International Journal of Applied Science and Mathematics
Volume 4, Issue 2, ISSN (Online): 2394-2894

Analysis of Three Server Queues with Stalling

R Sivasamy*, K Thaga and L.L. Mokatlhe
*Corresponding author email id: ramasamysr@mopipi.ub.bw
Date of publication (dd/mm/yyyy): 18/03/2017

***Abstract* — This article analyses a Markovian queueing system $M/(M_1,M_2,M_3)/3/(B_1,B_2)$ with stalling. It stalls customers ofqueue-1 into a finite buffer $B_1$of maximum size 'K< ∞'and accommodates all other waiting customers ofqueue-2 into an infinite buffer B_2. There are three heterogeneous servers labelled as 'S_1, S_2 and S_3' with exponential rates μ_1, μ_2, and μ_3 respectively where $\mu_1>\mu_2>\mu_3$. Arrivals occur according to a Poisson process with mean arrival rateλ. Main focus is on the steady state queue length distribution which consists of two stages.Instage-1, 'Queue Length' process of the finite $M/(M_1,M_2,M_3)/3/(B_1,3)$ system is formulated as a Quasi-Birth and Death process in a three dimensional finite state space. The stationary probability vector of the queue length is obtained using matrix analyticalmethods. In stage-2, using analytical methods, stage-1 results obtained on the finite queue length are linked to the infinite queue length process of the $M/ (M_1,M_2,M_3)/ 3/(B_1,B_2)$ system subject to a condition $\rho=(\lambda/\mu)<1$ where $\mu= \mu_1+ \mu_2+\mu_3$. Further steady state expressions are found for some of the performance measures such as the expected queue length, the probability that each server is busy etc. Results for some special systems have been obtained from these computational methods. A numerical study is then carried out to support the advantages of the proposed methodology.**

***Keywords* — Generator Matrix, Heterogeneous Servers, Mean Queue Length, Stationary Probability Vector and Quasi-Birth-Death Process.**

I. INTRODUCTION

This paper describes a Markovian queueing system $M/(M_1,M_2, M_3)/3/(B_1,B_2)$ with stalling. It is operated by three heterogeneous servers called S_1, S_2 and S_3. Service time distribution of the server S_j follows exponential distribution with mean $(1/\mu_j)$ for j=1, 2, and 3 which satisfy a condition $\mu_1>\mu_2>\mu_3$. Inter-arrival times also follow an exponential distribution with mean $(1/\lambda)$. Assumption is that these inter-arrival times and service times are independently distributed random variables.

A. Queue Discipline

An arriving customer occupies S_1 server if idle or otherwise joins the queue-1 given that the queue-1 formation in buffer B_1 of size 'K< ∞' is not full. Non-empty buffer B_1 feeds customers one by one to the server S_1 at each time epoch when the S_1 server finishes a service.

If an arrival occurs at a time point when queue-1 is full (i.e. possible if server S_1 is busy) and there are more than one idle servers, it occupies that idle server with the lowest number or otherwise joins the free server if any. If the system has (K+3) or more number of customers at a customer's arrival epoch, it joins the buffer B_2 to form queue-2.

When the server S_1 finishes a service, it serves the next customer in queue-1 (if there is one; otherwise it idles) and the customer at the top of queue-2 (if there is one) leaves queue-2 and joins queue-1. At each service completion epoch either of the slow servers S_2 and S_3, serves the next customer in queue-2 (if there is one; otherwise it idles). Thus queue-2 feeds both queue-1 and the slow servers S_2 and S_3 whichever could first accept the head-of-the-line customer of queue-2.

Under parallel configuration, S_1 works faster than S_2 and S_2 works faster than S_3. Considered here is that customers are of informed types. Hence customers have to wait in buffer B_1 at times even when the slow servers S_2 and S_3 are free, until B_1 becomes full. Buffer B_2 accommodates all other arriving customers at time instances when B_1 is full and all three servers are busy i.e. when the system size exceeds (K+3). It is assumed that waiting customers form queue-1 and queue-2 according to their order of arrival. One aim of this paper is to compare the steady state results of the Markovian queueing system $M/(M_1,M_2, M_3)/3/(B_1,B_2)$ with that of the steady state characteristics of the system $M/M/3/(B_1,B_2)$.

The queue discipline that governs how each customer on its arrival epoch makes decision and how customers are buffered while waiting to be dispatched is described depending on the following 'Five' mutually exclusive events E_j for j=1 2, 3, 4 and 5 observed by a monitoring device of server S_1 and the buffer B_1:

E_1: Server S_1 is idle or S_1 just completes a service
E_2: Server S_1 is busy and queue-1 (0< queue-1<K) is not full (*i.e.* B_1 has less than K customers)
E_3: queue-1 has exactly K customers and both slow servers S_2 and S_3 are free
E_4: queue-1 has exactly K and one of the two slow servers S_2 and S_3 is alone free
E_5: the system has (K+3) or more number of customers

Decision of an arriving customer that can happen during an infinitesimal period subsequent to E_j will be as below:

(i) E_1 event, joins server S_1 if queue-1 is empty
(ii) E_2 event, joins queue-1
(iii) E_3 event, joins server S_2
(iv) E_4 event, joins the free server who is available among S_2 and S_3
(v) E_5 event, joins queue-2

Dispatching rule for a customer either from queue-1>0 or from queue-2>0that can happen during an infinitesimal period just after a service completion by any one of the servers 'S_1, S_2 and S_3' will be as below:

(i) E_1 event, a customer is dispatched from queue-1 to server S_1
(ii) E_2 event, a customer is dispatched from queue-1 to server S_1
(iii) E_3 event, a customer is dispatched from queue-2 to server S_2
(iv) E_4 event, a customer is dispatched from queue-2 to the slow server who, just, completes her service

(v) E_5 event, a customer is dispatched either from queue-1 to the fast sever S_1 and then queue-1 accepts a customer from queue-2 instantaneously or from queue-2 to the slow server (S_2 or S_3) who, just, completes her service

It is observed that server S_1 will be continuously busy as long as queue-1 has at least one stalled customer. Further, if a customer arrives at an instance when the length of queue-1 is K and both of the two slow servers are idle, then it occupies idle server S_2.

The graphical representation of the $M/(M_1,M_2, M_3)/3/(B_1,B_2)$ queue is drawn in Figure1.

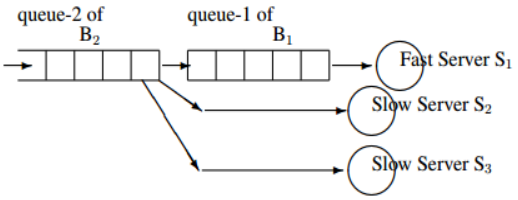

Figure1: Three Server Buffered Queue with Stalling

A typical methodology is developed to obtain the stationary queue length distribution in two stages: Stage-1 deals with the finite capacity queue M/ (M_1, M_2, M_3/3/ (B_1, 3).Formulating the queue length (queue +service) process as a QBD processes, steady state results to state probabilities and mean queue length have been obtained using matrix-analytical methods. For the determination of the `K+3' boundary characteristics of stage-1 substantial effort is taken by solving the equations satisfied by sub-matrices of the generator matrix**Q** of the QBD process. Exploiting the Markov property of the QBD process, stage-1 results are linked to the stage-2 which helps to obtain the whole queue length distribution of $M/(M_1,M_2, M_3/3/(B_1,B_2)$ queue with staling and other characteristics as compact and closed form expressions. Results of the proposed queueing system with stalling have viable applications in the area of computer networks and manufacturing industries.

Reference [1] studied waiting lines with heterogeneous servers where the new customer is dispatched to any server if all the servers are busy. In [2], Krishnamoorthi investigated a Poisson queue with one fast and one slow server operating under 'Fist Come Fist Served' queue discipline. An optimal control policy was established in [3] for a queuing system with two heterogeneous Servers.

In [4]-[5], Rubinovitch studied the problem of a heterogeneous two channel queueing systems with no-stalling and stalling respectively and gave conditions when to discard the slow server. The literature about fast and slow server queueing systems is growing steadily. A simple approach for installing slow server issues is used in [6] while the slow server problem is discussed in [7] for uninformed customers.An efficient way of managing queues with heterogeneous servers is explored in [8]. In [9]-[10], Singh studied a simple queue with two heterogeneous servers with infiniteand finite waiting spaces. His emphasis was on comparing the two-server heterogeneous $M/M_n/2/N$ and homogeneous M/M/2/N systems. Using the optimal service rates, the average characteristics of the heterogeneous system are minimized, and their improvement over the corresponding homogeneous system characteristics is established. Authors of [11] have provided analysis for an M/G/2 queue operating under FCFS (First Come First Served) queue discipline.

Unless a multi-server queuing system is mechanically controlled, the case of heterogeneous servers is more applicable in practice. If the servers have equal service rates, the situation is referred to as a queuing system with homogeneous servers, or otherwise as a queuing system with heterogeneous servers.

B. Motivating Factors

The motivating example for this work is power management in data centres, where we have a fixed power budget P and a server farm consisting of three servers. Wise decision is needed on how much power to allocate to each server, so as to minimize overall mean response time for jobs arriving at the server farm. It isremarked that themore power we allocate to a server, the faster it runs (the higher its frequency), and subject to some maximum possible frequency and some minimum power level needed just to turn the server on.

There is a monitor that specifies the relationship between the quanta of power allocated to a server and the speed (frequency) at which it runs. To answer the question on how to allocate power, we need to think about whether we prefer many slow servers (allocate just a little power to every server) or a few fast ones (distribute all the power among a small number of servers). In this application, the three server problem which is presented in this article could be used to optimally answer our question under a wide variety of parameter settings. For similar applications on power allocation problem from a CPU (Central Processing Unit), details are found in Sivasamy et al. [12]-[13]. In [14], several threshold properties of a Poisson queue with a T-policy are studied.

In section II, matrix analytic method is used to analyse the M /M_n/3/(B_1, 3) queues with staling. Section III deals with $M/(M_1,M_2,M_3)/3/(B_1,B_2)$ with stalling for informed customers. Section IVpresents a comparative study between the results of $M/M_n/3$ and M/M/3 queues. Section V provides a summary of results and future scope.

II. Modified Loss System: M /M_N/3/(B_1, 3) Queues

Let $X_1(t)$ be the number of customers available in the buffer $B_1 \cup B_2$ and with the fast server S_1 at time t>= 0. Also let $X_j(t)$ be the number of customers present with Server-j for j=2 and 3 at time t>= 0. The vector process **X(t)** =($X_1(t)$, $X_2(t)$, $X_3(t)$: $t \geq 0$) defined on the Cartesian product space $\mathbb{S}$ ={0, 1,2,…K+1) }x(0,1)x(0,1)} is aperiodic and positive recurrent if $\rho = \lambda/(\mu_1+ \mu_2+ \mu_3)<1$.

A. Generator Matrix Q

Assume that buffer B_2 is removed or empty from the

system described in Figure1. Thus the queue length (queue + service) process **X (t)** of the modified loss system M/ $(M_1,M_2,M_3)/3/(B_1,3)$ forms a quasi-birth-and-death (QBD) process on a finite portioned form of three dimensional space $\mathbb{S}$:

$\mathbb{S}$ ={ L(n) : n=0, 1, …, K+1} , where

L(n)=((n,0,0), (n,1,0), (n,0,1), (n,1,1)) ...*(1)*

Define: $\lambda_1=\lambda+\mu_1$, $\lambda_2=\lambda+\mu_2$, $\lambda_3=\lambda+\mu_3$, $\lambda_{12}=\lambda+\mu_1+\mu_2$, $\lambda_{13}=\lambda+\mu_1+\mu_3$, $\lambda_{23}=\lambda+\mu_2+\mu_3$, $\lambda_{123}=\lambda+\mu_1+\mu_2+\mu_3$and $\mu=\mu_1+\mu_2+\mu_3$. Then the infinitesimal generator matrix **Q** of the QBD process is then given by

$$\mathbf{Q}=\begin{array}{c|ccccccc} & L(1) & L(1) & L(1) & \dots & L(K-1) & L(K) & L(K+1) \\ \hline L(0) & \mathbf{A}_1^{(0)} & \mathbf{A}_0^{(0)} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ L(1) & \mathbf{A}_2^{(1)} & \mathbf{A}_1^{(1)} & \mathbf{A}_0^{(1)} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ L(2) & \mathbf{0} & \mathbf{A}_2^{(2)} & \mathbf{A}_1^{(2)} & \mathbf{A}_0^{(2)} & \dots & \mathbf{0} & \mathbf{0} \\ . & \mathbf{0} & \mathbf{0} & . & . & . & \dots & \mathbf{0} \\ . & \mathbf{0} & \mathbf{0} & \mathbf{0} & . & . & . & \mathbf{0} \\ L(K) & \mathbf{0} & \mathbf{0} & \mathbf{0} & .. & \mathbf{A}_2^{(K)} & \mathbf{A}_1^{(K)} & \mathbf{A}_0^{(K)} \\ L(K+1) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_2^{(K+1)} & \mathbf{A}_1^{(K+1)} \end{array}$$

All component matrices $A_j^{(\cdot)}$for j=0, 1 and 2 of the generator **Q** are given by

$$\mathbf{A}_1^{(0)}=\begin{array}{c|cccc} & (0,0,0) & (0,1,0) & (0,0,1) & (0,1,1) \\ \hline (0,0,0) & -\lambda & 0 & 0 & 0 \\ (0,1,0) & \mu_2 & -\lambda_2 & 0 & 0 \\ (0,0,1) & \mu_3 & 0 & -\lambda_3 & 0 \\ (0,1,1) & 0 & \mu_3 & \mu_2 & -\lambda_{23} \end{array}$$

For n=0,1,…,(K),

$$\mathbf{A}_0^{(n)}=\begin{array}{c|cccc} & (n+1,0,0) & (n+1,1,0) & (n+1,0,1) & (n+1,1,1) \\ \hline (n,0,0) & \lambda & 0 & 0 & 0 \\ (n,1,0) & 0 & \lambda & 0 & 0 \\ (n,0,1) & 0 & 0 & \lambda & 0 \\ (n,1,1) & 0 & 0 & 0 & \lambda \end{array}$$

For n=1,2, … (K),

$$\mathbf{A}_1^{(n)}=\begin{array}{c|cccc} & (n,0,0) & (n,1,0) & (n,0,1) & (n,1,1) \\ \hline (n,0,0) & -\lambda_1 & 0 & 0 & 0 \\ (n,1,0) & \mu_2 & -\lambda_{12} & 0 & 0 \\ (n,0,1) & \mu_3 & 0 & -\lambda_{13} & 0 \\ (n,1,1) & 0 & \mu_3 & \mu_2 & -\lambda_{123} \end{array}$$

and for n=K+1

$$\mathbf{A}_1^{(K+1)}=\begin{array}{c|cccc} & (K+1,0,0) & (K+1,1,0) & (K+1,0,1) & (K+1,1,1) \\ \hline (K+1,0,0) & -\lambda_1 & \lambda & 0 & 0 \\ (K+1,1,0) & \mu_2 & -\lambda_{12} & 0 & \lambda \\ (K+1,0,1) & \mu_3 & 0 & -\lambda_{13} & \lambda \\ (K+1,1,1) & 0 & \mu_3 & \mu_2 & -\mu \end{array}$$

For n=1, 2, …,(K+1)

$$\mathbf{A}_2^{(n)}=\begin{array}{c|cccc} & (n-1,0,0) & (n-1,1,0) & (n-1,0,1) & (n-1,1,1) \\ \hline (n,0,0) & \mu_1 & 0 & 0 & 0 \\ (n,1,0) & 0 & \mu_1 & 0 & 0 \\ (n,0,1) & 0 & 0 & \mu_1 & 0 \\ (n,1,1) & 0 & 0 & 0 & \mu_1 \end{array}$$

All other '**0**' entries of **Q** matrix are zero matrices

B. Stationary Probability Vector

Let $\mathbf{\Pi}_n=(\pi_{n,0,0},\pi_{n,1,0},\pi_{n,0,1},\pi_{n,1,1})$ be a four component row vector of probability values of the states occupied by the QBD process **X(t)** in the long run.Alsoeach

$$\pi_{n,i,j}=\lim_{t\to\infty}P\big(\mathbf{X}(t)=(n,i,j)\big)$$

exists and is defined for n=0,1,…,K+1, i=0,1 and j=0,1. Thus

$$\mathbf{\Pi}=(\mathbf{\Pi}_0,\mathbf{\Pi}_1,...,\mathbf{\Pi}_{K+1})$$

represents the stationary probability vector of the QBD process **X (t)**. Let $\mathbf{1}_{4(K+1)}$ be a column vector of unit values and $\mathbf{e}=(1,1,1,1)^T$ be another column vector of size 4. This $\mathbf{\Pi}$ vector can be determined by solving

$$\mathbf{\Pi}\,\mathbf{Q}=\mathbf{0}$$

subject to the normalizing condition

$$\mathbf{\Pi}\,\mathbf{1}=\sum_{n=0}^{K+1}\mathbf{\Pi}_n\,\mathbf{e}=1$$

Following a modified procedure by cutting the levels off starting from the upper level L(K+1) and moving down to the lowest level L(0) for the generator matrix **Q** developed from that of Latouche and Ramasamy[15], one can show that the stationary probability vector $\mathbf{\Pi}$ has a matrix-geometric solution given by

$$\mathbf{\Pi}_0\,\mathbf{U}_0=\mathbf{0}$$

$$\mathbf{\Pi}_n=\mathbf{\Pi}_{n-1}\,\mathbf{R}^{(n)} \text{ for n=1, 2, ..., (K+1)}$$

$$\mathbf{\Pi}\,\mathbf{1}=\sum_{n=0}^{K+1}\mathbf{\Pi}_n\,\mathbf{e}=1$$

where $\mathbf{U}^{(K+1)}=\mathbf{A}_1^{(K+1)}$,

$\mathbf{U}^{(n)}=\mathbf{A}_1^{(n)}+\mathbf{A}_0^{(n)}(-\mathbf{U}^{(n+1)})^{-1}\mathbf{A}_2^{(n+1)}$ for n= K, K-1 , …, 1 and 0

and $\mathbf{R}^{(n)}=\mathbf{A}_0^{(n-1)}(-\mathbf{U}^{(n)})^{-1}$ for n=1, 2,… ,(K+1).

It is remarked that the final vector ' $\mathbf{\Pi}$ ' is normalised by $(\mathbf{\Pi}\,\mathbf{1})^{-1}\mathbf{\Pi}$. The last equation of the system $\mathbf{\Pi}\,\mathbf{Q}=\mathbf{0}$ or $\mathbf{\Pi}_K\,A_0^{\,K}+\mathbf{\Pi}_{K+1}\,A_1^{\,K+1}=\mathbf{0}$ is given by

$\lambda\,\pi_{K11}+\lambda\,\pi_{(K+1)10}+\lambda\,\pi_{(K+1)01}=(\mu_1+\mu_2+\mu_3)\,\pi_{(K+1)11}\ldots$ (4)

Rewriting (4), it is seen that

$\pi_{(K+1)11}=\rho\,(\pi_{K11}+\pi_{(K+1)10}+\pi_{(K+1)01})$ …(5)

Further the marginal distribution $\{a_n= P(X_1(t)=n: n=0, 1, \ldots, (K+1)\}$ of $X_1(t)$ that represents the queue length process in an $M/M_n/1/(K+1)$ loss system (see Baily(1957)) is given by

$$a_n=\sum_{i,j=0}^{1}\pi_{n,i,j}=\frac{(1-\rho_1)\rho_1^{\,n}}{1-\rho_1^{(K+2)}},\quad n=0,1,2,\ldots,(K+1) \qquad \ldots(6)$$

where $\rho_1=\dfrac{\lambda}{\mu_1}$.

The values of the marginal probabilities

b_0= P(both slow servers S_2 and S_3 are idle),

b_1= P(slow servers S_2 alone is busy) ,

b_2= P(slow servers S_3 alone is busy) , and

b_{12}= P(Both slow servers S_2 and S_3 are busy)

are given by

$$b_0=\sum_{n=0}^{K+1}\pi_{n00},\ b_1=\sum_{n=0}^{K+1}\pi_{n10},\ b_2=\sum_{n=0}^{K+1}\pi_{n01},$$

$$\text{and } b_{12}=\sum_{n=0}^{K+1}\pi_{n11},\ b_0+b_1+b_2+b_{12}=1$$

Thus the expected number of customers in the system, say $L_{(K+2)}$, is obtained as

$$L_{(K+1)}=\sum_{n=0}^{K+1}n\,a_n+b_1+b_2+2\,b_{12} \qquad \ldots(7)$$

Computational Complexity: In the finite case (i.e. $K<\infty$), elements of $U^{(i)}$ matrices depend on the levels since there is a finite upper boundary (K+1) in the generator Q. Further this algorithm suffers from a computational complexity which amounts to O(4(K+1)).

III. M/(M1,M2 M3)/2/(B1,B2) QUEUE WITH STALLING INTO B1 BUFFER

Replacing the vector process **X(t)** =($X_1(t)$, $X_2(t)$, $X_3(t)$: $t \geq 0$) defined on the space $\mathbb{S}$ ={0,1,2,…K+1) }x(0,1)x(0,1)} by **Y**(t)=($Y_1(t)$,$Y_2(t)$, $Y_3(t)$: $t \geq 0$) process defined on the full space W= $\mathbb{S}$ U{K+4,K+5,…,∞} to monitor the transitions of queue with stalling described in Fig.1.

A. *Linking of Infinite Queue Length with Finite Queue*

For this extension, there exists a proportionality constant, say β, expected to be a function of K such that

$\alpha_{nij}=\beta\ \pi_{nij}$ …(8)

for n=0,1,2,...,(K+1), i=0,1 and j=0,1

$$p_n\sum_{i,j=0}^{1}\alpha_{nij}\ \text{for}\quad 0,1,2,\ldots,(K+1) \qquad \ldots(9)$$

$p_n=\alpha_{(K+1)11}\rho^{n-(K+3)}$ for n>(K+3) …(10)

Hence the normalizing condition

$$\sum_{n=0}^{K+1}p_n+\sum_{n=K+4}^{\infty}p_n=1,$$

Now, using the facts $\sum_{n=0}^{K+1}a_n=\sum_{n=0}^{K+1}\sum_{i,j=0}^{1}\pi_{n,i,j}=1$, and $\pi_{(K+1)11}=\rho\,(\pi_{K11}+\pi_{(K+1)10}+\pi_{(K+1)01})$ given by(5) , it is obtained that

$$\beta=\frac{1-\rho}{1-\rho+\rho\,\pi_{(K+1)11}} \qquad \ldots(11)$$

Further, fraction D_2 of the time Server-2 is busy is given by

$D_2=\beta[b_1]$ …(12)

Fraction D_3 of the time Server-3 is busy is given by

$D_3=\beta[b_2]$ …(13)

Fraction D_{23} of the time both Server-2 and Server-3 are busy is given by

$D_{23}=\beta[b_{12}+(\rho\pi_{(K+1)11})/(1.0-\rho)]$ …(14)

Fraction D_0 of the time both Server-2 and Server-3 are idle is given by

$D_0=\beta b_0$ …(15)

Fraction of the time Server-1 is idle is $D_{a0}=\beta\,[a_0]$ and Fraction of the time Server-1 is busy is

$D_{a1}=\beta[1-a_0]+\beta\{(\rho\pi_{(K+1)11})/(1.0-\rho)]$ …(16)

The mean number E(L) of customers in the system is given by

$E(L)=\beta[L_{(K+1)}+\pi_{(K+1)11}\rho\{(K+4)(1-\rho)+\rho\}/(1-\rho)^2]$ …(17)

B. *Distribution of the System Size*

Let the probability of finding the system size (queue + service) at 'n' beq_n for n=0,1,2,…,∞. Then this distribution $\{q_n\}$ of system size can be obtained from $\{\alpha_{nij}\}$ values as follows:

q_0= P(idle system)=$\beta\pi_{00}$

q_1= P(system size is 1)= $\beta\,(\pi_{100}+\pi_{010}+\pi_{001})$

q_n= P(system size is n)=$\beta(\pi_{n00}+\pi_{n-110}+\pi_{n-101}+\pi_{n-211})$ for n=2, 3, 4, …(K+1)

q_{K+2}= P(system size is (K+2))

$=\beta(\pi_{(K+1)10}+\pi_{(K+1)01}+\pi_{K11})$

q_{K+3}= P(system size is (K+3))= $\beta\,\pi_{(K+1)11}$

$=\rho\,q_{K+2}$ by using (5)

q_n= P(system size is n)=

$q_{K+2}\rho^{n-(K+2)}$ for n=(K+3),(K+4) …, ∞ …(18)

An alternative method of finding the mean number of customers in the system is $E(L)=\sum_{n=0}^{\infty}n\,q_n$

$$=\sum_{n=0}^{K+2}n\,q_n+q_{K+2}[K+3)-(K+2)\rho]\rho/(1-\rho)^2 \qquad \ldots(19)$$

By assigning K=0 in (18) and (19) one can deduce the corresponding steady state distribution of the number of customers in the three server heterogeneous system of $M/M_i/3$ queues and its mean value respectively as studied by Singh[9].

IV. Comparison of M/M$_N$/3 and M/M/3 Queues

In order to compare the steady state results of homogeneous M/M/3 and non-homogenous M/M$_n$/3 queueing systems, a criterion is suggested[1] as below:

If μ_n (n=1, 2, 3) are the service rates of the three servers S_1, S_2 and S_3 of the heterogeneous system M/M$_n$/3, then the service rate for each server of the homogeneous system M/M/3 is the average service rate $\mu = \frac{\mu_1 + \mu_2 + \mu_3}{3}$.

For an illustration, the input parameter values are selected at random as λ=0.9384, μ_1=0.38, μ_2=0.37, and μ_3=0.27. It is noted that ρ=0.92 since μ=0.34.Corresponding output relating to a few steady state characteristics of both systems for this specific case are computed numerically and reported in Table 1 and Table 2.

Table 1: Results for NH and H systems when λ=0.9384, μ_1=0.38, μ_2=0.37, and μ_3=0.27.

	NH	H	NH	H
K=3	h_0=0.0221	h_0=0.0217	b_0=0.2021 b_1=0.2165 b_2=0.1482 b_{12}=0.4332	b_0=0.1963 b_1=0.2409 b_2=0.1227 b_{12}=0.4400
K=2	h_0=0.0421	h_0=0.0286	b_0=0.1869 b_1=0.2091 b_2=0.1475 b_{12}=0.4565	b_0=0.1822 b_1=0.2348 b_1=0.1226 b_{12}=0.4604
K=1	h_0=0.0284	h_0=0.0231	b_0=0.1608 b_1=0.1945 b_2=0.1432 b_3=0.5015	b_0=0.1603 b_1=0.2237 b_2=0.1209 b_{12}=0.4951
Condition	h_0(NH) > h_0(H)		b_0(NH)>b_0(H), b_1(NH)<b_1(H),	

Such numerical results of the two systems are distinguished using short symbols NH and H for the results of heterogeneous and homogeneous systems respectively.

Table 2: Mean System Size for NH and H systems when λ=0.9384, μ_1=0.38, μ_2=0.37, and μ_3=0.27.

	NH	H	NH	H
K=3	E(a)=1.12 E(b)=1,23 2.35	E(a)=1.51 E(b)=1.24 2.75	E(L)=7.78	E(L)=10.13
K=2	E(a)=0.99 E(b)=1,27 2.26	E(a)=1.24 E(b)=1.28 2.52	E(L)=9.90	E(L)=11.23
K=1	E(a)=0.80 E(b)=1.34 2.14	E(a)=0.92 E(b)=1.34 2.26	E(L)=11.82	E(L)=12.21
	Observed Condition		E(L)(NH)<E(L)(H)	

If the mean number of customers E(L)= $\sum_{n=0}^{\infty} n\ q_n$ of the heterogeneous (NH) system is smaller than the corresponding homogeneous H system when both of the systems operate under the same ρ<1 across the values of K>0, it is concluded that the former system NH performs better than the latter system. The following conditions ensure that a three server heterogeneous system performs better than the corresponding homogeneous system:

(i) q_0(NH) > q_0(H)
(ii) b_0(NH) >b_0(H), b_1(NH) < b_1(H)
(iii) [E(a)+E(b)](NH)< [E(a)+E(b)](H)

where E(a)= $\sum_{n=0}^{K+1} n\ a_n$ and E(b)=b_1+b_2+2b_{12}

Since from the matrix geometric solution $\mathbf{\Pi}$ and hence from the stationary distribution {q_n: n=0,1,…,∞}, it is not easy to obtain scalar expressions for the components of $\mathbf{\Pi}$ as functions λ and ρ , the above conditions (i) to (iii) have been established numerically. There is little scope to see three server homogeneous systems in real life applications as compared with heterogeneous type of queueing problems. Nevertheless, heterogeneous class of queueing models have grater scope in dynamic routing of messages arriving at the buffers and then dispatched to one of the computer systems belonging to a communication network for its transmission towards a destination (see[3])

V. Conclusion

This paper analyses a Markovian queuing system M/M$_n$/3/(B$_1$,B$_2$) with stalling.In this work, both scalar analytical and matrix analytical methods have been used for studying the prosed queueing systems with heterogeneous service times. If μ_i, is the service rate of the ith(heterogeneous) server for i= 1, 2 and 3, then a queue discipline of threshold type is formulated to stall the arriving customers in queue-1 and to accommodate the waiting customers in queue-2 with a despatching rule from the two queues to the servers. For the case of informed customers,describing the 'Queue Length' process of the system M/ (M$_1$,M$_2$,M$_3$)/3/(B$_1$,3) as a QBD process, stationary probability distribution of the queue length and performance measures such as the expected queue length, the probability that each server is busy have been obtained. Further,assuming ρ=(λ/μ)<1 where λ=mean arrival rate, and μ= μ_1+ μ_2+μ_3, all these results are linked to the infinite queue length process of M/(M$_1$,M$_2$,M$_3$)/3/(B$_1$,B$_2$) system.

To support advantages of these results, a numerical study is then carried out. Conditions are established for a better performance of the three server heterogeneous system over the three server homogeneous system.

There is much scope to extend this queueing model with staling of customers by allowing the fast server to render service in batches or allowing the faster server to provide service with a general service time distribution.

Acknowledgment

The authors would like to thank the authorities at the University of Botswana for providing the infrastructure facilities for this piece.

References

[1] Gumbel, H.(1960). Waiting lines with heterogeneous servers, *Operations Research*, vol.8, no.4, pp.504-511
[2] Krishnamoorthi, B.(1963).On Poisson queue with two heterogeneous servers. *Operations Research*, 2, No. 3 321-330.

[3] Lin, W. and Kumar,P.R.,(1984). Optimal control of a queuing system with two heterogeneous Servers. IEEE Transactions on Automatic Control, 29, pp.696-703.

[4] Rubinovitch, M.(1985a) The slow server problem, J.Appl.prob. , 22, 205-213.

[5] Rubinovitch, M.(1985b) The slow server problem: A queue with stalling, J. Appl. Prob., 22, 879-892.

[6] Abou-El-Ato, M. O.and A. L.Shawky (1999) A simple Approach for the Slow Server Problem,*Commum.fac.Univ.Ank,SeriesA,V.*48., pp 1-6.

[7] Fabricio Bandeira Cabari (2005). The slow server problem for uninformed Customers. Queuing systems, 50, 353-370. Bailey, N.T.J.(1957). Further result in the non-equilibrium Theory of a Simple Queue. *J.R. Statist.soc.,* B19, 326-333.

[8] Kim, J.H., Ahn H.S and R. Righter(2011). Managing queues with heterogeneous servers. *Journal of Applied Probability*,48,No2435-2452.

[9] Singh, V.P.(1971). Markovian queues with three heterogeneous servers. AIIE Transations, vol.3, no.1, pp.45-48.

[10] Singh, V.P.(1976). A heterogeneous system with finite waiting space. Journal of Engineering Mathematics, Vol. 10, No. 2,, pp 125-134

[11] Sivasamy, R., Daaman, O.A. and S. Sulaiman, (2015).An M/G/2 Queue subject to a minimum violation of the FCFS queue discipline, *European Journal of Operational Research*, 240, pp 140-146.

[12] Sivasamy, R., and K. Thaga, (2016a). Discrete Time Queues operated by Two Heterogeneous Servers under 'Fisrst Come First Served' Queue Discipline, *Proceedings of Dynamic Systems and Applications*, 7, pp 1-10

[13] Sivasamy, R., Paulraj, G., Kalaimani, S., and N. Thillaigovindan(2016b). A two server Poission Queue Operating under FCFS Discipline with an 'm' Policy, *Singapore SG* January 07-08, 2016,18 part 1

[14] Xuelu Zhang, Jinting Wamg, Tien Van Do (2015).Threshold properties of the M/M/1 queue under T-policy with applications, Applied Mathematics and Computation, volume 261, 15, pages 284-301

[15] Latouche. G, and V. Ramaswami (1999). Introduction to Matrix Analytic Methods in Stochastic Modeling, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

AUTHOR'S PROFILE

Prof. Sivasamy was born in Chidambaram town , India on 19 February 1953. Sivasamy received his B.Sc degree (Mathematics) from Madras University (1974) and his M.Sc.(Statistics, 1976) and Ph.D. in Statistics(1989) from Annamalai University, India. Sivasamy's teaching interests include stochastic processes and their applications in Operations research, Queues, Reliability theory, Survival Analysis, Inventory control theory and statistical process control. His primary research interests are in the field of stochastic processes and optimization problems.

Professor Sivasamy joins University of Botswana as an Associate Professor of the Statistics Department in February 2008 and was promoted to the rank of full professor in 2013. Prior to coming to University of Botswana, he was a faculty member at Annamalai University from 1978 to 2010; he entered as Lecturer in Statistics on 14-07-1978 and retired as head of the department of Statistics on 06-08-2010,As a member of the academic council, he prepared good number of policies of the Undergraduate and/or Graduate courses, Research Policy and Student Affairs Policy. He organised Seminars and Symposia. He was also Principal investigator of a UGC major project. He won an award called "**Prof. P. V. Sukhatme**" award in 2013 for his teaching.He is a friendly teacher with number of research papers to his credit,

Prof R Sivasamy served as the Head of the Department of Statistics from 2002 to 2010, by virtue of which he was the member of the Senate of the Annamalai University: he recommended curricula and degrees for approval as and when academic policies required Senate approval. As a member of the academic council, heprepared good number of policies of the Undergraduate and/or Graduate courses, Research Policy and Student Affairs Policy. Currently, Prof Sivasamy is a member of (i) International Biometric Society (IBS), (ii) Indian Science Congress Association (ISCA), and (iii) Indian Society for Probability and Statistics (ISPS) and (iv) Botswana Statistics Association (BOSA).

Professor Keoagile Thaga was born in Serowe, Botswana on the 2nd January 1964. He received his BA degree (Statistics) in 1991 from the University of Botswana, BSc and PhD (both in Statistics) in 1994 and 2004 respectively from the University of Manitoba, Winnipeg, Canada. His teaching interest includes statistical quality control and applied statistics

Prof. Thaga's primary research interest is Statistical Process Control with particular emphasis on the development of new control charts used for detecting small shifts in the process quality characteristics.

He joined the University of Botswana in August 1991 as a Staff Development Fellow and was appoint to a lecturer position in June 1994 after completing his MSc degree. He was promoted to the ranks of Senior Lecture and Associate Professor in 2006 and 2010 respectively. He was further promoted to the rank of full professor in February 2016.

Prof. Thaga is currently the Deputy Dean in the Faculty of Social Sciences. He has also served for six years as HOD-Statistics and as a graduate coordinator in the department for four years and played a key role in developing the PhD programme which started in August 2010. He has supervised two PhD students, 2 Master's and many undergraduate students in the department. He is a member of the International Biometric Society (IBS) and Botswana Statistics Association (BOSA).

Dr. Lucky Mokgatlhe was born in Francistown on 12th August 1967. Lucky received a BA (Statistics) from University of Botswana in May 1991, an MS (Statistics) from Iowa State University, USA in 1995, an MA (Applied Statistics) from Pittsburgh, Pennsylvania (USA) in 2000 and a PhD in Statistics with emphasis in survival analysis from University of Free State, South Africa in 2003.

Dr. Mokgatlhe has teaching interest in survival analysis, categorical data analysis, linear models and biostatistics. His primary research area is in modelling HIV data using linear and survival models, use ROC curves to determine cutoff points for metabolic syndrome and study of disease prevalence in Botswana.

Dr. Lucky Mokgatlhe joined the Department of Statistics at the University of Botswana as a staff development fellow in 1992 and was appointed to lecturer position on completing his master's degree in 1995. In 2014 he was promoted to Senior Lecturer position and in 2016 was appointed as Head of the Department of Statistics. Dr. Mokgatlhe has taught both undergraduate and graduate courses.

He is presently the Principal Investigator in the Development of Biostatistics Curricula D71 grant awarded by Fogarty and a key personnel in the P20 Biostatistics Training grant awarded by NIH, both awarded 2016. He has supervised two PhD students who have graduated. He is a member of the Biometric Society and Botswana Statistics Association.