Steady State Analysis Of (M/M/1):(Fcfs/K/∞) Queueing Network With Feedback Retrial And Blocking

S.Shanmugasundaram[1], C.Baby[2]

[1]Department of Mathematics, Government Arts College (Autonomous), Salem-636007, Tamil Nadu, India, Email : sundaramsss@hotmail.com
[2]Department of Mathematics, Government Arts College (Autonomous), Salem-636007, TamilNadu, India, Email: babyakcika2015@gmail.com

Received: 17.04.2024 Revised : 18.05.2024 Accepted: 20.05.2024

ABSTRACT
In this research, we study a single server retrial feedback queueing network with blocking. If there is room, external consumers join the queue with probability s; if not, they go to orbit with probability 1-s; if they receive service at node, either the node with probability 1-p or the node with probability p is reached. Similarly, he can choose to leave the system with probability q or join node with probability 1-q after receiving service in node. In the case of node, after obtaining the service (feedback), he either moves to node with the probability 1-r or leaves the system with the probability r.Here, we look at the length of the queue and the waiting time for each of the four nodes. We also examine the scenario both with and without blocking in the steady state. To verify that the model is accurate, numerical examples are provided.

Keywords: Customers, Server, Queueing Network, Feedback, Retial, Blocking

INTRODUCTION
A frequent word for a waiting line is "queue." It happens every time a customer arrives and has to wait in line to be served at one or more service locations. It arises when there are more customers arriving than there are being served in a given amount of time.Different scenarios of queuing up are: patients waiting for doctor's clinic, planes arriving in an airport for landing etc.,
Queueing theory was first introduced by Agner Krarup Erlang [2] ,who was a Mathematician, Statistician and an Engineer, his first paper served as foundation of queueing theory.
When a queueing system's operation is dependent on time, it is referred to as being in a transient state.A queueing system is considered to be in a steady state when its operational characteristics are time-independent.
The queueing systems we have seen so far consist of a single service facility with one or more servers; however, the queueing systems we encounter in real-world scenarios are frequently isolated but participate in structured systems known as queueing networks, which are networks of service facilities where clients receive service at some or all of the facilities. For instance, a factory may have multiple queues connected by the logical flow of the production process.
Queueing network was introduced by James R Jackson in 1957.A queueing network can be divided into three categories. There are two categories of queuing networks: closed and open. Open networks take traffic from outside sources and route it to outside locations. The population of closed networks moves constantly, hopping between queues without ever leaving the system.
Retrial queues, also called orders with repeated requests, are produced by the queuing theory, which postulates that new users may retry for service at a later time if they discover the server is full. Classical retrial policies and continuous retrial policies are the two categories of retrial queues. Since consumers in orbit behave independently of one another, the number of customers in orbit defines the retrial rate for classical retrial policy. Recurring consumers form an orbiting queue under a continual retrial policy, and the first-ranking customers can request a service following a randomly determined retrial time. Kosten L [6] established the retry queue in 1947. Shan Gao and Tao Li Liyuan Zhang talked about the traditional retrial policy and the M/M/1 retrial queue with working vacation interruption. [11]. A study of the primary findings and techniques of the theory of retrial queues, focussing on Markovian single and multi-channel systems, has been evaluated by Falin G.I. [3].
When a customer completes their service at one queue and then enters the station again for another round of service, this is known as a feedback network or cyclic network. The customer enters the station

for service and leaves the station after the service is finished. In the event that the client is not happy with the service, he goes back to the same queue (instant feedback); if he goes to a different queue, feedback is delayed. Finch [4] proposed the idea of feedback in his study, Cyclic Queues with Feedback. Markovian queueing network with feedback on a single server was examined by Shanmugasundaram S and Vanitha S [7]. Sreekala M.S , Manoharan M [9] describe how system dynamics was used as a central part of a complex system of fertility health care to improve its performance using queueing network.
Blocking can usually happen in a network of queues when some or all of the queues have finite buffer capacities. The flow of customers from that node will be stopped if every server on the destination node is in use and every waiting area is packed with waiting customers. Takahashi Y, Miyahara H, and Hasegawa T described an approximation approach for investigating open restricted queueing networks with poisson arrival and exponential service time. [10]. Following the previously mentioned idea, Sreekala M.S. and Manoharan M. [8] conducted a node-by-node decomposition analysis by including effective arrival rate and service rate. Mathematical studies were employed by Weiss E.N. and McClain J.O. [12] to analyze queues with blocking.

An explanation of the model

In this study, we investigate an open queuing network featuring feedback, blocking, four single server nodes, finite and indefinite capacity. The arrival of consumers is controlled by a homogeneous Poisson process with an arrival rate of λ. In this case, the relationship between the service rate and connected nodes and the service time is exponential. In the event that node one's server is idle, customers receive service instantly; external customers either wait in the orbit, where they have a probability of 1-s, or they enter the non-trivial group with a probability of s.The customers wait in the orbit till the server in node one gets free, the orbit is of infinite size and the retrial rate is α. Once the service is finished at node one, where customers arrive with an arrival rate of λ, they are either routed to node two with a probability of p or entered into node three with a probability of 1-p. For consumers from the orbit to node one, the retry rate is equal to nα, where n is the total number of customers in the orbit. If the service is finished with a probability of q or if it is routed to node three with a probability of 1-q, customers exit the system. Customers are either routed to node 4 (instant feedback) with probability 1-r, or if the service is finished in node 3, quit the system with probability r. Customers enter node one with chance 1 if the service is finished at node 4.In this model node is of infinite capacity, where as node , node and node follows finite capacity.Here all the four stations follows (M/M/1):(FCFS/K/∞) schedule.
Figure 1 displays the diagrammatic representation of the aforementioned paradigm.

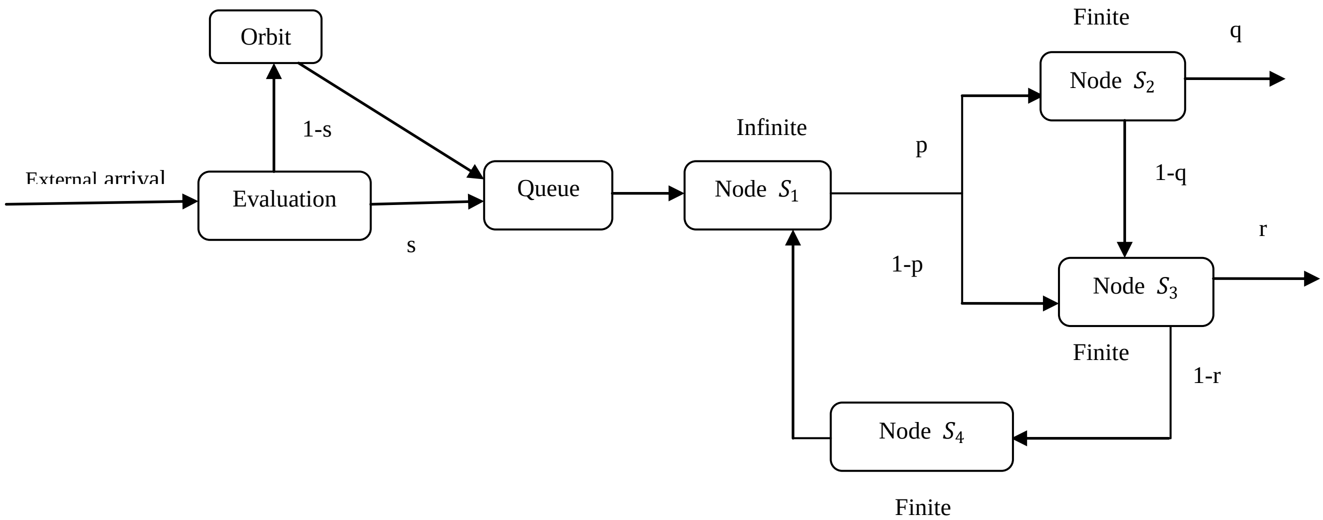

Fig 1. blocking network for queuing

Because there is a limited amount of waiting space between nodes in a network, blocking happens when nodes block node to node. Every node's arrival rate is determined by a poisson process at a specific rate. Congestion that manifests in a specific node when blocking occurs. It might have an impact on every upstream node's level of congestion. Using the idea of effective service time, service time is adjusted to corporate block between nodes. [9] .In our model blocking occurs at the flows and because node two, node three and node four have finite capacities, If the immediate node has limitless capacity, blocking won't happen. The flow is blocking free because node one has infinite capacity, then the remaining flows are modeled. The types of congestion associated in our model are summarized in the table 1

Table 1. Congestion types for each flow

Flow	Cause of congestion	Facing station	Congestion type
S_1 to S_2 S_1 to S_3	S_2 is full S_3 is full	S_1	Classic Congestion
S_2 to S_3 S_3 to S_4	S_3 is full S_4 is full	S_2 S_3	Blocking
S_4 to S_1	Not applicable	Not applicable	No Congestion

Steady-state Analysis

We first analysis the present queueing network without considering blocking between the stations, Next, we apply blocking to the queueing network modification. Regarding the examination of queueing models in a steady state, we employ the steady state methodology ([5] and [1]). Our steps are as follows.

Steady State Analysis without Blocking

The routing matrix in our model has limitless space, and analysis is written as

i,j = 1,2,3,4 , where is likelihood of routing from station i to station j.

The network's overall arrival rate, internal arrival rate, and exterior arrival rate are indicated from traffic equations based on the routing probabilities in Figure 1

For node 1 : $\lambda_1 = \lambda s + n\alpha + P_{41}\lambda_4 = \lambda s + n\alpha + \lambda_4$ (1)

For node 2 : $\lambda_2 = \lambda_1 P_{12} = p\lambda_1$ (2)

For node 3 : $\lambda_3 = \lambda_1 P_{13} + \lambda_2 P_{23} = (1-p)\lambda_1 + (1-q)\lambda_2$ (3)

For node 4 : $\lambda_4 = \lambda_3 P_{34} = (1-r)\lambda_3$ (4)

Solving the above equations, we get

$$\lambda_1 = \frac{\lambda s + n\alpha}{r(1-pq)+pq} \quad (5)$$

$$\lambda_2 = \frac{p(\lambda s + n\alpha)}{r(1-pq)+pq} \quad (6)$$

$$\lambda_3 = \frac{(1-pq)(\lambda s + n\alpha)}{r(1-pq)+pq} \quad (7)$$

$$\lambda_4 = \frac{(1-pq)(1-r)(\lambda s + n\alpha)}{r(1-pq)+pq} \quad (8)$$

We assume that all the stations are of infinite capacity. Thus, by using the M/M/1/∞ model, each station may be solved independently. The effect of congestion between any of the system's stations is not taken into account while analyzing the steady state.

Average time spent in line

Expected queue length at station i is

$$L_i^q = \frac{\rho_i^2}{1-\rho_i}, \ i = 1,2,3,4 \quad (9)$$

Where $\rho_i = \frac{\lambda_i}{\mu_i} < 1$

$$L_1^q = \frac{\rho_1^2}{1-\rho_1} = \frac{\lambda_1^2}{\mu_1(\mu_1-\lambda_1)} = \frac{(\lambda s+n\alpha)^2}{\mu_1(r+pq-pqr)[\mu_1(r+pq-pqr)-(\lambda s+n\alpha)]} \quad (10)$$

$$L_2^q = \frac{\rho_2^2}{1-\rho_2} = \frac{\lambda_2^2}{\mu_2(\mu_2-\lambda_2)} = \frac{p^2(\lambda s+n\alpha)^2}{\mu_2(r+pq-pqr)[\mu_2(r+pq-pqr)-p(\lambda s+n\alpha)]} \quad (11)$$

$L_3^q = \frac{\rho_3^2}{1-\rho_3}$

$= \frac{\lambda_3^2}{\mu_3(\mu_3-\lambda_3)}$

$= \frac{(1-pq)^2(\lambda s+n\alpha)^2}{\mu_3(r+pq-pqr)[\mu_3(r+pq-pqr)-(1-pq)(\lambda s+n\alpha)]}$ (12)

$L_4^q = \frac{\rho_4^2}{1-\rho_4}$

$= \frac{\lambda_4^2}{\mu_4(\mu_4-\lambda_4)}$

$= \frac{[(1-r)(1-pq)]^2(\lambda s+n\alpha)^2}{\mu_4(r+pq-pqr)[\mu_4(r+pq-pqr)-(1-r)(1-pq)(\lambda s+n\alpha)]}$ (13)

Average queue delay

Anticipated constant state waiting duration to access at the station

$W_i^q = \frac{L_i^q}{\lambda_i}$

$W_i^q = \frac{\rho_i^2}{\lambda_i(1-\rho_i)}$, i = 1,2,3,4 (14)

Average queue lengths for respective nodes are as follows

$W_1^q = \frac{L_1^q}{\lambda_1}$

$= \frac{(\lambda s+n\alpha)}{\mu_1[\mu_1(r+pq-pqr)-(\lambda s+n\alpha)]}$ (15)

$W_2^q = \frac{L_2^q}{\lambda_2}$

$= \frac{p(\lambda s+n\alpha)}{\mu_2[\mu_2(r+pq-pqr)-p(\lambda s+n\alpha)]}$ (16)

$W_3^q = \frac{L_3^q}{\lambda_3}$

$= \frac{(1-pq)(\lambda s+n\alpha)}{\mu_3[\mu_3(r+pq-pqr)-(1-pq)(\lambda s+n\alpha)]}$ (17)

$W_4^q = \frac{L_4^q}{\lambda_4}$

$= \frac{(1-r)(1-pq)(\lambda s+n\alpha)}{\mu_4[\mu_4(r+pq-pqr)-(1-r)(1-pq)(\lambda s+n\alpha)]}$ (18)

Analysis of steady states using blocking

Blockages at a particular station may have an impact on block levels at all upstream stations when stations reach their finite capacity. We modify Jackson's approach in this model to handle hold-related station collaboration. Using the idea of effective service time $\frac{1}{\mu_1}$, this analysis examines the queueing network with blocking [10]. The duration of service in unblocked networks is.The mean effective service time at station i is denoted as $\frac{1}{\mu_1}$, and we assume that the effective service time when blocking occurs follows an exponential distribution.

Generally speaking, effective waiting times are convex combinations, and effective service time is given by

$\frac{1}{\bar{\mu_i}} = P_{i0}\left(\frac{1}{\mu_i}\right) + \sum_j P_{ij}\left(\frac{1}{\mu_i} + W_j\right)$

Where P_{i0} is the routing probability of customers leaving the system from i^{th} station without facing any wait, P_{ij} represent the routing probability from i^{th} station to j^{th} station. In our model node two (S_2) and node three (S_3) faces blocking, the effective service time corresponding to S_2 and S_3 are as follows

$\frac{1}{\bar{\mu_2}} = P_{20}\left(\frac{1}{\mu_2}\right) + r_{23}\left(\frac{1}{\mu_2} + W_3^q\right)$ (19)

$\frac{1}{\bar{\mu_3}} = P_{30}\left(\frac{1}{\mu_3}\right) + r_{34}\left(\frac{1}{\mu_3} + W_4^q\right)$ (20)

Using the above equations with (9) and (14), we get steady state mean queue lengths and waiting times in terms of effective service time for each station.

Analysis of Stations

By applying the node by node decomposition approximation technique, we analyze the steady state of each station. Updated arrival and service parameters are examined independently because the network is divided into individual nodes. Consequently, the network is solved independently and separately beginning at node four and finishing at node one using the single node decomposition approach. Here, each finite node that adheres to M/M/1/∞ undergoes a steady state analysis utilizing the single node decomposition approximation approach.

Steady state analysis at $\mathbf{S_4}$

In our model, customers from node S_3 is directed to node S_4 (feedback), which is finite capacity node, but its downstream node is S_1 which is infinite, so S_4 does not face blocking.

Effective service to station S_4 is

$$\frac{1}{\overline{\mu_4}} = \frac{1}{\mu_4}$$

So, queue length and queue delay are obtained by using (9) and (14). Node S_4 receives customers only from node S_3.Queue length at station S_4, which is steady state, number of blocked customers at S_3 waiting to enter S_4 is denoted by

$$L_{34}^q = L_4^q\left[\frac{\lambda_{34}}{\lambda_4}\right]$$
$$= L_4^q\left[\frac{\lambda_4}{\lambda_4}\right]$$
$$L_{34}^q = L_4^q \qquad (21)$$

Steady state analysis at $\mathbf{S_3}$

Here, there are two ways to enter station S_3 , one is from station S_1 and another is from station S_2 , station S_2 experience blocking if station S_3 is full, so the queue length an queue delay for station S_3 are obtained by solving (9) and (14) interms of effective service time expressed in (20).The following equations yield the length of the line to enter each station.

$$L_{13}^q = r_{13}L_3^q$$
$$= L_3^q\left[\frac{\lambda_{13}}{\lambda_3}\right]$$
$$= L_3^q\left[\frac{1-p}{1-pq}\right] \qquad (22)$$
$$L_{23}^q = r_{23}L_3^q$$
$$= L_3^q\left[\frac{\lambda_{23}}{\lambda_3}\right]$$
$$= L_3^q\left[\frac{p(1-q)}{1-pq}\right] \qquad (23)$$

Steady state analysis at $\mathbf{S_2}$

Station S_2 receives customers from station S_1. The queue length (9) and queue delay (14) of station S_2 in terms of effective service time are calculated using (19). So, number of customers at S_1 waiting to enter S_2 is given by

$$L_{12}^q = L_2^q\left[\frac{\lambda_{12}}{\lambda_2}\right]$$
$$= L_2^q\left[\frac{p\lambda_1}{\lambda_2}\right]$$
$$= L_2^q \qquad (24)$$

Steady state analysis at $\mathbf{S_1}$

Station S_1 receives customers from external and from station S_4 . Since it has infinite capacity, S_1 does not face blocking. Effective service to station S_1 is

$$\frac{1}{\overline{\mu_1}} = \frac{1}{\mu_1}$$

So, queue length an queue delay are obtained by using (9) and (14).The queue length at station S_1 which is steady state, number of customers waiting to enter station S_1 is denoted by

$$L_{41}^q = L_1^q\left[\frac{\lambda_{41}}{\lambda_1}\right]$$
$$= L_1^q\left[\frac{L_{41}\lambda_1}{\lambda_1}\right]$$
$$= L_1^q[(1-r)(1-pq)] \qquad (25)$$

Numerical Examples

The steady state probabilities and the performance measures with and without blocking are computed for the given parameter values

$\lambda = 0.4, s = 0.5, n = 2, \alpha = 0.3, p = 0.5, q = 0.3, r = 0.5, \frac{1}{\mu_1} = 0.6, \frac{1}{\mu_2} = 0.7, \frac{1}{\mu_3} = 0.5, \frac{1}{\mu_4} = 0.6$

Table 2. Performance measures with and without blocking

Station	Performance Indicator	With blocking	Without blocking
S_1	L_1^q	4.152	4.152
	L_{41}^q	1.764	1.764
	W_1^q	2.985	2.985
S_2	L_2^q	4.4	0.470
	W_2^q	6.321	0.675
S_3	L_3^q	2.863	0.853
	W_3^q	2.422	0.733
	L_{23}^q	1.179	0.351
	L_{13}^q	1.684	0.501
S_4	L_4^q	0.195	0.195
	W_4^q	0.329	0.329

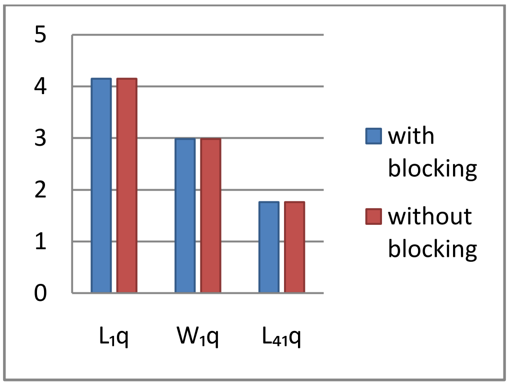

Fig 2: Analysis of Station S_1

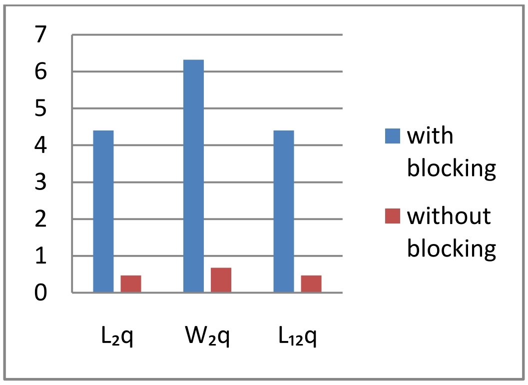

Fig 3. Analysis of Station S_2

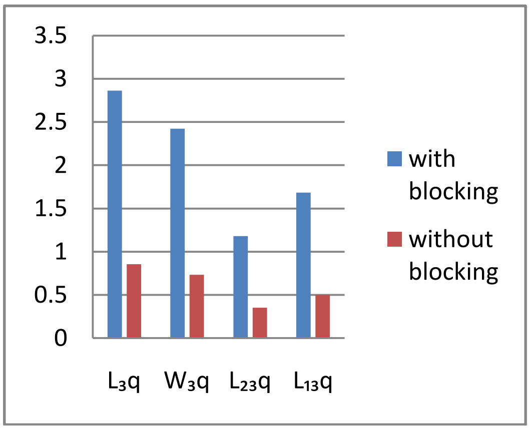

Fig 4. Analysis of Station S_3

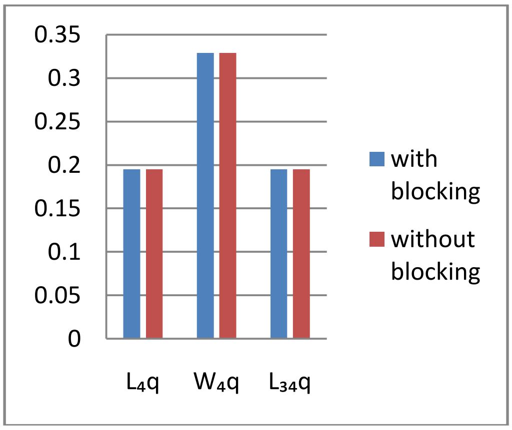

Fig 5. Analysis of Station S_4

CONCLUSION

Here, we examine a queuing network with four nodes that experiences feedback and retrials. For each of the four nodes, we determine the length of the queue and the wait time. S_1, S_2, S_3 and S_4 with and without blocking (steady state). The numerical example shows that the queue length and waiting time for node S_1 and node S_4 are the same with and without blocking, but they differ for node S_2 and node S_3.It shows the correctness of the study.

REFERENCES

[1] Bose S.K., "An introduction to queueing Systems", Kluwer academic / Plenum publishers, New York, 2002.

[2] Erlang A.K., "The Theory of Probabilities and Telephone Conversations", Nyt Jindsskriff Math, B 20, 33-39, (1909).
[3] Falin G.I, "A survey of retrial queues", Queueing system 7, pp.127-167 (1990).
[4] Finch, P. D. (1959). "Cyclic queues with feedback".Journal of Royal Statistical Society, Series-B, Vol.21, pp.153-157.
[5] Gross.G and Harris.C, "Fundamentals of queueing theory", John Wiley and Sons, 1998.
[6] Kosten L.,"On the influence of repeated calls in the theory of probabilities of blocking", De Ingenieur 59, pp.1-25 (1947).
[7] Shanmugasundaram .S and Vanitha.S ,"Single server Markovian queueing network with feedback", Stochastic Modeling and Applications ,Vol. 24 No.1, 15-23,(2020)
[8] Sreekala M.S. and Manoharan M., "Analysis of Restricted Queueing Networks – A Blocking Approach", Journal of Statistical Science and Application, Oct 2016, Vol. 4, No. 09-10, 220-230
[9] Sreekala M.S. and Manoharan M.,"A queueing network model with feedback and its application in healthcare", International Journal of new technologies in science & Engineering Vol.3 Issue 12, pp.78-89 (2016).
[10] Takahashi.Y, Miyahara.H and Hasegawa.T, "An approximation method for open restricted queueing networks", Operation Research 28 (1980), 594 – 602.
[11] Tao Li, Liyuan Zhang and Shan Gao "Performance of an M/M/1 Retrial Queue with Working Vacation Interruption and Classical Retrial Policy", Advances in Operations Research, Volume 2016 |Article ID 4538031 | https://doi.org/10.1155/2016/4538031.
[12] Weiss E.N. and McClain J.O., "Administrative days in acute care facilities: A queuing-analytic approach", Operations Research 35 (1) (1987) , 35 – 44 .