International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 10, October-2021

Occupied Resources in A System With Resource Sharing Between Service-based Customers

Toky Basilide Ravaliminoarimalalason
Ecole Doctorale en Sciences et Techniques de l'Ingénierie et de l'Innovation,
University of Antananarivo,
Antananarivo, Madagascar

Falimanana Randimbindrainibe
Ecole Doctorale en Sciences et Techniques de l'Ingénierie et de l'Innovation,
University of Antananarivo
Antananarivo, Madagascar

***Abstract*— This paper presents analytical and generalized expressions of the amounts of occupied resources in a system providing several services to its customers. This system can be assimilated to a multi-service queue. We considered that the required amount of resources for each service are deterministic, then random variables amount of resources is studied in second part. Analysis carried out on the resource occupations have made possible the dimensioning of the total amount of necessary resources that must be deployed on the queue server. They are to serve the customers up to a certain quality of service in terms of the probability of congestion. Reported to a single-service case, we found from the analytical expressions established in this work the probability of congestion in the Fry-Molina traffic model. The obtained results are compared to digital simulations of case studies similar to a real system.**

***Keywords*— Queueing; Dimensioning; Multi-service; Resource occupation; Resource sharing**

I. INTRODUCTION

The dimensioning of the amount of resources necessary for a system guaranteeing a certain quality of service began in the era of Erlang [1]. He established a famous Erlang B formula to determine the number required channels for a telephone network [1][2]. This formula only considers a single service (voice service) provided by a system (telephone network) to users who request it. Its formula is derived from the study of a finite capacity M/M/*n*/*n* queue. The results of his studies have been developed and extended by many researchers like Palm [3], Vaulot & Chaveau [4], Joys [5], Iversen [6] and many others.

Then Iversen [7] studied the case of a multi-service system by evaluating the congestion rate of such a system with access control. Elastic traffic cases have been developed on multi-service queueing systems by Hanczewski et al. [8][9]. In [8], they proposed a multi-service queue model with SD-FIFO (State Dependent FIFO) discipline. This system allocates its resources to each class of service in a balanced fairness way. They further proposed a generalization of multi-service queues used in elastic traffic [10].

Other studies on a fairness resource sharing have been put forward by Bonald et al. [11] where all the resources are available to all service classes and to all the users who present themselves in front of the server. An extension of Haddad and Mazumdar [12] assesses a certain congestion in such a system.

Hanczewski et al. [13] have proposed another type of discipline called cFIFO (continuous FIFO) for the multi-service queue. They approximated the convolutional model of a cFIFO multi-service queuing system while allowing variable rate based on the amounts of unused resources in the server.

In all these literatures, whether it is in the sharing of resources available to all the customers present in the system, or either with variable rate, or elastic traffic, the problem arises on the total amount of resources that the multi-service server queue has for its customers. This article proposes a generalization of a multi-service queue, with fixed or random amount of resources required for each service class, in order to determine the capacity required in the server to provide a certain quality of service. Our study is based on the analytical determinations of resource occupations in such a generalized system. It is also an extension of the work done in [14] and [15].

II. MODEL DESCRIPTION

Let be a system sharing its resources through customers requesting services. Each customer can request a service *i* requiring a fixed resource R_i from the system for a random average duration μ_i. The arrival rate of customers requesting service *i* from the system is described by a Poisson process of intensity λ_i.

By denoting by *N* the total number of services that this system can provide to his customers, assuming that the arrivals of customers of each service are independent, we can say that the arrival to this system forms a Poisson process whose intensity λ is expressed by $\lambda = \lambda_1 + ... + \lambda_N = \sum_{i=1}^{N} \lambda_i$. The rate p_i of customers requesting service *i* is equal to $p_i = \lambda_i / \lambda$ for each *i*. These two expressions derive from the superposition property of independent Poisson processes [16].

Let S_i be the distribution of the service *i* duration requested by a customer. It is a strictly positive real random variable with mean μ_i. The duration of service requested by any customer will be denoted by *S*. From the total probability formula, we can express that the cumulative distribution function $F_S(s)$ of *S* is equal to $F_S(s) = \sum_{i=1}^{N} p_i F_{S_i}(s)$, where $F_{S_i}(s)$ is the cumulative distribution function of S_i. We will denote by $F_S^c(s)$ its complementary cumulative distribution function.

The system is then equivalent to a queue, with a single server with a total of *R* resources. Customers arrive at this queue according to a Poisson process of intensity λ. The

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 10, October-2021

service duration of any customer is a random variable S with an average μ. These customers will use an amount R_i of resources from the queue server depending on the type i of service requested. If customers are being served by the system, using resources r, the sum of all the amounts of resources used by these customers, another arriving customer wanting to use an amount of resources less than or equal to $R-r$ will be served immediately. If the amount of resources requested by the arriving customer is greater than $R-r$, then it will be queued waiting for a served customer to complete and release resources.

If the initial amount R of resources available in the server is infinite, there is a M/GI/∞ queue (in terms of customers). Customers who arrive will be immediately served.

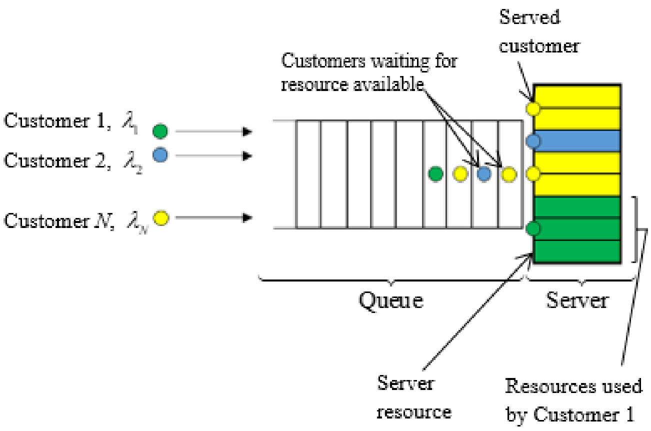

Fig. 1. System description.

III. ANALYTICAL EXPRESSIONS

A. Sharing of infinite resources

We will express the total amount of resources used by the customers described in paragraph 2. Note by π_n the probability that the system serve n customers.

Theorem

π_n follows a Poisson distribution with parameter $\lambda\int_0^{+\infty} F_S^c(s).ds$, i.e. :

$$\pi_n = \frac{\lambda^n}{n!}.\left(\int_0^{+\infty} F_S^c(s)ds\right)^n .\exp\left(-\lambda.\int_0^{+\infty} F_S^c(s)ds\right) \quad (1)$$

Proof

Let $N(t)$ be the number of customers being served at instant t, assuming that $N(0)=0$, and $\pi_n(t) = P(N(t)=n)$ the probability that the system serves n customers at instant t.

On the one hand, we denote by $(A_t)_{t\geq 0}$ the process of customer arrivals to the queue, and $(T_n)_{n\geq 0}$ their arrival times, i.e. the instants of the process. Conditionally on the counting measure $A_t = m$ of the process, the random variables $T_1,...,T_m$ are arranged in ascending order, and uniformly over the interval $(0,t)$. On the other hand, the probability that a customer will have a service duration greater than s is $P(S \geq s) = 1 - F_S(s) = F_S^c(s)$. Then, the probability that a customer arriving uniformly between the interval $(0,t)$ and still being served at time t is equal to $p_t = \frac{1}{t}\int_0^t F_S^c(s)ds$.

Conditionally on $A_t = m$, the number $N(t)$ of customers still served at time t is therefore equal to n with a probability:

$$\binom{m}{n} p_t^n (1-p_t)^{m-n}$$

This is due to the fact that n customers are still being served (probability p_t) at instant t, and therefore $m-n$ customers have finished their services before instant t (probability $1-p_t$).

Then, by applying the formula of the total probabilities for all the possible values of m, we obtain:

$$\pi_n(t) = P(N(t)=n) = \sum_{m=n}^{+\infty} P(A_t = m).P(N(t)=n \mid A_t = m)$$

$$= \sum_{m=n}^{+\infty} \frac{(\lambda t)^m}{m!}.e^{-\lambda t}.\binom{m}{n}.p_t^n(1-p_t)^{m-n}$$

$$= \sum_{m=n}^{+\infty} \frac{(\lambda t)^m}{m!}.e^{-\lambda t}.\frac{m!}{n!(m-n)!}.p_t^n(1-p_t)^{m-n}$$

$$= \frac{p_t^n}{n!} e^{-\lambda t} \sum_{m=n}^{+\infty} \frac{(\lambda t)^m}{(m-n)!}.(1-p_t)^{m-n} = \frac{p_t^n.(\lambda t)^n}{n!} e^{-\lambda t} \sum_{m=n}^{+\infty} \frac{(\lambda t)^{m-n}}{(m-n)!}.(1-p_t)^{m-n}$$

$$= \frac{(\lambda t p_t)^n}{n!} e^{-\lambda t} \sum_{k=0}^{+\infty} \frac{(\lambda t)^k}{k!}.(1-p_t)^k = \frac{(\lambda t p_t)^n}{n!} e^{-\lambda t} \sum_{k=0}^{+\infty} \frac{(\lambda t - \lambda t p_t)^k}{k!}$$

$$= \frac{(\lambda t p_t)^n}{n!} e^{-\lambda t} e^{\lambda t - \lambda t p_t} = \frac{(\lambda t p_t)^n}{n!} e^{-\lambda t p_t}$$

$$= \frac{\left(\lambda \int_0^t F_S^c(s)ds\right)^n}{n!}.\exp\left(-\lambda \int_0^t F_S^c(s)ds\right)$$

Equilibrium is obtained when t tends towards infinity, i.e.:

$$\pi_n = \lim_{t\to\infty} \pi_n(t)$$

$$= \frac{\left(\lambda \int_0^{+\infty} F_S^c(s)ds\right)^n}{n!}.\exp\left(-\lambda \int_0^{+\infty} F_S^c(s)ds\right)$$

□

By replacing by the expressions of λ and F_S within our system providing N services to its customers, we therefore have:

$$\pi_n = \frac{\left(\sum_{i=1}^{N} \lambda_i\right)^n}{n!}.\left(\int_0^{+\infty} \sum_{i=1}^{N} F_{S_i}^c(s)ds\right)^n \quad (2)$$

$$.\exp\left(-\left(\sum_{i=1}^{N} \lambda_i\right).\int_0^{+\infty} \sum_{i=1}^{N} F_{S_i}^c(s)ds\right)$$

Conditionally on the number k of customers that the server is serving, we will note r_k the amount of resources used by these customers. Let $P(R \leq r)$ be the cumulative distribution

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 10, October-2021

function of the probability that a total amount r of resources of this system is used by the customers it serves.

Thus, we have:

$$P(R \le r) = \sum_{k=0}^{+\infty} \pi_k . P(r_k \le r) \quad (3)$$

where $P(r_k \le r)$ denotes the cumulative distribution function of the probability that the amount r of resources occupied by the k customers is r_k.

Now we are assuming that the system serves k customers. The cumulative distribution function of the probability that the total amount of occupied resources is r is $P(r_k \le r)$.

A customer requesting a service i ($i = 1,..., N$) uses a resource R_i. There are N-tuples of coefficients $(k_1,...,k_N)$ verifying $\sum_{i=1}^{N} k_i . R_i \le r$ such that $k_1 + ... + k_N = k$. k_i indicates the number of customers using the service i among k the customers served by the system.

$$P(r_k \le r) = \sum_{\substack{k_1,...,k_N \\ k_1 R_1 + ... + k_N R_N \le r}} \binom{k}{k_1,...,k_N} . p_1^{k_1} ... p_N^{k_N} \quad (4)$$

where p_i denotes the probability that a randomly chosen customer, from among k served customers, uses the service i, i.e. $p_i = \lambda_i / \lambda$.

and $\binom{k}{k_1,...,k_N} = \frac{k!}{k_1! k_2! ... k_N!}$.

The expression of the cumulative distribution function of the occupied resources in the system is therefore given by:

$$P(R \le r) = \sum_{k=0}^{+\infty} \pi_k . \left(\sum_{\substack{k_1,...,k_N \\ k_1 R_1 + ... + k_N R_N \le r}} \binom{k}{k_1,...,k_N} . p_1^{k_1} ... p_N^{k_N} \right) \quad (5)$$

where:

$$\pi_n = \frac{\left(\sum_{i=1}^{N} \lambda_i \right)^n}{n!} . \left(\int_0^{+\infty} \sum_{i=1}^{N} F_{S_i}^c (s) ds \right)^n$$

$$. \exp\left(-\left(\sum_{i=1}^{N} \lambda_i \right) . \int_0^{+\infty} \sum_{i=1}^{N} F_{S_i}^c (s) ds \right)$$

In the following, we will simplify the writing of the equation (5) by:

$$P(R \le r) = \sum_{k=0}^{+\infty} \sum_{\substack{\vec{k_i} \\ \vec{k_i}\vec{R_i} \le r}} \left(\pi_k . \binom{k}{\vec{k_i}} . \prod_i p_i^{k_i} \right)$$

By adopting the notations $\vec{k_i} = k_1,...,k_N$, $\vec{R_i} = R_1,...,R_N$ and $\vec{k_i}\vec{R_i} = k_1 R_1 + ... + k_N R_N$.

B. Sharing of finite resources

For the case of finite resources, it is obvious to think of a M/GI/C queue where C denotes a certain finite capacity of the system. But the analytical resolution of this queue is still an open problem, and the solutions proposed so far are only approximations [17][18][19]. In addition, in our case, we have a variable capacity depending on the customers served by the system and their requirements in terms of amount of resources, so this will further worsen the complexity of our problem.

In our research, we focus on dimensioning the amount of required resources to satisfy customers in their uses of the services provided by the system.

Being given that the system has a capacity C in terms of total amount of available resources it can allocate to customers. If the requirement of the next customer is greater than the amount of available resources in the system, we can say that there is a blocking phenomenon. Two scenarios can arise: either the customer in question is rejected, or it is put on hold. In both these cases, the capacity in terms of amount of resources is insufficient.

On the one hand, if we are going to deploy a lot of resources. Some of them may not be used. It is therefore a waste of resources. In some cases, the availability of resources is chargeable and it will therefore be a loss in terms of investment. On the other hand, the deployment of few resources can lose some customers as we described in the previous blocking phenomenon. It is therefore necessary to determine the right total amount of resources to put in place. For this, a certain tolerance threshold must be adopted such that the probability of rejection or of placing customers on hold is less than this tolerance.

Let's assume that we have deployed an infinite amount of resources. For this, the probability of using an amount of at most r resources is given by equation (5) recalled below:

$$P(R \le r) = \sum_{k=0}^{+\infty} \sum_{\substack{\vec{k_i} \\ \vec{k_i}\vec{R_i} \le r}} \left(\pi_k . \binom{k}{\vec{k_i}} . \prod_i p_i^{k_i} \right)$$

In other words, the probability of using an amount of resources greater than C is given by:

$$B = P(R > C) = \sum_{k=0}^{+\infty} \sum_{\substack{\vec{k_i} \\ \vec{k_i}\vec{R_i} > C}} \left(\pi_k . \binom{k}{\vec{k_i}} . \prod_i p_i^{k_i} \right)$$

We can say that this is the probability of blocking the system of capacity C because since from this amount of resources C, it can no longer serve more customers. It will be called probability of congestion.

For dimensioning with a tolerance ε, the objective is therefore to determine a capacity C such that:

$$B = P(R > C) = \sum_{k=0}^{+\infty} \sum_{\substack{\vec{k_i} \\ \vec{k_i}\vec{R_i} > C}} \left(\pi_k . \binom{k}{\vec{k_i}} . \prod_i p_i^{k_i} \right) \le \varepsilon \quad (6)$$

C. What happens in case of single service system and unit resource for each customer?

In this part, we will see the case of a system offering a single service in which each served customer uses only a single amount of resource. For the case of exponential service duration with mean $1/\mu$, we will see that from equations (5) and (6), we end up with the BCH (Busy-Calls Held) traffic model of Fry-Molina [20][21][22].

The number of services provided by the system is then $N=1$. The customer arrival rate follows a Poisson process of intensity $\lambda_1=\lambda$. Since the service duration is exponential, the cumulative distribution function of the random variable S of service duration is given by $F_S(s)=1-\exp(-\mu s)$. The rate of customers requesting service 1 is therefore $p_1=1$.

The cumulative distribution function of the total amount of occupied resources, from equation (5), is given by:

$$P(R\le r)=\sum_{k=0}^{+\infty}\pi_k\cdot\left(\sum_{\substack{k_1,\dots,k_N\\ k_1R_1+\dots+k_NR_N\le r}}\binom{k}{k_1,\dots,k_N}.p_1^{k_1}\dots p_N^{k_N}\right)$$

$$=\sum_{k=0}^{+\infty}\pi_k\cdot\left(\sum_{\substack{k_1\\ k_1\le r}}\binom{k}{k_1}.p_1^{k_1}\right)=\sum_{k=0}^{+\infty}\pi_k.1_{k\le r}(k)=\sum_{k=0}^{r}\pi_k$$

where:

$$\pi_k=\frac{\left(\sum_{i=1}^{N}\lambda_i\right)^k}{k!}.\left(\int_0^{+\infty}\sum_{i=1}^{N}F_{S_i}^c(s)ds\right)^k .\exp\left(-\left(\sum_{i=1}^{N}\lambda_i\right).\int_0^{+\infty}\sum_{i=1}^{N}F_{S_i}^c(s)ds\right)$$

$$=\frac{\lambda^k}{k!}.\left(\int_0^{+\infty}\sum_{i=1}^{N}F_{S_i}^c(s)ds\right)^k .\exp\left(-\lambda.\int_0^{+\infty}\sum_{i=1}^{N}F_{S_i}^c(s)ds\right)$$

$$=\frac{\lambda^k}{k!}.\left(\int_0^{+\infty}\exp(-\mu s)ds\right)^k.\exp\left(-\lambda.\int_0^{+\infty}\exp(-\mu s)ds\right)$$

$$=\frac{\lambda^k}{k!}.\left(\frac{1}{\mu}\right)^k.\exp\left(-\frac{\lambda}{\mu}\right)$$

$$=\frac{\rho^k}{k!}\exp(-\rho)$$

By denoting $\lambda/\mu=\rho$.

Then,

$$P(R\le r)=\sum_{k=0}^{r}\pi_k=\sum_{k=0}^{r}\frac{\rho^k}{k!}\exp(-\rho)=\exp(-\rho)\sum_{k=0}^{r}\frac{\rho^k}{k!}.$$

Thus, equation (6) becomes:

$$P(R>C)=\sum_{k=0}^{+\infty}\sum_{\substack{\vec{k_i}\\ \vec{k_i}\vec{R_i}>C}}\left(\pi_k\cdot\binom{k}{\vec{k_i}}.\prod_i p_i^{k_i}\right)\le\varepsilon$$

$$=\exp(-\rho)\sum_{k=C}^{+\infty}\frac{\rho^k}{k!}\le\varepsilon$$

By replacing $\exp(-\rho)$ with $\sum_{k=0}^{+\infty}\frac{\rho^k}{k!}$, we find the congestion rate $\sum_{k=C}^{+\infty}\frac{\rho^k}{k!}/\sum_{k=0}^{+\infty}\frac{\rho^k}{k!}$ posed by Fry-Molina in their traffic model..

We note that this model, which is an alternative to the Erlang B formula [1] [2], is used in certain dimensioning of telephony in North America. It is observed that this model reflects much more the reality of observed traffics compared to the Erlang model [23] [24].

D. Generalization for random amounts of resources requested by customers

This time, we consider that the amounts of resources requested by service *i* customers are no longer deterministic as in the previous paragraphs. Let's assume that the service *i* customers request a random amount of resource following a probability distribution L_i with mean μ_i. We will denote by $G_i(r)$ the cumulative distribution function of the amount of resources requested by a customer using service *i*.

We still assume an infinite amount of resources available from the system server. According to the methodology that we adopt in paragraph 3.1, to determine the cumulative distribution function $P(R\le r)$ of the resources occupation of the system, it is necessary to determine a cdf $P(R_k\le r)$ conditional on *k* customers served by the system.

We decompose this number *k* of served customers into the number k_i of customers per service such as $k=k_1+\dots+k_N$. We can have $\binom{k}{k_1,\dots,k_N}=\frac{k!}{k_1!\dots k_N!}$ possible combinations of these customers. For each customer, the probability that he uses an amount of resource less than or equal to *x* is equal to $G_i(x)$ if he requested service *i* (or with a probability of p_i in other words).

The probability that all customers then use a total amount of resources less than or equal to *x* for a given vector $(k_1,\dots,k_N)$ is expressed by:

$$\binom{k}{k_1,\dots,k_N}\sum_{x_{11}+\dots+x_{Nk_N}=x}\underbrace{\left(p_1G_1(x_{11})\dots p_1G_1(x_{1k_1})\right)}_{k_1\text{ users of service 1}}\dots\underbrace{\left(p_NG_N(x_{N1})\dots p_NG_N(x_{Nk_N})\right)}_{k_N\text{ users of service }N} \quad (7)$$

where x_{ij} denotes the amount of resources used by *j*-th customer of service *i*.

Equation (7) can be rewritten as a discrete convolution product:

$$\binom{k}{k_1,...,k_N}\sum_{x_{11}+...+x_{Nk_N}=x}\underbrace{\left(p_1G_1(x_{11})...p_1G_1(x_{1k_1})\right)}_{k_1 \text{ users of service } 1}...\underbrace{\left(p_NG_N(x_{N1})...p_NG_N(x_{Nk_N})\right)}_{k_N \text{ users of service } N}$$

$$=\binom{k}{k_1,...,k_N}\underbrace{\left(p_1G_1(x)*...*p_1G_1(x)\right)}_{k_1 \text{ consecutive convolution products}}*...*\underbrace{\left(p_NG_N(x)*...*p_NG_N(x)\right)}_{k_N \text{ consecutive convolution products}}$$

$$=\binom{k}{k_1,...,k_N}p_1^{k_1}...p_N^{k_N}.G_1^{*k_1}(x)*...*G_N^{*k_N}(x)$$

where $f(t)*g(t)=\sum_{t_1+t_2=t}f(t_1)g(t_2)=\sum_{s=0}^{+\infty}f(s)g(t-s)$ for two functions $f(t)$ and $g(t)$ with positive variable *t*, and $f^{*n}(t)=f(t)*...*f(t)$ the *n* successive convolution products of the function $f(t)$ by itself.

Thus, with all the possible values of the vector $(k_1,...,k_N)$ and all the values of $x\le r$, we obtain the cdf $P(R_k\le r)$ conditionally to *k* customers served by the system:

$$P(R_k\le r)$$

$$=\sum_{k_1+...+k_N=k}\sum_{x\le r}\binom{k}{k_1,...,k_N}p_1^{k_1}...p_N^{k_N}.G_1^{*k_1}(x)*...*G_N^{*k_N}(x) \quad (8)$$

$$=\sum_{k_1+...+k_N=k}\sum_{x\le r}\binom{k}{k_1,...,k_N}\prod_{i=1}^{N}p_i^{k_i}.\prod_{i=1}^{*N}G_i^{*k_i}(x)$$

where $\prod_{i=1}^{*N}f_i(x)=f_1(x)*...*f_N(x)$ for a family of functions $\{f_i(x)\}_{i=1,...,N}$.

And the cdf of the amount of occupied resources is given by:

$$P(R\le r)$$

$$=\sum_{k=0}^{+\infty}\left(\pi_k.\sum_{k_1+...+k_N=k}\sum_{x\le r}\binom{k}{k_1,...,k_N}\prod_{i=1}^{N}p_i^{k_i}.\prod_{i=1}^{*N}G_i^{*k_i}(x)\right) \quad (9)$$

For the case of a finite capacity of resources *C*, we can deduct from equation (9) the probability of the system being blocked:

$$B=P(R>C)$$

$$=\sum_{k=0}^{+\infty}\left(\pi_k.\sum_{k_1+...+k_N=k}\sum_{x>C}\binom{k}{k_1,...,k_N}\prod_{i=1}^{N}p_i^{k_i}.\prod_{i=1}^{*N}G_i^{*k_i}(x)\right)$$

It is recalled that this probability must be less than a tolerance ε for the determination of the sufficient capacity of the system.

Remark:

When the amount of resources requested by customers, regardless of the service used, is an independent and identically distributed random variable with a cdf $G(r)$, equation (9) can be written as:

$$P(R\le r)$$

$$=\sum_{k=0}^{+\infty}\left(\pi_k.\sum_{k_1+...+k_N=k}\sum_{x\le r}\binom{k}{k_1,...,k_N}G^{*k}(x).\prod_{i=1}^{N}p_i^{k_i}\right)$$

IV. RESULTS AND DISCUSSIONS

Let's take a case study of resources sharing between customers arriving at a system of a capacity *C* which will be determined. The system can provide one of the *N* = 4 services to each customer. The arrivals of these customers to this system form Poisson processes of respective intensity $\lambda_1=2,\ \lambda_2=3,\ \lambda_3=2.5,$ and $\lambda_4=5$ arrivals per unit of time for the service *i* = 1, 2, 3, and 4. By the principle of superposition of the Poisson processes, these arrivals form a single process with intensity $\lambda=\sum_{i=1}^{4}\lambda_i=12.5$.

The amounts of resources required for these services are respectively $R_1=4,\ R_2=3,\ R_3=2,$ et $R_4=1$ resources.

The service durations are assumed to be random following an exponential distribution with respective means $1/\mu_1=1/3,\ 1/\mu_2=1/4,\ 1/\mu_3=1/5,$ and $1/\mu_4=1/6$ units of time for the services *i* = 1 to 4.

From equation (5), the cumulative distribution function of the occupied resources in the infinite capacity system is equal to:

$$P(R\le r)=\sum_{k=0}^{+\infty}\pi_k.\left(\sum_{\substack{k_1,...,k_4\\ k_1R_1+...+k_4R_4\le r}}\binom{k}{k_1,...,k_4}.p_1^{k_1}...p_4^{k_4}\right)$$

Such that $p_i=\lambda_i/\lambda$ denotes the probability that the customer arriving at the system will use the service *i*.

And:

$$\pi_k=\frac{\lambda^k}{k!}.\left(\int_0^{+\infty}\sum_{i=1}^{4}e^{-\mu_is}ds\right)^k.\exp\left(-\lambda.\int_0^{+\infty}\sum_{i=1}^{4}e^{-\mu_is}ds\right)$$

$$=\frac{\lambda^k}{k!}.\left(\sum_{i=1}^{4}\frac{1}{\mu_i}\right)^k.\exp\left(-\lambda.\sum_{i=1}^{4}\frac{1}{\mu_i}\right)$$

Curves in Figures 2a and 2b represent this cumulative distribution function:

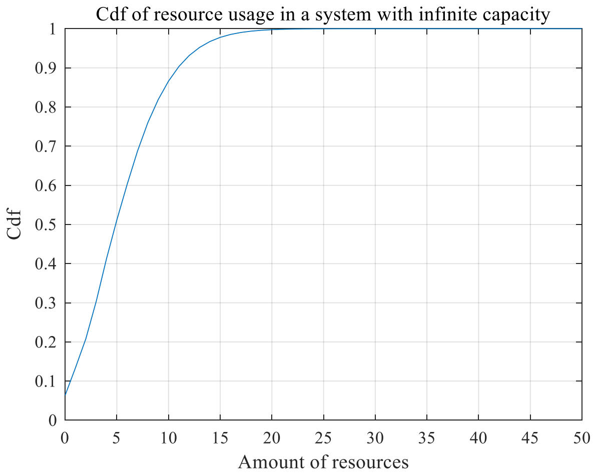

Fig. 2a. Cdf of the amount of occupied resources for an infinite capacity

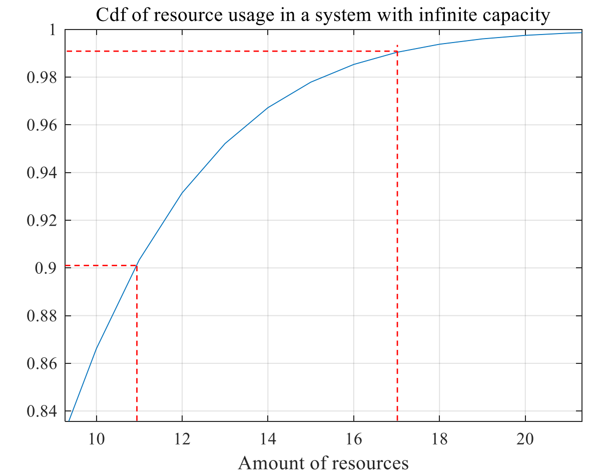

Fig. 2b. Cdf of the amount of occupied resources for an infinite capacity (zoomed in)

It is obvious from Figure 2a that the deployment of more than 20 resources does not influence the system. We deduce that the probability that the amount of occupied resources is greater than or equal to 20 is very small (close to zero). It is therefore not necessary to provide more than 20 resources to the system since not all of them will be used.

From these curves and equation (6), we can conclude a probability of congestion of 10% when the system has 11 resources, and this probability becomes 1% with 17 resources.

We used MATLAB - Simulink to simulate these two values of total amount of resources to verify these congestion probabilities.

Model presented in Figure 3 has been used. The customers (users) are generated from the "User of service *i*" blocks following a respective rate λ_i described above.

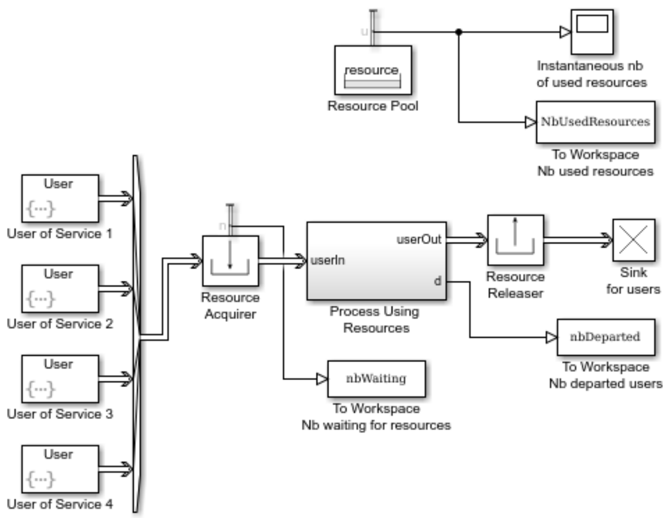

Figure 3. Resource sharing model for a system having four service-classes for its users

The probability of congestion is calculated from the percentage of time during which the amount of occupied resources equals to the capacity of the system. We then use the parameter `NbUsedResources` in the below algorithm to determine this percentage.

```
count = 0;
capacity = 11;
t = 0:0.01:2000;
nbUsedResourcesResampled =
resample(NbUsedResources,t,'zoh');
usedResRsmpl = nbUsedResourcesResampled.Data;
[timeResRsmpl,y]=size(usedResRsmpl);
for i = 1:timeResRsmpl
   if usedResRsmpl(i,1)>= capacity
      count = count + 1;
   end
end
pCongestion = count/timeResRsmpl;
disp('Capacity = 11');
fprintf('The percentage of time the capacity
is reached : %0.3g \n', pCongestion);
```

We found the following percentages for a capacity of 11 resources, and for a capacity of 17 resources at 5000 simulation time units.

```
>>
Capacity = 11
The percentage of time the capacity is
reached : 0.1073

>>
Capacity = 17
The percentage of time the capacity is
reached : 0.0166
```

These are the probabilities that any user, arriving at a random moment, will find all the resources occupied.

We then found the theoretical results of our model explained by the curves of Figures 2a and 2b which are respectively 10% and 1% for capacities 11 and 17. These results validate the model thus established.

As a concrete example, the model can be used for the dimensioning of codes (resources) of a third generation mobile communication network (3G network) which share these codes through its customers [25] [26]. This type of network provides multiple services, such as:

- AMR voice service: 2 codes needed
- VP service: 8 codes needed
- PS R99 DL service: 8 codes needed
- HSDPA service: 16 codes needed

These results remain valid for the case of non-discrete resources. In this case, the amount of resources initially available in the system is a strictly positive real number. The amounts of resources necessary for each service i can also be strictly positive real numbers (not necessarily a non-zero natural number). The random variable designating the total amount of resources occupied by the customers is then a positive real random variable. And the expressions of equations (5) and (6) are already carried with the cumulative distribution functions which characterize the real random variables.

V. CONCLUSION

A multi-service system shares its resources through its customers according to each customer's requirements for each service requested. They arrive on this system according to Poisson processes characteristic of each service-class. The system is assimilated with a multi-service M/GI/C queue in terms of the number of customers. Two cases of capacities in terms of amount of resources were considered: the infinite case which allowed us to evaluate the distribution of occupancy of the queue resources, and the finite case from which we derived a probability of congestion from the infinite case. The analytical expressions established have made it possible for us to dimension the system by determining the amount of resources necessary for the server with an admissible probability of congestion. From the theoretical result and the case study presented, we found that it is quite possible to determine this necessary amount since from a certain value of the capacity, the characteristic of the system in terms of the number of occupied resources does not show much variation. We have shown that the result of the Fry-Molina model is found from our model and from our result in the single-service case.

REFERENCES

[1] Brockmeyer, E.; Halstrom, H.; Jensen, A.; Erlang, A.K. “The life and works of A.K. Erlang”. Copenhagen Telephone Co, 1948
[2] Erlang, A.K. “Løsning af nogle Problemer fra Sandsynlighedsregningen af Betydning for de automatiske Telefoncentraler” (Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges). Elektroteknikeren. 1917, vol 13, pp 5-13.
[3] Palm, C. “Table of the Erlang Loss Formula”. Telefonaktiebolaget LM Ericsson, Stockholm, 1947.
[4] Joys, L.A. “Comments on the Engset and Erlang Formulae for Telephone Traffic Losses”. Thesis. Report TF No. 25,/71, Research Establishment, The Norwegian Telecommunications Administration. 1971.
[5] Iversen, V.B. “The A-formula”. Teleteknik, English ed. 1980, vol 23 n°2, pp 64-79.
[6] Vaulot, E.; Chaveau, J. “Extension de la formule d’Erlang au Cas où le Trafic est Fonction du Nombre d’Abonnés Occupés” (Extension of the Erlang formula to the case where the traffic depends on the number of busy subscribers). Annales de Télécommunications. 1949, vol 4 n°8, pp 319-324.
[7] Iversen, V.B. “The exact evaluation of multi-service loss system with access control”. Teleteknik, English ed. 1987, vol 31 n°2, pp 56-61.
[8] Hanczewski, S.; Stasiak, M.; Weissenberg, J. “A queueing model of a multi-service system with state-dependent distribution of resources for each class of calls”. IEICE Trans. Commun. 2004, vol E97-B n°8, pp 1592-1605.
[9] Hanczewski, S.; Stasiak, M.; Weissenberg, J. “The queueing model of a multiservice system with dynamic resource sharing for each class of calls”. Computer Networks, CN 2013, Communications in Computer and Information Science. 2013, vol 370, pp 436-445
[10] Hanczewski, S.; Kmiecik, D.; Stasiak, M.; Weissenberg, J. “Multi-service queuing system with elastic traffic”. Proc. IEICE Gen. Conf. 2016, BS-3-18
[11] Bonald, T.; Proutière, A.; Virtamo, J. “A queueing analysis of max-min fairness, proportional fairness and balanced fairness”. Queueing Systems. 2006, vol 53 n°1, pp 65-84
[12] Haddad, J.-P.; Mazumdar, R.R. “Congestion in large balanced multirate networks”. Queueing Systems. 2013, vol 74 n°2, pp 333-368
[13] Hanczewski, S.; Kaliszan, A.; Stasiak, M. “Convolution Model of a Queueing System with the cFIFO Service Discipline”. Hindawi Publishing Corporation, Mobile Information Systems. 2016, vol 2016, Article ID 2185714, pp 1-15
[14] Ravaliminoarimalalason, T. B; Randimbindrainibe, F. “Distribution of Occupied Resources on A Discrete Resources Sharing in A Queueing System”. International Journal of Engineering Research and Technology. 2021, vol 10 n°1, pp 638-643
[15] Ravaliminoarimalalason, T. B; Randimbindrainibe, F.; Rakotomalala M., “Distribution of occupied resources on a fractional resource sharing in a queueing system”. Int. J. of Operational Research. 2021. DOI 10.1504/IJOR.2021.10039107
[16] Crane, H.; McCullagh, P. “Poisson superposition processes”. Journal of Applied Probability. 2015, vol 52 n°4, pp 1013-1027
[17] Hokstad, Per. “Approximations for the M/G/m Queue”. Operations Research. 1978, vol 26 n°3, pp 510–523
[18] Tijms, H. C.; Van Hoorn, M. H.; Federgruen, A. “Approximations for the Steady-State Probabilities in the M/G/c Queue”. Advances in Applied Probability. 1981, vol 13 n°1, pp 186–206
[19] Ma, B. N. W.; Mark, J. W. “Approximation of the Mean Queue Length of an M/G/c Queueing System”. Operations Research. 1995, vol 43 n°1, pp 158-165.
[20] Molina, E.C. “The Theory of Probability Applied to Telephone Trunking Problems”. The Bell System Technical Journal. 1922, vol 1 n°2, pp 69–81
[21] Fry, T.C. “Probability and its Engineering Uses”. New York, 1928
[22] Molina, E.C. “Application of the Theory of Probability to Telephone Trunking Problems”. The Bell System Technical Journal. 1927, vol 6 n°3, pp 461–494
[23] Iversen, V.B. “Teletraffic engineering handbook”, ITU–D SG 2/16 & ITC, 2001
[24] Iversen, V.B. “Teletraffic engineering and network planning”, DTU Library, 2015
[25] Kasapović, S.; Sarajlić, N. “OVSF code assignment in UMTS networks”. 20th International Crimean Conference Microwave & Telecommunication Technology. 2010, pp 429-432
[26] Bernhard, H.W.; Seidenberg, P.; Althoff, M.P. “UMTS : The Fundamentals”. John Wiley & Sons, 2003