Performance analysis of unreliable single server queueing system With glue period

S.R. Anantha Lakshmi¹* S.Shymala ² and K.Suresh³

Abstract
This paper deals with single server queueing system with the new feature of so-called glue periods or reservation periods. When the glue periods are exponentially distributed, we obtain an explicit expression for the probabilities $P(m,n)$ and the equations for the average number of customers in the system. An unreliable system with processing customers and breakdown and two parallel systems with a common infinite buffer is considered.

Keywords
unreliable system, glue period, breakdowns.

AMS Subject Classification
60K25, 60K30, 90B22.

1 Department of Mathematics, Easwari Engineering College, Chennai-600089, India.
2 Department of Mathematics, Government Arts College, Tiruvannamalai-606603, India.
3 Department of Mathematics, SKP Engineering College, Tiruvannamalai-606611, India.

*Corresponding author: ¹sr.ananthalakshmi@gmail.com; ²subramaniyanshyamala@gmail.com; ³suresh4maths@gmail.com

Table of Contents

1 Introduction .................................................. 257
2 Model Description and Analysis .......................... 258
3 Unreliable System ........................................ 259
4 A Fast and a Slow System ................................. 260
References .................................................. 261

1. Introduction

Queueing systems have wide applications in the fields of performance analysis of a wide range of systems in advanced computer networks, telecommunication systems and traffic management on high-speed communication networks. It is essential to have analytically tractable models and techniques for queueing analysis when the simulations are not feasible either due to time or storage complexity. A significant impact on computer network performance is expected to be made by the network complexity, an increasing number of customers, customer behavior and the retrial phenomenon.

In real life circumstances, the server failures which lead to service interruptions are quite common. Server failures to a large extent heavily influence the performance measures of unreliable queueing systems. The areas have been studied over decades to analyze the server failures due to Server failures. Recently, Choudhury and Ke [1], Choudhury and Tadj [2], Dimitriou [3], Dimitriou and Langaris [4], Falin [6], Ke [9], Ke et al.[10], Lee et al.[14], Yang et al.[19], and others considered the unreliable queueing frameworks with different highlights wherein one of the fundamental presumptions is that a failed server is sent for repair at the repair shop and present clients in the framework should trust that the server will be fixed without being served. Be that as it may, in down to earth circumstances, the framework ought to be furnished with a substitute (backup) server in anticipation of conceivable fundamental server disappointments. The substitute server renders administrations to clients while the primary server is fixed. The administration rate of the substitute server is not the same as (presumably lower than) that of the fundamental server. At the moment of the fix fruition, the primary server comes back to the framework and ends up accessible. This is the idea of the working breakdowns previously presented by Kalidass and Ramanath [13]. By enabling the substitute server to give benefits, the framework has the ability to deal with crises which may happen amid the fix time frame and the framework usage is improved. Moreover, the working breakdown administration can diminish grumblings from the clients who should trust that the fundamental server will be, fixed and decreases the expense of holding up clients or jobs. Along these lines, the working breakdown benefit is a progressively sensible fix approach for unreliable queueing frameworks.

The queueing frameworks in which debacles happen
have been broadly connected to demonstrate the total server disappointments. Consider a storage area network (SAN), which is one of the mass stockpiling arrangements with the goal that high unwavering quality and rapid system administrations are given. A SAN is made out of at least one separates associated with the capacity gadgets through exchanging gadgets, for example, centers, switches, and scaffolds. Since the association control between the capacity gadgets and customers is the real job of SAN servers, they are the practical objectives of the disseminated denial-of-service attacks (DDoS attacks). DDoS assaults make all system assets and information inaccessible to its expected customers, and it might be difficult to recuperate all the pulverized information. For another precedent, in email contact focuses, a clearing wonder some of the time happens and makes the framework be vacant. Krishna Kumar and Arivudainambi [11] and Krishna Kumar et al.[12] have talked about an M/M/1 queuing framework with disasters. They have gotten unequivocal articulations for the transient and steady state probabilities of the framework estimate and related execution measures. Parthasarathy and Sudesh [15] have considered an M/M/C queuing framework with N-approach.

This paper is given to the examination of a problematic single-server queuing frameworks with multiple client types Towsley and Tripathi [17] contemplated the M/M/1 queue with disasters (DST) so as to depict the conduct of distributed database systems with site failure. This investigation has been reached out to the M/G/1/DST queue by Jain and Sigman [8] and to the GI/M/1/DST queue by Yang and Chae [18]. Yechiali [20] talked about the M/M/1/DST queue with client impatience where the stationary probabilities of the framework state are inferred. In progression, Economou and Kapodistria [5] researched the M/M/1/DST queue in which clients end up restless and perform synchronized deserting of their administration because of the absence of the server. The queuing frameworks with disasters are portrayed by the marvel in which the event of catastrophes decimates every single unfinished job as well as separates the machine (processor). In that sense, disasters are alluded to as stochastic clearing [16]. The queuing frameworks in which disasters happen have been widely applied to show the entire server.

An essential element in the present paper is that of purported glue or reservation periods. The reservation period is enacted just before the landing of the server at a station. Amid a glue period, clients landing at the station (either fresh introductions or retrying clients) stick in the line of that station and will be served amid the accompanying administration time of that station, while amid some other period, arriving clients at the station join the circle of that station and will retry after an exponentially distributed time. One inspiration for contemplating reservation periods is the accompanying. A complex innovation that one may endeavor to add to the utilization of fiber delay loops in optical systems administration is fluctuating the speed of light by changing the refractive file of the fiber loop. Utilizing a higher refractive file in a little piece of the circle one can accomplish 'moderate light', which suggests backing off the bundles. This component is, in our model, fused as glue periods, where we back off the bundles touching base toward the finish of the fiber loop just before the server arrives, with the goal that they don’t need to retry yet get served amid the ensuing administration time frame. Not limiting ourselves to optical systems, one can likewise decipher a glue period as a reservation period, i.e., a period in which clients can reserve a spot at a station for administration in the consequent administration time of that station. In our model, the reservation time frame promptly goes before the administration time frame and could be viewed as the last piece of a switchover period.

Whatever remains of this paper is sorted out as pursues. In segment II we clarify the model with reservation period. In area III we think about an unreliable framework with preparing clients and breakdowns. We think about two parallel clients in segment IV. Numerical outcomes are given to delineate the impact of system parameters on the execution measures.

### 2. Model Description and Analysis

Consider an $M/M/1$ queuing system with unlimited capacity waiting room for customers to wait. The service is provided by a single server who serves the customer on a first come first serve basis. The service times are independently and identically distributed exponential R.V's with rate $\mu$. Let the arrival of customers follow a Poisson process with rate $\lambda$. Once the server completes the service when there are no customers the system is turned off as soon as the system is empty. When a new customer arrives the system is turned on again. For stability we require $\rho = \frac{\lambda}{\mu} < 1$.

If $m = 1$, the system is in up state.

- $m = 0$, the system is in its glue period.

In each state the rate of in flow is equal to the rate of outflow.

The balance equations are:

$$p(0,0)\lambda = p(1,1)\mu. \quad (2.1)$$

$$p(n,0)(\lambda + \theta) = p(n-1,0)\lambda, \quad n = 1,2,... \quad (2.2)$$

$$p(1,1)(\lambda + \mu) = p(1,0)\theta + p(2,1)\mu. \quad (2.3)$$

$$p(n,1)(\lambda + \mu) = p(n,0)\theta + p(n+1,1)\mu + p(n-1,1)\lambda, \quad n = 2,3,... \quad (2.4)$$

These balance equations can be solved as follows: Equation (2.2) is a first order difference equation which has a solution of the form,

$$p(n,0) = p(0,0)z^n, \quad n = 0,1,2,... \quad (2.5)$$
Performance analysis of unreliable single server queueing system With glue period — 259/262

with

\[ z = \frac{\lambda}{\lambda + \theta}. \]

The fraction of time the system is not processing is given by

\[ \sum_{n=0}^{\infty} p(n,0) = 1 - \rho \]

hence with (2.5)

\[ p(0,0) = (1 - \rho)(1 - z) \]

To find the probabilities \( p(n,1) \) we substitute (2.5) into (2.4) yielding

\[ p(n,1)(\lambda + \mu) = (1 - \rho)(1 - z)z^n\theta + p(n+1,1)\mu + p(n-1,1)\lambda. \tag{2.6} \]

This is a second order inhomogeneous recurrence relation. Its general solution is given by

\[ p(n,1) = A(1 - \rho)(1 - z)z^n + B_1x_1^n + B_2x_2^n, \quad n = 1, 2, \ldots \tag{2.7} \]

The first term is a particular solution to the inhomogeneous relation, where

\[ A = \theta z \frac{1}{z - \frac{\mu}{\lambda + \theta}} = \frac{\lambda + \theta}{\mu - \lambda - \theta}. \]

For convenience we assume \( \theta \neq \mu - \lambda \), so that the denominator does not vanish. The second term is the general solution of the homogeneous equation in which the coefficients \( B_1 \) and \( B_2 \) are free and \( x_1 \) and \( x_2 \) are the roots of

\[ (\lambda + \mu)x = \mu x^2 + \lambda. \]

Hence \( x_1 = \frac{\lambda}{\mu} = \rho \) and \( x_2 = 1 \).

Since the probabilities \( p(n,m) \) have the add up to one, we set \( B_2 = 0 \). Substitution of (2.5) and (2.7) into equation (2.3) yields \( B_1 = -A(1 - \rho)(1 - z) \). Since the equilibrium equations are dependent the remaining equation in (0,0) is also satisfied. Now we have found

\[ p(n,0) = (1 - \rho)(1 - z)^n, \quad n = 0, 1, 2, \ldots \]

\[ p(n,1) = A(1 - \rho)(1 - z)(z^n - \rho^n), \quad n = 0, 1, 2, \ldots \]

From the equilibrium distribution we can derive the mean number of customers in the system yielding,

\[ E(L) = \sum_{n=0}^{\infty} n[p(n,0) + p(n,1)] \]

\[ = (1 - \rho)(1 - z) \left[ \frac{z}{(1 - z)^2} + \frac{Az}{(1 - z)^2} - \frac{Ap}{(1 - \rho)^2} \right], \]

and by Little’s law we obtain the mean production lead time

\[ E(S). \]

We derive the distribution of the production lead time by conditioning on the state seen on arrival and using PASTA we get,

\[ P(s > t) = \sum_{n=0}^{\infty} p(n,0)P(Y + X_1 + \ldots + X_{n+1} > t) \]

\[ + \sum_{n=0}^{\infty} p(n,1)P(X_1 + \ldots + X_{n+1} > t), \]

where \( Y \) is an exponential (residual) glue period with mean \( \frac{1}{\theta} \) and \( X_1, X_2, \ldots, X \) are exponential processing times with mean \( \frac{1}{\mu} \). Substituting the expressions for \( p(n,0) \) and \( p(n,1) \) and using that a geometric sum of exponential random variables is again exponential, we get

\[ P(s > t) = \frac{1}{\mu - \lambda - \theta} \left\{ \mu (1 - \rho) e^{-\mu t} - \theta e^{\mu(1-\rho)t} \right\} \tag{2.8} \]

The mean and distribution of the production lead time decides whether to keep the system continuously in operation or to turn off when there are no customers.

The means \( E(L) \) and \( E(S) \) can be determined by combining the PASTA property and Little’s law. Based on PASTA we know that the average number of customers in the system seen an arriving job equals \( E(L) \) and each of them has a (residual) a processing time with mean \( \frac{1}{\mu} \). With probability \( 1 - \rho \) the system is not in on arrival so that the customer also has to wait for the glue period with mean \( \frac{1}{\theta} \). Further the customer has to wait for its own processing time. Hence

\[ E(S) = (1 - \rho) \frac{1}{\theta} + E(L) \frac{1}{\mu} + \frac{1}{\mu} \]

and together with \( E(L) = \lambda E(S) \) we find

\[ E(S) = \frac{1}{1 - \rho} + \frac{1}{\theta}. \]

So the mean glue period is the extra mean delay caused by turning off the system when there are no customers.

### 3. Unreliable System

In this section we consider an unreliable system processing customers. The system breaks down at random instants whether it is processing or not. To obtain some insight in the effects of the break downs we formulate the following simple model.

At the system customers arrive according to a Poisson stream with rate \( \lambda \). The processing times are exponential with mean \( \frac{1}{\mu} \). The time between two break downs is exponentially distributed with mean \( \frac{1}{\eta} \). The repair time is also exponentially distributed with mean \( \frac{1}{\mu} \).

The state of the system can be characterized by the pair \((n,m)\) where \( n \) is the number of customers in the system and \( m \) denotes the state of the system. If \( m = 1 \), the system is in
We first try to find solutions of (3.3) and (3.5) of the form

\[ p(n,0) = c_1 z^n \]

\[ p(n,1) = c_2 z^n \]

where \( v_i \) is given by (3.6) with \( z = z_i, i = 1, 2 \) and the coefficients \( c_1 \) and \( c_2 \) are yet to be determined. These coefficients follow from the boundary conditions (3.2) and (3.4) together with the normalization condition. Inserting (3.7) into (3.2) yields,

\[ \frac{c_1}{z_1} + \frac{c_2}{z_2} = 0. \]

when \( c_1 \) and \( c_2 \) satisfy this relation, it can be verified that equation (3.4) is also satisfied. From the condition that the probabilities add up to one, we obtain

\[ \frac{c_1 (1 + v_1)}{1 - z_1} + \frac{c_2 (1 + v_2)}{1 - z_2} = 1. \]

The two equations \( c_1 \) and \( c_2 \) have as solution \( c_1 = c_1 z_1 \), \( c_2 = -c_2 z_2 \), where

\[ C = \left\{ \frac{z_1 (1 + v_1)}{1 - z_1} - \frac{z_2 (1 + v_2)}{1 - z_2} \right\}^{-1}. \]

This completes the determination of the probabilities \( p(n,m) \).

Based on the solution for \( p(n,m) \) it is easy to find the mean number of customers in the system,

\[ E(L) = C \left\{ \frac{z_1^2 (1 + v_1)}{(1 - z_1)^2} - \frac{z_2^2 (1 + v_2)}{(1 - z_2)^2} \right\}, \]

and \( E(S) \) follows by Little’s law. The result show that the variation of the availability of the system is essential to the behavior of the system.

### 4. A Fast and a Slow System

Consider two parallel systems with a common (infinite) buffer where customers arrive according to Poisson stream with rate \( \lambda \). The processing times are exponentially distributed with mean \( \frac{1}{\mu} \), on system 1 and \( \frac{1}{\mu'} \), on system 2 \((\mu_1 > \mu_2)\). Customers are processed in order of arrival. For stability we assume \( \rho = \frac{\lambda}{\mu_1 + \mu_2} < 1 \).

If an arriving customer finds an idle fast system he will join it. If he finds a busy fast system with \( n \) customers and an idle slow system, then he will go to the slow system if \( \frac{n+1}{\mu_1} > \frac{1}{\mu_2} \) and otherwise he will wait for the fast system. So the individual optimal policy is characterized by the integer part of \( \frac{\mu_1}{\mu_2} \), say \( N \). Customers wait for the fast system as long as the number of customers in the system is less or equal to \( N \).

The two systems have the same behavior of the system.

The state of the system can be described by the pair \( (n,m) \) where \( n \) is the number of customers in the system and \( m \) denotes the state of system 2 which is the slower one. If \( m = 1 \) system 2 is busy and if \( m = 0 \) system 2 is idle.

The possible states are the pairs \( (n,0) \) with \( n \) running from 0 to \( N \), and the pairs \( (n,1) \) with \( n = 1, 2, \ldots \).

The equilibrium probabilities satisfy the following set of equations.

\[ p(0,0) \lambda = p(1,0) \mu_1 + p(1,1) \mu_2. \]
The equations (4.2) form a second order recurrence relation in \(p(n,0)\) with the probabilities \(p(n+1,1)\) as inhomogeneous terms. The general solution is

\[
p(n,0) = D_1 p_1^n + D_2 - c_1 z_1^{n+1} - c_2 z_2^{n+1}, \quad n = 0, 1, 2, ..., N
\]

where \(p_1 = \frac{\lambda}{\mu_1}\). The first two terms are the general solution of the homogeneous relation, the other ones are the particular solution of the inhomogeneous relation. The constants \(D_1\) and \(D_2\) follow from the boundary condition (4.1) and (4.3). The first condition yields \(D_2 = 0\) and

\[
D_1 p_1^{N+1} = c_1 z_1^{N+2} + c_2 z_2^{N+2}
\]

hence (4.12) gives

\[
D_1 = A \left( \frac{\lambda - \mu_1 z_2}{{\lambda - \mu_1 z_1}} z_1^{N+2} - \frac{\lambda - \mu_1 z_1}{\lambda - \mu_1 z_1} z_2^{N+2} \right) \left( \frac{\rho}{\rho_1} \right)^{N+1}
\]

we have to solve \(A\) from the normalization condition. Substituting (4.8), (4.9) and (4.13) we get

\[
A \frac{p_1^{N+2}}{1-p} + D_1 \frac{1-p_1^{N+1}}{1-p_1} = 1.
\]

Insertion the expression for \(D_1\) it follows that

\[
A = \left[ \frac{p_1^{N+2}}{1-p} + \frac{1-p_1^{N+1}}{1-p_1} \right]^{-1} \left( \frac{\lambda - \mu_1 z_2}{\lambda - \mu_1 z_1} z_1^{N+2} - \frac{\lambda - \mu_1 z_1}{\lambda - \mu_1 z_1} z_2^{N+2} \right) \left( \frac{\rho}{\rho_1} \right)^{N+1}.
\]

The equilibrium probabilities are given by

\[
p(n,0) = D_1 p_1^n - c_1 x_1^{n+1} - c_2 x_2^{n+1}, \quad n = 0, 1, 2, ..., N.
\]

\[
p(n,1) = c_1 x_1^n + c_2 x_2^n, \quad n = 1, 2, ..., N.
\]

\[
p(n,1) = A \rho^n, \quad n = N + 1, N + 2, ...
\]

where the constants \(A, c_1, c_2, D_1\) satisfy (4.15), (4.12) and (4.14) respectively.

From these expressions we can derive the mean number of customers in the system and mean time spent in the system.

References


Performance analysis of unreliable single server queueing system With glue period — 262


