UNIVERSITÀ CA' FOSCARI DI VENEZIA
Dipartimento di Informatica
Technical Report Series in Computer Science

Rapporto di Ricerca CS-2007-1

Febbraio 2007

Simonetta Balsamo, Andrea Marin

# On representing multiclass M/M/k queue by Generalized Stochastic Petri Net

Dipartimento di Informatica, Università Ca' Foscari di Venezia
Via Torino 155, 30172 Mestre–Venezia, Italy

# On representing multiclass M/M/k queue by Generalized Stochastic Petri Net

Simonetta Balsamo, Andrea Marin

Dipartimento di Informatica
Università Ca' Foscari di Venezia
Via Torino 155, 30172 Venezia Mestre, Italy
{balsamo,marin}@dsi.unive.it

**Abstract.** In this paper we study the relations between multi-class BCMP-like service stations and generalized stochastic Petri nets (GSPN). Representing queuing discipline with GSPN models is not easy. We focus on representing multi-class queuing systems with different queuing disciplines by defining appropriate finite GSPN models. Note that queuing discipline in general affect performance measures in multi-class systems. For example, BCMP-like service centers with First Come First Served (FCFS) and with Last Come First Served with Preemptive Resume (LCFSPr) have a (different) product-form solution under different hypotheses. We define structurally finite GSPNs equivalent to the multi-class M/M/k queuing system with FCFS, LCFSPR, Processor Sharing (PS) and Infinite Servers (IS). Equivalence holds in terms of steady state probability function and average performance measure. The main idea is to define a finite GSPN model that simulates the behavior of a given queue discipline with some appropriate random choice. Moreover, we prove that the combination of the introduced equivalent models has a closed-form steady state probability by the $M \implies M$ property. We consider queuing systems with both a single server with load dependent service rate, and multiple servers with constant service rate.

## 1 Introduction

Queuing theory and (Generalized) Stochastic Petri Nets are important classes of stochastic models used to evaluate system performances. Queuing systems have been widely applied to represent resource contention systems where a set of customers competes for resource usage. However, basic queuing systems cannot model the synchronization between concurrent activities. Stochastic Petri nets (SPN) can be naturally used to represent systems with synchronization and concurrency and to perform both qualitative and quantitative analysis. An important problem in system performance evaluation based on performance models is the efficiency of the solution algorithms, i.e. the ability to define classes of models that can be analyzed by methods with a limited space and time computational complexity. Many results have been proposed in literature which give efficient solutions of some types of stochastic models under certain conditions.

Single queuing system models have been widely analyzed by considering various arrival distributions, service time distributions, classes of users and scheduling disciplines. The single server queue is analyzed [14], [10], [13], [19] and several results have been derived for special cases. Queuing networks (QN) extend and combine various queuing systems to represent more complex systems. QN models can be represented by defining an associated stochastic process that, under some exponential and independence assumptions, is a continuous-time Markov Chain (CTMC) process. Although the stationary state probability solution of the associated Markov Process, under stability conditions, can be easily defined, the computation can soon become unfeasible due to the high computational complexity. However, under certain assumptions QNs can be efficiently analyzed by applying the product-form theorem [5], which defines the steady state probability function as product of functions of each single service center state. Product-form QNs can be analyzed by efficient algorithms (e.g. [18], [8], [7], [6]) that yield a low polynomial computational cost.

SPNs, which are defined in terms of a set of places and a set of exponentially timed transitions connected by arcs, and a marking, which is the state of the net, are represented by a CTMC process whose state space is the set of all possible markings of the net. Specifically the computation of performance measures is inefficient because it requires to calculate the reachability set, which depends on the initial marking and whose size grows exponentially with the number of places of the net and the number of tokens in the initial marking. Henderson et al. introduced the idea of product-form for SPN in [12] and [11] simplifying the computation of the performance indexes. However, the algorithms for product-form SPNs still require to check conditions on the reachability set. The class of Generalized Stochastic Petri Nets (GSPN) allows modelling nets with both exponentially timed and immediate transitions introducing more flexibility. The underlying process of a GSPN can be defined as a semi-Markov process. Product-form for GSPN has been studied by Balbo et al. in [3] and it is based on techniques to reduce the problem to the product-form SPN theorem.

Investigating the relations between classes of queuing models and GSPN models is an interesting problem and it has been considered by some recent research literature ([2], [20], [4]). Most of these works focus on the relations between QNs and SPNs or GSPNs and derive some equivalence results. However a little attention has been devoted to the problem of representing various types of scheduling disciplines of queuing systems by GSPNs. To the best of our knowledge, representing scheduling disciplines in multiclass models with finite GSPNs is still an open problem. In [20] the authors introduce a comparison between QN models and SPN models based on the representation of multiclass features by colored Petri nets. However the differences between different scheduling disciplines are not analyzed. Balbo et al. in [2] combine GSPN and product-form QN by replacing subsystem in a low-level model with their flow equivalents models. Still little attention is devoted to scheduling disciplines. In [4] the authors observe how they can map each service station of a BCMP QN to a complex GSPN. The GSPN model depends on the scheduling disciplines but it has an infinite number of places and transitions for the FCFS and LCFSPR stations. Then they give a finite and remarkably compact representation by a GSPN equivalent to the detailed model. The compact representation holds the product-form conditions for GSPN showed in [3] but it does not distinguish different queuing disciplines by mapping everything in the PS discipline. Thus it is not possible to define on the GSPN equivalent conditions to the ones depending on the service center scheduling discipline of the BCMP theorem. On the other hand the detailed representations of queuing discipline yield non product-form GSPN models equivalent to BCMP QN.

In this paper we present an equivalence result between two types of stochastic models. We propose a finite GSPN representation of a set of queuing systems with various scheduling disciplines. According to the BCMP-type service centers we analyze First Come First Served (FCFS), Last Come First Served with preemptive resume (LCFSPR), Processor Sharing (PS) and Infinite Servers (IS) scheduling disciplines. The main idea behind these results is a probabilistic model of the queue, i.e., all the customers of the same class wait in the same place and when a server becomes free the customer which gets the service is chosen in a probabilistic way similarly to what happens with the random queuing discipline. In the LCFSPR discipline, we also choose probabilistically the customer that looses the server when a new customer arrives to the system.

The advantage of having a finite representation which is different for the various scheduling disciplines is twofold: first it makes the analysis easier. Second it does not require the definition of new semantic for the GSPN according to the queuing disciplines. Thus existing analysis or simulation tools can be used with the GSPN nets defined in this work. The proposed results are interesting because they allow the representation of an M/M/k queue with various queuing disciplines by a compact GSPN, which is equivalent to the queuing system in term of steady state queue length distribution. A practical consequence can be that it can extend a GSPN simulator or analyzer for analyzing multiclass queue systems. The only requirement is that the tool is able to model state-dependent firing rates for timed transitions and state-dependent weights for immediate transitions. There is no need to support the colored model extension to represent different classes. One could also integrate a GSPN analyzer by a functionality that identifies net structures equivalent to different scheduling disciplines M/M/k queuing systems and then it can apply the closed form steady state formula. For the FCFS and PS disciplines, the GSPN structure complexity, i.e. the number of places and transictions, grows linearly just with the number $R$ of classes of users, for the LCFSPR it grows like $\mathcal{O}(R^2)$.

We give a GSPN model for the queuing disciplines considered in the BCMP theorem [5]. An open problem and a possible further research is the analysis of the GSPN models obtained by combining these blocks. Possibly, under some assumptions, it is possible to define equivalence between the GSPN steady state probability function and the BCMP queuing network steady state probability function.

The paper is structured as follows. Section 2 briefly reviews the GSPN models recalling formalism we chose, Section 3 reviews some results of the queuing systems theory used later in the paper. In Sections 4, 5 we introduce the GSPNs respectively equivalent to the FCFS and LCFS multiclass M/M/k queue. Section 6 discuss the GSPN models for both PS scheduling and IS systems. The proof of some theorems are given in appendix. Finally, Section 8 provides some concluding remarks.

## 2 Generalized Stochastic Petri Nets

In this section we briefly recall the Generalized Stochastic Petri Nets (GSPN). We consider the notation for GSPN introduced in [15]. In order to allow marking dependent probabilities for solving conflicts among immediate transitions we use the techniques discussed in [9]. Then in the next section we shall present GSPN models equivalent to queuing systems under various assumptions. Let us define a marked Stochastic Petri Net which consists of a 8-tuple as follows:

$$GSPN = (\mathcal{P}, \mathcal{T}, I(\cdot,\cdot), O(\cdot,\cdot), H(\cdot,\cdot), \Pi(\cdot), w(\cdot,\cdot), \mathbf{m_0})$$

where:

- $\mathcal{P} = \{P_1, \ldots, P_M\}$ is the set of $M$ places,
- $\mathcal{T} = \{t_1, \ldots, t_N\}$ is the set of $N$ transitions (both immediate and timed),
- $I(t_i, p_j) : \mathcal{T} \times \mathcal{P} \to \mathbb{N}$ is the input function, $1 \le i \le N$, $1 \le j \le M$,
- $O(t_i, p_j) : \mathcal{T} \times \mathcal{P} \to \mathbb{N}$ is the output function, $1 \le i \le N$, $1 \le j \le M$,
- $H(t_i, p_j) : \mathcal{T} \times \mathcal{P} \to \mathbb{N}$ is the inhibition function, $1 \le i \le N$, $1 \le j \le M$,
- $\Pi(t_i) : \mathcal{T} \to \mathbf{N}$ is a function that specifies the priority of transition $t_i$, $1 \le i \le N$,
- $\mathbf{m} \in \mathbb{N}^M$ denotes a marking or state of the net, where $m_i$ represents the number of tokens in place $P_i$, $1 \le i \le N$,
- $w(t_i, \mathbf{m}) : \mathcal{T} \times \mathbb{N}^M \to \mathbb{R}$ is the function which specifies for each timed transition $t_i$ and each marking $\mathbf{m}$ a state dependent firing rate, and for immediate transitions a state dependent weight,
- $\mathbf{m_0} \in \mathbf{N}^M$ represents the initial state of the GSPN, i.e. the number of tokens in each place at the initial state.

We consider ordinary nets, i.e., functions $I, O$ and $H$ take values in $\{0, 1\}$. For each transition $t_i$ let us define the input vector $\mathbf{I}(t_i)$, the output vector $\mathbf{O}(t_i)$ and the inhibition vector $\mathbf{H}(t_i)$ as follows: $\mathbf{I}(t_i) = (i_1, \ldots, i_M)$ where $i_j = I(t_i, P_j)$, $\mathbf{O}(t_i) = (o_1, \ldots, o_M)$ where $o_j = O(t_i, P_j)$ and $\mathbf{H}(t_i) = (h_1, \ldots, h_M)$ where $h_j = H(t_i, P_j)$. Function $\Pi(t_i)$ associates a priority to transition $t_i$. If $\Pi(t_i) = 0$ then $t_i$ is a timed transition, i.e., it fires after an exponentially distributed firing time with mean $1/w(t_i, \mathbf{m})$, where $\mathbf{m}$ is the marking of the net. If $\Pi(t_i) > 0$ then $t_i$ is an immediate transition and its firing time is zero. We say that transition $t_a$ is enabled by marking $\mathbf{m}$ if $m_i \ge I(t_a, p_i)$ and $m_i < H(t_a, p_i)$ for each $i = 1, \ldots, M$ and no other transition of higher priority is enabled. We consider just two priority levels, 0 and 1. Hence when an immediate transition is enabled all the timed ones are disabled. The firing of transition $t_i$ changes the state of the net from $\mathbf{m}$ to $\mathbf{m}-\mathbf{I}(t_i)+\mathbf{O}(t_i)$. The reachability set $RS(\mathbf{m_0})$ of the net is defined as the set of all markings that can be reached in zero or more firings from $\mathbf{m_0}$. We say that marking $\mathbf{m}$ is tangible if it enables only timed transitions and it is vanishing otherwise. For a vanishing marking $\mathbf{m}$ let $\mathcal{T}_\alpha$ be the set of enabled immediate transitions. Then the firing probability for any transition $t_i \in \mathcal{T}_\alpha$ and any state $\mathbf{m}$ is denoted by $p(t_i, \mathbf{m})$ and it is defined as follows:

$$p(t_i, \mathbf{m}) = \frac{w(t_i, \mathbf{m})}{\sum_{t_j \in \mathcal{T}_\alpha} w(t_j, \mathbf{m})}. \tag{1}$$

Given a tangible marking $\mathbf{m}$ the transition with the lowest associated stochastic time fires.

A GSPN is represented by a graph with the following conventions:

- timed transitions are white filled boxes,
- immediate transitions are black filled boxes,
- places are circles,
- if $I(t_i, p_j) > 0$ we draw an arrow from $p_j$ to $t_i$ labelled with $I(t_i, p_j)$,
- if $O(t_i, p_j) > 0$ we draw an arrow from $t_i$ to $p_j$ labelled with $O(t_i, p_j)$,
- if $H(t_i, p_j) > 0$ we draw an circle ending line from $p_j$ to $t_i$ labelled with the value of $H(t_i, p_j)$,
- the marking $\mathbf{m}$ is represented by a set of $m_j$ filled circles representing the tokens in place $p_j$ for each $j = 1, \ldots, M$.

For ordinary nets we do not use labels for the arrows.

GSPN analysis consists in finding the steady state probability for each tangible marking of the reachability set. Some analysis techniques are presented in [15]. Under general assumptions, the stochastic process generated by the dynamic behavior of a standard SPN is a CTMC process. Mean state sojourn times are computed from the mean transition delays of the net. For GSPNs the distribution of the sojourn time in any marking can be expressed as a negative exponential and deterministically zero distributions for tangible and vanishing markings, respectively. Thus the marking process can be studied as a semi-Markov random process.

The GSPN models introduced in this paper present marking processes which allow us to easily reduce the semi-Markov process to a CTMC. In fact whenever a vanishing marking is reached, the next marking is tangible. Thus we can simply obtain a CTMC whose states are the tangible states of the original process and the transition rates are computed weighting the transitions rates of the original process with the firing probabilities of the immediate transitions. Hence the mean sojourn times in the tangible states of the original semi-Markov process and the mean sojourn times of the CTMC are the same.

Finally let us introduce some other notations: let $\mathbf{e_i}$ be an $M$-dimensional vector with all zero components but the $i$-th which is 1. We use the lower case $t$ to name immediate transitions, the upper case $T$ to name timed transitions, $\tilde{t}$ to name a generic timed or immediate transition.

# 3 Single Queuing systems with different classes of customers

In this section we briefly recall single queuing systems with different classes of customers classifying them on the number of servers and scheduling disciplines. Let us consider an open queuing system with external arrivals, a queue, a set of identical servers and a set of $R$ customer classes. The queuing system is shown in Figure 1. Customers of class $r$ arrive at the system according to a Poisson process with rate $\lambda_r$ and require an exponentially distributed random service time with parameter $\mu_r$, $r = 1, \ldots, R$. The system has a set of independent servers, possibly infinite.

For single class queuing systems some results in terms of steady state probability hold for any scheduling discipline that is work-conservative and independent from the service time [10], [14]. These results can be extended to multiclass queuing systems although they depend on the scheduling discipline. We consider the following disciplines: First Come First Server (FCFS), Last Come First Server with Preemptive Resume (LCFSPR), Processor Sharing (PS). The steady

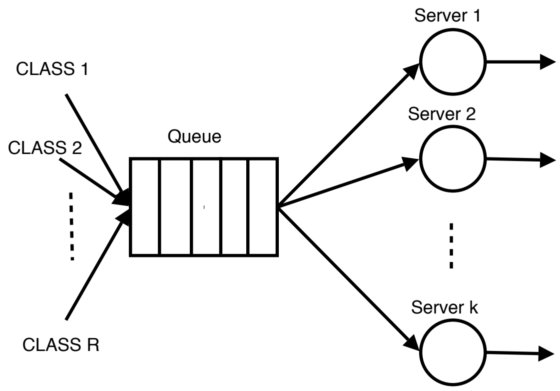


**Fig. 1.** An M/M/k multiclass queuing system.

state probability of a M/M/k multiclass system with a specific queuing discipline and constant service rate is equivalent to the steady state probability of a M/M/1 multiclass system with the same queuing discipline and an appropriate load-dependent service rate. If all the customer service times are identical, i.e., $\mu_r = \mu$ for $r = 1, \ldots, R$, the load dependent service rate $\mu(j)$, where $j$ is the number of customers at the system, is defined as follows:

$$\mu(j) = \begin{cases} j\mu & \text{if } j \le k \\ k\mu & \text{if } j > k \end{cases} \tag{2}$$

If the stability condition $\sum_{i=1}^{R} \frac{\lambda_r}{k\mu} < 1$ holds, then we can evaluate the stationary queue length distribution of the multiclass M/M/k system for any scheduling discipline by the corresponding M/M/1 load dependent system. Let $\pi'(\mathbf{n})$ denote the steady state probability of the M/M/k system, with $\mathbf{n} = (n_1 \dots n_R)$, i.e., the probability of finding $n_i$ customers of class $i$ for $i = 1, \dots, R$ in the system. Then we can write:

$$\pi'(\mathbf{n}) = \pi'_0 \prod_{i=1}^{R} \lambda_i^{n_i} \frac{(\sum_{i=1}^{r} n_i)!}{\prod_{i=1}^{R} n_i!} \prod_{j=1}^{\sum_{i=1}^{R} n_i} \frac{1}{\mu(j)}, \tag{3}$$

where $\pi'_0$ is the probability of finding the system empty.

When mean service rates for different customer classes are not identical, i.e., $\mu_i \neq \mu_j$ for $i \neq j$ for some couple $i$, $j$, then the load dependent service rate function $\mu_r(\mathbf{n})$, for any class $r$ and state $\mathbf{n}$, is defined as follows:

$$\mu_r(\mathbf{n}) = \frac{n_r}{n} \min(n, k)\mu_r, \quad n = \sum_{i=1}^{R} n_i. \tag{4}$$

The following steady state probability holds for LCFSPR and PS queuing disciplines:

$$\pi'(\mathbf{n}) = \pi'_0 \frac{\sum_{i=1}^{R} n_i)!}{\prod_{i=1}^{R} n_i!} \prod_{i=1}^{R} \lambda_i^{n_i} \prod_{i=1}^{R} \Big(\frac{1}{\mu_i}\Big)^{n_i} \prod_{i=1}^{\sum_{j=1}^{R} n_j} \frac{1}{\min(k, i)}. \tag{5}$$

The stability condition is $\exists \mathbf{k} : \forall \mathbf{n} > \mathbf{k}[\sum_{r=1}^{R} \frac{\lambda_r}{\mu_r(\mathbf{n})} < 1]$.

The BCMP theorem [5] considers service centers with single servers and state dependent service time. For FCFS service stations the service time can depend only on the total number of customers in the system. Let $n = \sum_{r=1}^{R} n_r$ and $x(n)$ be an arbitrary positive function of $n$, representing the service rate when there are $n$ customers at the service center relative to the service rate when $n = 1$. Then the steady state probability function is:

$$\pi'(\mathbf{n}) = \pi'_0 \frac{n!}{\prod_{i=1}^{R} n_i!} \prod_{i=1}^{R} \lambda_i^{n_i} \Big(\frac{1}{\mu}\Big)^{n} \prod_{i=1}^{n} \frac{1}{x(i)}. \tag{6}$$

For LCFSPR and PS systems, BCMP theorem considers another state dependent service rate. Let $y_r(n_r)$ be an arbitrary positive function of $n_r$ which denotes the service rate of class $r$ customers at service center $i$ relative to the service rate when there is one class $r$ customer at service center $i$ i.e. $\mu_r$. Then the steady state probability function is:

$$\pi'(\mathbf{n}) = \pi'_0 \frac{n!}{\prod_{i=1}^{R} n_i!} \prod_{i=1}^{R} \lambda_i^{n_i} \prod_{r=1}^{R} \Big[\Big(\frac{1}{\mu_r}\Big)^{n_r} \prod_{a=1}^{n_r} \frac{1}{y_r(a)}\Big]. \tag{7}$$

Note that these various forms of state dependent service rates can be combined. For example the steady state probability (5) can be obtained combining equations (6) and (7) by setting $x(n) = \frac{\min(n,k)}{n}$ and $y_r(n_r) = n_r$.

## 4 Representing M/M/k/FCFS queue by GSPN

In this section we define a GSPN that represents an $R$-multiclass M/M/k/FCFS queue. Then we prove that the GSPN model is equivalent to the queuing system in terms of the steady state probability. Given the M/M/k/FCFS models defined as in Section 3 let us define the model called *GSPN-1*.

**Definition 1 (GSPN-1).** *According to GSPN definition given in Section 2:*

- $\mathcal{P} = \mathcal{P}_q \cup \mathcal{P}_s \cup \{P_{2R+1}\}$ *with* $\mathcal{P}_q = \{P_1, \ldots, P_R\}$ *and* $\mathcal{P}_s = \{P_{R+1}, \ldots, P_{2R}\}$,
- $\mathcal{T} = \mathcal{T}_w \cup \mathcal{T}_q$ *where* $\mathcal{T}_q = \{t_1, \ldots, t_R\}$ *and* $\mathcal{T}_w = \{T_{R+1}, \ldots, T_{2R}\}$,
- *function* $\Pi$ *defined as follows:*

$$\Pi(\tilde{t_i}) = \begin{cases} 0 & \textit{if} \;\; R+1 \leq i \leq \; 2R \\ 1 & \textit{if} \;\; 1 \leq i \leq R \end{cases} ,$$

- *input and output vectors for transition* $t_i$, $1 \leq i \leq R$: $\mathbf{I}(t_i) = \mathbf{e_i} + \mathbf{e_{2R+i}}$ *and* $\mathbf{O}(t_i) = \mathbf{e_{R+i}}$. *Input and output vector for transition* $T_{R+i}$: $\mathbf{I}(T_{R+i}) = \mathbf{e_{R+i}}$ *and* $\mathbf{O}(T_{R+i}) = \mathbf{e_{2R+1}}$,
- $\mathbf{H}(t_i) = (0, \ldots, 0)$ *for all* $t_i \in \mathcal{T}$,
- $w(T_{R+i}, \mathbf{m}) = m_{R+i}\mu$ *for* $1 \leq i \leq R$ *and* $w(t_i, \mathbf{m}) = m_i$ *for* $1 \leq i \leq R$,
- $\mathbf{m_0} = (0, \ldots, 0, k)$ .

*Tokens arrive to places* $P_i$, $1 \leq i \leq R$ *according to Poisson stochastic processes.*

Figure 2 illustrates the graphical representation of GSPN-1 model where $t_1, \ldots, t_R$ are immediate transitions and $T_{R+1}, \ldots, T_{2R}$ are exponential transitions.

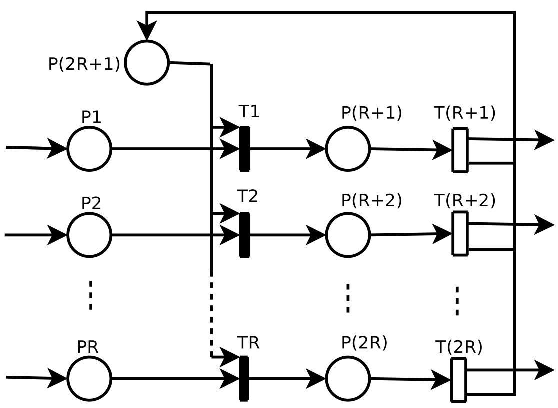


**Fig. 2.** Graphical representation of model GSPN-1

Let $\mathbf{m}$ be a valid vanishing state of the GSPN-1, and let $\mathcal{T}_a \subseteq \mathcal{T}_q$ be the set of immediate transitions enabled by $\mathbf{m}$, then the probability of firing of $t_i \in \mathcal{T}_a$

can be written as:

$$p(t_i, \mathbf{m}) = p_i(\mathbf{m}) = \frac{m_i}{\sum_{j \in \{j | t_j \in \mathcal{T}_a\}} m_j} \tag{8}$$

We shall now derive a closed form solution for the steady state probability of GSPN-1 model by considering the set of reachable markings $\mathbf{m} = (m_1, \ldots, m_{2R+1})$. This is given by Lemma 1. Then we introduce a state aggregation by defining the aggregate state $\mathbf{n} = (n_1, \ldots, n_R)$ where $n_i = m_i + m_{R+i}$, $1 \leq i \leq R$. This state corresponds to the number of customers of class $i$ in the queuing model. Theorem 1 provides the closed form solution for model GSPN-1 in terms of aggregated stationary probability of state $\mathbf{n}$. Finally the GSPN-1 model is shown to be equivalent to the M/M/k FCFS multiclass queuing system in terms of stationary probability.

**Lemma 1.** *Let $\mathbf{m} = (m_1, \ldots, m_{2R+1})$ be a reachable tangible state of the GSPN-1. Then if the stability condition holds, the stationary state probability can be written as follows:*

$$\pi(\mathbf{m}) = \pi_0 \prod_{i=1}^{R} \lambda_i^{m_i + m_{R+i}} \frac{(\sum_{i=R+1}^{2R} m_i)!}{\prod_{i=R+1}^{2R} m_i!} \frac{(\sum_{i=1}^{R} m_i)!}{\prod_{i=1}^{R} m_i!} \prod_{j=1}^{\sum_{i=1}^{2R} m_i} \frac{1}{\mu(j)}. \tag{9}$$

*where $\pi_0$ is a normalizing constant and $\mu(j)$ is the function defined by (2).*

The proof is given in appendix A and is based on verifying the set of the CTMC global balance equations.

**Theorem 1.** *Consider model GSPN-1 and let $n_i = m_i + m_{R+i}$, $1 \leq i \leq R$ and $\mathbf{n} = (n_1 \ldots, n_R)$ be an aggregated state. Let $\pi_a(\mathbf{n})$ be the steady state probability of $n_i$ for $i = 1, \ldots, R$. Then we can write:*

$$\pi_a(\mathbf{n}) = \pi_0 \frac{(\sum_{i=1}^{R} n_i)!}{\prod_{i=1}^{R} n_i!} \prod_{i=1}^{r} \lambda_i^{n_i} \prod_{i=1}^{\sum_{i=1}^{R} n_i} \frac{1}{\mu(i)} \quad \forall \mathbf{n} \in \mathbb{N}^R. \tag{10}$$

*Proof.* In order to derive equation (10) we prove that:

$$\pi_a(\mathbf{n}) = \sum_{\substack{\mathbf{m} | m_i + m_{R+i} = n_i \\ 1 \leq i \leq R}} \pi(\mathbf{m}), \tag{11}$$

for $\mathbf{n} \in \mathbb{N}^R$ and $\mathbf{m}$ in the reachability set of model GSPN-1. Consider the two following cases: case 1) $\sum_{i=1}^{R} n_i \geq k$ and case 2) $\sum_{i=1}^{R} n_i < k$.

**Case 1:** $\sum_{i=1}^{R} n_i \geq k$. Consider any combination of $j_i$ with $1 \leq i \leq r$ and $0 \leq j_i \leq n_i$. Then the right-hand side of equation (11) by using formula (9) can be written as follows:

$$\sum_{\substack{j_1+\ldots+j_R=k\\ j_i\le n_i}} \pi(n_1-j_1,n_2-j_2,\ldots,n_R-j_R,j_1,\ldots,j_R,0)$$

$$=\pi_0\prod_{j=1}^{R}\lambda_j^{n_j}\prod_{j=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(j)}\sum_{\substack{j_1+\ldots+j_r=k\\ j_i\le n_i}}\frac{k!}{\prod_{i=1}^R j_i!}\frac{(\sum_{i=1}^R n_i-k)!}{\prod_{i=1}^R(n_i-j_i)!}$$

$$=\pi_0\prod_{j=1}^{R}\lambda_j^{n_j}\prod_{j=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(j)}\frac{(\sum_{i=1}^R n_i-k)!}{\prod_{i=1}^R n_i!}k!\sum_{\substack{j_1+\ldots+j_R=k\\ j_i\le n_i}}\frac{\prod_{i=1}^R n_i!}{\prod_{i=1}^R(n_i-j_i)!}\frac{1}{\prod_{i=1}^R j_i!}$$

$$=\pi_0\prod_{j=1}^{R}\lambda_j^{n_j}\prod_{j=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(j)}\frac{(\sum_{i=1}^R n_i-k)!}{\prod_{i=1}^R n_i!}k!\sum_{\substack{j_1+\ldots+j_R=k\\ j_i\le n_i}}\prod_{i=1}^{R}\binom{n_i}{j_i},$$

where the last sum satisfies the Vandermonde convolution, thus we can write:

$$\pi_0\prod_{j=1}^{R}\lambda_j^{n_j}\prod_{j=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(j)}\frac{(\sum_{i=1}^R n_i-k)!}{\prod_{i=1}^R n_i!}k!\binom{\sum_{i=1}^R n_i}{k}$$

$$=\pi_0\prod_{j=1}^{R}\lambda_j^{n_j}\prod_{j=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(j)}\frac{(\sum_{i=1}^R n_i)!}{\prod_{i=1}^n n_i!},$$

which is formula 10.

**Case 2:** $\sum_{i=1}^R n_i < k$, that corresponds to the behavior of the queuing system where all the customers are being served and in GSPN-1 every place $P_i$ with $1 \le i \le R$ is empty. Note that $n_i = m_{R+i}$, so by equation (9) we can write:

$$\pi(0,\ldots,0,m_{R+1},\ldots,m_{2R},l)=\pi_0\prod_{i=1}^{R}\lambda_i^{n_i}\prod_{i=1}^{\sum_{i=1}^R n_i}\frac{1}{\mu(i)}\frac{(\sum_{i=1}^R n_i)!}{\prod_{i=1}^R n_i!},$$

that yields formula (10) and this ends the proof. ◻

**Corollary 1.** *The M/M/k queuing system with FCFS discipline, R customer classes, arrival rates* $\lambda_i$, $1 \le i \le R$, *single server rate* $\mu$ *and steady state probability* $\pi'(\mathbf{n})$ *is equivalent to the GSPN-1 in terms of steady state probability, i.e.,* $\pi_a(\mathbf{n}) = \pi'(\mathbf{n})$ *for all* $\mathbf{n} \in \mathbb{N}^R$ *where* $\pi_a(\mathbf{n})$ *is the aggregated probability of GSPN given by formula (10)*

*Proof.* It follows immediately from equation (3) and Theorem 1. ◻

Note that GSPN-1 model represents the M/M/k multiclass system when the service rate is independent from the class of the customers in service. Thus it can not be used to represent LCFSPR or PS scheduling disciplines. For example consider the system with LCFSPR queue, single server, and different average service rates for each class. Then we show now a counterexample to prove that the steady state given by queuing theory does not satisfy the GBE for the GSPN. This implies that GSPN-1 catches somehow the FCFS semantic (by not allowing preemption).

*Example 1.* As example, consider a LCFSPR M/M/1 queue with two classes of customers with average service time $1/\mu_1$ and $1/\mu_2$. From queuing theory we can write the steady state probability as follows:

$$\pi'(n_1, n_2) = \pi'(0,0)\lambda_1^{n_1}\lambda_2^{n_2}\frac{(n_1+n_2)!}{n_1!n_2!}\frac{1}{\mu_1^{n_1}}\frac{1}{\mu_2^{n_2}}.$$

Suppose to represent this system by GSPN-1 associating different firing rates to transitions $T_3$ and $T_4$: $w(T_3, \mathbf{m}) = m_3\mu_1$ and $w(T_4, \mathbf{m}) = m_4\mu_2$. We calculate the effective arrival rate to reachable tangible state $\mathbf{m} = (0, m_2, 1, 0, 0)$, with $m_2 > 0$. The adjacent states are $\mathbf{m_1} = (0, m_2 - 1, 1, 0, 0)$, $\mathbf{m_2} = (1, m_2, 1, 0, 0)$, $\mathbf{m_3} = (1, m_2, 0, 1, 0)$, thus the effective arrival rate to state $\mathbf{m}$ is:

$$\begin{aligned}\pi(\mathbf{m})\Big[\frac{1}{\lambda_2}\frac{m_2}{m_2}\mu_2\lambda_2 + \lambda_1(m_2+1)\frac{1}{\mu_1}\mu_1\frac{1}{m_2+1} + \lambda_2(m_2+1)\frac{1}{\mu_2}\mu_2\frac{1}{m_2+1}\Big] \\ = \pi(\mathbf{m})\Big[\mu_2 + \lambda_1 + \lambda_2\Big].\end{aligned}$$

The effective leaving rate for state $\mathbf{m}$ is clearly $\pi(\mathbf{m})(\mu_1 + \lambda_1 + \lambda_2)$, so the GBE on state $\mathbf{m}$ is satisfied if $\mu_1 = \mu_2$.

Finally note that GSPN-1 can as well simulate a single server FCFS service station with an BCMP-like load dependent service rate. We can state the following lemma:

**Lemma 2.** *Let $m = \sum_{i=1}^{2R} m_i$, $x(m)$ be an arbitrary positive function, $\mathbf{m_0} = \mathbf{e_{2R+1}}$ and let the firing rate of transition $T_{R+1}, \ldots, T_{2R}$ be $w(T_{R+r}) = x(m)\mu$, $1 \leq r \leq R$. Then if $\mathbf{m}$ is a tangible reachable marking the steady state probability function is:*

$$\pi(\mathbf{m}) = \pi_0 \prod_{i=1}^{R} \lambda_i^{m_i + m_{R+i}} \frac{(\sum_{i=1}^{R} m_i)!}{\prod_{i=1}^{R} m_i!}\Big(\frac{1}{\mu}\Big)^{\sum_{i=1}^{2R} m_i} \prod_{j=1}^{\sum_{i=1}^{2R} m_i} \frac{1}{x(j)}. \quad (12)$$

The proof is given in appendix. By defining $n_i = m_i + m_{R+i}$ we can aggregate the states and we can prove that the steady state probability $\pi_a$ of the aggregated CTMC is identical to $\pi'$ defined by equation (6) of BCMP theorem. The net structure complexity is linear on $R$, the number of customer classes.

## 5 Representing M/M/k/LCFSPR queue by GSPN

In this section we introduce a GSPN which can be considered equivalent, for steady state probability, to a multiclass M/M/k queue with LCFS with preemptive resume scheduling discipline. As we consider just exponentially distributed service times, we do not consider the problem of representing the resume. We provide a model for this queue system whose structure is finite and depends only on the number of classes of customers, i.e., not on the number of servers. A trivial solution could be obtained by recalling that the steady state formula for a LCFSPR queue is equal to the Processor Sharing one so that we could use the same GSPN representation. On the other hand we want to provide a model which semantically simulate closer the LCFS queue.

**Definition 2 (GSPN-2).** *According to GSPN definition given in Section 2:*

- $\mathcal{P} = \mathcal{P}_q \cup \mathcal{P}_w \cup \mathcal{P}_a \cup \{P_{3R+1}\}$ *where* $\mathcal{P}_q = \{P_1, \ldots, P_R\}$ *and* $\mathcal{P}_w = \{P_{R+1}, \ldots, P_{2R}\}$ *and* $\mathcal{P}_a = \{P_{2R+1}, \ldots, P_{3R}\}$,
- $\mathcal{T} = \mathcal{T}_q \cup \mathcal{T}_w \cup \mathcal{T}_f \cup \mathcal{T}_g$ *where* $\mathcal{T}_q = \{t_1, \ldots, t_R\}$ *and* $\mathcal{T}_w = \{T_{R+1}, \ldots, T_{2R}\}$ *and* $\mathcal{T}_f = \{t_{2R+1}, \ldots, t_{3R}\}$ *and* $\mathcal{T}_g = \{t_{ij},\ 1 \leq i, j \leq R\}$,
- *function* $\Pi$ *is defined as follows:*

$$\Pi(\tilde{t}) = \begin{cases} 1 & \text{if } \tilde{t} \in \mathcal{T}_q \cup \mathcal{T}_f \cup \mathcal{T}_g \\ 0 & \text{if } \tilde{t} \in \mathcal{T}_w \end{cases} ,$$

- *Let* $1 \leq i, j \leq R$. *The input and output vectors of* $t_i \in T_q$: $\mathbf{I}(t_i) = \mathbf{e}_i + \mathbf{e}_{3R+1}$ *and* $\mathbf{O}(t_i) = \mathbf{e}_{R+i}$. *The input and output vectors for* $T_{R+i} \in \mathcal{T}_w$: $\mathbf{I}(t_{R+i}) = \mathbf{e}_{R+i}$ *and* $\mathbf{O}(t_{R+i}) = \mathbf{e}_{3R+1}$. *The input and output vectors for* $t_{2R+i} \in \mathcal{T}_f$: $\mathbf{I}(t_{2R+i}) = \mathbf{e}_{2R+i} + \mathbf{e}_{3R+1}$ *and* $\mathbf{O}(t_{2R+i}) = \mathbf{e}_{R+i}$. *The input and output vectors for* $t_{ij} \in \mathcal{T}_g$: $\mathbf{I}(t_{ij}) = \mathbf{e}_{2R+i} + \mathbf{e}_{R+j}$ *and* $\mathbf{e}_j + \mathbf{e}_{R+i}$,
- $\mathbf{H}(t_i) = (0, \ldots, 0)$ *for* $t_i \in \mathcal{T}_q \cup \mathcal{T}_w \cup \mathcal{T}_f$ *and* $\mathbf{H}(t_{ij}) = \mathbf{e}_{3R+1}$ *for* $t_{ij} \in \mathcal{T}_g$,
- *for* $1 \leq i, j \leq R$ *let* $w(T_{R+i}, \mathbf{m}) = m_{R+i}\mu_i$, $w(t_i, \mathbf{m}) = m_i$, $w(t_{2R+i}, \mathbf{m}) = 1$ *and* $w(t_{ij}, \mathbf{m}) = m_{R+j}$,
- $\mathbf{m_0} = (0, \ldots, 0, k)$.

*Tokens arrive to places* $P_{2R+i}$, $1 \leq i \leq R$, *according to Poisson stochastic processes.*

Figure 3 shows a graphical model for $R = 2$ classes LCFSPR queue where dotted lines are introduce for the sake of readability and they do not have ant particular meaning. Note that when a token arrives to the place $P_{2R+i}$ it is temporally (i.e. the state is vanishing) stored in $P_{2R+i}$ and we have two cases:

- there is at least one free server, i.e. $m_{3R+1} > 0$, thus the customer goes immediately in service. This is modelled by the immediate transition set $\mathcal{T}_f$
- all the servers are busy, i.e. $m_{3R+1} = 0$, so a customer is preempted and put in queue and the new customer goes in service. This is modelled by $R^2$ transitions, $\mathcal{T}_g$. The inhibitor arcs are needed to avoid pre-emption when there is at least one free server.

**Lemma 3.** *Consider the sets of immediate transitions* $\mathcal{T}_q, \mathcal{T}_f, \mathcal{T}_g$. *Any two transitions belonging to two different sets cannot be simultaneously enabled. Moreover any two transitions of* $\mathcal{T}_f$ *cannot be enabled simultaneously, and if* $t_{ia} \in \mathcal{T}_g$ *is enabled then just transitions* $t_{ib} \in \mathcal{T}_g$ *with* $1 \leq b \leq R$ *can be enabled.*

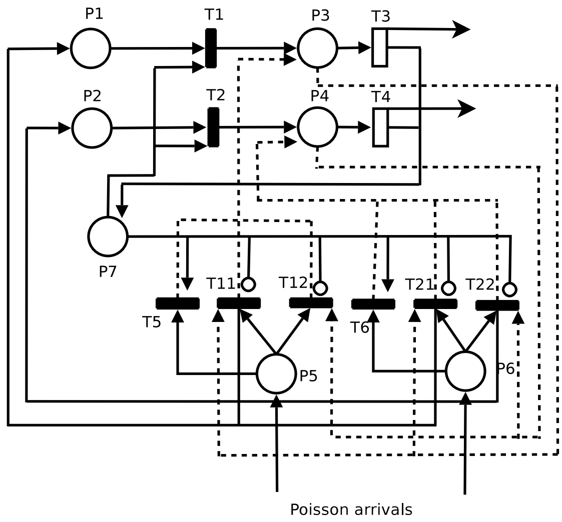


**Fig. 3.** Graphical representation of model GSPN-2.

The proof immediately derives by GSPN-2 structure.

As consequence to lemma 3 we can solve the conflicts on immediate transitions with just one simple function. When one or more transitions of $\mathcal{T}_q$ are enabled, the probability of firing for the $i$-th transition is:

$$p(t_i, \mathbf{m}) = p_i(\mathbf{m}) = \frac{m_i}{\sum_{l=1}^{R} m_l}. \tag{13}$$

When one or more transitions of $\mathcal{T}_g$ are enabled, the probability of firing is:

$$p(t_{ij}, \mathbf{m}) = p_{ij}(\mathbf{m}) = \frac{m_{R+j}}{\sum_{l=1}^{R} m_{R+l}}. \tag{14}$$

Now we can state a main lemma for model GSPN-2 representation:

**Lemma 4.** *Let $\mathbf{m} = (m_1, \ldots, m_{3R+1})$ be a reachable tangible marking of GSPN-2 model. Then if the stability condition holds, the stationary state probability can be written as follows:*

$$\pi(\mathbf{m}) = \pi_0 \prod_{i=1}^{R} \lambda_i^{m_i + m_{R+i}} \frac{(\sum_{i=1}^{R} m_i)!}{\prod_{i=1}^{R} m_i!} \frac{(\sum_{i=1}^{R} m_{R+i})!}{\prod_{i=1}^{R} m_{R+i}!} \prod_{i=1}^{R} \left(\frac{1}{\mu_i}\right)^{m_i + m_{R+i}} \cdot \prod_{j=1}^{\sum_{i=1}^{2R} m_i} \frac{1}{\min(j,k)}. \tag{15}$$

*where $\mu_i$ is the average service rate for one customer of class $i$ when there are no other customers in the system, $k$ is the number of servers, $\pi_0$ is a normalizing constant.*

The proof is given in appendix A.

**Theorem 2.** *Consider model GSPN-2 and let* $n_i = m_i + m_{R+i}$, $1 \le i \le R$ *and* $\mathbf{n} = (n_1, \ldots, n_R)$ *be an aggregated state. Let* $\pi_a(\mathbf{n})$ *be the steady state probability of* $n_i$ *for* $i = 1, \ldots, R$. *Then we can write:*

$$\pi_a(\mathbf{n}) = \pi_0 \frac{(\sum_{i=1}^R n_i)!}{\prod_{i=1}^R n_i!} \prod_{i=1}^R \lambda_i^{n_i} \prod_{i=1}^R \Big(\frac{1}{\mu_i}\Big)^{n_i} \prod_{i=1}^{\sum_{i=1}^R n_i} \frac{1}{\min(k,i)} \quad \forall \mathbf{n} \in \mathbf{N}^R. \tag{16}$$

*Proof.* The proof is based on the Vandermonde formula and it is similar the one given for Theorem 1. □

**Corollary 2.** *The M/M/k queuing system with LCFSPR discipline, R customer classes, arrival rates* $\lambda_i$, *single server rate* $\mu_i$ *for class i customers and steady state probability* $\pi'(\mathbf{n})$ *is equivalent to model GSPN-2 in terms of steady state probability, i.e.,* $\pi_a(\mathbf{n}) = \pi'(\mathbf{n})$ *for all* $\mathbf{n} \in \mathbf{N}^R$, *where* $\pi_a(\mathbf{n})$ *is the aggregated probability of GSPN given by formula (16). The normalizing constant is* $\pi_0 = \pi(0, \ldots, 0, k) = \pi'(0, \ldots, 0, k)$.

*Proof.* It follows immediately from queuing theory and Theorem 2. □

The net GSPN-2 can as well simulate a single server LCFSPR service station with a BCMP-like load dependent service rate. We can state the following lemma:

**Lemma 5.** *Let* $m'_r = m_r + m_{R+r}$, $y_r(m'_r)$ *an arbitrary positive function,* $\mathbf{m}_0 = \mathbf{e_{3R+1}}$ *and let the firing rate of transitions* $T_{R+1}, \ldots, T_{2R}$ *be* $w(T_{R+r}) = y_r(m'_r)$. *The if* $\mathbf{m}$ *is a reachable tangible marking, the steady state probability function is:*

$$\pi(\mathbf{m}) = \pi_0 \prod_{i=1}^R \lambda_i^{m_i + m_{R+i}} \frac{(\sum_{i=1}^R m_i)!}{\prod_{i=1}^R m_i!} \prod_{r=1}^R \Big[\Big(\frac{1}{\mu_r}\Big)^{m_r + m_{R+r}} \prod_{a=1}^{m_r + m_{R+r}} \frac{1}{y_r(a)}\Big]. \tag{17}$$

The proof is given in appendix.

By defining $n_i = m_i + m_{R+i}$ we can aggregate the states and we can prove that the steady state probability $\pi_a$ of the aggregated CTMC is identical to probability $\pi'$ defined by equation (7). For what concern the net structure complexity, the number of places grows as $\mathcal{O}(R)$ and the number of transitions grows as $\mathcal{O}(R^2)$.

# 6 Representing M/M/k/PS queue and M/M/∞ queue by GSPN

The processor sharing discipline can be easily represented considering that the $k$ processors are shared among the users in the system. Different classes of users can have different average time services, but all modelled by exponentially distributed random variables. We can think that the $k$ servers are shared among the $R$ classes in proportion to the number of customers of the classes.

**Definition 3 (GSPN-3).** *Let us define the model GSPN-3 as follows:*

- $\mathcal{P} = \{P_1, \ldots, P_R\}$,
- $\mathcal{T} = \{T_1, \ldots, T_R\}$,
- $\Pi(T_i) = 1$ *for each* $T_i \in \mathcal{T}$,
- $\mathbf{I}(T_i) = \mathbf{e}_i$ *and* $\mathbf{O}(T_i) = (0, \ldots, 0)$ *for each* $T_i \in \mathcal{T}$,
- $\mathbf{H}(T_i) = (0, \ldots, 0)$ *for each* $T_i \in \mathcal{T}$,
- $w(T_i, \mathbf{m}) = \frac{m_i}{m} \min(k, m)$ *where* $m = \sum_{j=1}^{R} m_j$ *for each* $T_i \in \mathcal{T}$,
- $\mathbf{m_0} = (0, \ldots, 0)$.

Figure 4 shows a graphical representation of the GSPN-3 model. Note that this

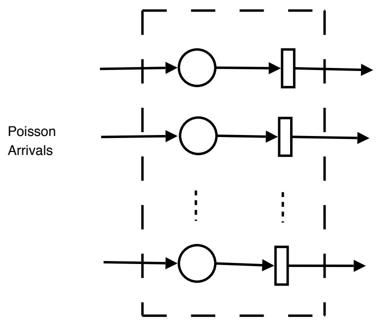


**Fig. 4.** Graphical representation of model GSPN-3

model is equivalent to a queuing system with PS discipline and one server with load-dependent exponential service time to simulate the multi-server feature. Therefore it immediately follows the theorem:

**Theorem 3.** *Consider model GSPN-3. Then if stability condition holds the stationary state probability can be written as follows:*

$$\pi(\mathbf{m}) = \pi_0 \frac{(\sum_{i=1}^{R} m_i)!}{\prod_{i=1}^{R} m_i!} \prod_{i=1}^{R} \lambda_i^{m_i} \prod_{i=1}^{R} \left(\frac{1}{\mu_i}\right)^{m_i} \prod_{i=1}^{\sum_{i=1}^{R} m_i} \frac{1}{\min(k, i)},$$

*where $\mu_i$ is the average service rate for one customer of class $i$ when there are no other customers in the system, $k$ is the number of servers, $\pi_0$ is a normalizing constant.*

This model is similar to the compact model introduced in [4], the only difference is that we allow a whole state dependent firing rate thus we don't need a place whose tokens represent the total number of customers in the system.

Model GSPN-3 can easily represent also the IS center. It suffices to set the firing rates of each transition $T_i$ as $m_i\mu_i$, $1 \leq i \leq R$.

## 7 M ⇒ M property on the GSPN representation

Markov implies Markov property is introduced and studied by Muntz [17]. In that paper he shows that if a queuing system with Poisson arrivals presents departures according to a Poisson process (M⇒M property) then a combination of service centers of this type in a queuing network has a product-form solution. As we are considering GSPNs we will prove that a combination of GSPN-1, GSPN-2 and GSPN-3 models still holds a closed-form steady state probability by defining appropriate traffic processes over the CTMC associated to each of the models and using the results given in [16] which generalize Muntz's work. We now briefly review Melamed's results limited to a CTMC in steady state. Consider an ergodic CTMC with state space $\varGamma$ and a set of traffic transitions denoted by $\varTheta_1, \ldots, \varTheta_R$, where $\varTheta_i \subseteq \varGamma \times \varGamma$, $\varTheta_i \neq \emptyset$. Let us define $K_i(t)$ as the process which counts the number of transitions $(\alpha, \beta) \in \varTheta_i$ up to $t$. Let $m_i = \sum_{\gamma \in \varGamma} \sum_{\eta \in \varTheta_i(\cdot,\gamma)} \pi(\eta)\xi(\eta \to \gamma)$ and for each state $\gamma \in \varGamma$ let $m_i(\gamma) = \sum_{\eta \in \varTheta_i(\cdot,\gamma)} \pi(\eta)\xi(\eta \to \gamma)$ where $\varTheta_i(\cdot, \gamma) = \{\beta | (\beta, \gamma) \in \varTheta_i\}$ and $\xi(\eta \to \gamma)$ is the transition rate between states $\eta$ and $\gamma$.
Then we can state that $K_i(t)$ are mutually independent Poisson processes if and only if the following equation holds:

$$\forall \gamma \in \varGamma, \quad \sum_{i=1}^{R} m_i(\gamma) = \pi(\gamma) \sum_{i=1}^{R} m_i \tag{18}$$

We aim to study the departure traffic processes from our models. Take for example model GSPN-1, we can define $R$ traffic processes as follows:

$$\varTheta_i = \{(\mathbf{m}', \mathbf{m}) : |\mathbf{m}'|_i = |\mathbf{m}|_i + 1\}, \; i = 1, \ldots, R, \tag{19}$$

where $|\mathbf{m}|_i = m_i + m_{R+i}$. In our case, in order to prove that $K_i(t)$ are independent Poisson processes when there are Poisson arrivals, it suffices to prove that:

$$\forall \gamma \in \varGamma, \quad \sum_{\eta \in \varTheta_i(\cdot,\gamma)} \pi(\eta)\xi(\eta \to \gamma) = \lambda_i \pi(\gamma), \tag{20}$$

In appendix we prove that this condition holds for GSPN-1, GSPN-2 and GSPN-3 models by defining appropriate traffic processes. As observed in [16] this property of the CTMC is equivalent to the $M \implies M$ given by Muntz thus it assures that a BCMP-like composition of these GSPN models holds a closed-form steady state probability function. Random switches between the blocks and user class switches can be easily modelled by immediate transitions.

## 8 Final remarks

In this paper we have shown how to represent multi-class single queuing systems by structurally finite GSPN for various queuing disciplines. For each of the BCMP center types we have introduced a GSPN model whose steady state probability, aggregating on the number of customers in the system for each class, is equal to the correspondent single queue service center. Hence the two models are equivalent in terms of steady state distribution and average performance indexes. The main advantages of our representation are the following.

– We define a finite GSPN model. The abstraction level of the GSPN model allows the representation of the queuing behavior without introducing a high level of details in the state specification. We distinguish the customers waiting in the queue from those being served without taking in account the arrival order. This allows, as well as a finite representation, a steady state probability which is less detailed than the proposed in [4] which considers the single station detailed representation with the order of the customers in the queue, similarly to the BCMP paper [5]. On the other side the models we propose are more detailed than those which just consider the total number of customers in a center as the compact models of [4].
– The FCFS and the LCFSPR (or PS) scheduling disciplines have different GSPN representations and the FCFS can not be used to represent the other ones if the service depends on the customer class. The GSPN models simulate the corresponding queuing system even if their semantic is different.

The main idea of the definition of the GSPN models is the way we represent the customer of the queuing system which gets the free server and the customer which looses a server in case of preemption. In both cases we model the customer choice of a class $i$ with a random selection, according to the probability proportional to the number of customers in queue (or being served) of that class over the total number of customers in queue (number of servers).

The M⇒M property allows us to state that a combination of GSPN-1, GSPN-2 and GSPN-3 models similar to the service centers combination in BCMP networks, has a simple closed form steady state probability. In [1] authors define a queuing center isomorphic to GSPN-1 and show how it can be embedded in a BCMP queuing network so that the steady state probability function of the network does not change. In the GSPN formalism probabilistic routing can be easily simulated by introducing a block with a place and an immediate transition for each possible route just after the timed transitions of the models.

Further research deals with the extension of the proposed LCFSPR model to Coxian service time distributions and the definition of algorithms to identify GSPN which are compositions of models GSPN-1, GSPN-2, GSPN-3 and in order to obtain efficiently a set of significant performace indexes.

# A Appendix

Hereafter, for the sake of simplicity, by referring to the relations between the queueing system and GSPN models, we refer to and mix the notation of queuing theory and Petri nets. For example, we can say *waiting customers* to refer to tokens in place $P_1, .., P_R$ of GSPN-1 or GSPN-2, *customers being served* to refer to the tokens in $P_{R+1}, \ldots, P_{2R}$ of GSPN-1 or GSPN-2, *customer arrival* to refer to token arrival, *free servers* to refer to the presence of tokens in the place $P_{2R+1}$ or $P_{3R+1}$ in the GSPN-1 or GSPN-2 respectively. In our view this description should make the following proofs easier to understand.

### A.1 Proof of Lemma 1

*Proof.* We consider four cases.

Case 1) Suppose that all servers are busy and at least one customer is waiting, so $m_{2R+1} = 0$ and $m_i > 0$ for some $i = 1, \ldots, R$. Consider state $\mathbf{m} = (m_1, \ldots, m_R, m_{R+1}, \ldots, m_{2R}, 0)$. Clearly we have that $\sum_{i=R+1}^{2R} m_i = k$ that is the number of servers. Consider the tangible markings from which $\mathbf{m}$ is reachable possibly through the firing of immediate transitions combined with timed ones and the corresponding transition rates::

- $\mathcal{M}_\alpha = \{\mathbf{m'} = \mathbf{m} - \mathbf{e_i} | m_i > 0\}$ with rate $\lambda_i$
- $\mathcal{M}_\beta = \{\mathbf{m'} = \mathbf{m} + \mathbf{e_i} | m_{R+i} > 0\}$ with rate $(m_{R+i})\mu p_i(\mathbf{m'})$;
- $\mathcal{M}_\gamma = \{\mathbf{m'} = \mathbf{m} + \mathbf{e_i} - \mathbf{e_{R+i}} + \mathbf{e_{R+j}} | m_{R+i} > 0, j \neq i\}$ with rate $(m_{R+j} + 1)\mu p_i(\mathbf{m'})$.

The set of markings reachable from $\mathbf{m}$ and the corresponding transition rates can be classified as follows:

- $\mathcal{M}_a = \{\mathbf{m'} = \mathbf{m} + \mathbf{e_i}\}$ with rate $\lambda_i$;
- $\mathcal{M}_b = \{\mathbf{m'} = \mathbf{m} - \mathbf{e_i} | m_{R+i} > 0, m_i > 0\}$ with rate $(m_{R+i})\mu p_i(\mathbf{m})$;
- $\mathcal{M}_c = \{\mathbf{m'} = \mathbf{m} - \mathbf{e_i} + \mathbf{e_{R+i}} - \mathbf{e_{R+j}} | m_{R+j} > 0, m_i > 0, i \neq j\}$ with rate $(m_{R+j})\mu p_i(\mathbf{m})$.

We prove that equation (9) satisfies the global balance equations (GBE) by checking that the effective arrival rate to a state due to a new customer is equal to the effective leaving rate due to a job completion, and the effective arrival rate due to a job completion is equal to the effective leaving due to new customer arrival:

$$\sum_{\mathbf{m'} \in M_\alpha} \pi(\mathbf{m'})\xi(\mathbf{m'}, \mathbf{m}) = \pi(\mathbf{m}) \sum_{\mathbf{m'} \in \mathcal{M}_b \cup \mathcal{M}_c} \xi(\mathbf{m}, \mathbf{m'}) \wedge$$
$$\sum_{\mathbf{m'} \in \mathcal{M}_\beta \cup \mathcal{M}_\gamma} \pi(\mathbf{m'})\xi(\mathbf{m'}, \mathbf{m}) = \pi(\mathbf{m}) \sum_{\mathbf{m'} \in \mathcal{M}_a} \xi(\mathbf{m}, \mathbf{m'}) \implies$$
$$\sum_{\mathbf{m'} \in \mathcal{M}_\alpha \cup \mathcal{M}_\beta \cup \mathcal{M}_\gamma} \pi(\mathbf{m'})\xi(\mathbf{m'}, \mathbf{m}) = \pi(\mathbf{m}) \sum_{\mathbf{m'} \in \mathcal{M}_a \cup \mathcal{M}_b \cup \mathcal{M}_c} \xi(\mathbf{m}, \mathbf{m'})$$

where $\xi(\mathbf{a}, \mathbf{b})$ represents the transition rate from marking $\mathbf{a}$ to marking $\mathbf{b}$.

The effective leaving rate from state $\mathbf{m}$ due to completion of a job is:

$$\pi(\mathbf{m})\Big[\sum_{m_{j+k} > 0} (m_{j+k})\mu\Big] = \pi(\mathbf{m})(k\mu).$$

The effective arrival rate to state $\mathbf{m}$ due to arrivals to the system is:

$$\sum_{\mathbf{m}'\in\mathcal{M}_\alpha}\pi(\mathbf{m}')\xi(\mathbf{m}',\mathbf{m}) = \pi(\mathbf{m})\Big[\sum_{i\in X}(\frac{1}{\lambda_i}\frac{m_i}{\sum_{j=1}^R m_j}\mu(\sum_{j=1}^{2R} m_j)\lambda_i)\Big]$$
$$= \pi(\mathbf{m})(k\mu)\Big[\sum_{i\in X}\frac{m_i}{\sum_{j=1}^R m_j}\Big] = \pi(\mathbf{m})(k\mu),$$

where $X = \{i|m_i > 0, 1 \le i \le R\}$. Note that by hypothesis $\sum_{i=1}^{2R} m_i \ge k$ thus:

$$\mu(\sum_{i=1}^{2R} m_i) = k\mu.$$

Consider state $\mathbf{m}$ and set $Y = \{j|m_{R+j} > 0, 1 \le j \le R\}$. The effective leaving rate from state $\mathbf{m}$ due to arrivals of customers to the system is $\pi(\mathbf{m})\sum_{i=1}^R \lambda_i$. The effective arrival rate to state $\mathbf{m}$ is:

$$\pi(\mathbf{m})\Big[\sum_{j\in Y}\lambda_j\frac{(\sum_{i=1}^R m_i)+1}{m_j+1}\frac{1}{\mu((\sum_{i=1}^{2R} m_i)+1)}m_{R+j}\mu\frac{m_j+1}{\sum_{i=1}^R m_i+1}$$
$$+\sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j\frac{m_{R+i}}{m_{R+j}+1}\frac{(\sum_{g=1}^R m_g)+1}{m_i+1}\frac{1}{\mu((\sum_{g=1}^{2R} m_g)+1)}(m_{R+j}+1)\mu$$
$$\cdot\frac{m_i+1}{(\sum_{g=1}^R m_g)+1}\Big]$$
$$= \pi(\mathbf{m})\Big[\sum_{j\in Y}\lambda_j\frac{1}{k\mu}m_{R+j}\mu + \sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j m_{R+i}\frac{1}{k\mu}\mu\Big]$$
$$= \pi(\mathbf{m})\Big[\sum_{j\in Y}\lambda_j\frac{m_{R+j}}{k} + \sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j\frac{m_{R+i}}{k}\Big]$$
$$= \pi(\mathbf{m})\Big[\sum_{i=1}^R\lambda_i\Big].$$

Case 2) Consider now the case when all servers are busy, but places $P_1,\ldots,P_R$ are empty. Consider a generic state $\mathbf{m} = (0,\ldots,0,m_{R+1},\ldots,m_{2R},0)$. The markings from which $\mathbf{m}$ is reachable are classified as:

- $\mathcal{M}_\alpha = \{\mathbf{m}' = \mathbf{m} - \mathbf{e_{R+i}} + \mathbf{e_{2R+1}}|m_{R+i} > 0\}$ with rate $\lambda_i$;
- $\mathcal{M}_\beta = \{\mathbf{m}' = \mathbf{m} + \mathbf{e_i}|m_{R+i} > 0\}$ with rate $(m_{R+i})\mu p_i(\mathbf{m}')$;
- $\mathcal{M}_\gamma = \{\mathbf{m}' = \mathbf{m} + \mathbf{e_i} - \mathbf{e_{R+i}} + \mathbf{e_{R+j}}|m_{R+i} > 0, j \neq i\}$ with rate $(m_{R+j} + 1)\mu p_i(\mathbf{m}')$.

The markings reachable from $\mathbf{m}$ and their effective rates are the same as in case 1). Let $Y$ be the set defined in case 1). We now prove that the effective arrival rate to state $\mathbf{m}$ from markings in $\mathcal{M}_\alpha$ is equal to the effective leaving rate from $\mathbf{m}$ due to job completion which is $\pi(\mathbf{m})(k\mu)$.

$$\pi(\mathbf{m})\Big[\sum_{i\in Y}(\frac{1}{\lambda_i}\frac{m_{R+i}}{k}\mu(k)\lambda_i)\Big] = \pi(\mathbf{m})\Big[\sum_{i\in Y} m_{R+i}\mu\Big] = \pi(\mathbf{m})(k\mu).$$

We prove now that the effective arrival rate from the markings in $\mathcal{M}_\beta \cup \mathcal{M}_\gamma$ is equal to the effective leaving rate from $\mathbf{m}$.

$$\pi(\mathbf{m})\Big[\sum_{j\in Y}(\lambda_j \frac{1}{\mu(k+1)} m_{R+j}\mu) + \sum_{j=1}^{R}\sum_{\substack{i\in Y\\ i\neq j}} \lambda_j \frac{m_{R+i}}{m_{R+j}+1}\frac{1}{\mu(k+1)}(m_{R+j}+1)\mu\Big]$$

$$= \pi(\mathbf{m})\Big[\sum_{j\in Y}\lambda_j\frac{m_{R+j}}{k} + \sum_{j=1}^{R}\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j\frac{m_{R+i}}{k}\Big] = \pi(\mathbf{m})\Big[\sum_{j=1}^{R}\lambda_j\Big].$$

Case 3) Assume that there are not tokens in places $P_1, \dots, P_R$ and at least one server is free and at least one is busy, that is $0 < m_{2R+1} < k$. Then the effective leaving rate from $\mathbf{m}$ is simply $\pi(\mathbf{m})[\mu\sum_{j=1}^{R} m_{R+j} + \sum_{j=1}^{R}\lambda_j]$. The states from which $\mathbf{m}$ is reachable and the corresponding rates are:

- $\mathcal{M}_\alpha = \{\mathbf{m'} = \mathbf{m} - \mathbf{e_{R+i}} + \mathbf{e_{2R+1}} | m_{R+i} > 0\}$ with rate $\lambda_i$;
- $\mathcal{M}_\beta = \{\mathbf{m'} = \mathbf{m} + \mathbf{e_{R+i}} - \mathbf{e_{2R+1}}\}$ with rate $(m_{R+i}+1)\mu$.

Let $Y$ be defined as in case 1). We now prove that the effective arrival rate to state $\mathbf{m}$ is equal to the effective leaving rate from $\mathbf{m}$.

$$\begin{aligned}
&\pi(\mathbf{m})\Big[\sum_{i\in Y}\frac{1}{\lambda_i}\frac{m_{R+i}}{\sum_{j=1}^{R} m_{R+j}}\mu(\sum_{j=1}^{R} m_{R+j})\lambda_i \\
&\quad + \sum_{i=1}^{R}\lambda_i\frac{\sum_{j=1}^{R} m_{R+j}+1}{m_{R+i}+1}\frac{1}{\mu(\sum_{j=1}^{R} m_{R+j}+1)}(m_{R+i}+1)\mu\Big] \\
=&\pi(\mathbf{m})\Big[\sum_{i\in Y}\frac{1}{\lambda_i}\frac{m_{R+i}}{\sum_{j=1}^{R} m_{R+j}}(\sum_{j=1}^{R} m_{R+j})\mu\lambda_i \\
&\quad + \sum_{i=1}^{R}\lambda_i\frac{\sum_{j=1}^{R} m_{R+j}+1}{m_{R+i}+1}\frac{1}{(\sum_{j=1}^{R} m_{R+j}+1)\mu}(m_{R+i}+1)\mu\Big] \\
=&\pi(\mathbf{m})\Big[(\sum_{j=1}^{R} m_{R+j})\mu + \sum_{i=1}^{R}\lambda_i\Big].
\end{aligned}$$

Case 4) Assume that the system is empty, that is $m_{2R+1} = k$. This case is trivial. The effective leaving rate from $\mathbf{m}$ is clearly $\pi(\mathbf{m})[\sum_{i=1}^{R}\lambda_i]$. The effective arrival rate to $\mathbf{m}$ can just be due to a job completion and it is easy to show that is is equal to the leaving rate.

$\square$

## A.2 Proof of Lemma 2

*Proof.* The proof verifies that equation (12) satisfies the set of GBE for the CTMC associated to the net. We consider two cases.

Case1) Let

$$\mathbf{m} = (m_1, \ldots, m_R, 0, \ldots, 0) + \mathbf{e_{R+i}}$$

be a reachable tangible marking. State $\mathbf{m}$ can be reached from states $\mathbf{m}' = \mathbf{m} + \mathbf{e_{r+j}} - \mathbf{e_{r+i}} + \mathbf{e_i}$ due to a job completion with rate $x(1 + \sum_{r=1}^{2R} m_r)\mu p_i(\mathbf{m}')$. So the total effective arrival rate due to a job completion is:

$$\pi(\mathbf{m})\Big[\sum_{j=1}^{R} \lambda_j \frac{1 + \sum_{r=1}^{R} m_r}{m_i + 1} \frac{1}{x(1 + \sum_{r=1}^{2R} m_r)\mu} x(1 + \sum_{r=1}^{2R} m_r)\mu \frac{m_i + 1}{1 + \sum_{r=1}^{2R} m_r}\Big]$$
$$= \pi(\mathbf{m}) \sum_{j=1}^{R} \lambda_j,$$

which is the total leaving rate from $\mathbf{m}$ due to arrivals to the system. State $\mathbf{m}$ can be reached from states $\mathbf{m}' = \mathbf{m} - \mathbf{e_j}$ where $j \in X = \{j | m_j > 0, 1 \leq j \leq R\}$ with rate $\lambda_j$. So the effective arrival rate due to an arrival to the system is:

$$\pi(\mathbf{m})\Big[\sum_{j \in X} \frac{1}{\lambda_j} \frac{m_j}{\sum_{k=1}^{R} m_k} x(\sum_{k=1}^{2R} m_k)\mu\lambda_j\Big] = \pi(\mathbf{m})\Big[x(\sum_{k=1}^{2R} m_k)\mu\Big],$$

which is identical to the total leaving rate due to a job completion.

Case 2) If $m_{2R+1} = 1$ the proof is trivial.

□

## A.3 Proof of Lemma 4

*Proof.* The proof is similar to the proof of Lemma 1. We prove that equation (15) satisfies the GBE of the CTMC associated to GSPN-2 by verifying the partial balance. We consider again four cases.

Case 1) Take a reachable marking $\mathbf{m}$ with $m_{3R+1} = 0$ (i.e. all servers are busy), and there's at least one token in one of the places $P_i$, with $1 \leq i \leq R$ (i.e. there is at least one customer in queue). Let $X = \{i | m_i > 0, 1 \leq i \leq R\}$ and $Y = \{i | m_{R+i} > 0, 1 \leq i \leq R\}$. We verify that the effective arrival rate to state $\mathbf{m}$ due to a job completion is equal to the effective leaving rate from $\mathbf{m}$ due to a customer arrival and that the effective arrival rate to state $\mathbf{m}$ due to a customer arrival is equal to the effective leaving rate due to a job completion.

State $\mathbf{m}$ is reachable due to an arrival event from states $\mathbf{m}' = \mathbf{m} - \mathbf{e_i}$ with $i \in X \cap Y$ with rate $\lambda_i p_{ii}(\mathbf{m}')$ or from states $\mathbf{m}'' = \mathbf{m} - \mathbf{e_i} + \mathbf{e_{R+i}} - \mathbf{e_{R+j}}$ with

$i \in X$, $j \in Y$ and $i \neq j$ with rate $\lambda_j p_{ji}(\mathbf{m}'')$ Hence we can write:

$$\pi(\mathbf{m})\Big[\sum_{i\in X\cap Y}\frac{1}{\lambda_i}\frac{m_i}{\sum_{s=1}^R m_s}\mu_i k\lambda_i m_{R+i}k$$

$$+\sum_{j\in Y}\sum_{\substack{i\in X\\ i\neq j}}\frac{1}{\lambda_j}\frac{m_i}{\sum_{s=1}^R m_s}\frac{m_{R+j}}{m_{R+i}+1}\mu_j k\lambda_j\frac{m_{R+i}+1}{k}\Big]$$

$$=\pi(\mathbf{m})\Big[\frac{1}{\sum_{s=1}^R m_s}\sum_{j\in X\cap Y}m_j m_{R+j}\mu_j+\frac{1}{\sum_{s=1}^R m_s}\sum_{j\in Y}\mu_j m_{R+j}\sum_{\substack{i\in X\\ j\neq i}}m_i\Big]$$

$$=\pi(\mathbf{m})\Big[\frac{1}{\sum_{s=1}^R m_s}\sum_{j=1}^R\mu_j m_{R+j}\sum_{i\in X}m_i\Big]=\pi(\mathbf{m})\Big[\sum_{j=1}^R\mu_j m_{R+j}\Big],$$

which is identical to the effective leaving rate from $\mathbf{m}$ due to a job completion.

State $\mathbf{m}$ is reachable, due to job completion, from states $\mathbf{m'} = \mathbf{m} + \mathbf{e_i}$ with rate $m_{R+i}\mu_i p_i(\mathbf{m}')$ and from states $\mathbf{m}'' = \mathbf{m} + \mathbf{e_i} - \mathbf{e_{R+i}} + \mathbf{e_{R+j}}$ with rate $m_{R+j}\mu_j p_i(\mathbf{m}'')$. Hence the effective arrival rate due to a job completion can be written as:

$$\pi(\mathbf{m})\Big[\sum_{i\in Y}\lambda_i\frac{(\sum_{j=1}^R m_j)+1}{m_i+1}\frac{1}{k\mu_i}m_{R+i}\mu_i\frac{m_i+1}{\sum_{j=1}^R m_j+1}$$

$$+\sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j\frac{(\sum_{k=1}^R m_k)+1}{m_i+1}\frac{m_{R+i}}{m_{R+j}+1}\frac{1}{k\mu_j}(m_{R+j}+1)\mu_j\frac{m_i+1}{(\sum_{k=1}^R m_k)+1}\Big]$$

$$=\pi(\mathbf{m})\Big[\sum_{i\in Y}\lambda_i\frac{m_{R+i}}{k}+\sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\lambda_j\frac{m_{R+i}}{k}\Big]=\pi(\mathbf{m})\Big[\sum_{i=1}^R\lambda_i\Big],$$

which is identical to the effective leaving rate form state $\mathbf{m}$ due to a customer arrival to the system.

Case 2) Consider now state $\mathbf{m}$ with $m_{3R+1} = 0$ and $m_i = 0$ with $1 \leq i \leq R$ (i.e. no customers in queue). The leaving rates from $\mathbf{m}$ for a job completion or an arrival are the same as case 1. State $\mathbf{m}$ is reachable, due to a job completion, form states $\mathbf{m}' = \mathbf{m} + \mathbf{e_i} - \mathbf{e_{R+i}} + \mathbf{e_{R+j}}$ with $i \in Y$ and $j \neq i$ with rate $(m_{R+j} + 1)\mu_j p_{ji}(\mathbf{m}')$ and from states $\mathbf{m}'' = \mathbf{m} + \mathbf{e_i}$ with $i \in Y$ and rate $m_{R+i}\mu_i p_{ii}(\mathbf{m}'')$. Hence the effective arrival rate to $\mathbf{m}$ due to a job completion can be calculate as follows:

$$\pi(\mathbf{m})\Big[\sum_{i\in Y}\lambda_i\frac{1}{k\mu_i}m_{R+i}\mu_i+\sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}\frac{m_{R+i}}{m_{R+j}+1}\frac{1}{k\mu_j}(m_{R+j}+1)\mu_j\Big]=$$

$$=\pi(\mathbf{m})\Big[\frac{1}{k}\sum_{i\in Y}\lambda_i m_{R+i}+\frac{1}{k}\sum_{j=1}^R\sum_{\substack{i\in Y\\ i\neq j}}m_{R+i}\lambda_j\Big]=\pi(\mathbf{m})\Big[\sum_{j=1}^R\lambda_j\Big],$$

which is identical to the effective leaving rate from $\mathbf{m}$ due to an arrival to the system. State $\mathbf{m}$ is reachable, due to an arrival to the system, from states $\mathbf{m'} = \mathbf{m} - \mathbf{e_i} + \mathbf{e_{3R+1}}$ with $i \in Y$ with rate $\lambda_i$. Hence the effective arrival rate to state $\mathbf{m}$ due to a customer arrival can be calculated as follows:

$$\pi(\mathbf{m})\Big[\sum_{i\in Y}\frac{1}{\lambda_i}\frac{m_{R+i}}{\sum_{s=1}^{R}m_{R+s}}\mu_i k\lambda_i\Big] = \pi(\mathbf{m})\Big[\sum_{i\in Y}m_{R+i}\mu_i\Big] = \pi(\mathbf{m})\Big[\sum_{i=1}^{R}m_{R+i}\mu_i\Big],$$

which is identical to the effective leaving rate from $\mathbf{m}$ due to a job completion.

Case 3) Consider the case of $\mathbf{m}$ with $1 \leq m_{3R+1} = b < k$. State $\mathbf{m}$ is reachable, due to an arrival, from states $\mathbf{m'} = \mathbf{m} - \mathbf{e_{R+i}} + \mathbf{e_{3R+1}}$ with rate $i \in Y$ with rate $\lambda_i$, thus the effective arrival rate due to a customer arrival is:

$$\pi(\mathbf{m})\Big[\sum_{i\in Y}\frac{1}{\lambda_i}\frac{m_{R+i}}{\sum_{s=1}^{e}m_{R+s}}\mu_i(k-b)\lambda_i\Big] = \pi(\mathbf{m})\Big[\sum_{i=1}^{R}m_{R+i}\mu_i\Big],$$

which is identical to the effective leaving rate from $\mathbf{m}$ due to a job completion. State $\mathbf{m}$ is reachable, due to a job completion, from states $\mathbf{m'} = \mathbf{m} + \mathbf{e_{R+i}} - \mathbf{e_{3R+1}}$ with rate $(m_{R+i}+1)\mu_i$, thus the effective arrival rate due to a job completion is:

$$\pi(\mathbf{m})\Big[\sum_{i=1}^{R}\lambda_i\frac{(\sum_{s=1}^{R}m_{R+s}}{m_{R+i}+1}\frac{1}{\mu_i}\frac{1}{k-b+1}\Big](m_{R+i}+1)\mu_i\Big] = \pi(\mathbf{m})\sum_{i=1}^{R}\lambda_i,$$

which is identical to the effective leaving rate from $\mathbf{m}$ due to an arrival to the system.

Case 4) considers $m_{3R+1} = k$, i.e. when the system is empty, and it is trivial.
□

### A.4 Proof of Lemma 5

*Proof.* The proof verifies that equation (17) satisfies the set of GBE for the CTMC associated to the net. We consider two cases.

Case 1) Take the tangible reachable marking $\mathbf{m} = (m_1, \ldots, m_R, 0, \ldots, 0) + \mathbf{e_{R+i}}$. State $\mathbf{m}$ can be reached from states $\mathbf{m'} = \mathbf{m} + \mathbf{e_{r+j}} - \mathbf{e_{r+i}} + \mathbf{e_i}$ due to a job completion with rate $y_j(m_j + m_{R+j})\mu_j p_i(\mathbf{m'})$. So the effective arrival rate to $\mathbf{m}$ due to a job completion is:

$$\pi(\mathbf{m})\Big[\sum_{j=1}^{R}\lambda_j\frac{\sum_{k=1}^{R}m_k+1}{m_i+1}\frac{1}{y_j(m_j+1)\mu_j}j_j(m_j+1)\mu_j\frac{m_i+1}{\sum_{k=1}^{R}m_k+1}\Big]$$
$$= \pi(\mathbf{m})\sum_{j=1}^{R}\lambda_j,$$

which is identical to the total leaving rate from $\mathbf{m}$ due to arrivals to the system. State $\mathbf{m}$ can be reached from state $\mathbf{m'} = \mathbf{m} + \mathbf{e_{r+j}} - \mathbf{e_j} - \mathbf{e_{R+ki}}$ where $j \in X =$

$\{j|m_j > 0, 1 \leq j \leq R\}$ with rate $\lambda_j$. So the effective arrival rate to $\mathbf{m}$ due to an arrival to the system is:

$$\pi(\mathbf{m})\Big[\sum_{j\in X}\frac{1}{\lambda_i}\frac{m_j}{\sum_{k=1}^{R} m_k}y_i(m_i+m_{R+i})\mu_i\lambda_i\Big] = \pi(\mathbf{m})\Big[y_i(m_i+m_{R+i})\mu_i\Big],$$

which is identical to the total leaving rate from $\mathbf{m}$ due to a job completion.

Case 2) If $m_{3R+1} = 1$ the proof is trivial.

□

## A.5 Proof of Independence of the departure processes from GSPN models

### GSPN-1:

*Proof.* Consider model GSPN-1. Let $\Theta_i = \{(\mathbf{m}', \mathbf{m}) : |\mathbf{m}'|_i = |\mathbf{m}|_i + 1\}$ for $i = 1, \ldots, R$. In order to verify equation (20) consider a generic tangible markinf $\mathbf{m}$. We consider the following cases: 1) $m_{2R+1} = 0$, and 2) $m_{2R+1} > 0$.

Case 1) Let $\mathbf{m}$ be a reachable tangible state with $m_{2R+1} = 0$, then:

$$\begin{aligned}\Theta_i(\cdot, \mathbf{m}) = \{\mathbf{m}'|\mathbf{m}' = \mathbf{m} + \mathbf{e_{R+i}} + \mathbf{e_j} - \mathbf{e_{R+j}},\ m_{R+j} > 0, j \neq i\} \\ \cup \{\mathbf{m}'|\mathbf{m}' = \mathbf{m} + \mathbf{e_i}, m_{R+i} > 0\}, \quad 1 \leq i, j \leq R.\end{aligned}$$

Let $\mathrm{sgn}(m_i)$ be the indicator function defined as follows: $\mathrm{sgn}(m_i) = 1$ when $m_i > 0$ and 0 otherwise, and let $Y = \{j|m_{R+j} > 0,\ 1 \leq j \leq R\}$. The left hand side of equation (20) can be rewritten as follows:

$$\begin{aligned}&\pi(\mathbf{m})\Big[\sum_{\substack{j\in Y\\ j\neq i}}\lambda_i\frac{1+\sum_{a=1}^{R} m_a}{m_j+1}\frac{m_{R+j}}{m_{R+i}+1}\frac{1}{k\mu}(m_{R+i}+1)\mu\frac{m_j+1}{1+\sum_{a=1}^{R} m_a} \\ &+ \mathrm{sgn}(m_i)\lambda_i\frac{1+\sum_{a=1}^{R} m_a}{m_i+1}\cdot\frac{1}{k\mu}m_{R+i}\mu\frac{m_i+1}{1+\sum_{a=1}^{R} m_a}\Big] \\ &= \pi(\mathbf{m})\Big[\lambda_i\Big(\sum_{\substack{j\in Y\\ j\neq i}}\frac{m_{R+j}}{k}+\frac{m_{R+i}}{k}\Big)\Big] = \pi(\mathbf{m})\lambda_i,\end{aligned}$$

which gives the right hand side of equation (20).

Case 2) Le $\mathbf{m}$ be a reachable tangible state with $m_{2R+1} > 0$. Then $\Theta_i(\cdot, \mathbf{m}) = \{\mathbf{m} + \mathbf{e_{R+i}} - \mathbf{e_{2R+1}}\}$. Thus equation (20) holds, in fact:

$$\pi(\mathbf{m})\Big[\lambda_i\frac{1+\sum_{a=1}^{R} m_{R+a}}{m_{R+i}+1}\frac{1}{(1+\sum_{a=1}^{R} m_{R+a})\mu}(1+m_{R+i})\mu\Big] = \pi(\mathbf{m})\lambda_i.$$

This proves that the traffic processes associated to $\Theta_i$, $1 \leq i \leq R$, are pointwise independent Poisson processes, i.e., the departure processes for each class of customers are Poisson independent processes under independent Poisson arrivals.

□

**GSPN-2:**

*Proof.* Consider model GSPN-2. In this model $|\mathbf{m}|_i = m_i + m_{R+i}$, with $1 \leq i \leq R$ and let $\Theta_i = \{(\mathbf{m}', \mathbf{m}) : |\mathbf{m}'|_i = |\mathbf{m}|_i + 1\}$. In order to verify equation (20) conside a generic tangible state $\mathbf{m}$. We consider the following cases: 1) $m_{3R+1} = 0$, and 2) $m_{3R+1} > 0$.

Case 1) Let $\mathbf{m}$ be a reachable tangible state with $m_{3R+1} = 0$, then:

$$\begin{aligned}\Theta_i(\cdot, \mathbf{m}) = \{\mathbf{m}' | \mathbf{m}' = \mathbf{m} + \mathbf{e_{R+i}} + \mathbf{e_j} - \mathbf{e_{R+j}},\ m_{R+j} > 0, j \neq i\} \\ \cup \{\mathbf{m}' | \mathbf{m}' = \mathbf{m} + \mathbf{e_i} | m_{R+i} > 0\} \quad 1 \leq i, j \leq R.\end{aligned}$$

Let $\mathrm{sgn}(m_i)$ be the indicator function defined as follows: $\mathrm{sgn}(m_i) = 1$ when $m_i > 0$ and $\mathrm{sgn}(m_i) = 0$ otherwise, and let $Y = \{j | m_{R+j} > 0,\ 1 \leq j \leq R\}$. The left hand side of equation (20) can be rewritten as follows:

$$\begin{aligned}&\pi(\mathbf{m})\Big[\sum_{\substack{j\in Y\\ j\neq i}} \lambda_i \frac{1+\sum_{a=1}^{R} m_a}{m_j+1} \frac{m_{R+j}}{m_{R+i}+1} \frac{1}{k\mu_i}(m_{R+i}+1)\mu_i \frac{m_j+1}{1+\sum_{a=1}^{R} m_a} \\ &+ \mathrm{sgn}(m_i)\lambda_i \frac{1+\sum_{a=1}^{R} m_a}{m_i+1} \frac{1}{k\mu_i} m_{R+i}\mu_i \frac{m_i+1}{1+\sum_{a=1}^{R} m_a}\Big] \\ =&\pi(\mathbf{m})\Big[\lambda_i\Big(\sum_{\substack{j\in Y\\ j\neq i}} \frac{m_{R+j}}{k} + \frac{m_{R+i}}{k}\Big)\Big] = \pi(\mathbf{m})\lambda_i,\end{aligned}$$

which is the right hand side of equation (20).

Case 2) Le $\mathbf{m}$ be a reachable tangible state with $m_{3R+1} > 0$. Then $\Theta_i(\cdot, \mathbf{m}) = \{\mathbf{m} + \mathbf{e_{R+i}} - \mathbf{e_{3R+1}}\}$. Thus equation (20) holds, in fact:

$$\pi(\mathbf{m})[\lambda_i \frac{1+\sum_{a=1}^{R} m_a}{m_{R+i}+1} \frac{1}{(m_{R+i}+1)\mu_i}(m_{R+i}+1)\mu_i] = \pi(\mathbf{m})\lambda_i.$$

□

**GSPN-3**

*Proof.* Consider model GSPN-3. Let $\Theta_i = \{(\mathbf{m}', \mathbf{m}) | \mathbf{m}' = \mathbf{m} + \mathbf{e_i}, 1 \leq i \leq R\}$. Proving that equation (20) holds is trivial. □

## References


1. Afshari, P. V., Bruell, S. C., and Kain, R. Y. Modeling a new technique for accessing shared buses. In *Proceedings of the Computer Network Performance Symposium* (New York, NY, USA, 1982), ACM Press, pp. 4–13.
2. Balbo, G., Bruell, S. C., and Ghanta, S. Combining queueing network and generalized stochastic Petri nets for the solution of complex models of system behavior. *IEEE Transactions on Computers 37* (1998), 1251–1268.

3. Balbo, G., Bruell, S. C., and Sereno, M. Product form solution for Generalized Stochastic Petri Nets. *IEEE Transactions on Software Engineering 28* (2002), 915–932.
4. Balbo, G., Bruell, S. C., and Sereno, M. On the relations between BCMP Queueing Networks and Product Form Solution Stochastic Petri Nets. *10th International Workshop on Petri Nets and Performance Models, 2003. Proceedings* (2003), 103–112.
5. Baskett, F., Chandy, K. M., Muntz, R. R., and Palacios, F. G. Open, closed, and mixed networks of queues with different classes of customers. *J. ACM 22*, 2 (1975), 248–260.
6. Bruell, S. C., G.Balbo, and Afshari, P. V. Mean Value Analysis of mixed, multiple class BCMP networks with load dependent service stations. *Performance Evaluation 4* (1984), 241–260.
7. Buzen, J. P. Computational algorithms for closed queueing networks with exponential servers. *Commun. ACM 16*, 9 (1973), 527–531.
8. Chandy, K. M., and Sauer, C. H. Computational algorithms for product form queueing networks. *Commun. ACM 23*, 10 (1980), 573–583.
9. Chiola, G., Marsan, M. A., Balbo, G., and Conte, G. Generalized stochastic Petri nets: a definition at the net level and its implications. *IEEE Transactions on software engineering 19*, 2 (1993), 89–107.
10. Cohen, J. W. *The single server queue.* Wiley-Interscience, 1969.
11. Coleman, J. L., Henderson, W., and Taylor, P. G. Product form equilibrium distributions and a convolution algorithm for Stochastic Petri nets. *Performance Evaluation 26* (1996), 159–180.
12. Henderson, W., Lucic, D., and Taylor, P. G. A net level performance analysis of Stochastic Petri Nets. *J. Austral. Math. Soc. Ser. B 31* (1989), 176–187.
13. Kant, K. *Introduction to Computer System Performance Evaluation.* McGraw-Hill, 1992, ch. 4 and 9.
14. Kleinrock, L. *Queueing Systems*, vol. 1 (Theory). John Wiley and Sons, 1975.
15. Marsan, M. A., Balbo, G., Conte, G., Donatelli, S., and Franceschinis, G. *Modelling with generalized stochastic Petri nets.* Wiley, 1995.
16. Melamed, B. On Poisson traffic processes in discrete-state markovian systems with applications to queueing theory. *Advances in Applied Probability*, 11 (1979), 218–239.
17. Muntz, R. R. Poisson departure processes and queueing networks. Tech. Rep. IBM Research Report RC4145, Yorktown Heights, New York, 1972.
18. Resiser, M., and Lavenberg, S. S. Mean Value Analysis of closed multichain queueing network. *JACM 27*, 2 (1980), 313–320.
19. Taylor, H. M., and Karlin, S. *An Introduction To Stochastic Modeling*, 3rd ed. Academic Press, 1998, ch. IX.
20. Vernon, M., Zahorjan, J., and Lazowska, E. D. A comparsion of performance Petri Nets and queueing network models. *Proceedings 3rd Intern. Workshop on Modelling Techniques and Performance Evaluation* (1987), 181–192.