

Research Article

Product-Form Solutions for Integrated Services Packet Networks and Cloud Computing Systems

Wanyang Dai

Department of Mathematics and State Key Laboratory of Novel Software Technology, Nanjing University, Nanjing 210093, China

Correspondence should be addressed to Wanyang Dai; nan5lu8@netra.nju.edu.cn

Received 25 February 2014; Revised 26 August 2014; Accepted 27 August 2014; Published 2 October 2014

Academic Editor: Joao B. R. Do Val

Copyright © 2014 Wanyang Dai. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We iteratively derive the product-form solutions of stationary distributions for a type of preemptive priority multiclass queueing networks with multiserver stations. This type of queueing systems can typically be used to model the stochastic dynamics of some large scale backbone networks with multiprocessor shared-memory switches or local (edge) cloud computing centers with parallel-server pools. The queueing networks are Markovian with exponential interarrival and service time distributions. The obtained iterative solutions can be used to conduct performance analysis or as comparison criteria for approximation and simulation studies. Numerical comparisons with existing Brownian approximating model (BAM) related to general interarrival and service times are provided to show the effectiveness of our current designed algorithm and our previous derived BAM. Furthermore, based on the iterative solutions, we can also give some analysis concerning network stability for some cases of these queueing systems, which provides some insight for more general study.

1. Introduction

At present, integrated services packet networks (ISPN) are widely used to transport a wide range of information such as voice, video, and data. It is foreseeable that this integrated services pattern will be one of the major techniques in the future cloud computing based communication systems and Internet (see, e.g., Mullender and Wood [1]). The introduction of concept and architecture of cloud computing can be found in, for example, Mell and Grance [2] and Rhoton and Haukioja [3]. A possible cloud computing based telecommunication network architecture (i.e., a large-scale network infrastructure as a service) is designed in Figure 1, where an end-user may require service (or services) from single local cloud computing center or multiple local and remote cloud computing centers. Among these centers, they communicate each other by using core switching network system (note that the switching system itself can also be independently viewed and handled as a cloud computing system with multiple service pools).

The speed and efficiency of core switching network systems are the bottleneck in realizing high speed owing to the drastic improvement in transmission speed and reliability of optical fiber. In an ISPN network, information is partitioned into packets depending on the employed protocol such as Internet protocol (IP). For the purpose of transmission, each packet consists of user's data payload to be transmitted, a header containing control information (e.g., source and destination addresses, packet type, priority, etc.), and a checksum used for error control. The high-speed ISPN networks require fast packet switches to move packets along their respective virtual paths. The switches are computers with processing and storage capabilities. The main objective of a switch is to route an incoming packet arriving along a particular input link to a specified output link. More precisely, once the incoming packet is entirely received, the switch examines its header and selects the next outgoing link. In other words, packets are transmitted in a store-and-forward manner.

Various pieces of information can be classified into a fixed number of different types with separate requirements of quality of service (e.g., different end-to-end delays and packet loss ratios). Real time traffic packets with stringent delay constraint (e.g., interactive audio/video) are endowed service priority. In the meanwhile, it is imperative to size buffers for

End

users



End

FIGURE 1: An integrated services cloud computing based network.

service

nonreal time traffic packets (e.g., data), which can tolerate higher delay but demand much smaller packet loss ratios. Hence, efficient switching and buffer management schemes are needed for switches.

Currently, three basic techniques (see, e.g., Tobagi [4] and Rao and Tripathi [5]) are designed to realize the switching function: space-division, shared-medium, and sharedmemory. In our switches, the shared-memory technique is employed, which is comprised of a single dual-ported memory shared by all input and output lines. Packets arriving along input lines are multiplexed into a single stream that is fed to the common memory for storage. Inside the memory, packets are organized into different output buffers, one for each output line. In each output buffer, memory can be further divided into priority output queues, one for each type of traffic. In the meantime, an output stream of packets is formed by retrieving packets from the output queues sequentially, one per queue. Among different traffic types, packets are retrieved according to their priorities. Inside each type, packets are retrieved under the first-in first-out (FIFO) service discipline. Then, the output stream is demultiplexed and packets are transmitted on the output lines. There are some other ways to deal with the output queues such as processor sharing among output lines (see, e.g., Tobagi [4] and Rao and Tripathi [5]). We will address these issues elsewhere (see, e.g., Dai [6]).

In addition to switching, queueing is another main functionality of packet switches. The introduction of queueing function is owing to multiple packets arriving simultaneously from separate input lines and owing to the randomness of packet arrivals from both outside routing and internal routing of the network. There are three possibilities for queueing in a switch: *input queueing*, *output queueing*, and *shared-memory queueing* (see, e.g., Schwarz [7] and Rao and Tripathi [5]). Shared-memory-queueing mechanism is employed in our switches since it has numerous advantages with respect to switching and queueing functions can be implemented together by controlling the memory read and write properly. Furthermore, modifying the memory read and write control



FIGURE 2: A two-type and four-class network with two multiprocessor switches. Type 1 includes class 1 and class 2. Type 2 consists of class 3 and class 4. Type 1 owns the service priority.



FIGURE 3: A two-stage switch with four dual-ported shared-memory switching modules.

circuit makes the shared-memory switch sufficiently flexible to perform functions such as priority control and other operations.

Nevertheless, no matter what technology is employed in implementing the switch, it places certain limitations on the size of the switch and line speeds. Currently, two possible ways can be used to build large switches to match the transmission speed of optical fiber. The first one is to adopt parallel processors (see, e.g., Figure 2). The second one is to interconnect many switches (known as switch modules) in a *multistage* configuration to build large switches (see, e.g., Figure 3). The remaining issue is how to reasonably allocate resources of these switches and efficiently evaluate the system performance.

The statistical characteristics of packet interarrival times and packet sizes have a major impact on switch hardware and software designs owing to the consideration of network performance. How to more effectively identify packet traffic patterns is a very active and involved research field (see, e.g., Nikolaidis and Akyildiz [8] and Dai [6, 9, 10]). Independent and identical distribution (i.i.d.) is the popular assumption for these times and packet sizes. Doubly stochastic renewal process introduced in Dai [6] is the latest definition and generalization related to input traffic and service processes for a wireless network under random environment. The effectiveness of these characteristics is supported by recent discoveries in Cao et al. [11] and Dai [9, 10].

In all circumstances, it is imperative to find productform solutions for those queueing networks under suitable conditions to conduct performance analysis or provide comparison criteria to show the effectiveness of approximation and/or simulation studies (see, e.g., Dai et al. [6, 12–15]). Furthermore, note that the stochastic dynamics of the backbone networks with multiprocessor shared-memory switches and the local (edge) cloud computing centers with parallel-server pools in Figure 1 can both be described by a multiclass queueing networks with parallel servers at each station. Hence, in this paper, we aim to derive the productform solutions iteratively for one particular type of related queueing networks, that is, a type of particular preemptive priority multiclass queueing networks when the interarrival and service times are exponentially distributed. In addition, we also aim to provide some numerical comparisons with existing Brownian approximating model (BAM) related to general interarrival and service times to show the effectiveness of our current designed algorithm and our former derived BAM.

Next, we provide some review of the existing literature associated with the current study. Comparisons between the existing studies and our current discussion are also presented.

Under a general Whittle network framework, the product-form solutions are presented in Serfozo [16] for some multiclass networks, which include those with sectordependent (e.g., Example 3.3) and class-station-dependent service rates (e.g., BCMP networks in Section 3.3, which are introduced by Baskett et al. [17]). Without considering the interaction among different stations, the distinguishing feature of these networks is that a job's service rate at a station may consist of two intensities: one referred to as station intensity is a function of the total queue length at the station and the other one referred to as *class intensity* is a function of the queue length in the same class as the job being served. Some specific single-class queueing systems of BCMP networks in Section 3.3 of Serfozo [16] are revisited in Harrison [18, 19] by developing some Reversed Compound Agent Theorem (RCAT) based method. The corresponding product-form and non-product-form solutions are derived. Nevertheless, as claimed in Harrison [18, 19], heavier notation will be involved as long as a multiclass BCMP network is concerned.

Although our networks are of a form of those networks with multiple types of units as introduced in Serfozo [16], our networks are beyond those with sector-dependent and classstation-dependent service rates as introduced previously. More precisely, for a station in our networks, the station intensity is not only a function of the total queue length but also a function of combinations of queue lengths of various classes. The class intensity depends on not only the queue length of itself and/or the total queue length but also the numbers of jobs in other classes at the station. Therefore, how to find suitable function Φ and how to prove our service rates to be Φ -balanced as defined in Chapter 3 of Serfozo [16] are not obvious. Moreover, how to apply the RCAT based method developed in Harrison [18, 19] to our multiclass network cases is also not trivial. Thus, in this paper, we use the method of solving Kolmogorov (balance) equations to

get the product-form solutions iteratively, which are more engineering and computationally friendly. Furthermore, by this method, we can also give some analysis concerning network stability for some cases of these systems, which provides some insight for more general study.

The remainder of this paper is organized as follows. The open priority multiclass queueing network associated with high-speed ISPN is described in Section 2. Our main results including product-form solutions and performance comparisons are presented in Section 3. Numerical comparisons are given in Section 4. The proofs of our main theorems are provided in Section 5. The conclusion of the paper is presented in Section 6.

2. The Queueing Network Model

Note that the stochastic dynamics of the backbone core networks with multiprocessor shared-memory switches and the local (edge) cloud computing centers with parallel-server pools in Figure 1 can both be described by multiclass queueing networks with parallel servers at each station. Therefore, we consider a queueing network that has J multiserver stations. Each station indexed by $j \in \{1, ..., J\}$ owns c_i servers and has an infinite capacity waiting buffer. In the network, there are *I* job types. Each type consists of *J* job classes that are distributed at different stations. Therefore, the network is populated by K (= IJ) job classes which are labeled by $k \in \{1, \dots, K\}$. Upon the arrival of a job of a type from outside the network, it may only receive service for part of *J* classes and may visit a particular class more than once (but at most finite many times). Then, it leaves the network (i.e., the network is open). At any given time during its lifetime in the network, the job belongs to one of the classes and changes classes at each time a service is completed. All jobs within a class are served at a unique station and more than one class might be served at a station (so-called *multiclass* queueing network). The ordered sequence of classes that a job visits in the network is named *a route*. Interrouting among different job types is not allowed throughout the entire network.

We use $\mathscr{C}(j)$ to denote the set of classes belonging to station *j*. Let s(k) denote the station to which class *k* belongs. We implicitly set j = s(k) when *j* and *k* appear together. Associated with each class *k*, there are two i.i.d sequences of random variables (r.v.), $u_k = \{u_k(i), i \ge 1\}$ and $v_k = \{v_k(i), i \ge 1\}$, an i.i.d sequence of *K*-dimensional random vectors, $\phi^k = \{\phi^k(i), i \ge 1\}$, and two real numbers, $\alpha_k \ge 0$ and $m_k = 1/\mu_k > 0$. We suppose that the 3*K* sequences

$$u_1, \dots, u_K, \quad v_1, \dots, v_K, \quad \phi^1, \dots, \phi^K \tag{1}$$

are mutually independent. The initial r.v.s $u_k(1)$ and $v_k(1)$ have means $E[u_k(1)] = 1/\alpha_k$ and $E[v_k(1)] = m_k$, respectively. For each $i \in \{1, 2, ...\}$, $u_k(i)$ denotes the interarrival time between the (i - 1)th and the *i*th *externally* arrival job at class k. Furthermore, $v_k(i)$ denotes the service time for the *i*th class k job. In addition, $\phi^k(i)$ denotes the routing vector of the *i*th class k job. We allow $\alpha_k = 0$ for some classes $k \in \mathcal{C} \equiv \{k : \alpha_k \neq 0\}$. Then, it follows that α_k and μ_k are the external arrival rate and service rate for class k, respectively. We assume that

the routing vector $\phi^k(i)$ takes values in $\{e_0, e_1, \dots, e_K\}$, where e_0 is the *K*-dimensional vector of all 0's. For $l = 1, \dots, K$, e_l is the *K*-dimensional vector with *l*th component 1 and other components 0. When $\phi^k(i) = e_l$, the *i*th job departing class *k* becomes a class *l* job. Let $p_{kl} = P\{\phi^k(i) = e_l\}$ be the probability that a job departing class *k* becomes a class *l* job (of the same type). Thus, the corresponding $K \times K$ matrix $P = (p_{kl})$ is *routing matrix* of the network. Furthermore, the matrix

$$Q = I + P' + (P')^2 + \cdots$$
 (2)

is finite; that is, (I - P') is invertible with $\mathcal{Q} = (I - P')^{-1}$ since the network is open. The symbol ' on a vector or a matrix denotes the transpose and I denotes the identity matrix.

We use λ_k for $k \in \{1, ..., K\}$ to denote the overall arrival rate of class k, including both external arrivals and internal transitions. Then, we have the following *traffic equation*:

$$\lambda_k = \alpha_k + \sum_{l=1}^K \lambda_l p_{lk},\tag{3}$$

or in its vector form: $\lambda = \alpha + P' \lambda$ (all vectors in this paper are to be interpreted as column vectors unless explicitly stated otherwise). Note that the unique solution λ of (3) is given by $\lambda = @\alpha$. For each λ_k , if there is a long-run average rate of flow into the class which equals the long-run average rate out of that class, this rate will equal λ_k . Furthermore, we define the *traffic intensity* ρ_i for station *j* as follows:

$$\rho_j = \sum_{k \in \mathscr{C}(j)} \frac{\lambda_k}{c_j \mu_k},\tag{4}$$

where ρ_j with $\rho_j \le 1$ is also referred to as the nominal fraction of time that station *j* is nonidle.

The order of jobs being served at each station is dictated by a service discipline. In the current research, we restrict our attention to static buffer priority (SBP) service disciplines under which the classes at each station are assigned a fixed rank (with no ties). In our queueing network, each type of jobs is assigned the same priority rank at every station where it possibly receives service. When a server within a station switches from one job to another, the new job will be taken from the leading (or longest waiting) job at the highest rank nonempty class at the server's station. Within each class, it is assumed that jobs are served on the first-in first-out (FIFO) basis. We suppose that the discipline employed is *nonidling*; that is, a server is never idle when there are jobs waiting to be served at its station. We also assume that the discipline is *preemptive resume*; that is, when a job arrives at a station with all servers busy and if the job is with a higher rank than at least one of the jobs currently being served, one of the lower rank job services is interrupted; when there is a server available to the interrupted service, it continues from where it left off. For convenience and without loss of generality, we use consecutive numbers to index the classes that have the same priority rank at stations 1 to J. In other words, the highest priority classes for station 1 to J are indexed by 1 to J, the second highest priority classes are indexed by J + 1 to $2J, \ldots$, and the lowest priority classes are indexed by K-J+1 to K. An example of such a two-station network is given in Figure 2. In the network, type 1 traffic possibly requires class 1 and class 2 services, type 2 traffic possibly requires class 3 and class 4 services, and classes in type 1 have the higher priority at their corresponding stations.

Finally, we define the cumulative arrival, cumulative service, and cumulative routing processes by the following sums:

$$U_{k}(n) = \sum_{i=1}^{n} u_{k}(i), \qquad V_{k}(n) = \sum_{i=1}^{n} v_{k}(i),$$

$$\Phi^{k}(n) = \sum_{i=1}^{n} \phi^{k}(i),$$
(5)

where the *l*th component $\Phi_l^k(n)$ of $\Phi^k(n)$ is the cumulative number of jobs to class *l* for the first *n*th jobs leaving class *k* with n = 1, 2, ... and $k, l \in \{1, ..., K\}$. Then, we define

$$E_{k}(t) \equiv \max\{n \ge 0, U_{k}(n) \le t\},\$$

$$S_{k}(t) \equiv \max\{n \ge 0, V_{k}(n) \le t\},\$$

$$A_{k}(t) \equiv E_{k}(t) + \sum_{l=1}^{K} \Phi_{k}^{l}(S_{l}(t)).$$
(6)

Note that $E_k(t)$ denotes the total number of external arrivals to class k in time interval [0,t]. $S_k(t)$ represents the total number of class k jobs which have finished service in [0,t]. $A_k(t)$ is the total arrivals to class k in [0,t] including both external arrivals and internal transitions.

3. Steady-State Queue Length Distributions

We use $Q_k(t)$ to denote the number of class k jobs in station j = s(k) with $j \in \{1, 2, ..., J\}$ and $k \in \{1, 2, ..., K\}$ at time t. It is called the *queue length process* for class k jobs. For convenience, let $Q_{(i,j)}(t)$ and $n_{(i,j)}$ be the (j-i+1)-dimensional queue length process and (j-i+1)-dimensional state vectors, respectively. They are given by

$$Q_{(i,j)}(t) = (Q_i(t), \dots, Q_j(t)), \qquad n_{(i,j)} = (n_i, \dots, n_j)$$
 (7)

for j > i and $i, j \in \{1, 2, ..., K\}$ and nonnegative integers $n_i, ..., n_j$. Then, we use $P_{n_{(1,K)}}(t)$ to denote the probability of system at state $n_{(1,K)} = (n_1, ..., n_K)$ and at time *t*; that is,

$$P_{n_{(1,K)}}(t) = P\left\{Q_{(1,K)}(t) = n_{(1,K)}\right\}.$$
(8)

Furthermore, let

$$P_{n_{(1,K)}} = P\left\{Q_{(1,K)}\left(\infty\right) = n_{(1,K)}\right\}$$
(9)

denote the corresponding steady-state probability of system at state (n_1, \ldots, n_K) if the network exists in a stationary distribution. In addition, let $P_{n_{i_{j+k}}}^{s(iJ+k)}$ denote the probability at state n_{iJ+k} for class iJ + k with $i \in \{0, 1, \ldots, I-1\}$ and $k \in \{1, 2, \ldots, J\}$.

Mathematical Problems in Engineering

Under the usual convention, let $x \wedge y$ denote the smaller one of any two real numbers x and y. Let $x \vee y$ denote the larger one of x and y; that is, $x \wedge y \equiv \min\{x, y\}$ and $x \vee$ $y \equiv \max\{x, y\}$. Then, for each $k \in \{1, ..., K\}$, we have the following notation:

$$a(x, y) \equiv (x \wedge (c_{s(k)} - y)) \vee 0.$$
⁽¹⁰⁾

Finally, for i > 0, define

$$\kappa(n_{iJ+k}) \equiv \begin{cases} 0 & \text{if } n_{iJ+k} = 0, \\ \sum_{u=0}^{i-1} n_{uJ+k} < c_{s(k)} \\ \times \prod_{u=0}^{i-1} P_{n_{uJ+k}}^{s(uJ+k)} & \text{if } n_{iJ+k} > 0. \end{cases}$$
(11)

Furthermore, for i = 0, define $\kappa(n_k) \equiv n_k \wedge c_{s(k)}$.

Theorem 1 (steady-state distribution). Assume that all of the service times and external interarrival times are exponentially distributed with rates as before and the traffic intensity $\rho_j < 1$ for all $j \in \{1, ..., J\}$. Furthermore, suppose that

$$\frac{\lambda_{iJ+k}}{\kappa\left(n_{iJ+k}\right)\mu_{iJ+k}} < 1 \tag{12}$$

for $n_{iJ+k} > 0$, $i \in \{0, ..., I-1\}$, and $k \in \{1, ..., J\}$. Then, for each $h \in \{1, 2, ..., I\}$, the steady-state distribution is given by the following product form:

$$P_{n_{(1,h_f)}} = \prod_{i=0}^{h-1} \prod_{k=1}^{f} P_{n_{if+k}}^{s(if+k)}.$$
(13)

More precisely, for $n_{iI+k} \ge 1$ *,*

$$P_{n_{iJ+k}}^{s(iJ+k)} = P_0^{s(iJ+k)} \prod_{r_{iJ+k}=1}^{n_{iJ+k}} \frac{\lambda_{iJ+k}}{\kappa(r_{iJ+k})\,\mu_{iJ+k}},$$
(14)

and, for $n_{iJ+k} = 0$, the initial distribution $P_0^{s(iJ+k)}$ is determined by

$$P_0^{s(iJ+k)} = \frac{1}{1 + \sum_{n_{iJ+k}=1}^{\infty} \prod_{r_{iJ+k}=1}^{n_{iJ+k}} (\lambda_{iJ+k}/\kappa(r_{iJ+k})\,\mu_{iJ+k})}.$$
 (15)

The following theorem relates condition (12) to primitive interarrival time and service rates for some cases of these systems, which provides some insight for more general study.

Proposition 2 (network stability condition). Under the exponential assumptions as stated in Theorem 1, if $\rho_j < 1$, for each $j \in \{1, ..., J\}$, the stability condition (12) holds for the following networks.

Net I. Multiclass networks with single-server stations, that is, the number c_j of servers is one for all stations $j \in \{1, ..., J\}$ while the number I of job types can be arbitrary.

Net II. The number c_j *of servers can be arbitrary for all stations* $j \in \{1, ..., J\}$ while the number of job types equals two (I = 2).

TABLE 1: Numerical tests for network stability condition (12).

$1 < \lambda_k/\mu_k < 2, \rho_{s(k)} < 1$									
$\lambda_{2J+k}/\kappa(n_{2J+k})\mu_{2J+k}$	λ_k/μ_k	λ_{J+k}/μ_{J+k}	$\lambda_{2J+k}/\mu_{2J+k}$						
0.16748	1.10	0.30	0.10						
0.98814	1.10	0.30	0.59						
0.99111	1.50	0.30	0.19						
0.85890	1.90	0.05	0.04						

We conjecture that $\rho_j < 1$ for $j \in \{1, ..., J\}$ implies the condition (12) for our general network. Nevertheless, owing to complex computation involved, the corresponding analytical illustration is not a trivial task.

Example 3. Consider a network with three job types (I = 3) and at least one station j = s(k) having two servers $(c_{s(k)} = 2 \text{ for } k \in \{1, ..., J\})$ while other stations have at most two servers $(c_{s(l)} \leq 2 \text{ for } l \in \{1, 2, ..., J\} \setminus \{k\})$. For a station s(k) with three servers, the condition (12) can be explicitly expressed as follows:

$$\left(\frac{\lambda_{2J+k}}{\mu_{2J+k}}\right) \times \left(1 + \frac{\lambda_k}{\mu_k} + \frac{\lambda_{J+k}}{2\mu_{J+k}} - \left(\frac{\lambda_k}{2\mu_k}\right)^2 - \frac{1}{4}\left(\frac{\lambda_k}{\mu_k}\right)^3 + \frac{\lambda_k\lambda_{J+k}}{4\mu_k\mu_{J+k}}\right) \times \left(2\left(1 - \frac{\lambda_k}{2\mu_k} - \frac{\lambda_{J+k}}{2\mu_{J+k}}\right)\right) \times \left(1 + \frac{\lambda_k}{\mu_k} + \frac{\lambda_{J+k}}{2\mu_{J+k}} - \left(\frac{\lambda_k}{2\mu_k}\right)^2 - \frac{1}{4}\left(\frac{\lambda_k}{\mu_k}\right)^3 + 2\left(1 - \frac{\lambda_k}{2\mu_k}\right)\frac{\lambda_k\lambda_{J+k}}{4\mu_k\mu_{J+k}}\right)^{-1} < 1$$
(16)

for $n_{2J+k} \ge 2$. Under $\rho_j < 1$, it is easy to see that the above inequality is true if $\lambda_k/\mu_k \le 1$. Numerical tests in Table 1 have also been conducted and show that the inequality is true even for $1 < \lambda_k/\mu_k < 2$, but the corresponding analytic demonstration could be nontrivial. The detailed illustration of the example will be provided at the end of this paper.

Remark 4. The justifications of Theorem 1 and Proposition 2 are postponed to Section 5. Instead, we will first use these results as comparison criteria to illustrate the effectiveness of the diffusion approximation models developed in Dai [13] and answer the question on when these approximation models can be employed.

4. Numerical Comparisons

First of all, we note that partial results presented in this section were briefly reported in the short conference version (Dai [20]) of this paper. More precisely, we consider a network with single-server station and under preemptive

priority service discipline. By employing the exact solutions developed in previous sections, we conduct performance comparisons between these product-form solutions and the approximating ones of Brownian network models.

Brownian network models, which are also known as semimartingale reflecting Brownian motions (SRBM), have been widely employed as approximating models for multiclass queueing networks with general interarrival and service time distributions when the traffic intensity defined in (4) is suitably large or close to one (see, e.g., Dai [13]). The effectiveness of the Brownian network models has been justified in numerical cases by comparing their solutions with either exact solutions or simulation results (see, e.g., Dai [13], Chen and Yao [21], and Dai et al. [22]). Thus, it is meaningful to illustrate the correctness of our newly derived formula by comparing it with the Brownian network models.

For the network with exponential interarrival and service time distributions, by using Theorem 1, we get the steadystate mean queue length for each class $k \in \{1, ..., K\}$ as

$$EQ_{k}(\infty) = \sum_{n_{k}=0}^{\infty} n_{k} P_{n_{k}}^{s(k)} = \frac{\lambda_{k} m_{k}}{1 - \sum_{i=1}^{k} \lambda_{i} m_{i}}$$

$$= \frac{\alpha_{k} m_{k} / (1 - p_{kk})}{1 - \sum_{i=1}^{k} \alpha_{i} m_{i} / (1 - p_{ii})}$$
(17)

and the expected total time (sojourn time) a job had to spend in the system as

$$ET_{k}(\infty) = \frac{m_{k}}{1 - \sum_{i=1}^{k} \lambda_{i} m_{i}} = \frac{m_{k}}{1 - \sum_{i=1}^{k} \alpha_{i} m_{i} / (1 - p_{ii})}.$$
 (18)

For the network with general interarrival and service time distributions, owing to the nature of our network routing structure, the higher priority classes are independent of lower ones. Then, it follows from the studies in Dai [13], Harrison [23], and Chen and Yao [21] that the steady-state mean sojourn time and mean queue length for each class k can iteratively be calculated with respect to priority order as follows:

$$ET_{k}(\infty) = \frac{EW_{k}(\infty) + m_{k}}{1 - \sum_{i=1}^{k-1} \lambda_{i}m_{i}}, \qquad EQ_{k}(\infty) = \lambda_{k}ET_{k}(\infty),$$
(19)

where $EW_k(\infty)$ is the steady-state mean total workload for all classes $i \in \{1, ..., k\}$ and k = 1, ..., K. More precisely, it is given by

$$EW_k(\infty) = \frac{\sigma_k^2}{2|\theta_k|},\tag{20}$$

where

$$\theta_{k} = (1 - p_{kk}) \left(\sum_{i=1}^{k} \lambda_{i} m_{i} - 1 \right),$$

$$\sigma_{k}^{2} = (1 - p_{kk})^{2} \sum_{i=1}^{k} \left(\lambda_{i} b_{i} + \frac{m_{i}^{2} \left(\alpha_{i}^{3} a_{i} + \lambda_{i} p_{ii} \left(1 - p_{ii} \right) \right)}{\left(1 - p_{ii} \right)^{2}} \right),$$
(21)

and a_i , b_i are the variances of interarrival and service time sequences for each class *i*.

In our numerical comparisons, we consider an exponential network with K = 3. For this case, our corresponding data are listed in Table 2. In the table, ERROR = SRBM – EXACT and RATIO = (|ERROR|/EXACT) * 100%, $E\widehat{Q}(\infty) = \sum_{i=1}^{3} EQ_i(\infty)$ and $E\widehat{T}_k(\infty) = \sum_{i=1}^{k} T_i(\infty)$ for k = 1, 2, 3. From the table, we can see that SRBMs are more reasonable approximations of their physical queueing counterparts when traffic intensities for the lowest priority jobs are relatively large.

5. Proofs of Theorem 1 and Proposition 2

5.1. Proof of Theorem 1. For convenience, we introduce some additional notations. Let $\Delta E_k(t)$, $\Delta S_k(t)$, and $\Delta A_k(t)$ be defined by

$$\Delta E_{k}(t) \equiv E_{k}(t + \Delta t) - E_{k}(t),$$

$$\Delta S_{k}(t) \equiv S_{k}(t + \Delta t) - S_{k}(t),$$

$$\Delta A_{k}(t) \equiv A_{k}(t + \Delta t) - A_{k}(t),$$

(22)

which denote the cumulative external arrivals to class k, the cumulative number of jobs finished services at class k, and the total arrivals to class k in $[t, t + \Delta t]$. Then, we can justify Theorem 1 by induction as in the following several steps.

Step 1. We consider the steady-state distribution for the highest rank classes with each index $k \in \{1, 2, ..., J\}$. In this case, the type index i = 0 is used in Theorem 1. Owing to the preemptive-resume service discipline and the class routing structure, we know that these *J* classes form a Jackson network. Then, by the theorem in Jackson [24], we have the following product form:

$$P_{n_{(1,J)}} = P_{n_1}^{s(1)} P_{n_2}^{s(2)} \cdots P_{n_J}^{s(J)},$$
(23)

where $P_{n_k}^{s(k)}$ denotes the steady-state probability at state n_k for class $k \in \{1, 2, ..., J\}$ at station s(k). More precisely, it is given by

$$P_{n_{k}}^{s(k)} = \begin{cases} \frac{P_{0}^{s(k)} (\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!}, & \text{if } 0 \le n_{k} < c_{k}, \\ \frac{P_{0}^{s(k)} (\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}! (n_{k})^{n_{k}-c_{k}}}, & \text{if } n_{k} \ge c_{k} \end{cases}$$
(24)

with $P_0^{s(k)}$ being determined by equation $\sum_{n_k=0}^{\infty} P_{n_k}^{s(k)} = 1$; that is,

 $P_0^{s(k)}$

$$= \left(\sum_{n_{k}=0}^{c_{s(k)}-1} \frac{1}{n_{k}!} \left(\frac{\lambda_{k}}{\mu_{k}}\right)^{n_{k}} + \frac{\left(\lambda_{k}/\mu_{k}\right)^{c_{s(k)}}}{c_{s(k)}!\left(1-\lambda_{k}/c_{s(k)}\mu_{k}\right)}\right)^{-1}.$$
 (25)

Step 2. We derive the steady-state distribution for the highest rank and the second highest rank classes with $k \in$

	$K = 3, m_1 = 1, m_2 = 2, m_3 = 3, p_{11} = 0.1, p_{22} = 0.2, p_{33} = 0.3$									
	λ_1	λ_2	λ_3	ρ	$E\widehat{Q}(\infty)$	$E\widehat{T}_1(\infty)$	$E\widehat{T}_2(\infty)$	$E\widehat{T}_{3}(\infty)$		
EXACT	0.20	0.20	0.06	0.78	2.0682	1.2500	6.2500	14.8864		
SRBM	0.20	0.20	0.06	0.78	2.2410	1.2500	6.4722	17.0253		
ERROR					-0.1728	0.0000	0.2222	2.1389		
RATIO					8.35%	0.00%	3.56%	14.37%		
EXACT	0.20	0.20	0.085	0.855	3.0086	1.2500	6.2500	21.9397		
SRBM	0.20	0.20	0.085	0.855	3.2349	1.2500	6.4722	24.0785		
ERROR					-0.2263	0.0000	0.2222	2.1389		
RATIO					7.52%	0.00%	3.56%	9.75%		
EXACT	0.20	0.20	0.115	0.945	7.5227	1.2500	6.2500	55.7955		
SRBM	0.20	0.20	0.115	0.945	7.8131	1.2500	6.4722	57.9344		
ERROR					-0.2904	0.0000	0.2222	2.1389		
RATIO					3.86%	0.00%	3.56%	3.83%		
EXACT	0.10	0.15	0.115	0.745	1.9641	1.1111	4.4444	12.8758		
SRBM	0.10	0.15	0.115	0.745	2.0944	1.1111	4.5432	13.8804		
ERROR					-0.1303	0.0000	0.0988	1.0046		
RATIO					6.64%	0.00%	2.22%	7.80%		
EXACT	0.10	0.15	0.15	0.85	3.6111	1.2500	4.4444	21.1111		
SRBM	0.10	0.15	0.15	0.85	3.7766	1.2500	4.5432	22.1157		
ERROR					-0.1655	0.0000	0.0988	1.0046		
RATIO					4.58%	0.00%	2.22%	4.76%		
EXACT	0.10	0.15	0.185	0.955	12.9445	1.2500	4.4444	67.7778		
SRBM	0.10	0.15	0.185	0.955	13.1451	1.2500	4.5432	68.7824		
ERROR					-0.2007	0.0000	0.0988	1.0046		
RATIO					1.55%	0.00%	2.22%	1.48%		
EXACT	0.45	0.22	0.02	0.95	6.0182	1.8182	20.0000	61.8182		
SRBM	0.45	0.22	0.02	0.95	6.3818	1.8182	20.7273	72.0000		
ERROR					-0.3636	0.0000	0.7273	10.1818		
RATIO					6.04%	0.00%	3.64%	16.47%		
EXACT	0.45	0.22	0.03	0.98	9.3182	1.8182	20.0000	151.8181		
SRBM	0.45	0.22	0.03	0.98	9.7836	1.8182	20.7273	161.9999		
ERROR					-0.4655	0.0000	0.7273	10.1818		
RATIO					5.00%	0.00%	3.64%	6.71%		

TABLE 2: Performance comparisons for a priority multiclass network with three job types.

 $\{1, 2, ..., 2J\}$. Owing to the preemptive-resume service discipline and the job routing structure, we have

$$P_{n_{(1,2J)}}(t) = P_{n_{(1,J)}/n_{(J+1,2J)}}(t) P_{n_{(J+1,2J)}}(t)$$

$$= P_{n_{(1,J)}}(t) P_{n_{(J+1,2J)}}(t),$$
(26)

where the second equality follows from the preemptive assumption and $P_{n_{(1,J)}/n_{(J+1,2J)}}(t)$ is the conditional probability at time *t* for classes with $k \in \{1, ..., J\}$ at state $n_{(1,J)}$ in terms of classes with $k \in \{J + 1, ..., 2J\}$ at state $n_{(J+1,2J)}$; that is,

$$P_{n_{(1,J)}/n_{(J+1,2J)}}(t) = P\left\{Q_{(1,J)}(t) = n_{(1,J)} \mid Q_{(J+1,2J)}(t) = n_{(J+1,2J)}\right\}.$$
(27)

In order to get the steady state distribution for $P_{n_{(J+1,2J)}}(t)$, we consider each state $n_{(J+1,2J)}$ at time t for the second highest

rank class jobs. There are several ways in which the system can reach it. They can be summarized in the following formula:

$$\begin{split} P_{n_{(J+1,2J)}} \left(t + \Delta t \right) \\ &= \left(1 - \sum_{k=1}^{J} \alpha_{J+k} \Delta t \right. \\ &- \sum_{k=1}^{J} \sum_{n_1, \dots, n_J = 0}^{\infty} a \left(n_{J+k}, n_k \right) \mu_{J+k} \left(1 - p_{J+k,J+k} \right) \\ &\times P_{n_{(1,J)}} \left(t \right) \Delta t \left. \right) \end{split}$$

 $\times P_{n_{(J+1,2J)}}\left(t\right)$

$$+ \sum_{k=1}^{J} \left(\sum_{n_{1},\dots,n_{j}=0}^{\infty} a\left(n_{J+k}+1,n_{k}\right) \mu_{J+k} \times \left(1 - \sum_{l=1}^{J} p_{J+k,J+l}\right) P_{n_{(1,J)}}\left(t\right) \Delta t \right) \times P_{n_{J+1},\dots,n_{J+k}+1,\dots,n_{2J}}\left(t\right) \\+ \sum_{k=1}^{J} \left(n_{J+k} \wedge 1\right) \alpha_{J+k} \Delta t P_{n_{J+1},\dots,n_{J+k}-1,\dots,n_{2J}}\left(t\right) \\+ \sum_{r,s=1,s\neq r}^{J} \left(\sum_{n_{1},\dots,n_{j}=0}^{\infty} a\left(n_{J+r}+1,n_{r}\right) \times \mu_{J+r} p_{J+r,J+s} P_{n_{(1,J)}}\left(t\right) \right) \\\times \left(n_{J+s} \wedge 1\right) \Delta t P_{n_{J+1},\dots,n_{J+r}+1,\dots,n_{J+s}-1,\dots,n_{2J}}\left(t\right) \\+ o\left(\Delta t\right),$$
(28)

where $a(n_{J+k}, n_k) = (n_{J+k} \land (c_{s(k)} - n_k)) \lor 0$ as defined in (10). $o(\Delta t)$ is infinitesimal in terms of Δt ; that is, $o(\Delta t)/\Delta t \rightarrow 0$ as $\Delta t \rightarrow \infty$. Equation (28) can be illustrated in the following disjoint events.

Event \mathscr{A} . The *J*-dimensional queue length process $Q_{(J+1,2J)}(t)$ keeps at state $n_{(J+1,2J)}$ unchanging at times *t* and $t + \Delta t$; that is,

$$Q_{(J+1,2J)}(t) = n_{(J+1,2J)}, \qquad Q_{(J+1,2J)}(t+\Delta t) = n_{(J+1,2J)}.$$
(29)

This event involves the following two parts.

Part One. Suppose $\Delta S_{J+k}(t) = \Delta A_{J+k}(t) = 0$ for all $k \in \{1, 2, ..., J\}$; that is, no external jobs arrive to classes with indices belonging to $\{J + 1, ..., 2J\}$ during $[t, t + \Delta t]$ while no jobs finish their services, either because the jobs being served at time *t* require longer service times than Δt or because the services are blocked or interrupted by higher rank class jobs. Then, the probability for Part One can be expressed in the following product form:

$$P\left\{\sum_{k=1}^{J} \Delta S_{J+k}(t) = \sum_{k=1}^{J} \Delta A_{J+k}(t) = 0\right\}$$
$$= P\left\{\sum_{k=1}^{J} \Delta S_{J+k}(t) = \sum_{k=1}^{J} \Delta E_{J+k}(t) = 0\right\}$$
$$= \prod_{k=1}^{J} P\left\{\Delta S_{J+k}(t) = \Delta E_{J+k}(t) = 0 \mid \sum_{s=k}^{J} \Delta S_{J+s+1}(t) = \sum_{s=k}^{J} \Delta E_{J+s+1}(t) = 0\right\}$$

$$= \prod_{k=1}^{J} P \{ \mathscr{C}_{J+k}(t) \}$$

= $\prod_{k=1}^{J} (1 - \alpha_{J+k} \Delta t - a_{J+k}(t) \mu_{J+k} \Delta t + o(\Delta t))$
= $1 - \sum_{k=1}^{J} \alpha_{J+k} \Delta t - \sum_{k=1}^{J} a_{J+k}(t) \mu_{J+k} \Delta t + o(\Delta t),$
(30)

where $a_{I+k}(t)$ is defined to be

$$a_{J+k}(t) \equiv \left(Q_{J+k}(t) \land \left(c_{s(k)} - Q_k(t)\right)\right) \lor 0.$$
(31)

The event $\mathscr{C}_{J+k}(t)$ in the third equality of (30) is defined as follows:

$$\mathscr{C}_{J+k}(t) = \left\{ \Delta S_{J+k}(t) = \Delta E_{J+k}(t) = 0 \mid \\ \sum_{s=k}^{J} \Delta S_{J+s+1}(t) = \sum_{s=k}^{J} \Delta E_{J+s+1}(t) = 0 \right\}.$$
(32)

To explain the fourth equality of (30), we introduce more notations. Let $b_k(t)\Delta t$ denote the probability that $\Delta A_k(t) = 1$ for $k \in \{1, ..., J\}$, and let $c_{J+k}(t)\Delta t$ be the probability that $\sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 1$ for $k \in \{2, 3, ..., J\}$; that is,

$$b_{k}(t) \Delta t \equiv P \left\{ \Delta A_{k}(t) = 1 \right\},$$

$$c_{J+k}(t) \Delta t \equiv P \left\{ \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 1 \right\}.$$
(33)

These probabilities can be explicitly expressed in terms of the external arrival rates, service rates, and network states; for example,

$$b_{k}\left(t\right)\Delta t = \left(\alpha_{k} + \sum_{l \in \{1,\dots,J\}} p_{lk}\mu_{l}\left(c_{s\left(l\right)} \wedge Q_{l}\left(t\right)\right)\right)\Delta t.$$
(34)

Thus, by the independent assumptions on external arrival and service processes among different stations and classes, and for each $k \in \{2, 3, ..., J\}$, we have

$$P\left\{\mathscr{C}_{J+k}\left(t\right)\right\}$$
$$=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}P\left\{\mathscr{C}_{J+k}\left(t\right)\mid\Delta A_{k}\left(t\right)=n,\sum_{s=1}^{k-1}\Delta S_{J+s}\left(t\right)=m\right\}$$
$$\times P\left\{\Delta A_{k}\left(t\right)=n\right\}P\left\{\sum_{s=1}^{k-1}\Delta S_{J+s}\left(t\right)=n\right\}$$

$$= P \left\{ \mathscr{C}_{J+k}(t) \mid \Delta A_{k}(t) = 0, \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 0 \right\}$$

$$\times P \left\{ \Delta A_{k}(t) = 0 \right\} P \left\{ \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 0 \right\}$$

$$+ P \left\{ \mathscr{C}_{J+k}(t) \mid \Delta A_{k}(t) = 1, \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 0 \right\}$$

$$\times P \left\{ \Delta A_{k}(t) = 1 \right\} P \left\{ \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 0 \right\}$$

$$+ P \left\{ \mathscr{C}_{J+k}(t) \mid \Delta A_{k}(t) = 0, \sum_{s=1}^{k-1} \Delta S_{J+s}(t) = 1 \right\}$$

$$\times P \left\{ \Delta A_{k}(t) = 0 \right\} P \left\{ \sum_{s=0}^{k-1} \Delta S_{J+s}(t) = 1 \right\}$$

$$+ o \left(\Delta t \right)$$

$$= (1 - a_{J+k}(t) \mu_{J+k} \Delta t) (1 - \alpha_{J+k} \Delta t)$$

$$\times (1 - b_k(t) \Delta t) (1 - c_{J+k}(t) \Delta t)$$

$$+ (1 - ((Q_{J+k}(t) \wedge (c_{s(k)} - (Q_k(t) + 1)))) \vee 0) \mu_{J+k} \Delta t)$$

$$\times (1 - \alpha_{J+k} \Delta t)$$

$$\times (b_k(t) \Delta t - p_{kk} \mu_k (c_{s(k)} \wedge Q_k(t)) \Delta t)$$

$$\times (1 - c_{J+k}(t) \Delta t)$$

$$+ (1 - ((Q_{J+k}(t) \wedge (c_{s(k)} - Q_k(t))) \vee 0) \mu_{J+k} \Delta t)$$

$$\times (1 - \alpha_{J+k} \Delta t)$$

$$\times p_{kk} \mu_k (c_{s(k)} \wedge Q_k(t)) \Delta t (1 - c_{J+k}(t) \Delta t)$$

$$+ (1 - (((Q_{J+k}(t) + 1) \wedge (c_{s(k)} - Q_k(t))) \vee 0) \mu_{J+k} \Delta t)$$

$$\times (1 - \alpha_{J+k} \Delta t) (1 - b_k(t) \Delta t) c_{J+k}(t) \Delta t$$

$$+ o (\Delta t)$$

$$= 1 - \alpha_{J+k} \Delta t - a_{J+k}(t) \mu_{J+k} \Delta t + o (\Delta t).$$
(35)

For the case k = 1, one can use the similar way to check that the result in the above equation is also true.

Part Two. Suppose $\Delta S_{J+k}(t) = \Delta A_{J+k}(t) \ge 1$ for at least one $k \in \{1, 2, ..., j\}$; that is, the number (at least one) of jobs

that finished their services for class J + k in $[t, t + \Delta t]$ equals that of jobs that arrived at the class. It is easy to see that the probability corresponding part one is given by

$$P \{ \Delta S_{J+k}(t) = \Delta A_{J+k}(t) \ge 1$$

for at least one $k \in \{1, 2, ..., j\} \}$
$$= \sum_{k=1}^{J} a_{J+k}(t) \mu_{J+k} p_{J+k,J+k} \Delta t + o(\Delta t),$$

(36)

where $a_{I+k}(t)$ is defined in (31).

Then, it follows from (30) and (36) that the probability for Event *A* is given by

$$P \{Q_{(J+1,2J)}(t) = n_{(J+1,2J)}; Q_{(J+1,2J)}(t + \Delta t) = n_{(J+1,2J)}\}$$

$$= \sum_{n_1,\dots,n_J=0}^{\infty} P \{Q_{(J+1,2J)}(t + \Delta t) = n_{(J+1,2J)} \mid Q_{(1,2J)}(t) = n_{(1,2J)}\}$$

$$\times P \{Q_{(1,2J)}(t) = n_{(1,J)}\}$$

$$\times P \{Q_{(J+1,2J)}(t) = n_{(J+1,2J)}\} + o (\Delta t)$$

$$= \left(1 - \sum_{k=1}^{J} \alpha_{J+k} \Delta t - \sum_{k=1}^{J} \sum_{n_1,\dots,n_J=0}^{\infty} a (n_{J+k}, n_k) \mu_{J+k} (1 - p_{J+k,J+k}) P_{n_{(1,J)}}(t) \Delta t\right)$$

$$\times P_{n_{(J+1,2J)}}(t) + o (\Delta t).$$
(37)

Event \mathcal{B} . There is a $k \in \{1, ..., J\}$ such that $Q_{J+k}(t) = n_{J+1} + 1$, $Q_{J+k}(t + \Delta t) = n_{J+1}$ and for all $l \in \{1, ..., J\} \setminus \{k\}$, $Q_{J+l}(t) = Q_{J+l}(t + \Delta t) = n_{J+l}$. Similar to the discussion in Event \mathcal{A} , we can obtain the probability for Event \mathcal{B} as follows:

$$\begin{split} \sum_{k=1}^{J} P\left\{ Q_{J+k}\left(t\right) &= n_{J+k} + 1, \ Q_{J+k}\left(t + \Delta t\right) = n_{J+k}, \\ Q_{J+l}\left(t\right) &= Q_{J+l}\left(t + \Delta t\right) = n_{J+l}, \ l \in \{1, \dots, J\} \setminus \{k\} \right\} \\ &= \sum_{k=1}^{J} P\left\{ \Delta S_{J+k}\left(t\right) = 1, \Delta A_{J+k}\left(t\right) = 0, \\ \sum_{l=1, l \neq k}^{J} \Delta S_{J+l}\left(t\right) &= \sum_{l=1, l \neq k}^{J} \Delta A_{J+l}\left(t\right) = 0, \\ Q_{J+k}\left(t\right) &= n_{J+k} + 1, Q_{J+l}\left(t\right) = n_{J+l} \\ &\quad \forall l \in \{1, \dots, J\} \setminus \{k\} \right\} + o\left(\Delta t\right) \end{split}$$

$$\begin{split} &= \sum_{k=1}^{J} P \left\{ \sum_{l=1, l \neq k}^{J} \Delta S_{J+l} \left(t \right) = \Delta A_{J+l} \left(t \right) = 0, \Delta A_{J+k} \left(t \right) = 0, \\ &\quad Q_{J+k} \left(t \right) = n_{J+k} + 1, Q_{J+l} \left(t \right) = n_{J+l}, \\ &\quad l \in \{1, \dots, J\} \setminus \{k\} \mid \Delta S_{J+k} \left(t \right) = 1 \right\} \\ &\quad \times P \left\{ \Delta S_{J+k} \left(t \right) = 1 \right\} + o \left(\Delta t \right) \\ &= \sum_{k=1}^{J} \left(\sum_{n_1, \dots, n_J = 0}^{\infty} a \left(n_{J+k} + 1, n_k \right) \mu_{J+k} \right) \\ &\quad \times \left(1 - \sum_{l=1}^{J} p_{J+k, J+l} \right) P_{n_{(1,J)}} \left(t \right) \Delta t \right) \\ &\quad \times P_{n_{J+1}, \dots, n_{J+k} + 1, \dots, n_{2J}} \left(t \right) + o \left(\Delta t \right). \end{split}$$

$$(38)$$

Event \mathscr{C} . There is a $k \in \{1, ..., J\}$ such that $Q_{J+k}(t) = n_{J+k} - 1$, $Q_{J+k}(t + \Delta t) = n_{J+k}$ and for all $l \in \{1, ..., J\} \setminus \{k\}$, $Q_{J+l}(t) = Q_{J+l}(t + \Delta t) = n_{J+l}$. Then, we can get the probability for Event \mathscr{C} as follows:

$$\begin{split} \sum_{k=1}^{J} P \left\{ Q_{J+k} \left(t \right) &= n_{J+k} - 1, Q_{J+k} \left(t + \Delta t \right) = n_{J+k}, \\ Q_{J+l} \left(t \right) &= Q_{J+l} \left(t + \Delta t \right) = n_{J+l}, l \in \{1, \dots, J\} \setminus \{k\} \right\} \\ &= \sum_{k=1}^{J} P \left\{ \Delta S_{J+k} \left(t \right) = 0, \Delta E_{J+k} \left(t \right) = 1, \\ \sum_{l=1, l \neq k}^{J} \Delta S_{J+l} \left(t \right) &= \sum_{l=1, l \neq k}^{J} \Delta A_{J+l} \left(t \right) = 0, \\ Q_{J+k} \left(t \right) &= n_{J+k} - 1, \\ Q_{J+l} \left(t \right) &= n_{J+l} \ \forall l \in \{1, \dots, J\} \setminus \{k\} \right\} + o \left(\Delta t \right) \\ &= \sum_{k=1}^{J} \left(n_{J+k} \wedge 1 \right) \alpha_{J+k} \Delta t P_{n_{J+1}, \dots, n_{J+k} - 1, \dots, n_{2J}} \left(t \right) + o \left(\Delta t \right). \end{split}$$

(39)
Event
$$\mathcal{D}$$
. There exist $r, s \in \{1, ..., J\}$ such that $Q_{J+r}(t) = n_{J+r} + 1$, $Q_{J+r}(t+\Delta t) = n_{J+s}, Q_{J+s}(t) = n_{J+s} - 1$, and $Q_{J+s}(t+\Delta t) = n_{J+s}$
and for all $l \in \{1, ..., J\} \setminus \{r, s\}, Q_{J+l}(t) = Q_{J+l}(t + \Delta t) = n_{J+l}$.
Then, we can get the probability for Event \mathcal{D} as follows:

$$\sum_{r,s=1,s\neq r}^{J} P \{ Q_{J+r}(t) = n_{J+r} + 1, Q_{J+r}(t + \Delta t) = n_{J+r};$$

$$Q_{J+s}(t) = n_{J+s} - 1, Q_{J+s}(t + \Delta t) = n_{J+s};$$

$$Q_{J+l}(t) = Q_{J+l}(t + \Delta t) = n_{J+l}$$

$$\forall l \in \{1, \dots, J\} \setminus \{r, s\} \}$$

$$= \sum_{r,s=1,s\neq r}^{J} P\left\{ \Delta S_{J+r}(t) = 1, \Delta A_{J+r}(t) = 0; \Delta S_{J+s}(t) = 0, \\ \Delta A_{J+s}(t) = 1; \\ \sum_{l=1,l\neq r,s}^{J} \Delta S_{J+l}(t) = \sum_{l=1,l\neq r,s}^{J} \Delta A_{J+l}(t) = 0; \\ Q_{J+r}(t) = n_{J+r} + 1, Q_{J+s}(t) = n_{J+s} - 1, \\ Q_{J+l}(t) = n_{J+l} \ \forall l \in \{1,\ldots,J\} \setminus \{r,s\} \right\} + o(\Delta t)$$
$$= \sum_{r,s=1,s\neq r}^{J} \left(\sum_{n_1,\ldots,n_J=0}^{\infty} a(n_{J+r} + 1, n_r) \mu_{J+r} p_{J+r,J+s} P_{n_{(1,J)}}(t) \right)$$

$$\times (n_{J+s} \wedge 1) P_{n_{J+1}, \dots, n_{J+r}+1, \dots, n_{J+s}-1, \dots, n_{2J}} (t)$$

+ $o(\Delta t)$. (40)

Now we go back to discuss (28). After transferring the $P_{n_{(j+1,2)}}(t)$ from right to left, dividing by Δt , and taking the limit as Δt tends to zero, we get the following differential equations:

Next we show that the given distribution corresponding h = 2 in Theorem 1 is a steady-state solution of those equations described by (41). It is enough to demonstrate that the derivatives in the above equations are all made zero by setting $P_{n_{(1,2)}}(t) = P_{n_{(1,2)}}$, that is, to prove that

$$\begin{split} \left(\sum_{k=1}^{J} \alpha_{J+k} + \sum_{k=1}^{J} \sum_{n_{1},\dots,n_{J}=0}^{\infty} a\left(n_{J+k}, n_{k}\right) P_{n_{(1,J)}} \mu_{J+k}\right) P_{n_{(J+1,2J)}} \\ &= \sum_{k=1}^{J} \left(\sum_{n_{1},\dots,n_{J}=0}^{\infty} a\left(n_{J+k} + 1, n_{k}\right) P_{n_{(1,J)}} \mu_{J+k} \\ &\times \left(1 - \sum_{l=1}^{J} p_{J+k,J+l}\right)\right) \right) \\ &\times P_{n_{J+1},\dots,n_{J+k}+1,\dots,n_{2J}} \\ &+ \sum_{k=1}^{J} \left(n_{J+k} \wedge 1\right) \alpha_{J+k} P_{n_{J+1},\dots,n_{J+k}-1,\dots,n_{2J}} \\ &+ \sum_{r,s=1,s\neq r}^{J} \left(\sum_{n_{1},\dots,n_{J}=0}^{\infty} a\left(n_{J+r} + 1, n_{r}\right) P_{n_{(1,J)}} \mu_{J+r} p_{J+r,J+s}\right) \\ &\times \left(n_{J+s} \wedge 1\right) P_{n_{J+1},\dots,n_{J+r}+1,\dots,n_{J+s}-1,\dots,n_{2J}} \\ &+ \sum_{k=1}^{J} \sum_{n_{1},\dots,n_{J}=0}^{\infty} a\left(n_{J+k}, n_{k}\right) P_{n_{(1,J)}} \mu_{J+k} p_{J+k,J+k} P_{n_{(J+1,2J)}}. \end{split}$$

$$(42)$$

By the definition of $\kappa(\cdot)$ in Theorem 1, that is, for each $k \in \{1, \ldots, J\}$,

$$\kappa(n_{J+k}) = \sum_{n_1,\dots,n_J=0}^{\infty} a(n_{J+k}, n_k) P_{n_{(1,J)}},$$
(43)

then, we can rewrite (42) as follows:

$$\begin{pmatrix} \sum_{k=1}^{J} \alpha_{J+k} + \sum_{k=1}^{J} \kappa (n_{J+k}) \mu_{J+k} \end{pmatrix} P_{n_{(J+1,2J)}} \\ = \sum_{k=1}^{J} \kappa (n_{J+k} + 1) \mu_{J+k} \left(1 - \sum_{l=1}^{J} p_{J+k,J+l} \right) P_{n_{J+1},\dots,n_{J+k}+1,\dots,n_{2J}} \\ + \sum_{k=1}^{J} (n_{J+k} \wedge 1) \alpha_{J+k} P_{n_{J+1},\dots,n_{J+k}-1,\dots,n_{2J}} \\ + \sum_{r,s=1,s\neq r}^{J} \kappa (n_{J+r} + 1) \mu_{J+r} p_{J+r,J+s} (n_{J+s} \wedge 1) \\ \times P_{n_{J+1},\dots,n_{J+r}+1,\dots,n_{J+s}-1,\dots,n_{2J}} \\ + \sum_{k=1}^{J} \kappa (n_{J+k}) \mu_{J+k} p_{J+k,J+k} P_{n_{(J+1,2J)}}.$$

$$(44)$$

From the given distribution in Theorem 1, we have

$$\frac{P_{n_{J+1},\dots,n_{J+k}+1,\dots,n_{2J}}}{P_{n_{(J+1,2J)}}} = \frac{\lambda_{J+k}}{\kappa (n_{J+k}+1) \mu_{J+k}},$$

$$\frac{P_{n_{J+1},\dots,n_{J+k}-1,\dots,n_{2J}}}{P_{n_{(J+1,2J)}}} = \frac{\kappa (n_{J+k}) \mu_{J+k}}{\lambda_{J+k}},$$

$$\frac{P_{n_{J+1},\dots,n_{J+r}+1,\dots,n_{J+s}-1,\dots,n_{2J}}}{P_{n_{(J+1,2J)}}} = \frac{\kappa (n_{J+s}) \lambda_{J+r} \mu_{J+s}}{\kappa (n_{J+r}+1) \lambda_{J+s} \mu_{J+r}}.$$
(45)

Then, by (44) to (45), we know that it is sufficient to show

$$\sum_{k=1}^{J} \alpha_{J+k} + \sum_{k=1}^{J} \kappa (n_{J+k}) \mu_{J+k}$$

$$= \sum_{k=1}^{J} \left(1 - \sum_{l=1}^{J} p_{J+k,J+l} \right) \lambda_{J+k} + \sum_{k=1}^{J} \frac{\kappa (n_{J+k}) \alpha_{J+k} \mu_{J+k}}{\lambda_{J+k}} \quad (46)$$

$$+ \sum_{r,s=1}^{J} \frac{\kappa (n_{J+s}) \lambda_{J+r} \mu_{J+s} p_{J+r,J+s}}{\lambda_{J+s}}.$$

Owing to the routing structure and the traffic equation (3), we have the following observations:

$$\sum_{k=1}^{J} \left(1 - \sum_{l=1}^{J} p_{J+k,J+l} \right) \lambda_{J+k} = \sum_{k=1}^{J} \alpha_{J+k},$$

$$\sum_{r,s=1}^{J} \frac{\kappa (n_{J+s}) \lambda_{J+r} \mu_{J+s} p_{J+r,J+s}}{\lambda_{J+s}}$$

$$= \sum_{s=1}^{J} \left(\frac{\kappa (n_{J+s}) \mu_{J+s}}{\lambda_{J+s}} \right) \sum_{r=1}^{J} \lambda_{J+r} p_{J+r,J+s}$$

$$= \sum_{s=1}^{J} \kappa (n_{J+s}) \mu_{J+s} - \sum_{s=1}^{J} \frac{\kappa (n_{J+s}) \mu_{J+s} \alpha_{J+s}}{\lambda_{J+s}}.$$
(47)

Then, substituting (47) into (46), we get the necessary equality.

Next, we derive the initial distribution $P_0^{s(J+k)}$ corresponding to state $n_{J+k} = 0$. By the network stability conditions (12) and $\rho_j < 1$ for each $j \in \{1, \ldots, J\}$, it follows from $\sum_{n_{J+k}=0}^{\infty} P_{n_{J+k}}^{s(J+k)} = 1$ that the following initial distribution is well posed:

$$P_0^{s(J+k)} = \frac{1}{1 + \sum_{n_{J+k}=1}^{\infty} \prod_{r_{J+k}=1}^{n_{J+k}} (\lambda_{J+k} / \kappa(r_{J+k}) \mu_{J+k})}.$$
 (48)

Hence, we complete the proof of Theorem 1 for priority types $h \in \{1, 2\}$.

Step 3. To finish the induction procedure, in this step, we first suppose that the result described in Theorem 1 is true for all classes with priority rank $h \in \{1, ..., I-1\}$; then, we show that it is true for all classes with priority rank $h \in \{1, ..., I\}$. By

the similar illustration in getting (41), we have the following differential equations:

$$\begin{split} \dot{P}_{n_{(l-1)J+1,JJ}}(t) \\ &= -\sum_{k=1}^{J} \alpha_{(l-1)J+k} P_{n_{(l-1)J,IJ}}(t) \\ &- \sum_{k=1}^{J} \sum_{n_{1},\dots,n_{(l-1)J}=0}^{\infty} a \left(n_{IJ+k}, \sum_{u=0}^{I-2} n_{uJ+k} \right) \mu_{(I-1)J+k} \\ &\times (1 - p_{(I-1)J+k,(I-1)J+k}) \\ &\times P_{n_{(1,(l-1)J)}}(t) P_{n_{((l-1)J+k,IJ)}}(t) \\ &+ \sum_{k=1}^{J} \sum_{n_{1},\dots,n_{(l-1)J}=0}^{\infty} a \left(n_{(I-1)J+k} + 1, \sum_{u=0}^{I-2} n_{uJ+k} \right) \mu_{(I-1)J+k} \\ &\times \left(1 - \sum_{l=1}^{J} p_{(I-1)J+k,(I-1)J+l} \right) P_{n_{(1,(l-1)J)}}(t) \\ &\times P_{n_{(I-1)J+1},\dots,n_{(l-1)J+k}}(t) \\ &+ \sum_{k=1}^{J} \left(n_{(I-1)J+k} \wedge 1 \right) \alpha_{(I-1)J+k} \\ &\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}}(t) \\ &+ \sum_{r,s=1,s\neq r} \sum_{n_{1},\dots,n_{(l-1)J=0}}^{\infty} a \left(n_{(I-1)J+r} + 1, \sum_{u=0}^{I-2} n_{uJ+r} \right) \\ &\times \mu_{(I-1)J+r} p_{(I-1)J+r,(I-1)J+s} \\ &\times P_{n_{(I,(1-1)J)}}(t) \left(n_{(I-1)J+s} \wedge 1 \right) \\ &\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+s}}(t). \end{split}$$
(49)

Next, we show that the given distribution corresponding the lowest priority type in Theorem 1 is a steady-state solution of those equations described by (49). It suffices to demonstrate that the derivatives in the above equations are all made zero by setting $P_{n_{((I-1)J,IJ)}}(t) = P_{n_{((I-1)J,IJ)}}$, that is, to prove that

$$\sum_{k=1}^{J} \alpha_{(I-1)J+k} P_{n_{((I-1)J,IJ)}}$$

+ $\sum_{k=1}^{J} \sum_{n_1,\dots,n_{(I-1)J}=0}^{\infty} a\left(n_{(I-1)J+k}, \sum_{u=0}^{I-2} n_{uJ+k}\right)$
 $\times P_{n_{(1,(I-1)J)}} \mu_{(I-1)J+k} P_{n_{((I-1)J+1,IJ)}}$
= $\sum_{k=1}^{J} \sum_{n_1,\dots,n_{(I-1)J}=0}^{\infty} a\left(n_{(I-1)J+k} + 1, \sum_{u=0}^{I-2} n_{uJ+k}\right) P_{n_{(1,(I-1)J)}}$

$$\times \mu_{(I-1)J+k} \left(1 - \sum_{l=1}^{J} p_{(I-1)J+k,(I-1)J+l} \right)$$

$$\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}+1,\dots,n_{IJ}}$$

$$+ \sum_{k=1}^{J} \left(n_{(I-1)J+k} \wedge 1 \right) \alpha_{(I-1)J+k} P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}-1,\dots,n_{IJ}}$$

$$+ \sum_{r,s=1,s\neq r}^{J} \sum_{n_1,\dots,n_{(I-1)J}=0}^{\infty} a \left(n_{(I-1)J+r} + 1, \sum_{u=0}^{I-2} n_{uJ+r} \right)$$

$$\times P_{n_{(1,(I-1)J)}} \mu_{(I-1)J+r}$$

$$\times p_{(I-1)J+r,(I-1)J+s} \left(n_{(I-1)J+s} \wedge 1 \right)$$

$$\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+s}+1,\dots,n_{(I-1)J+s}-1,\dots,n_{IJ}}$$

$$+ \sum_{k=1}^{J} \sum_{n_1,\dots,n_{(I-1)J}=0}^{\infty} a \left(n_{(I-1)J+k}, \sum_{u=0}^{I-2} n_{uJ+k} \right) P_{n_{(1,(I-1)J)}}$$

$$\times \mu_{(I-1)J+k} p_{(I-1)J+k,(I-1)J+k} P_{n_{((I-1)J+1,IJ)}}.$$

$$(50)$$

By the definition of $\kappa(\cdot)$ in Theorem 1, we have

$$\kappa\left(n_{(I-1)J+k}\right) = \sum_{n_1,\dots,n_{(I-1)J}=0}^{\infty} a\left(n_{(I-1)J+k}, \sum_{u=0}^{I-2} n_{uJ+k}\right) P_{n_{(1,(I-1)J)}}.$$
(51)

Then, we can rewrite (50) as follows:

$$\begin{split} \left(\sum_{k=1}^{J} \alpha_{(I-1)J+k} + \sum_{k=1}^{J} \kappa\left(n_{(I-1)J+k}\right) \mu_{(I-1)J+k}\right) P_{n_{((I-1)J+1,IJ)}} \\ &= \sum_{k=1}^{J} \kappa\left(n_{(I-1)J+k} + 1\right) \mu_{(I-1)J+k} \left(1 - \sum_{l=1}^{J} p_{(I-1)J+k,(I-1)J+l}\right) \\ &\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}+1,\dots,n_{IJ}} \\ &+ \sum_{k=1}^{J} \left(n_{(I-1)J+k} \wedge 1\right) \alpha_{(I-1)J+k} P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}-1,\dots,n_{IJ}} \\ &+ \sum_{r,s=1,s\neq r}^{J} \kappa\left(n_{(I-1)J+k} + 1\right) \\ &\times \mu_{(I-1)J+r} p_{(I-1)J+r,(I-1)J+s} \left(n_{(I-1)J+s} \wedge 1\right) \\ &\times P_{n_{(I-1)J+1},\dots,n_{(I-1)J+r}+1,\dots,n_{(I-1)J+s}-1,\dots,n_{IJ}} \\ &+ \sum_{k=1}^{J} \kappa\left(n_{(I-1)J+k}\right) \mu_{(I-1)J+k} p_{(I-1)J+k,(I-1)J+k} P_{n_{((I-1)J+1,IJ)}}. \end{split}$$
(52)

From the given distribution in Theorem 1, we have

$$\frac{P_{n_{(I-1)J+1},\dots,n_{(I-1)J+k}+1,\dots,n_{2J}}}{P_{n_{((I-1)J+1,2J)}}} = \frac{\lambda_{(I-1)J+k}}{\kappa (n_{(I-1)J+k}+1) \mu_{(I-1)J+k}},$$

$$\frac{P_{n_{(I-1)J+1,2J)}}{P_{n_{((I-1)J+1,2J)}}} = \frac{\kappa (n_{(I-1)J+k}) \mu_{(I-1)J+k}}{\lambda_{(I-1)J+k}},$$

$$\frac{P_{n_{(I-1)J+1,\dots,n_{(I-1)J+r}+1,\dots,n_{(I-1)J+s}-1,\dots,n_{2J}}}{P_{n_{((I-1)J+1,2J)}}}$$

$$= \frac{\kappa (n_{(I-1)J+s}) \lambda_{(I-1)J+r} \mu_{(I-1)J+s}}{\kappa (n_{(I-1)J+r}+1) \lambda_{(I-1)J+s} \mu_{(I-1)J+r}}.$$
(53)

From (52) to (53), it is sufficient to show

$$\sum_{k=1}^{J} \alpha_{(I-1)J+k} + \sum_{k=1}^{J} \kappa \left(n_{(I-1)J+k} \right) \mu_{(I-1)J+k}$$

$$= \sum_{k=1}^{J} \left(1 - \sum_{l=1}^{J} p_{(I-1)J+k,(I-1)J+l} \right) \lambda_{(I-1)J+k}$$

$$+ \sum_{k=1}^{J} \frac{\kappa \left(n_{(I-1)J+k} \right) \alpha_{(I-1)J+k} \mu_{(I-1)J+k}}{\lambda_{(I-1)J+k}}$$

$$+ \sum_{r,s=1}^{J} \frac{\kappa \left(n_{(I-1)J+s} \right) \lambda_{(I-1)J+r} \mu_{(I-1)J+s} p_{(I-1)J+r,(I-1)J+s}}{\lambda_{(I-1)J+s}}.$$
(54)

Owing to the routing structure and the traffic equation (3), we have the following observations:

$$\sum_{k=1}^{J} \left(1 - \sum_{l=1}^{J} p_{(I-1)J+k,(I-1)J+l} \right) \lambda_{(I-1)J+k} = \sum_{k=1}^{J} \alpha_{(I-1)J+k},$$

$$\sum_{r,s=1}^{J} \frac{\kappa \left(n_{(I-1)J+s} \right) \lambda_{(I-1)J+r} \mu_{(I-1)J+s} p_{(I-1)J+r,(I-1)J+s}}{\lambda_{(I-1)J+s}}$$

$$= \sum_{s=1}^{J} \left(\frac{\kappa \left(n_{(I-1)J+s} \right) \mu_{(I-1)J+s}}{\lambda_{(I-1)J+s}} \right) \sum_{r=1}^{J} \lambda_{(I-1)J+r} p_{(I-1)J+r,(I-1)J+s}$$

$$= \sum_{s=1}^{J} \kappa \left(n_{(I-1)J+s} \right) \mu_{(I-1)J+s}$$

$$- \sum_{s=1}^{J} \frac{\kappa \left(n_{(I-1)J+s} \right) \mu_{(I-1)J+s} \alpha_{(I-1)J+s}}{\lambda_{(I-1)J+s}}.$$
(55)

Then, substituting (55) into (54), we get the necessary equality.

Next, we derive the initial distribution $P_0^{s((I-1)J+k)}$ corresponding to state $n_{(I-1)J+k} = 0$. By the network stability conditions (12) and $\rho_j < 1$ for each $j \in \{1, ..., J\}$, it follows

from $\sum_{n_{(I-1)J+k}=0}^{\infty} P_{n_{(I-1)J+k}}^{s((I-1)J+k)} = 1$ that the following initial distribution is well posed:

$$P_0^{s((I-1)J+k)} = \left(1 + \sum_{n_{(I-1)J+k}=1}^{\infty} \prod_{r_{(I-1)J+k}=1}^{n_{(I-1)J+k}} \frac{\lambda_{(I-1)J+k}}{\kappa(r_{(I-1)J+k})\mu_{(I-1)J+k}}\right)^{-1}.$$
 (56)

Hence, we complete the proof of Theorem 1.

5.2. Proof of Proposition 2

Net I. We justify the stability condition (12) for Net I by induction in terms of the number of job types; that is, h = 1, 2, ..., I.

First, we consider the case that h = 2. Note that $c_j = 1$ for all $j \in \{1, 2, ..., J\}$ and step one in the proof of Theorem 1, we have that, for $n_{I+k} > 0$ with $k \in \{1, 2, ..., J\}$,

$$\kappa\left(n_{J+k}\right) = P_0^{s(k)} = 1 - \frac{\lambda_k}{\mu_k}.$$
(57)

Therefore, condition (12) is true for h = 2. Furthermore, we have

$$P_0^{s(J+k)} = 1 - \frac{\lambda_{J+k}}{(1 - \lambda_k/\mu_k)\,\mu_{J+k}}.$$
(58)

Second, for $h \leq I - 1$, we suppose that

$$\kappa(n_{(h-1)J+k}) = 1 - \sum_{i=0}^{h-2} \frac{\lambda_{iJ+k}}{\mu_{iJ+k}}.$$
(59)

Hence, we have

$$P_0^{s((h-1)J+k)} = 1 - \frac{\lambda_{(h-1)J+k}}{\left(1 - \sum_{i=0}^{h-2} \lambda_{iJ+k} / \mu_{iJ+k}\right) \mu_{(h-1)J+k}}.$$
 (60)

From the induction assumptions (59) and (60), we know that (12) is true for $h \le I - 1$.

Finally, we show that (12) holds for h = I. In fact, from the definition of $\kappa(n_{(I-1)J+k})$, we know that, for $n_{(I-1)J+k} > 0$,

$$\kappa\left(n_{(I-1)J+k}\right) = \prod_{i=0}^{I-2} P_0^{s(iJ+k)} = 1 - \sum_{i=0}^{I-2} \frac{\lambda_{iJ+k}}{\mu_{iJ+k}}.$$
 (61)

Then, we see that (12) is true for h = I. Hence, we complete the proof of Net I.

Net II. Note that, for each $k \in \{1, ..., J\}$ and $n_{J+k} \ge c_{s(k)}$, we have

$$\begin{split} \kappa\left(n_{J+k}\right) \\ &= \sum_{n_{1},\dots,n_{J}=0}^{\infty} a\left(n_{J+k},n_{k}\right) P_{n_{(1,J)}} \\ &= \sum_{n_{k}=0}^{c_{s(k)}-1} a\left(n_{J+k},n_{k}\right) P_{n_{k}}^{s(k)} \\ &= P_{0}^{s(k)} \left(c_{s(k)} + (c_{s(k)}-1)\left(\frac{\lambda_{k}}{\mu_{k}}\right) + \dots + \frac{(\lambda_{k}/\mu_{k})^{c_{s(k)}-1}}{(c_{s(k)}-1)!}\right) \\ &= \left(\sum_{n_{k}=0}^{c_{s(k)}-1} (c_{s(k)}-n_{k}) \frac{(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!}\right) \\ &\times \left(\sum_{n_{k}=0}^{c_{s(k)}-1} \frac{(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!} + \frac{(\lambda_{k}/\mu_{k})^{c_{s(k)}}}{c_{s(k)}!\left(1-\lambda_{k}/c_{s(k)}\mu_{k}\right)}\right)^{-1} \\ &= \left(c_{s(k)} \left(1-\frac{\lambda_{k}}{c_{s(k)}\mu_{k}}\right) \sum_{n_{k}=0}^{c_{s(k)}-1} \frac{(c_{s(k)}-n_{k})\left(\lambda_{k}/\mu_{k}\right)^{n_{k}}}{c_{s(k)}n_{k}!}\right) \\ &\times \left(1+\sum_{n_{k}=0}^{c_{s(k)}-2} \left(\frac{1}{(n_{k}+1)!} - \frac{1}{c_{s(k)}n_{k}!}\right)\left(\frac{\lambda_{k}}{\mu_{k}}\right)^{n_{k}+1}\right)^{-1} \\ &= c_{s(k)} \left(1-\frac{\lambda_{k}}{c_{s(k)}\mu_{k}}\right). \end{split}$$

$$(62)$$

Furthermore, for $0 < n_{J+k} < c_{s(k)}$, we have

$$\kappa (n_{J+k})$$

$$= \sum_{n_1,\dots,n_j=0}^{\infty} a(n_{J+k}, n_k) P_{n_{(1,J)}}$$

$$= \sum_{n_k=0}^{c_{s(k)}-1} a(n_{J+k}, n_k) P_{n_k}^{s(k)}$$

$$= P_0^{s(k)} \left(n_{J+k} \left(1 + \frac{1}{1!} \left(\frac{\lambda_k}{\mu_k} \right) + \dots + \frac{1}{(c_{s(k)} - n_{J+k})!} \left(\frac{\lambda_k}{\mu_k} \right)^{c_{s(k)} - n_{J+k} + 1} + \frac{n_{J+k} - 1}{(c_{s(k)} - n_{J+k} + 1)!} \left(\frac{\lambda_k}{\mu_k} \right)^{c_{s(k)} - n_{J+k} + 1} + \dots + \frac{1}{(c_{s(k)} - 1)!} \left(\frac{\lambda_k}{\mu_k} \right)^{c_{s(k)} - 1} \right)$$

$$= \left(n_{J+k} \sum_{n_{k}=0}^{c_{s(k)}-n_{J+k}} \frac{(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!} + \sum_{n_{k}=c_{s(k)}-n_{J+k}+1}^{c_{s(k)}-1} \frac{(c_{s(k)}-n_{k})(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!} \right) \times \left(\sum_{n_{k}=0}^{c_{s(k)}-1} \frac{(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!} + \frac{(\lambda_{k}/\mu_{k})^{c_{s(k)}}}{c_{s(k)}!(1-\lambda_{k}/c_{s(k)})\mu_{k})} \right)^{-1} = \left(n_{J+k} \left(1 - \frac{\lambda_{k}}{c_{s(k)}\mu_{k}} \right) + \sum_{n_{k}=c_{s(k)}-n_{J+k}+1}^{c_{s(k)}-1} \frac{(c_{s(k)}-n_{k})(\lambda_{k}/\mu_{k})^{n_{k}}}{n_{k}!n_{J+k}} \right) \right) \times \left(1 + \sum_{n_{k}=0}^{c_{s(k)}-1} \left(\frac{1}{(n_{k}+1)!} - \frac{1}{c_{s(k)}n_{k}!} \right) \left(\frac{\lambda_{k}}{\mu_{k}} \right)^{n_{k}+1} \right)^{-1} \\ \ge n_{J+k} \left(1 - \frac{\lambda_{k}}{c_{s(k)}\mu_{k}} \right).$$
(63)

Then, we can see that (12) is true. Hence, we complete the proof of Proposition 2.

5.3. Illustration of Example 3. For each class $k \in \{1, 2, ..., J\}$, the initial distribution $P_0^{s(k)}$ can be calculated as follows, since $\rho_{s(k)} < 1$:

$$P_0^{s(k)} = \frac{1 - \lambda_k / 2\mu_k}{1 + \lambda_k / 2\mu_k}.$$
(64)

Then, for $n_{J+k} = 1$, we have

$$\kappa(n_{J+k}) = P_0^{s(k)} \left(1 + \frac{\lambda_k}{\mu_k}\right)$$

$$= \frac{\left(1 - \lambda_k/2\mu_k\right)\left(1 + \lambda_k/\mu_k\right)}{1 + \lambda_k/2\mu_k} \ge 1 - \frac{\lambda_k}{2\mu_k}.$$
(65)

Furthermore, for $n_{J+k} \ge 2$, we have

$$\kappa\left(n_{J+k}\right) = P_0^{s(k)}\left(2 + \frac{\lambda_k}{\mu_k}\right) = 2\left(1 - \frac{\lambda_k}{2\mu_k}\right). \tag{66}$$

Therefore, since $\rho_{s(k)} < 1$, we have

1

$$P_{0}^{s(J+k)} = \left(1 + \frac{\lambda_{J+k}}{\kappa(1)\,\mu_{J+k}} + \frac{\lambda_{J+k}}{\kappa(1)\mu_{J+k}} \frac{\lambda_{J+k}/\mu_{J+k}}{(1 - \lambda_{J+k}/\kappa(2)\mu_{J+k})\kappa(2)\mu_{J+k}}\right)^{-1} \quad (67)$$
$$= \frac{(1 - \lambda_{k}/2\mu_{k} - \lambda_{J+k}/2\mu_{J+k})(1 + \lambda_{k}/\mu_{k})}{1 + \lambda_{k}/2\mu_{k} + \lambda_{J+k}/2\mu_{J+k} - (1/2)(\lambda_{k}/\mu_{k})^{2}}.$$

Finally, we can compute $\kappa(n_{2J+k})$ for $n_{2J+k} \ge 2$ as follows:

$$\begin{aligned} c(n_{2J+k}) \\ &= 2P_0^{s(k)} P_0^{s(J+k)} + P_1^{s(k)} P_0^{s(J+k)} + P_0^{s(k)} P_1^{s(J+k)} \\ &= P_0^{s(k)} P_0^{s(J+k)} \left(2 + \frac{\lambda_k}{\mu_k} + \frac{\lambda_{J+k}}{\mu_{J+k}} \right) \\ &= \left(2 \left(1 - \frac{\lambda_k}{2\mu_k} - \frac{\lambda_{J+k}}{2\mu_{J+k}} \right) \left(1 + \frac{\lambda_k}{\mu_k} \right) \left(1 - \frac{\lambda_k}{2\mu_k} \right) \\ &\times \left(1 + \frac{\lambda_k}{2\mu_k} + \frac{\lambda_{J+k}}{2\mu_{J+k}} - \frac{1}{2} \left(\frac{\lambda_k}{\mu_k} \right)^2 \right) \left(1 + \frac{\lambda_k}{2\mu_k} \right) \right)^{-1} \\ &= \left(2 \left(1 - \frac{\lambda_k}{2\mu_k} - \frac{\lambda_{J+k}}{2\mu_{J+k}} - \left(\frac{\lambda_k}{2\mu_k} \right)^2 - \frac{1}{4} \left(\frac{\lambda_k}{\mu_k} \right)^3 \right) \\ &+ 2 \left(1 - \frac{\lambda_k}{2\mu_k} \right) \frac{\lambda_k \lambda_{J+k}}{4\mu_k \mu_{J+k}} \right) \\ &\times \left(1 + \frac{\lambda_k}{\mu_k} + \frac{\lambda_{J+k}}{2\mu_{J+k}} - \left(\frac{\lambda_k}{2\mu_k} \right)^2 - \frac{1}{4} \left(\frac{\lambda_k}{\mu_k} \right)^3 \right) \\ &+ 2 \left(1 - \frac{\lambda_k}{2\mu_k} \right) \frac{\lambda_k \lambda_{J+k}}{4\mu_k \mu_{J+k}} \right) \end{aligned}$$
(68)

Thus, if $\lambda_k / \mu_k \le 1$, from the condition $\rho_{s(k)} < 1$, we see that

$$\frac{\lambda_{2J+k}}{\kappa\left(n_{2J+k}\right)\mu_{2J+k}} \leq \frac{\lambda_{2J+k}}{2\left(1-\lambda_k/2\mu_k-\lambda_{J+k}/2\mu_{J+k}\right)\mu_{2J+k}} < 1;$$
(69)

that is, the condition (12) holds.

6. Conclusion

The research conducted in the paper is to iteratively derive the product-form solutions of stationary distributions for a 15

particular type of preemptive priority multiclass queueing networks with multiserver stations. This type of queueing systems can typically be used to model the stochastic dynamics of some large scale backbone networks with multiprocessor shared-memory switches or local (edge) cloud computing centers with parallel-server pools. The queueing networks are Markovian with exponential interarrival and service time distributions. The obtained iterative solutions can be used to conduct performance analysis or as comparison criteria for approximation and simulation studies. Numerical comparisons with existing Brownian approximating model related to general interarrival and service times are provided to show the effectiveness of our current designed algorithm and our former derived BAM. Furthermore, based on the iterative solutions, we also give some analysis concerning network stability for some cases of these queueing systems, which provides some insight for more general study.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Partial results, graphs, and tables are briefly summarized and reported in MICAI'06. This enhanced version with complete proofs of results and latest discussions is a journal version of the short conference report. This research is supported by National Natural Science Foundation of China with Grants nos. 10371053, 10971249, and 11371010.

References

- S. J. Mullender and T. L. Wood, "Cloud computing: new opportunities for telecom providers," *Bell Labs Technical Journal*, vol. 17, no. 2, pp. 1–4, 2012.
- [2] P. Mell and T. Grance, "The NIST definition of cloud computing," NIST Special Publication 800-145, National Institute of Standards and Technology, U.S. Department of Commerce, Gaithersburg, Md, USA, 2011.
- [3] J. Rhoton and R. Haukioja, *Cloud Computing Architected:* Solution Design Handbook, Recursive Limited, Wellington, New Zealand, 2013.
- [4] F. A. Tobagi, "Fast packet switch architectures for broadband integrated services digital networks," *Proceedings of the IEEE*, vol. 78, no. 1, pp. 133–167, 1990.
- [5] M. H. L. Rao and R. Tripathi, "FPGA implementation of reconfigurable switch architecture for next generation communication networks," *International Journal of Engineering and Technology*, vol. 4, no. 6, pp. 770–773, 2012.
- [6] W. Dai, "Optimal rate scheduling via utility-maximization for *J*-user MIMO Markov fading wireless channels with cooperation," *Operations Research*, vol. 61, no. 6, pp. 1450–1462, 2013.
- [7] M. Schwarz, Broadband Integrated Networks, Prentice Hall, Upper Saddle River, NJ, USA, 1996.
- [8] I. Nikolaidis and I. F. Akyildiz, "An overview of source characterization in ATM networks," in *Modeling and Simulation* of Computer and Communications Networks: Techniques, Tools,

and Tutorials, J. Walrand, K. Bachi, and G. Zobrist, Eds., pp. 123–150, Gordon and Breach Publishing, 1997.

- [9] W. Dai, "On the conflict of truncated random variable vs. heavy-tail and long range dependence in computer and network simulation," *Journal of Computational Information Systems*, vol. 7, no. 5, pp. 1488–1499, 2011.
- [10] W. Dai, "Heavy traffic limit theorems for a queue with Poisson ON/OFF long-range dependent sources and general service time distribution," *Acta Mathematicae Applicatae Sinica*, vol. 28, no. 4, pp. 807–822, 2012.
- [11] J. Cao, W. S. Cleveland, D. Lin, and D. X. Sun, "Internet traffic tends to poisson and independent as the load increases," Bell Labs Technical Report, 2001.
- [12] W. Dai, Brownian approximations for queueing networks with finite buffers: modeling, heavy traffic analysis and numerical implementations [Ph.D. thesis], Georgia Institute of Technology, Atlanta, Ga, USA, 1996.
- [13] W. Dai, "Diffusion approximations for multiclass queueing networks under preemptive priority service discipline," *Applied Mathematics and Mechanics*, vol. 28, no. 10, pp. 1331–1342, 2007.
- [14] J. G. Dai and W. Dai, "A heavy traffic limit theorem for a class of open queueing networks with finite buffers," *Queueing Systems*, vol. 32, pp. 5–40, 1999.
- [15] X. Shen, H. Chen, J. G. Dai, and W. Dai, "The finite element method for computing the stationary distribution of an SRBM in a hypercube with applications to finite buffer queueing networks," *Queueing Systems*, vol. 42, no. 1, pp. 33–62, 2002.
- [16] R. Serfozo, Introduction to Stochastic Networks, Springer, New York, NY, USA, 1999.
- [17] F. Baskett, K. M. Chandy, R. R. Muntz, and F. G. Palacios, "Open, closed and mixed networks of queues with different classes of customers," *Journal of the Association for Computing Machinery*, vol. 22, pp. 248–260, 1975.
- [18] P. G. Harrison, "Reversed processes, product forms and a nonproduct form," *Linear Algebra and Its Applications*, vol. 386, pp. 359–381, 2004.
- [19] P. G. Harrison, "Process algebraic non-product-forms," *Electronic Notes in Theoretical Computer Science*, vol. 151, no. 3, pp. 61–76, 2006.
- [20] W. Dai, "Product-form solutions for integrated services packet networks," in *Proceedings of the 5th Mexican International Conference on Artificial Intelligence (MICAI '06)*, pp. 296–308, IEEE, November 2006.
- [21] H. Chen and D. Yao, *Fundamentals of Queueing Networks*, Springer, New York, NY, USA, 2001.
- [22] J. G. Dai, D. H. Yeh, and C. Zhou, "The QNET method for re-entrant queueing networks with priority disciplines," *Operations Research*, vol. 45, no. 4, pp. 610–623, 1997.
- [23] J. M. Harrison, Brownian Motion and Stochastic Flow Systems, Wiley, New York, NY, USA, 1985.
- [24] J. R. Jackson, "Networks of waiting lines," *Operations Research*, vol. 5, pp. 518–521, 1957.



Advances in **Operations Research**



The Scientific World Journal







Hindawi

Submit your manuscripts at http://www.hindawi.com



Algebra



Journal of Probability and Statistics



International Journal of Differential Equations





Complex Analysis





Mathematical Problems in Engineering



Abstract and Applied Analysis



Discrete Dynamics in Nature and Society



International Journal of Mathematics and Mathematical Sciences





Journal of **Function Spaces**



International Journal of Stochastic Analysis

