# JOINT STATIONARY DISTRIBUTION OF QUEUES IN HOMOGENOUS $M|M|3$ QUEUE WITH RESEQUENCING

Ilaria Caraccio
University of Salerno
84084-Via Giovanni Paolo II, 132
Fisciano (SA) Italy
Email: icaraccio@unisa.it

Alexander V. Pechinkin
Rostislav V. Razumchik
Institute of Informatics Problems of RAS
Vavilova, 44-2,
119333, Moscow, Russia
Email: apechinkin@ipiran.ru,
rrazumchik@ieee.org

**KEYWORDS**

Resequencing, queueing system, joint distribution.

**ABSTRACT**

Resequencing issue is a crucial issue in simultaneous processing systems where the order of customers (jobs, units) upon arrival has to be preserved upon departure. In this paper stationary characteristics of $M/M/3/\infty$ queueing system with reordering buffer of infinite capacity are being analyzed. Noticing that customer in reordering buffer may form two separate queues, focus is given to the study of their size distribution. Expressions for joint stationary distribution are obtained both in explicit form and in terms of generating functions. Numerical example is presented.

## INTRODUCTION

Resequencing issue is a crucial issue in simultaneous processing systems where the order of customers (jobs, units, et.c.) upon arrival has to be preserved upon departure. Various analytical methods and models have been proposed to study the impacts of resequencing. A general survey of queueing theoretic methods and early models for the modeling and analysis of parallel and distributed systems with resequencing can be found in Boxma et al. (1994). Survey on the resequencing problem that covers period up to 1997 can be found in Dimitrov (1997). Queueing-theoretic approach to resequencing problem implies that the system under consideration is represented as interconnected queueing systems/networks. Following Leung et al. (2010) existing papers can be grouped into categories: papers that characterize the disordering process through a queueing system with several servers sharing a single queue (see, e.g. Agrawal and Ramaswami (1987)) and papers where disordering is modeled by a queueing system with several parallel servers and queues, and each server has its own dedicated queue (see, e.g. Ye Xia et al. (2008)). For a short survey of these two categories see e.g. Leung et al. (2010). Following this approach various problems setting have been considered and solved including distribution of number of packets in reordering buffer and in system under different assumptions about arrival and service process (see, e.g. Jain and Sharma (2011), Lelarge (2008), Chakravarthy (1998), Takine et al. (1994), De Nicola C. et al. (2013)); distribution and/or mean of the resequencing delay (see, e.g. Huisman and Boucherie (2002), Ding et al. (1991)), end-to-end (i.e. sender–receiver) delay (see, e.g. Chowdhury (1991)); large deviations of the queue size in reordering buffer (see, e.g. Gao et al. (2012)), asymptotics of the resequencing delay (see, e.g. Jun Li et al. (2010)), optimal allocation of customers to servers (Gogate and Panwar (1999), optimization issues (Dimitrov et al. (2002)). Among practical related papers one can also cite Leung et al. (2010), Zheng et al. (2010) andWen-Fen (2011), Li et al. (2010), Huisman and Boucherie (2001), Min Choi et al. (2012), Rubin et al. (1991).

In this paper we propose new problem statement for systems with resequencing that are modeled by multiserver queues followed with infinite resequencing buffer. New problem is motivated by noticing that customers awaiting in resequencing buffer may form separate queues. The most convenient way to explain how queues are separated in resequencing buffer is with the example. Consider a queueing system with three servers, infinite capacity main buffer and reordering buffer. Let the state of the system at some instant be as depicted in Fig. 1. In squares one can see customers' sequential numbers. White (black) squares in Fig. 1 mean that customers with these sequential numbers have received (have not yet received) service. Here one can distinguish two queues: one which is formed by customers awaiting customer no. 18 (queue #1), another is formed by customers awaiting customer no. 15 (queue #2). Three cases need to be considered.

*Case 1*. Now on if customer no. 21 is next to complete its service then it joins queue #1 and stays there until service of customer no. 18 is complete. Customer no. 22 joins idle server.

*Case 2*. If customer no. 15 is next to complete its service then it goes through queue #1 without waiting and joins queue #2. Meanwhile customer no. 22 joins idle server. As there is no customer in the system with sequential number smaller than any sequential number in

Proceedings 28th European Conference on Modelling and Simulation ©ECMS Flaminio Squazzoni, Fabio Baronio, Claudia Archetti, Marco Castellani (Editors)
ISBN: 978-0-9564944-8-1 / ISBN: 978-0-9564944-9-8 (CD)

queue #2, then all customers from queue #2 leave the system. Resequencing buffer "sees", that queue #2 is empty and moves its contents to queue #2. Now there are three options. First: if customer no. 18 is next to complete service, then it goes through queue #1 without waiting and joins queue #2. Customer no. 23 joins idle server. Again there is no customer in the system with sequential number smaller than any sequential number in queue #2. Thus all customers from queue #2 leave the system. Resequencing buffer becomes empty. Now if customer no. 21 is next to complete service, it leaves the system. If customer no. 22 is next to complete service, it goes through queue #1 without waiting and joins queue #2 where it waits for customer no. 21. Finally, if customer no. 23 is next to complete service, it joins queue #1 and does not proceed to queue #1 because it needs customer no. 22 to complete its service before both of them may join queue #2. Second: if customer no. 21 is next to complete service, then it goes through queue #1 again without waiting, joins queue #2 and waits there with other customer for service completion of customer no. 18. Third: if customer no. 22 is next to complete service, then customer no. 23 joins idle server, customer no. 22 joins queue #1 and stops there because "sees" gap between its sequential number and largest sequential number in queue #2. It waits there for customer no. 21.

*Case 3*. If customer no. 18 is the first to complete its service then it joins queue #1 and customer no. 22 joins idle server. Resequencing buffer "sees", that there is no gap in the middle of sequence and moves the content of queue #1 to queue #2 (queue #1 becomes empty). Now there are again three options. First: if customer no. 15 is next to complete service, then it goes through queue #1 without waiting, joins queue #2 and immediately (because the sequence is complete) leaves the system with all contents of queue #2. Second: if customer no. 21 is next to complete service, then it goes through queue #1 again without waiting, joins other customers in queue #2 that wait for service completion of customer no. 15. Third: if customer no. 22 is next to complete service, then it joins queue #1 and stops there, because "sees" gap between its sequential number and the largest sequence number in queue #2. The operation of the system proceeds along the line.

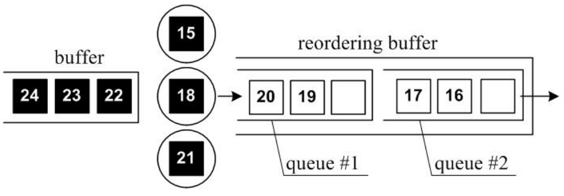


Figure 1: Scheme of the model

Clearly, when the number of server is $n$ there are $(n-1)$ queues in resequencing buffer. Sum their contents is the total number of customers in resequencing buffer. The main contribution of this paper are algorithm and probability generating function of joint stationary probabilities of the number of customers in buffer, queue #1 and queue #2.

The paper is organized as follows. In the next section we give detailed description of the system. Then we find joint stationary distribution both algorithm-wise and in terms of probability generating functions. The last section is devoted to numerical results.

## DESCRIPTION OF THE SYSTEM

Consider a queueing system with three servers, infinite capacity buffer, incoming poisson flow of customers (of intensity $\lambda$) and exponential distribution of service time at each server (with parameter $\mu$) and resequencing buffer (RB) of infinite capacity. Customers upon entering the system obtain sequential number and join buffer. Without loss of generality we suppose that the sequence starts from 1 and coincides with the row of natural numbers, i.e. the first customer upon entering the (empty) system receives number 1, the second — number 2 and so on and so forth. Customers leave the system strictly in order of their arrival (i.e. in the sequence order). Thus after customer's arrival it remains in the buffer for some time and then receives service when one of the servers becomes idle. If at the moment of its service completion there are no customers in the system or all other customers present at that moment in the queue and the rest two servers have greater sequential numbers it leaves the system. Otherwise it occupies one place in the RB. Customer from RB leaves it if and only if its sequential number is less than sequential numbers of all other customers present in system. Thus customers may leave RB in groups.

Let us call "$1^{st}$ level" customer the one which is in service and was the last to enter server; "$2^{nd}$ level" customer is the one which is in service and was the penultimate to enter server; finally, "$3^{rd}$ level" the customer is the one which is in service and was the first to enter server. If the number of busy servers is 3, then customers that entered RB between "$1^{st}$ level" and "$2^{nd}$ level" customer form queue #1; customers which entered RB between "$2^{nd}$ level" and "$3^{rd}$ level" customer form queue #2. If the number of busy servers is 2, then customers which entered RB after "$1^{st}$ level" customer form queue #1; customers which entered RB between "$1^{st}$ level" and "$2^{nd}$ level" customer form queue #2. When there is only one busy server all customers in RB form queue #2.

The operation of the considered queueing system can be completely described by Markov process $\zeta(t) = \{(\xi(t), \eta(t), \nu(t)), \quad t \geq 0\}$ with three components: $\xi(t)$ — number of customers in buffer and server at time $t$, $\eta(t)$ — number of customers in queue #1 of RB at time $t$, $\nu(t)$ — number of customers in queue #2 of RB at time $t$. In case $\xi(t) = 0$, the second and third component of $\zeta(t)$ are omitted; in case $\xi(t) = 1$, the second is omitted. The state space of $\zeta(t)$ is $\mathcal{X} = \{0\} \cup \{(1, i)\,,\ i \geq 0\} \cup \{(n, i, j)\,,\ n \geq 2,\ i \geq 0,\ j \geq 0\}$. Henceforth it is assumed that service and arrival processes are mutually independent and necessary and sufficient condition of stationarity $\rho/3 < 1$, where $\rho = \lambda/\mu$ holds for the system.

## STATIONARY JOINT PROBABILITY DISTRIBUTION

Note that the total number of customers in servers and buffer of the considered QS with resequencing coincides with the total number of customers in $M/M/3/\infty$ queue. Therefore, its stationary distribution $\{p_i,\ i \geq 0\}$, has the form

$$p_0 = \left( \sum_{i=0}^{2} \frac{\rho^i}{i!} + \frac{\rho^3}{2!\,(3-\rho)} \right)^{-1}, \quad (1)$$

$$p_i = \frac{\rho^i}{i!} p_0, \ \ i = \overline{1,3}, \quad (2)$$

$$p_i = \frac{\rho^i}{3!\,3^{i-3}} p_0 = \tilde{\rho}^{i-3} p_3, \ \ \tilde{\rho} = \frac{\rho}{3}, \ \ i \geq 4. \quad (3)$$

Provided that RB is empty when servers are idle, $p_0$ is also the probability of the considered system with resequencing to be empty.

Denote by $p_{n;i,j}$, $n \geq 3, i \geq 0, j \geq 0$, stationary probability of the fact that there are $n$ customers in servers and buffer, $i$ customers in queue #1 of RB, $j$ customers in queue #2 of RB. By $p_{n;i}$, $n \geq 3$, $i \geq 0$, denote stationary probability of the fact that there are $n$ customers in servers and buffer and $i$ customers in queue #1 of RB. Clearly $p_{n;i} = \sum_{j \geq 0} p_{n;i,j}$. Probabilities $p_{2;i,j}$, $i \geq 0$, $j \geq 0$ and $p_{2;i}$, $i \geq 0$, are defined by analogy. Finally, let $p_{1;i}$, $i \geq 0$, be stationary probability of the fact that there is only one busy server and $i$ customers reside in queue #2 of RB. Note that distribution $p_n$, $n \geq 0$, of the total number of customers in servers and buffer (which is defined by (1)–(3)) can be expressed as follows

$$p_1 = \sum_{i \geq 0} p_{1;i}, \quad p_n = \sum_{i \geq 0} \sum_{j \geq 0} p_{n;i,j}, \ \ n \geq 2.$$

In the next section we proceed to the derivation of the steady-state equations for the defined above probabilities.

## SYSTEM OF EQUILIBRIUM EQUATIONS

The derivation of steady-state equations we start with $p_{n;i}$, $n \geq 3$, — stationary probabilities of the fact that there are total of $n$ customers in servers and buffer and $i$ customers in queue #1 of RB. The easiest way to do this is to use global balance principle. Let $i = 0$. The set of states, corresponding to probability $p_{n;0}$ when $n \geq 3$ is $\cup_{j \geq 0} (n, 0, j)$. Mean rate out of this set is clearly $(\lambda + 3\mu) p_{n;0}$. Note that one can enter the set $\cup_{j \geq 0} (n, 0, j)$ either from set of states $\cup_{j \geq 0} (n-1, 0, j)$ (with mean rate $\lambda p_{n-1;0}$) or from $\cup_{i \geq 0} \cup_{j \geq 0} (n+1, i, j)$ (with mean rate $3\mu \frac{2\mu}{3\mu} p_{n+1} = 2\mu p_{n+1}$). Putting it altogether yields

$$(\lambda + 3\mu) p_{n;0} = \lambda p_{n-1;0} + 2\mu p_{n+1}, \ \ n \geq 3. \quad (4)$$

The derivation of other steady-state equations is done by analogy utilizing the properties of exponentially distributed random variables and thus omitted. Probabilities $p_{n;i}$, $n \geq 1$ satisfy

$$(\lambda + 3\mu) p_{n;i} = \lambda p_{n-1;i} + \mu p_{n+1;i-1}, \ \ n \geq 3, \ \ i \geq 1, \quad (5)$$

whereas for probabilities $p_{2;i}$, $i \geq 0$, it holds

$$(\lambda + 2\mu) p_{2;0} = \lambda p_1 + 2\mu p_3, \quad (6)$$

$$(\lambda + 2\mu) p_{2;i} = \mu p_{3;i-1}, \ \ i \geq 1. \quad (7)$$

Probabilities $p_{1;i}$, $i \geq 0$, satisfy

$$(\lambda + \mu) p_{1;0} = \lambda p_0 + \mu p_{2;0}, \quad (8)$$

$$p_{1;i}(\lambda + \mu) = \mu p_{2;i} + \mu \sum_{j=0}^{i-1} p_{2;i-j-1,j}, \ \ i \geq 1. \quad (9)$$

One can verify that for probabilities $p_{n;i,j}$, $n \geq 3$, $i \geq 0$, $j \geq 0$, the following equations hold:

$$(\lambda + 3\mu) p_{n;0,0} = \lambda p_{n-1;0,0} + \mu p_{n+1;0}, \ \ n \geq 3, \quad (10)$$

$$(\lambda + 3\mu) p_{n;0,j} = \lambda p_{n-1;0,j} + \mu p_{n+1;j} + \mu \sum_{k=0}^{j-1} p_{n+1;k,j-k-1}, \ \ n \geq 3, \ \ j \geq 1, \quad (11)$$

$$(\lambda + 3\mu) p_{n;i,j} = \lambda p_{n-1;i,j} + \mu p_{n+1;i-1,j}, \ \ n \geq 3, \ \ i \geq 1, \ \ j \geq 0. \quad (12)$$

Finally, probabilities $p_{2;i,j}$, $i \geq 0$, $j \geq 0$ satisfy

$$(\lambda + 2\mu) p_{2;0,0} = \lambda p_{1;0} + \mu p_{3;0}, \quad (13)$$

$$(\lambda + 2\mu) p_{2;0,j} = \lambda p_{1;j} + \mu p_{3;j} + \mu \sum_{k=0}^{j-1} p_{3;k,j-k-1}, \ \ j \geq 1, \quad (14)$$

$$(\lambda + 2\mu) p_{2;i,j} = \mu p_{3;i-1,j}, \ \ i \geq 1, \ \ j \geq 0. \quad (15)$$

The analysis of steady-state equations resulted in the development of simple algorithm for step-by-step computation of stationary joint probabilities $p_{n;i,j}$, $n \geq 2$, $i \geq 0$, $j \geq 0$, and $p_{n;i}$, $n \geq 1$, $i \geq 0$. The algorithm is given below.

For practical purposes it may be sometimes sufficient to know either only $\pi_{n;i}$, $n \geq 1$, $i \geq 0$, — stationary probabilities of the fact that total number of customers in servers and in buffer is $n$ and total number of customers in RB (sum of queue #1 and queue #2) is $i$, or only $\pi_i$, $i \geq 0$, — stationary probabilities of the fact that there are $n$ customers in total in the whole system (including buffer, servers, RB). These quantities can be calculated from joint probability distribution as follows

$$\pi_{1;i} = p_{1;i}, \ i \geq 0, \quad \pi_{2;i} = \sum_{j=0}^{i} p_{2;j,i-j}, \ \ i \geq 0,$$

$$\pi_{n;i} = \sum_{j=0}^{i} p_{n;j,i-j}, \ \ n \geq 3, \ \ i \geq 0,$$

$$\pi_0 = p_0, \quad \pi_1 = \pi_{1;0}, \quad \pi_2 = \pi_{1;1} + \pi_{2;0},$$

$$\pi_i = \pi_{1;i-1} + \pi_{2;i-2} + \sum_{j=3}^{i} \pi_{j;i-j}, \quad i \geq 3.$$

**Algorithm 1** Algorithm for calculation of stationary joint probability distribution

```
Initialize λ, μ;
for n ≥ 0 do
    Calculate p_n from Eq. (1), (2), (3);
end for
Calculate p_{2;0} from Eq. (6);
for n ≥ 3 do
    Calculate p_{n;0} from Eq. (4);
end for
for i ≥ 1 do
    Calculate p_{2;i} from Eq. (7);
    for n ≥ 3 do
        Calculate p_{n;i} from Eq. (5);
    end for
end for
Calculate p_{1;0} from Eq. (8);
Calculate p_{2;0,0} from Eq. (13);
for n ≥ 3 do
    Calculate p_{n;0,0} from Eq. (10);
end for
for i ≥ 1 do
    Calculate p_{2;i,0} from Eq. (15);
    for n ≥ 3 do
        Calculate p_{n;i,0} from Eq. (12);
    end for
end for
for i ≥ 2 do
    Calculate p_{1;i} from Eq. (9);
    Calculate p_{2;0,i} from Eq. (14);
    for n ≥ 3 do
        Calculate p_{n;0,i} from Eq. (11);
    end for
    for j ≥ 1 do
        Calculate p_{2;j,i} from Eq. (15);
        for m ≥ 3 do
            Calculate p_{m;j,i} from Eq. (12);
        end for
    end for
end for
```

In the next section we will dwell on the derivation of probability generating functions of the joint stationary distribution.

## GENERATING FUNCTIONS

Though the calculation of probabilities $p_{n;i,j}$, $n \geq 2$, $i \geq 0$, $j \geq 0$ and $p_{n;i}$, $n \geq 1$, $i \geq 0$ is just a matter of computational effort due to obtained above algorithm, performance characteristics (e.g. moments and/or correlation of queue lengths in RB) are not so straightforward to obtain. Below we show that in the considered case one can obtain expressions for probability generating functions (PGF) that ease the computation of various performance characteristics. Let us introduce the following PGF:

$$p_n(z) = \sum_{i=0}^{\infty} z^i p_{n;i}, \;\; 0 \leq z \leq 1, \;\; n \geq 1,$$

$$p_n(z_1, z_2) = \sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} z_1^{i_1} z_2^{i_2} p_{n;i_1,i_2}, \;\; 0 \leq z_1 \leq 1, \;\; 0 \leq z_2 \leq 1, \;\; n \geq 2,$$

$$P(u, z) = \sum_{n=3}^{\infty} u^{n-3} p_n(z), \;\; 0 \leq u \leq 1,$$

$$P(u, z_1, z_2) = \sum_{n=3}^{\infty} u^{n-3} p_n(z_1, z_2), \;\; 0 \leq u \leq 1.$$

If one puts $z_1 = z_2 = z$ in $P(u, z_1, z_2)$, then function $P(u, z, z)$ is the double PGF of the total number of customers in buffer and servers and total number of customers in RB when all three servers are busy. By analogy $p_n(z, z)$, $n \geq 2$, is the PGF of the total number of customers total number of customers in RB and probability of total $n$ customers in servers and buffer. In the following we will make use of PGF $P(u) = \sum_{n=3}^{\infty} u^{n-3} p_n$, $|u| \leq 1$, which, with respect to (1)–(3), equals $P(u) = p_3/(1 - \tilde{\rho} u)$.

Now, starting from equation (4), we will successively obtain relations for PGF defined above. Multiplying (4) and (5) by $z^i$ and summing up over all $i$ from 0 to infinity, we have:

$$(\lambda + 3\mu) p_n(z) = \lambda p_{n-1}(z) + \mu z p_{n+1}(z) + 2\mu p_{n+1}, \;\; n \geq 3. \quad (16)$$

If one multiplies the previous equations by $u^{n-3}$ and sums them up over all $n \geq 3$, then after collecting the common term, one obtains the relation for $P(u, z)$:

$$P(u, z) = \frac{\mu z p_3(z) - \lambda p_2(z) u - 2\mu [P(u) - p_3]}{\lambda u^2 - (\lambda + 3\mu) u + \mu z}. \quad (17)$$

Consider equations (6) and (7). Multiplication by $z^i$ and summation over all possible values of $i$, gives the relation for $p_2(z)$:

$$(\lambda + 2\mu) p_2(z) = \lambda p_1 + 2\mu p_3 + \mu z p_3(z). \quad (18)$$

The same manipulation with equations (8) and (9). leads to relation for $p_1(z)$:

$$(\lambda + \mu) p_1(z) = \lambda p_0 + \mu p_2(z) + \mu z p_2(z, z). \quad (19)$$

Multiplying (10)–(12) by $z_1^{i_1} z_2^{i_2}$ and summing equations over all possible values of $i_1$ and $i_2$, one gets relation for $p_n(z_1, z_2)$:

$$(\lambda + 3\mu) p_n(z_1, z_2) = \lambda p_{n-1}(z_1, z_2) + \mu p_{n+1}(z_2) + \mu z_2 p_{n+1}(z_2, z_2) + \mu z_1 p_{n+1}(z_1, z_2), \;\; n \geq 3.$$

In order to obtain equation for $P(u, z_1, z_2)$ one must multiply the previous relation by $u^{n-3}$ and again sum up over $n \geq 3$. It holds that

$$P(u, z_1, z_2) = \left[\lambda u^2 - (\lambda + 3\mu)u + \mu z_1\right]^{-1} \times \left[\mu p_3(z_2) + \mu z_2 p_3(z_2, z_2) + \mu z_1 p_3(z_1, z_2) - \mu z_2 P(u, z_2, z_2) - \mu P(u, z_2) - \lambda p_2(z_1, z_2)u\right]. \quad (20)$$

Finally, from (13)–(15) repeating the traditional manipulations one obtains relation for the last unknown PGF $p_2(z_1, z_2)$:

$$(\lambda + 2\mu)p_2(z_1, z_2) = \lambda p_1(z_2) + \mu p_3(z_2) + \mu z_2 p_3(z_2, z_2) + \mu z_1 p_3(z_1, z_2). \quad (21)$$

In order to obtain expressions for introduced PGF one has to solve system of equations (16)–(21). It can be done as follows.

By putting $z_1 = z_2 = z$ in (20) and (21) one yields the following two equations:

$$P(u, z, z) = \left[\lambda u^2 - (\lambda + 3\mu)u + 2\mu z\right]^{-1} \times \left[\mu p_3(z) + 2\mu z p_3(z, z) - \mu P(u, z) - \lambda p_2(z, z)u\right], \quad (22)$$

$$(\lambda + 2\mu)p_2(z, z) = \lambda p_1(z) + \mu p_3(z) + 2\mu z p_3(z, z). \quad (23)$$

Consider function $f_m(u, z)$, $m = 1, 2$, given by expression

$$f_m(u, z) = \lambda u^2 - (\lambda + 3\mu)u + m\mu z, \ \ m = 1, 2. \quad (24)$$

Denote by $u_m = u_m(z)$ minimal and by $\widehat{u}_m = \widehat{u}_m(z)$ — maximal solution of the equation $f_m(u, z) = 0$, $m = 1, 2$, i.e.

$$u_m = \frac{\lambda + 3\mu - \sqrt{(\lambda + 3\mu)^2 - 4m\lambda\mu z}}{2\lambda}, \quad \widehat{u}_m = \frac{\lambda + 3\mu}{\lambda} - u_m, \ \ m = 1, 2.$$

Note that $0 < u_m < 1$ for $0 \leq z \leq 1$ and $m = 1, 2$. Relations (17) and (22) can be rewritten with respect to (24) in the form

$$P(u, z) = \frac{\mu z p_3(z) - \lambda u p_2(z) - 2\mu[P(u) - p_3]}{f_1(u, z)}, \quad (25)$$

$$P(u, z, z) = [f_2(u, z)]^{-1} \times \left[2\mu z p_3(z, z) - \lambda u p_2(z, z) - \mu[P(u, z) - p_3(z)]\right]. \quad (26)$$

Denominator in (25) and (26) is zero at points $(u_1, z) = (u_1(z), z)$ and $(u_2, z) = (u_2(z), z)$. Since PGF $P(u, z, z)$ is analytic function in the domain $0 \leq z \leq 1$ then numerator must be zero at these points too. This leads to the following equations

$$\mu z p_3(z) - \lambda u_1 p_2(z) - 2\mu[P(u_1) - p_3] = 0, \quad (27)$$

$$2\mu z p_3(z, z) - \lambda u_2 p_2(z, z) - \mu[P(u_2, z) - p_3(z)] = 0. \quad (28)$$

Firstly we find PGF $P(u, z)$. Solution of equations (18) and (27)

$$(\lambda + 2\mu)p_2(z) - \mu z p_3(z) = \lambda p_1 + 2\mu p_3,$$

$$\mu z p_3(z) - \lambda u_1 p_2(z) = 2\mu[P(u_1) - p_3]$$

gives the expression for PGF $p_2(z)$ and $p_3(z)$ in the form:

$$p_2(z) = \frac{\lambda p_1 + 2\mu P(u_1)}{\lambda - \lambda u_1 + 2\mu},$$

$$p_3(z) = \frac{1}{\mu z}(\lambda u_1 p_2(z) + 2\mu[P(u_1) - p_3]).$$

Substitution the form of $p_2(z)$ and $p_3(z)$ in (25) and collecting the common terms, leads to expression for PGF $P(u, z)$:

$$P(u, z) = \frac{1}{(\widehat{u}_1 - u)(1 - \tilde{\rho}u_1)} \times \left(\frac{[(\lambda + 2\mu)p_{2;0} - \lambda p_1 \tilde{\rho} u_1]}{(\lambda + 2\mu - \lambda u_1)} + \frac{2\mu p_4}{\lambda(1 - \tilde{\rho}u)}\right).$$

Let us find the expression for $P(u, z, z)$. Solving system of equations (19), (23) and (28), one obtains the following expression for PGFs $p_1(z)$, $p_2(z, z)$ and $p_3(z, z)$:

$$p_2(z, z) = \left(\lambda - \lambda u_2 + 2\mu - \frac{\lambda\mu z}{\lambda + \mu}\right)^{-1} \times \left(\frac{\lambda}{\lambda + \mu}[\lambda p_0 + \mu p_2(z)] + \mu P(u_2, z)\right),$$

$$p_1(z) = \frac{1}{\lambda + \mu}[\lambda p_0 + \mu p_2(z) + \mu z p_2(z, z)],$$

$$p_3(z, z) = \frac{1}{2\mu z}(\lambda u_2 p_2(z, z) + \mu[P(u_2, z) - p_3(z)]).$$

If one substitutes expressions for $p_1(z)$, $p_2(z, z)$ and $p_3(z, z)$ into (26) then, after collecting the common terms, one finds $P(u, z, z)$:

$$P(u, z, z) = \frac{p_2(z, z)}{(\widehat{u}_2 - u)} + \frac{\mu}{\lambda(\widehat{u}_2 - u)(\widehat{u}_1 - u_2)} \times \left(P(u, z) + \frac{2\mu p_4 \tilde{\rho}}{\lambda(1 - \tilde{\rho}u_2)(1 - \tilde{\rho}u)(1 - \tilde{\rho}u_1)}\right).$$

The last PGF to find is $P(u, z_1, z_2)$. Denominator in (20) is zero at point $(u_1(z_1), z_1)$. Since PGF $P(u, z_1, z_2)$ is analytic function in the domain $0 \leq z_1 \leq 1$, $0 \leq z_2 \leq 1$ then numerator must vanish at this point. Hence it holds

$$\mu z_1 p_3(z_1, z_2) - \lambda u_1(z_1) p_2(z_1, z_2) = \mu[P(u_1(z_1), z_2) - p_3(z_2)] + \mu z_2[P(u_1(z_1), z_2, z_2) - p_3(z_2, z_2)]. \quad (29)$$

From relation (21) it follows that

$$\mu z_1 p_3(z_1, z_2) = (\lambda + 2\mu)p_2(z_1, z_2) - \lambda p_1(z_2) - \mu p_3(z_2) - \mu z_2 p_3(z_2, z_2).$$

Substitution of $\mu z_1 p_3(z_1, z_2)$ into (29), leads to the expression for $p_2(z_1, z_2)$:

$$p_2(z_1, z_2) = \left[\lambda + 2\mu - \lambda u_1(z_1)\right]^{-1} \times \left[\lambda p_1(z_2) + \mu P(u_1(z_1), z_2) + \mu z_2 P(u_1(z_1), z_2, z_2)\right].$$

Thus we have obtained all the unknown quantities in PGF $P(u, z_1, z_2)$ and it is determined completely. Its final expression is too cumbersome and thus omitted.

In the next section we proceed to numerical examples that depict the behaviour of queues in buffer and RB.

## NUMERICAL EXAMPLE

There are several quantities related to the number of customers in the system that may be of interest. They are mean and variance of the number of customers in queue #1 and queue #2, correlation between queue size in buffer and queue #1, between queue size in buffer and queue #2 and between queue #1 and queue #2. These quantities are depicted in Fig. 2 – Fig. 4. In all examples service rate $\mu = 1$.

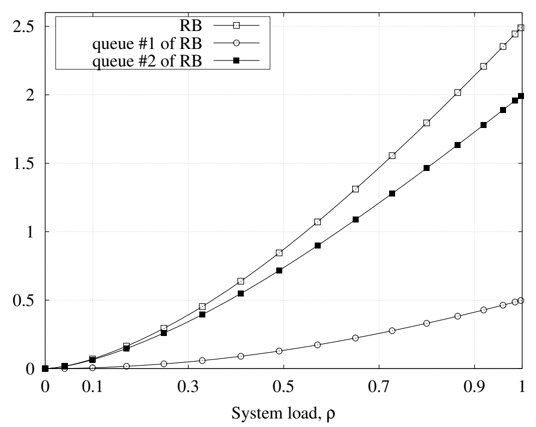


Figure 2: Mean number of customers

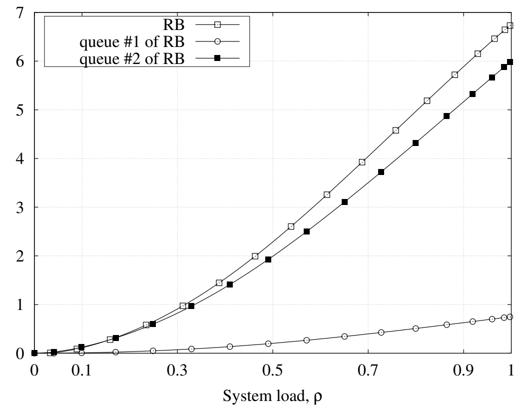


Figure 3: Variance of the number of customers

Interesting to notice from Fig. 4 that correlation between queue sizes is insignificant especially between queues of RB. This raises the question for further study about the presence of correlation between queues in RB

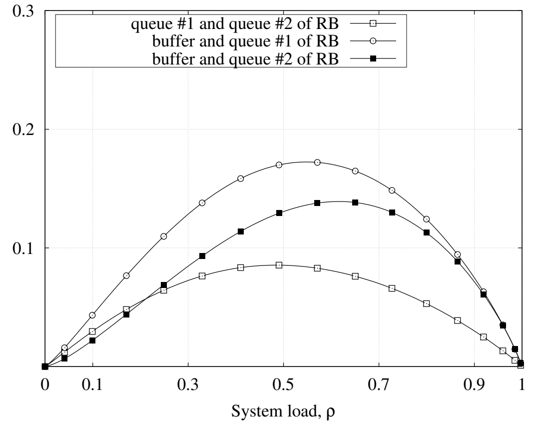


Figure 4: Coefficients of correlation

for larger values of $n$. The variance of the total number of customers in RB (see Fig. 3) is almost equal to the sum of variances for queue #1 and queue #2 of RB.

## SUMMARY

In this study, consideration is given to $M/M/3/\infty$ queueing system with resequencing buffer of infinite capacity. Noticing that customer in reordering buffer may form two separate queues, focus is given to the study of their size distribution. Results of the thorough analysis of joint stationary distribution (both explicit and in terms of generating functions) are presented. It is shown numerically that for the all possible range of load values correlation between any queues that are formed in the system is almost insignificant. Further study will be devoted to the analysis of joint stationary distribution of queues in reordering buffer in more complex systems with possibly arbitrary number of servers.

**Notes and Comments.** This work was partially supported in part by the Russian Foundation for Basic Research (grants 14-07-00041, 13-07-00223).

## REFERENCES

Agrawal, S., Ramaswamy, R. 1987. Analysis of the resequencing delay for M/M/m systems. Proceedings of the ACM SIGMETRICS conference on Measurement and modeling of computer systems. Pp. 27–35.

Baccelli, F., Gelenbe, E., Plateau, B. 1981. An end-to-end approach to the sequencing problem. Rapports de Recherche, INRIA.

Boxma, O., Koole, G., Liu, Z. 1994. Queueing-theoretic solution methods for models of parallel and distributed systems. Performance Evaluation of Parallel and Distributed Systems Solution Methods. CWI Tract 105 and 106. Pp. 1–24.

Chowdhury, S. 1991. Distribution of the total delay of packets in virtual circuits. Proceedings of the Tenth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 91). Vol. 2. Pp. 911–918.

Chakravarthy, S., Chukova, S., Dimitrov, B. Analysis of MAP/M/2/K queueing model with infinite resequencing buffer. Journal of Performance Evaluation. Vol. 31. Issue 3-4. Pp. 211–228.

Ding, Y., Sholl, H., Lipsky L. 1991. Approximations of the mean resequencing waiting time in M/GI/c systems. Proceedings of the 11th International Conference on Distributed Computing Systems. Pp. 28–35.

Dimitrov, B. 1997. Queues with resequencing. A survey and recent results. Proceedings 2-nd World Congress on Nonlinear Analysis, Theory, Methods, Applications. Vol. 30, No. 8. Pp. 5447–5456.

Dimitrov, B., Green Jr., D., Rykov, V., Stanchev, P. 2002. On performance evaluation and optimization problems in queues with resequencing. Advances in Stochastic Modelling, edt. J. Artalejo, A. Krishnamoorty. Notable publications. Pp. 55–72.

De Nicola, C., Pechinkin, A., Razumchik, R., 2013. Stationary Characteristics of Homogenous Geo/Geo/2 Queue with Resequencing in Discrete Time. Proceedings of the 27th European Conference on Modelling and Simulation. Pp. 594–600.

Gao, Y., Zhao,Y. Q. 2012. Large Deviations for Re-Sequencing Buffer Size. IEEE Transactions on Information Theory. Vol. 58. No. 2. Pp. 1003–1009.

Gogate, R., Panwar, S. 1999. Assigning customers to two parallel servers with resequencing. IEEE Communications Letters. Vol. 3. No. 4. Pp. 119–121.

Huisman, T., Boucherie, R.J. (2002) The sojourn time distribution in an infinite server resequencing queue with dependent interarrival and service times. Journal of Applied Probability. Vol. 39. No. 3. Pp. 590–603.

Huisman, T., Boucherie, R.J. (2001) Running times on railway sections with heterogeneous train traffic. Transportation Research Part B: Methodological. Vol. 35. No. 3.

Jun Li, Yifeng Zhou, Lamont, L., Minyi Huang, Zhao, Y.Q. 2010. Probabilistic Analysis of Resequencing Queue Length in Multipath Packet Data Networks. IEEE Global Telecommunications Conference (GLOBECOM 2010). Pp. 1–5.

Jain, M., Sharma, G.C. 2011. No passing multiserver queue with additional heterogeneous servers and inter-dependent rates. 5-th Canadian Conference in Applied Statistics. 20-th conference of the Forum for Interdisciplinary Mathematics -Interdisciplinary Mathematical Statistical Techniques. Concordia University, Montreal, Quebec, Canada.

Lelarge, M. 2008. Packet reordering in networks with heavy-tailed delays. Mathematical Methods on Operation Research. Vol. 67. Pp.341-371.

Leung, K., Li, V.O.K. 2010. A resequencing model for high-speed packet-switching networks. Journal Computer Communications. Vol. 33. Issue 4. Pp. 443–453.

Lelarge, M. 2008. Packet reordering in networks with heavy-tailed delays. Mathematical Methods on Operation Research. Vol. 67. Pp.341-371.

Li, J., Zhao, Y.Q. 2010. Modeling and Analysis of Resequencing Delay in Selective Repeat Automatic Repeat Request. Journal of networks. Vol. 5. No. 7. Pp. 792–799.

Matyushenko, S. I. 2010. Stationary characteristics of the two-channel queueing system with reordering customers and distributions of phase type. Informatics and its Applications. Vol. 4. Issue 4. Pp. 67–71. (in Russian).

Min Choi, Jong Hyuk Park, Young-Sik Jeong. 2012. Revisiting reorder buffer architecture for next generation high performance computing. The Journal of Supercomputing. Pp. 1–12.

Rubin, I., Zhang, Z. 1991. Message delay and queue-size analysia for circuit-switched TDMA systems. IEEE Transactions on Communications. Vol. 39. Issue 6. Pp. 905–914.

Takine, T., Ren, J., Hasegawa, T. 1994. Analysis of the Resequencing Buffer in a Homogeneous M/M/2 Queue. Performamce Evaluation. Vol. 19. Issue 4. Pp. 353–366.

Wen-Fen, L. 2011. An Analysis of Resequencing Queue Size at Receiver on Multi-Path Transfers. Proceedings of the International Conference on Internet Technology and Applications (iTAP). Pp. 1–4.

Ye Xia, Tse, D.N.C. 2008. On the large deviations of resequencing queue size: 2-M/M/1 Case. IEEE Transactions on information theory. Vol. 54. No. 9. Pp. 4107–4118.

Zheng, K., Jiao, X., Liu, M., Li, Z. 2010. An Analysis of Resequencing Delay of Reliable Transmission Protocols over Multipath. Proceedings of the IEEE International Conference on Communications (ICC). Pp. 1–5.

## AUTHOR BIOGRAPHIES

**ILARIA CARACCIO** received her degree in Mathematics in 2012 at University of Salerno. At the moment she is a Phd in Mathematics at University of Salerno. Her current research activity focuses on queuing theory. Her email address is `icaraccio@unisa.it`.

**ALEXANDER V. PECHINKIN** is a Doctor of Sciences in Physics and Mathematics and principal scientist at the Institute of Informatics Problems of the Russian Academy of Sciences, and a professor at the Peoples' Friendship University of Russia. He is the author of more than 150 papers in the field of applied probability theory. His email address is `apechinkin@ipiran.ru`.

**ROSTISLAV V. RAZUMCHIK** received his Ph.D. in Physics and Mathematics in 2011. Since then, he has worked as a senior researcher at the Institute of Informatics Problems of the Russian Academy of Sciences. His current research activities focus on queuing theory and adaptive strategies. His email address is `rrazumchik@ieee.org`