12}}{\Sigma_1 \Sigma_2} = C_{22}$, and the expression is equal to zero only when $\alpha$ attains the upper bound.

Proceeding as above, we have also shown that $\Delta(s, a) \geq 0$ for $a \in \{a_{00}, a_{01}, a_{02}, a_{10}, a_{20}\}$, with equality occurring only when $\mu_{11}\mu_{22} = \mu_{12}\mu_{21}$ or $\alpha = C_{22}$. This proves that $d_1(s) = d_0(s)$ for all $s \in S$. By Theorem 9.5.1 of Puterman [15] (which says that in a (weakly) communicating model policy iteration terminates with an optimal policy) this proves that the policy described in part (i) of Theorem 3.1 is optimal. The proof of the uniqueness of the optimal policy is similar to the uniqueness proof in Theorem 3.1 of Andradóttir and Ayhan [5].

We now consider the proof of part (ii) of Theorem 3.1. Note that if $\mu_{ij} > 0$ for all $i \in \{1, 2\}$ and $j \in \{1, 2\}$, then part (ii) of Theorem 3.1 follows directly from Theorem 4.2, but the proof below is valid even when $\mu_{ij} = 0$ for some $i \in \{1, 2\}$ and $j \in \{1, 2\}$. Without loss of generality, we again only provide the proof for the case with $\mu_{11}\mu_{22} \geq \mu_{21}\mu_{12}$, $\Sigma_1 > 0$, and $\Sigma_2 > 0$ (so that $C_{22} = \frac{\mu_{11}}{\Sigma_1} + \frac{\mu_{22}}{\Sigma_2}$).

Note that we again have a communicating Markov decision process (e.g., to go from state $s_1$ to some state $s_2$, assign both servers to station 1 for all $s < s_2$, both servers to station 2 for all $s > s_2$, and both servers to station 1 (2) if $s_2 = 0$ ($s_2 > 0$)). Thus, as in the proof of part (i), we use the policy iteration algorithm for communicating Markov decision processes. This time, as the initial policy of the policy iteration algorithm, we choose the expedite policy given as

$$d_0(s) = \begin{cases} a_{11} & \text{for } s = 0, \\ a_{22} & \text{for } 1 \leq s \leq B + 2, \end{cases}$$

corresponding to the expedite policy (which is one of the optimal policies as described in part (ii) of Theorem 3.1). Then

$$r(s, d_0(s)) = \begin{cases} 0 & \text{for } s = 0, \\ \alpha \Sigma_2 & \text{for } 1 \leq s \leq B + 2, \end{cases}$$

and

$$p(s'|s, d_0(s)) = \begin{cases} \frac{\alpha \Sigma_1}{q} & \text{for } s = 0, s' = 1, \\ 1 - \frac{\alpha \Sigma_1}{q} & \text{for } s = s' = 0, \\ \frac{\alpha \Sigma_2}{q} & \text{for } 1 \leq s \leq B + 2, s' = s - 1, \\ 1 - \frac{\alpha \Sigma_2}{q} & \text{for } 1 \leq s \leq B + 2, s = s', \end{cases}$$

where $q$ is the uniformization constant. Since the policy $(d_0')^\infty$ (corresponding to the decision rule $d_0'$) has a unichain structure, we again find a scalar $g_0'$ and a vector $h_0'$ solving (12) (with $d_0$ replaced by $d_0'$) subject to $h_0'(0) = 0$. Then

$$g_0' = T_{\text{expedite}} = \frac{\alpha \Sigma_1 \Sigma_2}{\Sigma_1 + \Sigma_2},$$

21
\( h'_0(0) = 0, \) and
\[
  h'_0(s) = \frac{qs \Sigma_2}{\Sigma_1 + \Sigma_2}
\]
for \( 1 \leq s \leq B + 2 \) constitute a solution to equation (12).

For the next step of the policy iteration algorithm, we again choose
\[
d'_1(s) \in \operatorname{arg \ max}_{a \in A_s} \left\{ r(s, a) + \sum_{j \in S} p(j | s, a) h'_0(j) \right\}, \quad \forall s \in S,
\]
setting \( d'_1(s) = d'_0(s) \) if possible and show that \( d'_1(s) = d'_0(s) \) for all \( s \in S \). In particular, for all \( s \in S \setminus \{0, B + 2\} \) and \( a \in A_s \setminus \{d'_0(s)\} \), we compute the differences
\[
  \Delta'(s, a) = r(s, d'_0(s)) + \sum_{j \in S} p(j | s, d'_0(s)) h'_0(j) - \left( r(s, a) + \sum_{j \in S} p(j | s, a) h'_0(j) \right)
\]
and show that the differences are non-negative.

For \( 1 \leq s \leq B + 1 \), we have that \( d'_0(s) = a_{22} \). In the interest of space, we will again specify \( \Delta'(s, a) \) only for actions \( a_{11}, a_{12}, \) and \( a_{21} \). With some algebra, we obtain \( \Delta'(s, a_{11}) = 0 \), which implies that both actions \( a_{11} \) and \( a_{22} \) can be used in states \( 1 \leq s \leq B + 1 \) (see the comment on the execution of the expedite policy after the statement of Theorem 3.1). Similarly,
\[
  \Delta'(s, a_{12}) = \frac{\alpha \Sigma_1 \Sigma_2 - 2 \mu_{11} \mu_{22} - \mu_{11} \mu_{12} - \mu_{21} \mu_{22}}{\Sigma_1 + \Sigma_2} > 0
\]
because \( \alpha > \frac{2 \mu_{11} \mu_{22} + \mu_{11} \mu_{12} + \mu_{21} \mu_{22}}{\Sigma_1 \Sigma_2} = C_{22} \). Finally,
\[
  \Delta'(s, a_{12}) = \frac{\alpha \Sigma_1 \Sigma_2 - 2 \mu_{12} \mu_{21} - \mu_{11} \mu_{12} - \mu_{21} \mu_{22}}{\Sigma_1 + \Sigma_2} > 0
\]
because \( \alpha > C_{22} = \frac{2 \mu_{12} \mu_{21} + \mu_{11} \mu_{12} + \mu_{21} \mu_{22}}{\Sigma_1 \Sigma_2} = \frac{q_{12}}{\Sigma_1} + \frac{q_{12}}{\Sigma_2} \).

We have also shown that \( \Delta'(s, a) > 0 \) for \( a \in \{a_{00}, a_{01}, a_{02}, a_{10}, a_{20}\} \). This proves that \( d'_1(s) = d'_0(s) \) for all \( s \in S \) and, hence, the expedite policy is optimal. The proof of the uniqueness of the optimal policy (subject to the interpretation that a fully collaborative policy can be executed in various ways as long as the two servers work in a team at all times) is again similar to the uniqueness proof in Theorem 3.1 of Andradóttir and Ayhan [5]. \( \square \)