Exact analysis of queueing networks with multiple job classes and blocking-after-service *

Ian F. Akyildiz and Chueng-Chiu Huang

School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received 21 November 1991; revised 2 November 1992

Two models of closed queueing networks with blocking-after-service and multiple job classes are analyzed. The first model is a network with N stations and each station has either type II or type III. The second model is a star-like queueing network, also called a central server model, in which the stations may have either type I or type IV, with the condition that the neighbors of these stations must be of type II or type III such that blocking will be caused only by this set of station types. Exact product form solutions are obtained for the equilibrium state probabilities in both models. Formulae for performance measures such as throughput and the mean number of jobs are also derived.

Keywords: Performance evaluation; queueing networks; finite buffer capacity; blocking; deadlock; equilibrium state probabilities; throughput; mean number of jobs.

1. Introduction

In recent years there has been an increased interest in the analysis of queueing networks with blocking. This is due to the realization that these queueing networks are useful in modelling computer systems, communication networks, and flexible manufacturing systems. The set of rules that dictate when a station becomes blocked and when it becomes unblocked is commonly referred to as the *blocking* mechanism. There are basically only a few blocking mechanisms that have been extensively studied in the literature, Akyildiz and Perros [3] and Onvural [13]. We consider the so-called *blocking-after-service* [3,13] (in short form, BAS) mechanism, i.e., when a job finishes service at a station and wants to enter a station which is full, it stays in the server of the source station, waiting for a space to be available in the destination station. This blocking policy is also known as type 1 blocking, transfer blocking, manufacturing blocking, production blocking and non-immediate blocking in the literature [13]. Several papers consider this blocking policy, e.g., Akyildiz [1,2], Akyildiz and von Brand [4], Balsamo et al. [6], Balsamo and Donatiello [7], Bocharov [9] and Onvural [12]. Akyildiz and von Brand [4] consider

* This work was supported by the National Science Foundation (NSF) under Grant No. CCR-90-11981.

queueing networks with BAS mechanism. In the first model they analyze a two-station model with different types of stations and a single class of jobs. They derive an exact product form solution for the equilibrium state probabilities. Akyildiz and von Brand [4] also analyze a second model with N stations, multiple classes of jobs and a limited number of jobs $K = \min_{i \in N}\{B_i\} + 1$, where B_i is the buffer capacity of station i. They obtain an exact product form solution for the models where the stations that will cause blocking are type I stations. Akyildiz and Liebeherr [5] determine necessary and sufficient conditions for deadlock-freedom in multi-job class queueing networks with BAS-mechanism.

In this paper we consider two models of closed queueing networks with BAS mechanism. In both models we assume that there are N stations and R classes of jobs. The first model is a general topology network with N stations where each station is either of type II or type III, Baskett et al. [8], Kelly [10,11]. The second model is a star-like queueing network (also known as a central server model) where the central server must be of type I or type IV and the neighbors of the central server must be either of type II or of type III. In the second model we assume that the blocking will be caused only by the central server. For both models we obtain exact product form solutions for equilibrium state probabilities. We also derive formulae for performance measures such as throughput and the mean number of jobs.

The paper is organized as follows: In section 2 we introduce common notations and definitions for both models. In section 3 we analyze the first model. In section 4 we investigate the central server model. In section 5 we conclude the paper.

2. Basic model description

Both models contain N stations, R job classes and K total number of jobs. The types of stations we consider in this paper are from the BCMP, Baskett et al. [8]:

- Type I: The service discipline is first-come-first-served (FCFS); all job classes have the same service time distribution and the service rate can be state-dependent where $\mu(k)$ will denote the service rate with k jobs.
- Type II: There is a single server and the service discipline is processor sharing (PS). Each job class may have a distinct service time distribution.
- Type III: The number of servers is greater than or equal to the maximum number of jobs which can be queued at the station; infinite servers (IS). Each job class may have a distinct service time distribution.
- Type IV: There is a single server and the queueing discipline is last-come-first-served preemptive-resume (LCFS-PR). Each job may have a distinct service time distribution.

Here we assume that all service times follow exponential distributions which depend on the station and on the class. Note that our results can easily be extended to general service time distributions for types II, III and IV. For the sake of simplicity here we give results only for exponential cases.

The transition probabilities are denoted by $p_{ir,js}$, $1 \leqslant i, j \leqslant N$ and $1 \leqslant r, s \leqslant R$, where a class r job departs from station i and visits station j and becomes a class s job. We also assume that the routing matrix $\mathbf{P} = [p_{ir,js}]$, is irreducible. We define

$$\alpha_{js} = \sum_{i=1}^{N} \sum_{r=1}^{R} \alpha_{ir} p_{ir,js}. \tag{1}$$

Let the vector $n = (x_1, \ldots, x_N)$ denote the number of jobs in N stations and their positions in each station, i.e., $x_i = (x_{i1}, \ldots, x_{in_i})$, where x_{il} denotes the class of the job in station i at the position l. Let n_i be the number of jobs in station i, and $n_i = n_{i1} + \ldots + n_{iR}$, where n_{ir} (for $r = 1, \ldots, R$) is the number of class r jobs in station i. We note that if station i is of type II or III, then there is no requirement for the order of the jobs; thus, we denote x_i as n_i, $n_i = (n_{i1}, \ldots, n_{iR})$.

A product form solution for equilibrium state probabilities exists when there is no blocking in the network, Baskett et al. [8], Kelly [10,11]:

$$p(n) = C \prod_{i=1}^{N} f_i(x_i), \tag{2}$$

where C is the normalization constant such that the sum of equilibrium state probabilities will be equal to one and $f_i(x_i)$ is defined by the type of station i,

$$f_i(x_i) = \begin{cases} \left(\frac{1}{\mu_i}\right) \prod_{l=1}^{n_i} \alpha_{ix_{il}} & \text{if station } i \text{ is type I}, \\ n_i! \prod_{r=1}^{R} \frac{1}{n_{ir}!} \left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_{ir}} & \text{if station } i \text{ is type II}, \\ \prod_{r=1}^{R} \frac{1}{n_{ir}} \left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_{ir}} & \text{if station } i \text{ is type III}, \\ \prod_{l=1}^{n_i} \left(\frac{\alpha_{ix_{il}}}{\mu_{ix_{il}}}\right) & \text{if station } i \text{ is type IV}. \end{cases} \tag{3}$$

We define T_1 as the set of type I or IV stations, and T_2 as the set of type II or III stations. Since we are considering the models with blocking, we need to define B_i as the capacity of station i (i.e., buffer and server capacity). In addition, to avoid the self-loops that cause *deadlock*, we assume that

$$p_{ii} = 0 \quad \text{for } i = 1, \ldots, N.$$

For convenience we define the following quantities for each station type. For type II and III stations, the service rate is dependent on the number of jobs in the station and on their classes but is independent of their order. Thus, we let $n_i = (n_{i1}, \ldots, n_{iR})$ for $i \in T_2$, where n_{ir} denotes the number of class r active jobs in station i. A job is said to be active if it is not blocked. Also we let $n_i = n_{i1} + \ldots + n_{iR}$ be the total number of active jobs in station i for $i \in T_2$.

Let b denote the number of blocked jobs. Thus, with the active jobs we have

$$K = \sum_{i=1}^{N} n_i + b, \tag{4}$$

with

$$0 \leqslant b \leqslant K - \min_{1 \leqslant i \leqslant N} \{B_i\}. \tag{5}$$

We assume that a job will choose its destination station and the class when it finishes its service. Considering the b blocked jobs we need to know their locations and their classes. We also assume that the blocked jobs will join that station which caused blocking according to *first-come-first-in* (FCFI) scheduling discipline. Thus, we let $y = (y_i)$ and $y_i = (y_{i1}, \ldots, y_{ib_i})$ if $b_i \geqslant 1$, otherwise $y_i = \emptyset$, where b_i denotes the number of jobs blocked by station i, and y_{il}, $1 \leqslant y_{il} \leqslant N$ denotes the location of the lth job blocked by station i. Similarly, $z = (z_i)$ and $z_i = (z_{i1}, \ldots, z_{ib_i})$, where z_{il}, $1 \leqslant z_{il} \leqslant R$, denotes the class of the lth blocked job. To have the complete information about the system we define the state space $S = \{(n, x, y, z)\}$, where $x = (x_i, i \in T_1)$ and $n = (n_i, i \in T_2)$. For a given state (n, x, y, z) the service rate of station i for $i \in T_2$ is $\mu_i(n_i) = \mu_{i1}(n_{i1}) + \ldots + \mu_{iR}(n_{iR})$, where $\mu_{ir}(n_{ir}) = n_{ir}\mu_{ir}$ if station i is of type III; if station i is of type II, then $\mu_{ir}(n_{ir}) = (n_{ir}/n_i)\mu_{ir}$. For $i \in T_1$ then the service rate is: if station i is of type I, then $\mu_i(x_i) = \mu_i$ for $n_i > 0$, otherwise it is equal to 0. If station i is of type IV, then $\mu_i(x_i) = \mu_{in_i}$ for $n_i \leqslant B_i$.

Let e_{ir} denote an $N \times R$ matrix with all elements zero except the element at the position of ith row and rth column with value 1.

3. General topology network with type II and III stations

In this section we consider a general topology network with only type II and III stations. We assume the following constraint in the model:

$$K < \min_{1 \leqslant i,j \leqslant N; p_{ij} > 0} \{B_i + B_j\}, \tag{6}$$

This constraint says that only one of the N stations may cause blocking at any time. The reason for making this assumption is to avoid more than two jobs moving at one event. Thus, the state space is reduced to $S = (n, y, z)$, where y and z are b by 1 vectors, because if there are any blocked jobs, then they are all blocked by the same station.

Now we consider the transitions of the jobs for a nonblocking state (n, y, z), where y and z are empty vectors $\emptyset$. In this case the system will reach the state $(n - e_{ir} + e_{js}, \emptyset, \emptyset)$ with rate $\mu_{ir}(n_{ir})p_{ir,js}$, if $n_j \neq B_j$. Otherwise, a class r job from station i will be blocked by station j and the transition will occur from the state $(n, \emptyset, \emptyset)$ to the state $(n - e_{ir}, y', , z')$, where $y' = (y_1)$ and $y_1 = i$. The same occurs for $z' = (z_1)$ and $z_1 = s$. The rate to reach this state is the same as before, i.e., $\mu_{ir}(n_{ir})p_{ir,js}$.

Now we consider the transitions from a blocking state. Suppose station k is full, i.e., $n_k = B_k$, and the current state is $(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z})$, for $b \geq 0$. Then a class r job leaves station i and moves to station j, for $i \neq k$, $j \neq k$, and becomes a class s job and the system reaches the state $(\boldsymbol{n} - \boldsymbol{e}_{ir} + \boldsymbol{e}_{js}, \boldsymbol{y}, \boldsymbol{z})$ with rate $\mu_{ir}(n_{ir})p_{ir,js}$. If a class r job moves from station i to station k and becomes a class s job, then the state $(\boldsymbol{n} - \boldsymbol{e}_{ir}, \boldsymbol{y} + y_{b+1}, \boldsymbol{z} + z_{b+1})$ and $y_{b+1} = i, z_{b+1} = s$ will be reached. Note that in this case the transition rate is $\mu_{ir}(n_{ir})p_{ir,js}$.

Let $\boldsymbol{y} = (y_1, \ldots, y_b)$, then the operator $\boldsymbol{y} + y_{b+1}$ will be $(y_1, \ldots, y_{b+1})$. Similar is also valid for the operator $\boldsymbol{z} + z_{b+1}$. When considering a class r job which moves from station k to station j and becomes a class s job with the transition rate $\mu_{kr}(n_{kr})p_{kr,js}$, then one of the first blocked jobs will move to station k at the same time. Thus, the new state is $(\boldsymbol{n} - \boldsymbol{e}_{kr} + \boldsymbol{e}_{js} + \boldsymbol{e}_{kz}, \boldsymbol{y}', \boldsymbol{z}')$, with $\boldsymbol{y}' = (y'_1, \ldots, y'_{b-1})$, $y'_l = y_{l+1}$, and $\boldsymbol{z} = (z'_1, \ldots z'_{b-1})$, $z'_l = z_{l+1}$, for $1 \leqslant l \leqslant b-1$.

From these state transitions we obtain the global balance equation for a state $(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z}) \in S$ if $n_i < B_i$ for $1 \leqslant i \leqslant N$, thus $\boldsymbol{y} = \emptyset$ and $\boldsymbol{z} = \emptyset$:

$$p(\boldsymbol{n}, \emptyset, \emptyset) \sum_{i=1}^{N} \sum_{r=1}^{R} \mu_{ir}(n_{ir}) = \sum_{i=1}^{N} \sum_{r=1}^{R} \sum_{j=1}^{N} \sum_{s=1}^{R} p(\boldsymbol{n} + \boldsymbol{e}_{ir} - \boldsymbol{e}_{js}, \emptyset, \emptyset) \mu_{ir}(n_{ir} + 1) p_{ir,js}, \tag{7}$$

and the global balance equation for a blocking state (suppose $n_k = B_k$):

$$\begin{aligned} p(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z}) \sum_{i=1}^{N} \sum_{r=1}^{R} \mu_{ir}(n_{ir}) &= \sum_{i=1, i \neq k}^{N} \sum_{r=1}^{R} \sum_{j=1, j \neq k}^{N} \sum_{s=1}^{R} p(\boldsymbol{n} + \boldsymbol{e}_{ir} - \boldsymbol{e}_{js}, \boldsymbol{y}, \boldsymbol{z}) \mu_{ir}(n_{ir} + 1) p_{ir,js} \\ &+ \sum_{j \neq k} \sum_{s=1}^{R} \sum_{r=1}^{R} \sum_{y_0=1}^{N} \sum_{z_0=1}^{R} p(\boldsymbol{n} + \boldsymbol{e}_{kr} - \boldsymbol{e}_{js} - \boldsymbol{e}_{kz_0}, \boldsymbol{y} + y_0, \boldsymbol{z} + z_0) \mu_{kr}(n_{kr}) p_{kr,js} \\ &+ \sum_{s=1}^{R} p(\boldsymbol{n} + \boldsymbol{e}_{y_b s}, \boldsymbol{y} - y_b, \boldsymbol{z} - z_b) \mu_{y_b s}(n_{y_b s} + 1) p_{y_b s, k z_b}. \end{aligned} \tag{8}$$

The first term on the right-hand side of (8) is the total flow-in rate to the state $(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z})$ from the state $(\boldsymbol{n} + \boldsymbol{e}_{ir} - \boldsymbol{e}_{js}, \boldsymbol{y}, \boldsymbol{z})$. We note that the moving job is from station i to station j, $i \neq k$, $j \neq k$. The second term on the right-hand side of (8) is the flow-in rate to the state $(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z})$ by a class r job moving out from station k, and in the mean time another blocked job, which is of class z_0, will move into station k from station y_0. As a consequence a transition occurs from the state $(\boldsymbol{n} + \boldsymbol{e}_{kr} - \boldsymbol{e}_{js} - \boldsymbol{e}_{kz_0}, \boldsymbol{y} + y_0, \boldsymbol{z} + z_0)$ to the state $(\boldsymbol{n}, \boldsymbol{y}, \boldsymbol{z})$.

For notational convenience we let y_0, z_0 denote the location and the class of the blocked job that will move into station k while a job leaves station k and joins station i. The third term in (8) is the flow-in rate to the state (n, y, z) by a joining job at station k.

THEOREM 1

The model has the following product form solution for equilibrium state probabilities:

$$p(n,y,z) = C\prod_{i=1}^{N} f_i(n_i)\left(\prod_{l=1}^{b}\frac{\sum_{t=1}^{R}\alpha_{y_lt}p_{y_lt,iz_l}}{\sum_{t=1}^{R}\mu_{it}(n_{it})}\right)^{I_{\{n_i=B_i\}}}, \tag{9}$$

where $f_i(n_i)$ is defined as

$$f_i(n_i) = \begin{cases} n_i!\prod_{r=1}^{R}\frac{1}{n_{ir}!}\left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_{ir}} & \text{if station } i \text{ is type II,} \\ \prod_{r=1}^{R}\frac{1}{n_{ir}!}\left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_{ir}} & \text{if station } i \text{ is type III,} \end{cases} \tag{10}$$

and C is a normalization constant,

$$C^{-1} = \sum_{(n,y,z)\in S}\prod_{i=1}^{N} f_i(n_i)\left(\prod_{l=1}^{b}\frac{\sum_{t=1}^{R}\alpha_{y_lt,iz_l}}{\sum_{t=1}^{R}\mu_{it}(n_{it})}\right)^{I_{\{n_i=B_i\}}}. \tag{11}$$

Proof

We omit the proof of the nonblocking states since it is exactly the same as in the BCMP or Kelly model, Baskett et al. [8] and Kelly [10,11]. For the global balance equation of blocking states we assume that the state k has $n_k = B_k$ and the number of blocked jobs is b, then for (9) and (10) we have

$$p(n+e_{ir}-e_{js},y,z) = p(n,y,z)\frac{\alpha_{ir}}{\mu_{ir}(n_{ir}+1)}\frac{\mu_{js}(n_{js})}{\alpha_{js}} \tag{12}$$

if $i \neq k$ and $j \neq k$.

Next we consider the departure event of a job leaving station $j \neq k$ and visiting station k

$$\begin{aligned} &p(n+e_{kr}-e_{js}-e_{kz_0},y+y_0,z+z_0) \\ &= p(n,y,z)\frac{\alpha_{kr}}{\mu_{kr}(n_{kr}+1)}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\frac{\mu_{kz_0}(n_{kz_0})}{\alpha_{kz_0}}\frac{\sum_{t=1}^{R}\alpha_{y_0t}P_{y_0t,kz_0}}{\sum_{t=1}^{R}\mu_{kt}(n_{kt})} \end{aligned} \tag{13}$$

and for the departure event of a job leaving station k and visiting station i

$$p(n+e_{y_bs},y-y_b,z-z_b) = p(n,y,z)\frac{\alpha_{y_bs}}{\mu_{y_bs}(n_{y_bs}+1)}\frac{\sum_{t=1}^{R}\mu_{kt}(n_{kt})}{\sum_{t=1}^{R}\alpha_{y_bt}P_{y_bt,kz_b}}. \tag{14}$$

By substituting (12)–(14) on the right-hand side of (8) separately and using (1), we obtain

$$\sum_{i=1,i\neq k}^{N}\sum_{r=1}^{R}\sum_{j=1,j\neq k}^{N}\sum_{s=1}^{R} p(n+e_{ir}-e_{js},y,z)\mu_{ir}(n_{ir}+1)p_{ir,js} = p(n,y,z)\sum_{j=1,j\neq k}^{N}\sum_{s=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\left[\sum_{i=1,i\neq k}^{N}\sum_{r=1}^{R}\alpha_{ir}p_{ir,js}\right], \tag{15}$$

$$\begin{aligned}
&\sum_{j\neq k}\sum_{s=1}^{R}\sum_{y_0=1}^{N}\sum_{z_0=1}^{R} p(n+e_{kr}-e_{js}-e_{kz_0},y+y_0,z+z_0)\mu_{kr}(n_{kr})p_{kr,js}\\
&= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\sum_{r=1}^{R}\sum_{z_0=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\frac{\alpha_{kr}}{\alpha_{kz_0}}\frac{\mu_{kz_0}(n_{kz_0})}{\sum_{t=1}^{R}\mu_{kt}(n_{kt})}p_{kr,js}\left[\sum_{y_0=1}^{N}\sum_{t=1}^{R}\alpha_{y_0t}p_{y_0t,kz_0}\right]\\
&= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\sum_{r=1}^{R}\left[\alpha_{kr}p_{kr,js}\sum_{z_0=1}^{R}\left(\frac{\mu_{kz_0}(n_{kz_0})}{\sum_{t=1}^{R}\mu_{kt}(n_{kt})}\right)\right]\\
&= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\left[\sum_{r=1}^{R}\alpha_{kr}p_{kr,js}\right].
\end{aligned}\tag{16}$$

By combining (15) and (16) we have

$$\begin{aligned}
(15)+(16) &= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\left[\sum_{i=1,i\neq k}^{N}\sum_{r=1}^{R}\alpha_{ir}p_{ir,js}+\sum_{r=1}^{R}\alpha_{kr}p_{kr,js}\right]\\
&= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\frac{\mu_{js}(n_{js})}{\alpha_{js}}\left[\sum_{i=1}^{N}\sum_{r=1}^{R}\alpha_{ir}p_{ir,js}\right]\\
&= p(n,y,z)\sum_{j\neq k}\sum_{s=1}^{R}\mu_{js}(n_{js}).
\end{aligned}\tag{17}$$

In (14) we substitute the third term of (8) and obtain

$$\begin{aligned}
&\sum_{s=1}^{R} p(n+e_{y_bs},y-y_b,z-z_b)\mu_{y_bs}(n_{y_bs}+1)p_{y_bs,kz_b}\\
&= p(n,y,z)\frac{\sum_{t=1}^{R}\mu_{kt}(n_{kt})}{\sum_{t=1}^{R}\alpha_{y_bt}p_{y_bt,kz_b}}\sum_{s=1}^{R}\alpha_{y_bs}p_{y_bs,kz_b}\\
&= p(n,y,z)\sum_{t=1}^{R}\mu_{kt}(n_{kt}).
\end{aligned}\tag{18}$$

Thus, with (17) and (18) we have $p(n,y)\sum_{j=1}^{N}\sum_{s=1}^{R}\mu_{js}(n_{js})$ which is equal to the right-hand side of (8).

Remark

The result (9) follows from the following argument: we consider the buffers occupied by the blocked jobs which are stored in type II or III stations only, as the extended buffers of the stations causing blocking. The existence of blocked jobs in type II or type III stations will not effect the service of the non-blocking jobs in these stations. That is the throughput rate of these two types of stations will be dependent on the number of nonblocked jobs in this station only. Thus, we can ignore the locations of the blocked jobs and the result (9) will follow. □

Now we derive the throughput of the system $\lambda = \sum_{i=1}^{N} \lambda_i$, where λ_i is the throughput of station i, and the mean number of jobs $\bar{k}_i$ in station i.

THEOREM 2

The *throughput* of station i is computed from

$$\lambda_i = \sum_{(n,y,z)\in S} p(n,y,z)\left[\sum_{r=1}^{R} \mu_{ir}(n_{ir})\left(1 - \sum_{k=1}^{N}\sum_{s=1}^{R} p_{ir,ks} I_{\{n_k=B_k\}}\right) + \sum_{k=1}^{N} I_{\{n_k=B_k\}} I_{\{y_1=i\}}\left(\sum_{r=1}^{R} \mu_{kr}(n_{kr})\right)\right]. \tag{19}$$

The *mean number of jobs* at station i is

$$\bar{k}_i = \sum_{(n,y,z)\in S} p(n,y,z)\left(\sum_{r=1}^{R} n_{ir} + \sum_{l=1}^{b} I_{\{y_l=i\}}\right). \tag{20}$$

Proof

For any given state (n, y, z) we consider the throughput of station i. Suppose the number of class r active jobs is n_{ir}, thus, the service rate is $\mu_{ir}(n_{ir})$. However, a job which completed its service in station i can leave station i only by choosing a nonfull station j, i.e., $n_j < B_j$ with probability $\sum_{s=1}^{R} p_{ir,js}$. When considering the blocked jobs, if the first blocked job in station i is blocked by station k, i.e., $y_1 = i$, then this job has the same rate with which it will leave station i and join the station k as the service rate of station k, i.e., $\sum_{r=1}^{R} \mu_{kr}(n_{kr})$. Thus, (19) follows.

Similarly, for a given state (n, y, z) we have the number of jobs in station i which will be equal to the active jobs $\sum_{r=1}^{R} n_{ir}$ plus the blocked jobs in this station, i.e., $\sum_{l=1}^{b} I_{\{y_l=i\}}$, so that (20) follows. □

4. Central server model with different types of stations

In this section we analyze central server models (star networks) with BAS mechanism. We put the constraint that the central server must be a station of type either I or IV and all other stations may be of type II or III. We further assume that only stations of type I or IV will cause blocking, and the neighbors of such stations must be of type II or III. The total number of jobs must satisfy the constraint

$$\min_{i \in T_1}\{B_i\} < K < \min_{i \in T_2}\{B_i\}$$

and $p_{ir,js} = 0$, if $i, j \in T_1$.

We note that $p_{ir,js} = 0$ if $i, j \in T_1$ and for all r, s, then (1) will become

$$\alpha_{js} = \sum_{i \in T_2} \sum_{r=1}^{R} \alpha_{ir} p_{ir,js} \quad \text{for } j \in T_1. \tag{21}$$

We denote T_A as the set of stations in T_1 if $n_i < B_i$, and T_B as the set of stations in T_1 with $n_i = B_i$. Note that $T_A \cup T_B = T_1$. In the following we give the global balance equation for a given state (n, x, y, z).

$$\begin{aligned}
&p(n,x,y,z)\left(\sum_{i \in T_1} \mu_i(x_i) + \sum_{i \in T_2} \sum_{r=1}^{R} \mu_{ir}(n_{ir})\right) \\
&= \sum_{j \in T_A} \sum_{i \in T_2} \sum_{r=1}^{R} p(n + e_{ir}, x - x_{jn_j}, y, z) \mu_{ir}(n_{ir}+1) p_{ir, jx_{jn_j}} \\
&+ \sum_{j \in T_B} \sum_{r=1}^{R} p(n + e_{y_{jb_j}r}, x, y - y_{jb_j}, z - z_{jb_j}) \mu_{y_{jb_j}r}(n_{y_{jb_j}r}+1) p_{y_{jb_j}r, jz_{jb_j}} \\
&+ \sum_{i \in T_2, j \in T_2} \sum_{s=1}^{R} \sum_{r=1}^{R} p(n + e_{ir} - e_{js}, x, y, z) \mu_{ir}(n_{ir}+1) p_{ir,js} \\
&+ \sum_{i \in T_A, j \in T_2} \sum_{s=1}^{R} \sum_{x_{i0}=1}^{R} p(n - e_{js}, x + x_{i0}, y, z) \mu_i(x_i + x_{i0}) p_{ix_{i0},js} \\
&+ \sum_{i \in T_B, j \in T_2} \sum_{s=1}^{R} \sum_{y_{i0}=1}^{N} \sum_{x_{i0}=1}^{R} p(n - e_{js}, x + x_{i0} - x_{iB_i}, y + y_{i0}, z + z_{i0}) \\
&\times \mu_i(x_i + x_{i0} - x_{iB_i}) p_{ix_{i0}\,js}.
\end{aligned} \tag{22}$$

We consider a given state (n, x, y, z), then the total flow-out rate from this state is $p(n,x,y,z)(\sum_{i \in T_1} \mu_i(x_i) + \sum_{i \in T_2} \sum_{r=1}^{R} \mu_{ir}(n_{ir}))$. In order to check the flow-in rates we consider the different events caused by these jobs: The first term on the right-hand side of (22) denotes the departure event of a class r job from station $i \in T_2$ to station $j \in T_A$ to class x_{jn_j}. The second term denotes the events of a class r job in station $y_{jb_j} \in T_2$ which completed its service and is joining station $j \in T_B$. However, since station j is full, $n_j = B_j$, this job is blocked and becomes a class z_{jb_j} job and joins the list of blocked jobs waiting for station j. The third term denotes the event of moving of a class r job from station $i \in T_2$ to station $j \in T_2$ and becoming a class s job. The last two terms describing the events of moving of a class x_{i0} job from station $i \in T_A, i \in T_B$ to station $j \in T_2$ and becoming a class s job. We note that x_{i0} denotes the class in which the job leaves station i. i_{i0} and z_{i0} denote the station and the class of the blocked job at the first position of the list of blocked jobs which will move to station i. Note that $z_{i0} = x_{iB_i}$.

THEOREM 3

The model has the following product form solution for equilibrium state probabilities:

$$p(n, x, y, z) = C \prod_{i \in T_1} f_i(x_i, y_i, z_i) \prod_{i \in T_2} f_i(n_i), \tag{23}$$

where $f_i(x_i, y_i, z_i)$ for $i \in T_1$ is defined as

$$f_i(x_i, y_i, z_i) = \begin{cases} \left(\frac{1}{\mu_i}\right)^{n_i} \prod_{l=1}^{n_i} \alpha_{ix_{il}} & \text{if } i \text{ is type I and } i \in T_A, \\ \left(\frac{1}{\mu_i}\right)^{n_i} \prod_{l=1}^{B_i} \alpha_{ix_{il}} \prod_{l=1}^{b_i} \frac{\sum_{t=1}^{R} \alpha_{y_{il}t} p_{y_{il}t, iz_{il}}}{\mu_i} & \text{if } i \text{ is type I and } i \in T_B, \\ \prod_{l=1}^{n_i} \frac{\alpha_{ix_{il}}}{\mu_{ix_{il}}} & \text{if } i \text{ is type IV and } i \in T_A, \\ \prod_{l=1}^{B_i} \frac{\alpha_{ix_{il}}}{\mu_{ix_{il}}} \prod_{l=1}^{b_i} \frac{\sum_{t=1}^{R} \alpha_{y_{il}t} p_{y_{il}t, iz_{il}}}{\mu_{ix_{iz_{il}}}} & \text{if } i \text{ is type IV and } i \in T_B\ , \end{cases} \tag{24}$$

and $f_i(n_i)$ for $i \in T_2$, is defined as

$$f_i(n_i) = \begin{cases} n_i! \prod_{r=1}^{R} \frac{1}{n_{ir}!} \left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_{ir}} & \text{if station } i \text{ is type II}, \\ \prod_{r=1}^{R} \frac{1}{n_{ir}!} \left(\frac{\alpha_{ir}}{\mu_{ir}}\right)^{n_i r} & \text{if station } i \text{ is type III}, \end{cases} \tag{25}$$

where C is a normalization constant

$$C^{-1} = \sum_{(n,x,y,z) \in S} \prod_{i \in T_1} f_i(x_i, y_i, z_i) \prod_{i \in T_2} f_i(n_i). \tag{26}$$

Proof

From (23)–(25) we have

$$p(n + e_{ir}, x - x_{jn_j}, y, z) = p(n, x, y, z) \frac{\mu_j(x_j)}{\alpha_{jx_{jn_j}}} \frac{\alpha_{ir}}{\mu_{ir}(n_{ir} + 1)}, \tag{27}$$

$$p(n + e_{y_{jb_j}r}, x - x_{jn_j}, y - y_{jb_j}, z - z_{jb_j}) = p(n, x, y, z) \frac{\mu_j(x_j)}{\sum_{t=1}^{R} \alpha_{y_{jb_j}t} p_{y_{jb_j}t, jz_{jb_j}}} \frac{\alpha_{y_{jb_j}r}}{\mu_{y_{jb_j}r}(n_{y_{jb_j}r} + 1)}, \tag{28}$$

$$p(n + e_{ir} - e_{js}, x, y, z) = p(n, x, y, z) \frac{\alpha_{ir}}{\mu_{ir(n_{ir}+1)}} \frac{\mu_{js(n_{js})}}{\alpha_{js}}, \tag{29}$$

$$p(n - e_{js}, x + x_{i0}, y, z) = p(n, x, y, z) \frac{\alpha_{ix_{i0}}}{\mu_i(x_i + x_{i0})} \frac{\mu_{js(n_{js})}}{\alpha_{js}}, \tag{30}$$

$$p(n - e_{js}, x + x_{i0} - x_{iB_i}, y + y_{i0}, z + z_{i0}) = p(n, x, y, z) \frac{\mu_{js(n_{js})}}{\alpha_{js}} \frac{\alpha_{ix_{i0}}}{\mu_i(x_i + x_{i0} - x_{iB_i})} \frac{\sum_{t=1}^{R} \alpha_{y_{i0}t} p_{y_{i0}t, ix_{B_i}}}{\alpha_{ix_{iB_i}}} \frac{\mu_i(x_i)}{\mu_i(z_i + z_{i0})}. \tag{31}$$

Note that the last term on the right-hand side of (31) is $\mu_i(x_i)/\mu_i(z_i + z_{i0}) = 1$, because if the station i is of type I, then the service rate is independent of the class. However, if the station is of type IV then $\mu_i(x_i) = \mu_{ix_{iB_i}}$. Since we assume the blocked jobs are FCFI, then $\mu_i(z_i + z_{i0}) = \mu_{iz_{i0}}$ and $x_{iB_i} = z_{i0}$. Now we substitute (27)–(31) on the right-hand side of (22) separately, and use (1) to obtain

$$\sum_{j \in T_A} \sum_{i \in T_2} \sum_{r=1}^{R} p(n + e_{ir}, x - x_{jn_j}, y, z) \mu_{ir}(n_{ir} + 1) p_{ir, jx_{jn_j}}$$

$$= p(n, x, y, z) \sum_{j \in T_A} \sum_{i \in T_2} \sum_{r=1}^{R} \frac{\mu_j(x_j)}{\alpha_{jx_{jn_j}}} \alpha_{ir} p_{ir, jx_{jn_j}}$$

$$= p(n, x, y, z) \sum_{j \in T_A} \mu_j(x_j), \tag{32}$$

$$\sum_{j \in T_B} \sum_{r=1}^{R} p(n + e_{y_{jb_j}r}, x, y - y_{jb_j}, z - z_{jb_j}) \mu_{y_{jb_j}r}(n_{y_{jb_j}r} + 1) p_{y_{jb_j}r, jz_{jb_j}}$$

$$= p(n, x, y, z) \sum_{j \in T_B} \sum_{r=1}^{R} \frac{\mu_j(x_j)}{\sum_{t=1}^{R} \alpha_{y_{jb_j}t} p_{y_{jb_j}t, jz_{jb_j}}} \alpha_{y_{jb_j}r} p_{y_{jb_j}r, jz_{jb_j}}$$

$$= p(n, x, y, z) \sum_{j \in T_B} \mu_j(x_j), \tag{33}$$

$$\sum_{i \in T_2, j \in T_2} \sum_{s=1}^{R} \sum_{r=1}^{R} p(n + e_{ir} - e_{js}, x, y, z) \mu_{ir}(n_{ir} + 1) p_{ir,js}$$

$$= p(n, x, y, z) \sum_{j \in T_2} \sum_{s=1}^{R} \frac{\mu_{js}(n_{js})}{\alpha_{js}} \sum_{i \in T_2} \sum_{r=1}^{R} \alpha_{ir} p_{ir,js}, \tag{34}$$

$$\sum_{i \in T_A, j \in T_2} \sum_{s=1}^{R} \sum_{x_{i0}=1}^{R} p(n - e_{js}, x + x_{i0}, y, z) \mu_i(x_i + x_{i0}) p_{ix_{i0}js}$$

$$= p(n, x, y, z) \sum_{j \in T_2} \sum_{s=1}^{R} \frac{\mu_{js}(n_{js})}{\alpha_{js}} \sum_{i \in T_A} \sum_{x_{i0}=1}^{R} \alpha_{ix_{i0}} p_{ix_{i0},js}, \tag{35}$$

$$\sum_{i \in T_B, j \in T_2} \sum_{s=1}^{R} \sum_{y_{i0}=1}^{N} \sum_{x_{i0}=1}^{R} p(n - e_{js}, x + x_{i0} - x_{iB_i}, y + y_{i0}, z + z_{i0}) \mu_i(x_i + x_{i0} - x_{iB_i}) p_{ix_{i0}js}$$

$$= p(n, x, y, z) \sum_{i \in T_B, j \in T_2} \sum_{s=1}^{R} \sum_{x_{i0}=1}^{R} \frac{\mu_{js}(n_{js})}{\alpha_{js}} \frac{\alpha_{ix_{i0}} p_{ix_{i0},js}}{\alpha_{ix_{iB_i}}} \sum_{y_{i0}=1}^{N} \sum_{t=1}^{R} \alpha_{y_{i0}t} p_{y_{i0}t, ix_{iB_i}}$$

$$= p(n, x, y, z) \sum_{j \in T_2} \sum_{s=1}^{R} \frac{\mu_{js}(n_{js})}{\alpha_{js}} \left(\sum_{i \in T_B} \sum_{x_{i0}=1}^{R} \alpha_{ix_{i0}} p_{ix_{i0},js} \right). \tag{36}$$

By combining (34)–(36) we have

$$(34) + (35) + (36) = p(n, x, y, z) \sum_{j \in T_2} \sum_{s=1}^{R} \frac{\mu_{js}(n_{js})}{\alpha_{js}} \left(\sum_{i=1}^{N} \sum_{r=1}^{R} \alpha_{ir} p_{ir,js} \right)$$

$$= p(n, x, y, z) \sum_{j \in T_2} \sum_{s=1}^{R} \mu_{js(n_{js})}. \tag{37}$$

The result follows by checking the flow-in rate equal to the flow-out rate. □

Remark

The result (23) is obtained by similar arguments as in (9). The blocked jobs are kept in an extended buffer at a station of type II or type III. The blocked jobs behavior is captured through the state description in our model.

THEOREM 4

The *throughput* of the ith station is computed from

$$\lambda_i = \sum_{(n,x,y,z)\in S} p(n,x,y,z) \begin{cases} \mu_i(x_i) & \text{for } i \in T_1\,, \\ \sum_{r=1}^{R} \mu_{ir}(n_{ir})(1 - \sum_{j\in T_B}\sum_{s=1}^{R} p_{ir,js}) & \\ \quad + \sum_{j\in T_B} \mu_j(x_j) I_{\{y_{j1}=i\}} & \text{for } i \in T_2\,. \end{cases} \tag{38}$$

The *mean number of jobs* in the ith station is

$$\bar{k}_i = \sum_{(n,x,y,z)\in S} p(n,x,y,z) \begin{cases} n_i & \text{for } i \in T_1\,, \\ \sum_{r=1}^{R} n_{ir} + \sum_{j\in T_B}\sum_{l=1}^{b_j} I_{\{y_{jl}=i\}} & \text{for } i \in T_2\,. \end{cases} \tag{39}$$

Proof

Consider a state (n,x,y,z). Since the jobs from stations of T_1 can never be blocked, thus the throughput of such station $i \in T_1$ is $\mu_i(x_i)$. For stations in T_2 we have the throughput rate as the service rate $\sum_{r=1}^{R} \mu_{ir}(n_{ir})$. However, the job may choose the next station in T_B and a blocking may occur. If station $i \in T_2$ has blocked jobs which have the first priority to get into the destination station that caused their blocking, then the throughput rate of those jobs from station i is equal to the service rate of the station which caused blocking. Thus (38) follows.

The *mean number of jobs* is also obtained from given a state (n,x,y,z). The number of jobs in station i for $i \in T_1$ is n_i. If $i \in T_2$, then the number of jobs in station i will include the active jobs of all classes $n_{i1} + \ldots + n_{iR}$ and the blocked jobs in that station. Thus, we take sum over the stations that cause blocking, $j \in T_B$, and check the vectors y_j if $y_{il} = i$, then the lth job is located in station i, so that the result (38) follows. □

5. Conclusions

We derived exact product form solutions for the equilibrium state probabilities and computed throughput and the mean number of jobs in two different models. We used a property of type II and III stations that the location of the blocked jobs has no effect on the service rates in those stations. Thus, we used these types of stations as kind of *optional* storage spaces for the blocking station. In other words, the system is treated as a virtual *nonblocking* system. By the well-defined state space, particularly the vector of the locations for the blocked jobs, we obtain the exact product form solutions. Another observation is that by the construction of the models we always have the constraint that no more than two jobs will move on a single departure event. The first model has a stronger constraint, that is only one of the stations will cause blocking at a time. However, for the second model we allow more than one station to be of type I or IV which may cause blocking at the same time. This is due to the constraint of the neighbors which will not cause blocking.

References

[1] I.F. Akyildiz, Exact product form solution for queueing networks with blocking, IEEE Trans. Comput. C-36 (1987) 122–125.

[2] I.F. Akyildiz, On the exact and approximate throughput analysis of closed queueing networks with blocking, IEEE Trans. Software Eng. SE-14 (1988) 62–71.

[3] I.F. Akyildiz and H.G. Perros, Special issue on queueing networks with finite capacity queues: introduction, Performance Evaluation 10 (3) (1989).

[4] I.F. Akyildiz and H. von Brand, Exact solutions for networks of queues with blocking-after-service, to appear in Theor. Comput. Sci. J. (1993).

[5] I.F. Akyildiz and J. Liebeherr, Optical deadlock free buffer allocation in multiple chain blocking networks of queues, in: *Proc. Int. Conf. on the Performance of Distributed Systems and Integrated Communication Networks*, September 10–12, 1991.

[6] S. Balsamo, V. De Nitto Persone and G. Iazeolla, Identity and reducibility properties of some blocking and non-blocking mechanisms in congested networks, in: *Flow Control of Congested Networks*, NATO ASI Series (Springer Verlag, 1987) pp. 243–254.

[7] S. Balsamo and L. Donatiello, On the cycle time distribution in a two-stage cyclic network with blocking, IEEE Trans. Software Eng. SE-15 (10) (1989).

[8] F. Baskett, K.M. Chandy, R.R. Muntz and F.G. Palacios, Open closed and mixed networks of queues with different classes of customers, J. ACM 22 (1975) 248–260.

[9] P.P. Bocharov, On the two-node queueing networks with finite capacity, in: *Proc. 1st Int. Workshop on Queueing Networks with Blocking* (North-Holland, 1989) pp. 105–125.

[10] F.P. Kelly, Networks of queues with customers of different types, J. Appl. Prob. 12 (1975) 542–554.

[11] F.P. Kelly, Networks of queues, Adv. Appl. Prob. 8 (1976) 416–432.

[12] R.O. Onvural, A note on the product form solutions of multiclass closed queueing networks with blocking, Performance Evaluation 10 (1989) 247–255.

[13] R.O. Onvural, A survey of closed queueing networks with finite buffers, ACM Comp. Surveys 22 (1990) 83–121.