FINITE CAPACITY QUEUEING SYSTEM WITH VACATIONS AND SERVER BREAKDOWNS

R.P. Ghimire and Ritu Basnet*

Department of Mathematics, School of Science, Kathmandu University, Dhulikhel (Nepal)

r12ghimire@yahoo.com, rbasnet07@ku.edu.np

*Corresponding Author

(Received: October 08, 2009 – Accepted in Revised Form: October 20, 2011)


Abstract This paper deals with finite capacity single server queuing system with vacations. Vacation starts at rate $\nu$ if the system is empty. Also, the server takes another vacation if upon his arrival to the system, finds the system empty. Customers arrive in the system in Poisson fashion at rate $\lambda_s$ during vacation, faster rate $f\lambda_s$ during active service and slower rate $s\lambda_s \geq 0$ during the breakdown. Customers are served exponentially with the rate $\mu$. Server breakdowns at rate $b$ and it is immediately repaired exponentially with the rate $r$. We derive the explicit formulas for queue length distribution, average queue length, average number of customers in the system and average waiting time for a customer in queue, and in the system. Numerical illustrations have been cited to show that the proposed model is practically sound.

Keywords Multiple–vacation, Server breakdowns, Repair, Generating functions, Average queue length.

1. INTRODUCTION

Several researchers [1-4] studied the queuing systems with server vacations and obtained various measures of performance. Takagi [5] studied M/G/1/N queue with server vacations and obtained the distributions of the unfinished work, the virtual waiting time and the real waiting time, etc. Wang [6] proposed an N-policy M/M/1 queueing system with server breakdowns and obtained analytic closed-form solutions. Takine and Sengupta [7] obtained the queue-length distribution and waiting time distribution of a single-server queue under the provision of service interruptions. Boxma et al. [8] studied the length of a vacation and steady-state workload distribution, both for single and multiple vacations. Ke [9] proposed M/G/1 queueing system with server vacations, startup and breakdowns and obtained the system total expected cost function per unit time under optimal control mechanism. Zhang et al. [10] studied the M/M/1/N queue with balking, reneging and server vacations and derived the matrix form solution of the steady-state probabilities and formulated a cost model to determine the optimal service rate. Wang et al. [11] made a comparative analysis between the exact results and the maximum entropy results, illustrated through the maximum entropy results that the maximum entropy principle approach is accurate enough for practical purposes. Jain and Agrawal [12] studied the M/M/1

IJE Transactions A: Basics Vol. 24, No. 4, November 2011 - 387
queueing system with multiple types of server breakdown under N-policy and provided the numerical results to demonstrate the effects of various parameters on the system performance characteristics. Ke and Wang [13] analyzed the operating characteristics for the heterogeneous batch arrival queue with startup and breakdown and performed a sensitivity analysis among the optimal value of N, specific values of system parameters, and the cost elements of M/M/1 queue. Srivastava and Jain [14] analyzed an optimal N-policy model for single Markovian queue with breakdown repair and state dependent arrival rate and obtained steady state for various operational characteristic and optimal value of N under a linear cost structure. Hur and Paik [15] analyzed the effect of different arrival rates on the N-policy of M/G/1 with server setup and derived the distribution function of the steady state queue length, waiting time using Laplace Stieltjes transform. Ke and Pearn [16] studied the management policy of M/M/1 queueing service system with heterogeneous arrivals under the N policy, in which the server is characterized by breakdowns and vacations and derived the distribution of the system size and mean queue length. Gray et al. [17, 18] studied the vacation queueing model with service breakdown under the assumption of different arrival rates and obtained formulas for queue length distribution and the mean queue length for M/M/1.

In this paper we investigate finite capacity queueing system with multiple vacations and server breakdowns. The server completely stops serving customers during a vacation and start serving whenever number of customers N (≥1) in the system. Once service starts, there can be an interruption due to server breakdown, and it is sent to repair facility. As soon as the repair process completes, the server starts to serve the same interrupted customer.

2. MATHEMATICAL MODEL AND ANALYSIS

- (0, i) is the state in which there are i customers in the queue and the server is on vacation, 0 ≤ i ≤ N. Its probability is \( p(0,i) \).
- (1, i) is the state in which there are i customers in the system during active service, 1 ≤ i ≤ N. Its probability is \( p(1,i) \).
- (2, i) is the state in which there are i customers in the system during repair process, 1 ≤ i ≤ N. Its probability is \( p(2,i) \).

The generating function for the queue length distribution is

\[
F(z) = F_0(z) + F_1(z) + F_2(z)
\]

where the partial generating functions are:

\[
F_0(z) = \sum_{i=0}^{N} p(0,i)z^i, \quad F_1(z) = \sum_{i=1}^{N} p(1,i)z^i, \quad F_2(z) = \sum_{i=1}^{N} p(2,i)z^i
\]

The balance equations for the queue length distribution are:

\[\lambda_0 p(0,0) = \mu p(1,1)\]

(2)

\[(\lambda + \nu) p(0,i) = \lambda_0 p(0,i - 1) \quad 1 \leq i < N\]

(3)

\[\nu p(0,N) = \lambda_0 p(0,N - 1)\]

(4)

\[(\lambda_i + \mu + b)p(1,1) = \nu p(0,1) + \mu p(1,2) + rp(2,1)\]

(5)

\[(\lambda_i + \mu + b)p(1,i) = \lambda_i p(1,i - 1) + \nu p(0,i) + \mu p(1,i + 1) + rp(2,i) \quad 2 \leq i < N\]

(6)

\[(\mu + b)p(1,N) = \lambda_i p(1,N - 1) + \nu p(0,N) + rp(2,N)\]

(7)

\[\lambda_i p(2,1) = \nu p(2,1)\]

(8)

\[\lambda_i p(2,i) = \nu p(2,i) + \lambda_i p(2,i - 1) \quad 2 \leq i < N\]

(9)

\[rp(2,N) = \mu p(2,N) + \lambda_i p(2,N - 1)\]

(10)

Equation (2) gives

\[p(1,1) = \frac{\lambda_0}{\mu} p(0,0)\]

(11)

From Equation (4), we have:
\[ p(0,N) = \frac{\lambda_0}{v} p(0,N - 1) \]  

Substituting (11) into (8), we obtain:
\[ p(2,1) = \frac{b}{\lambda_i + r} p(1,1) = \frac{b \lambda_0}{\mu(\lambda_i + r)} p(0,0) \]  

(13)

From Equation (3), we get:
\[ p(0,i) = p_0^i p(0,0) \quad 0 \leq i < N \]  

(14)

where \( p_0 = \frac{\lambda_0}{\lambda_i + v} \)

From Equations (12) and (14), we get:
\[ F_i(z) = \sum_{s=0}^{i} p(0,s)z^s = \left\{ \frac{1 - (0, z)^s}{1 - p_0 z} + \frac{\lambda_0}{\lambda_i + v} p_0^{i-1} z^{s-1} \right\} p(0,0) \]  

(15)

Multiplying Equation (6) by \( z^i \) and summing for \( i = 2,3,\ldots,N - 1 \)
\[ \left\{ b + \frac{b}{\lambda_i + v} \cdot \lambda_i \cdot (z - 1) \right\} F_i(z) = \left( \frac{\lambda_i + \mu + b}{p_0} \right) p(1,1)z \]  

(16)

Substituting the (17) into (16), we get
\[ \frac{(z - 1)Q(z)}{z(\lambda_i + r - \lambda_i z)} F_i(z) = \frac{(z - 1)p(z)}{z(\lambda_i + v)} p(0,0) + v \left\{ \frac{p_0^{i-1}(z - z^2) - p_0^{i-1} \Phi(z) + p_0 z^2 + p_0 z - z}{(1 - p_0)(1 - p_0 z)} \right\} p(0,0) \]  

(22)

where
\[ \Phi(z) = z^N + z^{N-1} + \ldots + z^2 \]  

(23)

\[ F_i(z) = \frac{(\lambda_i + r - \lambda_i z)^i}{Q(z)(1 - p_0 z)} \left\{ (1 - p_0)z^2 - (z - 1) \{ z + p_0 \Phi(z) \} \right\} p(0,0) \]  

(24)

For \( \lambda_i > 0 \), discriminant \( \Delta \) of the quadratic expression (19) satisfies
\[ \Delta \geq \lambda_i^2 b^2 + \lambda_i^2 \mu^2 + \lambda_i^2 r^2 - 2 \lambda_i \lambda_r \mu \cdot 2 \lambda_i \lambda_r \mu \cdot r + 2 \lambda_i \lambda_r \mu = \lambda_i^2 b^2 + (\lambda_i + \lambda_r - \lambda_i) > 0 \]

So, the equation \( Q(z) = 0 \) has two distinct real roots \( z_1 \) and \( z_2 \).

In order for the steady-state queue length distribution to exist, both roots of the equation \( Q(z) = 0 \) must be greater than 1. Since in \( Q(z) \), the coefficient of \( z^2 \) is positive, the two roots of \( Q(z) = 0 \) will be greater than 1 if \( Q(0) > 0 \) and \( Q'(0) < 0 \). Since \( Q(0) = \mu r - \lambda_i b - \lambda_i r > 0 \), we have that:
\[ \mu r > \lambda_i b + \lambda_i r \quad \text{or} \quad \frac{\lambda_i b + \lambda_i r}{\mu} < 1 \]  

(25)

The above equation implies that \( \mu > \lambda_i r \), so if (25) holds, then
\[ Q'(0) = \lambda_i (\lambda_i - \mu) - \lambda_i b - \lambda_i r < 0 \]

In order for the queue length distribution to exist, the R.H.S. of Equation (18) must vanish when \( z = 1 \). Since \( p(1,1) \) and \( p(2,1) \) are given, now we find \( \mu p(1,2) \)
Thus, if we assume that (25) holds, then the roots $Z_1$ and $Z_2$ of $Q(z) = 0$ will be greater than 1. From (23), we have for $z = 1$

$$\Phi(l) = N - l$$

and

$$\Phi'(l) = \frac{N(N+1)}{2} - l$$

Substituting Equations (15), (17) and (24) into (1), we get:

$$F(z) = \frac{1 - (\rho z) N}{1 - (\rho z)} + \rho z \frac{Q(z)(1 - \rho z)}{\mu(1 - \rho z) + \rho z \Phi(z)}$$

From (28) and the normalizing condition $F(1) = 1$, we obtain:

$$p(0,0) = \frac{(\mu r - \lambda b - \lambda r)(1 - \rho_0)}{\mu r - \lambda b - \lambda r + (b + r) \lambda_0 (1 - \rho_0^N)}$$

Now, assuming $\lambda_s > 0$ and substituting $\alpha = 1/Z_1$ and $\beta = 1/Z_2$, Equation (29) becomes:

$$p(0,0) = \frac{\mu r + \lambda_0 (1 - \alpha)(1 - \beta)(1 - \rho_0)}{\mu r - \lambda b - \lambda r + (b + r) \lambda_0 (1 - \rho_0^N)}$$

Using this in Equation (28), we obtain:

$$F(z) = R(z) \frac{(1 - \alpha)(1 - \beta)(1 - \rho_0)}{(1 - \rho_0)(1 - \rho_0 z)}$$

where

$$R(z) = \frac{1 - (\rho z) N + \rho z \Phi'(z)}{1 - (\rho z) - \frac{1}{\mu} + \frac{\lambda_0}{1 - \rho_0} + \frac{(b + r) \lambda_0 z}{1 - \rho_0^N} (1 - \rho_0)}$$

This reduces to $R(1) = 1$. In the case when there is no customer admitted in the queue during a repair process, $\lambda_s = 0$. Then, Equation (19) takes the form:

$$Q(z) = \frac{\mu r (1 - \rho z)}{\mu r - \lambda b - \lambda r + \lambda_0 (1 - \rho_0^N)}$$

Substituting it into Equation (28), we get:

$$F(z) = \frac{1 - (\rho z) N + \rho z \Phi'(z)}{1 - (\rho z) - \frac{1}{\mu} + \frac{\lambda_0}{1 - \rho_0} + \frac{(b + r) \lambda_0 z}{1 - \rho_0^N} (1 - \rho_0)}$$

Then:

$$p(0,0) = \frac{\mu r (1 - \rho_0)(1 - \rho_0)}{\mu r + \lambda_0 (b + r) (1 - \rho_0)}$$

and

$$F(z) = R(z) \frac{(1 - \rho_0)(1 - \rho_0)}{(1 - \rho_0 z)(1 - \rho_0^N)}$$

where

$$R(z) = \frac{1 - (\rho z) N + \rho z \Phi'(z)}{1 - (\rho z) - \frac{1}{\mu} + \frac{\lambda_0}{1 - \rho_0} + \frac{(b + r) \lambda_0 z}{1 - \rho_0^N} (1 - \rho_0)}$$

Equations (31) and (36) are the queue length distribution for $\lambda_s > 0$ and $\lambda_s = 0$ respectively.

If $\lambda_s = 0$ Expression (25) becomes the necessary and sufficient condition for the queue length distribution to exist; gives the utilization factor for M/M/1 queue which is independent from breakdown and repair rates.

For $\lambda_s > 0$, the mean queue length $L_q$ can be obtained by computing $F'(1)$ from (31) and (32):

$$L_q = \frac{\alpha}{1 - \alpha} + \frac{\beta}{1 - \beta} + \frac{\lambda_0 (\lambda_s - \mu - b)}{\mu r - \lambda b - \lambda r + \lambda_0} + \lambda_0 \frac{(b + r) \lambda_0 (1 - \rho_0^N)}{(b + r) (1 - \rho_0)}$$
The average number of customer in the system $L_s$ can also be obtained as:

$$L_s = L_q + \frac{\lambda_0}{\mu} + \frac{\lambda_f}{\mu} + \frac{\lambda_s}{\mu}$$  \hspace{1cm} (39a)$$

Average waiting times per customer in the queue and the system are:

$$W_q = \frac{L_q}{\lambda_0} + \frac{L_q}{\lambda_f} + \frac{L_q}{\lambda_s}$$ \hspace{1cm} (39b)$$

Respectively, and

$$W_s = L_s + \frac{1}{\mu}$$ \hspace{1cm} (39c)$$

For $\lambda_s = 0$, the mean queue length is determined from (36) and (37) as:

$$L_q = \frac{\lambda_f (b+r) - \mu r}{\mu (1-\rho_r)} + \frac{\rho_0}{\mu} \left(1 - \frac{1}{\mu(1-\rho_r)} \right)$$ \hspace{1cm} (40)$$

And the average number of customers in the system $L_s$ is:

$$L_s = L_q + \frac{\lambda_0}{\mu} + \frac{\lambda_f}{\mu}$$ \hspace{1cm} (41a)$$

Average waiting times in the queue and in the system are:

$$W_q = \frac{L_q}{\lambda_0} + \frac{L_q}{\lambda_f}$$ \hspace{1cm} (41b)$$

Respectively, and

$$W_s = L_s + \frac{1}{\mu}$$ \hspace{1cm} (41c)$$

3. Special cases

If $N \rightarrow \infty$, that is when the system capacity is infinite, then our result coincide with the result obtained by Gray et al. [17]. The mean queue length given by (38) and (40) reduces to the following forms:

**Case 1:** When $\lambda_s > 0$,

$$L = \frac{\alpha}{1-\alpha} + \frac{\beta}{1-\beta} + \frac{\rho_0}{1-\rho_0} + \frac{\lambda_s (b+r - \mu) + \lambda_f (b+r - \lambda_s) - \lambda_s r}{\mu r - \lambda_s b - \lambda_f r + \lambda_s (b+r)}$$

**Case 2:** When $\lambda_s = 0$,

$$L = \frac{\rho_0}{1-\rho_0} + \frac{\lambda_0 (b+r) - \mu r}{\mu (1-\rho_r)}$$

4. NUMERICAL RESULTS AND INTERPRETATIONS

In this section, we provide the numerical results for various performance indices using Equations (38) and (40). For the computation purpose, we fix

For $\lambda_s > 0$

**TABLE 1.** $N=25$, $\lambda_0=1$, $\lambda_f=1$, $\rho=1$, $r=3$, $b=2$

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>7</th>
<th>7.5</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_f$</td>
<td>$L_q$</td>
<td>$L_q$</td>
<td>$L_q$</td>
</tr>
<tr>
<td>2</td>
<td>1.5000</td>
<td>1.4439</td>
<td>1.3988</td>
</tr>
<tr>
<td>3</td>
<td>1.7667</td>
<td>1.6509</td>
<td>1.5641</td>
</tr>
<tr>
<td>4</td>
<td>2.3452</td>
<td>2.0588</td>
<td>1.8667</td>
</tr>
<tr>
<td>5</td>
<td>4.0833</td>
<td>3.0693</td>
<td>2.5238</td>
</tr>
<tr>
<td>6</td>
<td>19.1667</td>
<td>7.2667</td>
<td>4.5000</td>
</tr>
</tbody>
</table>

**TABLE 3.** $N=25$, $\lambda_0=1$, $\lambda_f=2$, $\lambda_s=1$, $\rho=1$, $r=3$ $b=2$

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_f$</td>
<td>$L_q$</td>
<td>$L_q$</td>
<td>$L_q$</td>
</tr>
<tr>
<td>3</td>
<td>1.3571</td>
<td>1.4259</td>
<td>1.5000</td>
</tr>
<tr>
<td>4</td>
<td>1.6667</td>
<td>1.8167</td>
<td>1.9881</td>
</tr>
<tr>
<td>5</td>
<td>1.4667</td>
<td>1.5619</td>
<td>1.6667</td>
</tr>
</tbody>
</table>

**TABLE 4.** $N=25$, $\lambda_0=1$, $\lambda_f=2$, $\lambda_s=1$, $\rho=1$, $b=2$

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_f$</td>
<td>$L_q$</td>
<td>$L_q$</td>
<td>$L_q$</td>
</tr>
<tr>
<td>3</td>
<td>9.1667</td>
<td>6.8810</td>
<td>5.7500</td>
</tr>
<tr>
<td>4</td>
<td>2.8333</td>
<td>2.5810</td>
<td>2.4167</td>
</tr>
<tr>
<td>5</td>
<td>1.9881</td>
<td>1.8761</td>
<td>1.8000</td>
</tr>
</tbody>
</table>

**TABLE 5.** $N=25$, $\lambda_0=1$, $\lambda_f=2$, $\lambda_s=1$, $\rho=1$, $b=2$ $r=3$

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_f$</td>
<td>$L_q$</td>
<td>$L_q$</td>
<td>$L_q$</td>
</tr>
<tr>
<td>3</td>
<td>6.8810</td>
<td>5.7500</td>
<td>4.7500</td>
</tr>
<tr>
<td>4</td>
<td>2.5810</td>
<td>2.4167</td>
<td>2.2500</td>
</tr>
<tr>
<td>5</td>
<td>1.8761</td>
<td>1.8000</td>
<td>1.6800</td>
</tr>
</tbody>
</table>
For $\lambda_s = 0$

<table>
<thead>
<tr>
<th>$\lambda_f$</th>
<th>$L_q$</th>
<th>$L_q$</th>
<th>$L_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.3500</td>
<td>1.3171</td>
<td>1.2899</td>
</tr>
<tr>
<td>3</td>
<td>1.5147</td>
<td>1.4505</td>
<td>1.4000</td>
</tr>
<tr>
<td>4</td>
<td>1.8333</td>
<td>1.6912</td>
<td>1.5882</td>
</tr>
<tr>
<td>5</td>
<td>2.5909</td>
<td>2.2000</td>
<td>1.9524</td>
</tr>
<tr>
<td>6</td>
<td>5.37500</td>
<td>3.6316</td>
<td>2.8182</td>
</tr>
</tbody>
</table>

**Table 4.** N=25, $\lambda_0 = 1$, $\lambda_f = 2$, $\nu = 1$, $r = 3$, $b = 2$

<table>
<thead>
<tr>
<th>$\mu$</th>
<th>$L_q$</th>
<th>$L_q$</th>
<th>$L_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu = 7$</td>
<td>2.7143</td>
<td>2.8000</td>
<td>2.8750</td>
</tr>
<tr>
<td>$\mu = 7.5$</td>
<td>1.8000</td>
<td>1.8571</td>
<td>1.9091</td>
</tr>
<tr>
<td>$\mu = 8$</td>
<td>1.5128</td>
<td>1.5556</td>
<td>1.5952</td>
</tr>
<tr>
<td>$\mu = 9$</td>
<td>1.3750</td>
<td>1.4091</td>
<td>1.4412</td>
</tr>
<tr>
<td>$\mu = 10$</td>
<td>1.2947</td>
<td>1.3231</td>
<td>1.3500</td>
</tr>
</tbody>
</table>

**Table 5.** N=25, $\lambda_0 = 1$, $\lambda_f = 2$, $\nu = 1$, $r = 3$

<table>
<thead>
<tr>
<th>$b$</th>
<th>$L_q$</th>
<th>$L_q$</th>
<th>$L_q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b = 1$</td>
<td>2.8750</td>
<td>2.8333</td>
<td>2.8000</td>
</tr>
<tr>
<td>$b = 1.5$</td>
<td>1.9091</td>
<td>1.8800</td>
<td>1.8571</td>
</tr>
<tr>
<td>$b = 2$</td>
<td>1.5952</td>
<td>1.5729</td>
<td>1.5556</td>
</tr>
<tr>
<td>$b = 2.5$</td>
<td>1.4412</td>
<td>1.4231</td>
<td>1.4091</td>
</tr>
<tr>
<td>$b = 3$</td>
<td>1.3500</td>
<td>1.3348</td>
<td>1.3231</td>
</tr>
</tbody>
</table>

**Table 6.** N=25, $\lambda_0 = 1$, $\lambda_f = 2$, $\nu = 1$, $r = 2$

Different system parameters are as follows:

Figure 1 displays the correlation between mean queue length ($L_q$) and fast arrival rate ($\lambda_f$) by varying the service rates. We also observe that the mean queue length ($L_q$) increases with the fast arrival rate ($\lambda_f$) whereas it decreases by increasing the service rates ($\mu$). Figure 2 exhibits the mean queue length ($L_q$) by varying service rate ($\mu$) and breakdown rate ($b$). It is seen that $L_q$ decreases with the increase of service rate ($\mu$), and $L_q$ increases with breakdown rate ($b$). Figure 3 demonstrates the mean queue length ($L_q$) decreases with the increase of service rate ($\mu$) and repair rate ($r$). When $\lambda_s > 0$
Figures 4, 5, 6 give the comparison of mean queue length ($L_q$) when no arrival during breakdown (i.e. $\lambda_s = 0$) with Figures 1, 2, 3 the mean queue length ($L_q$) when there may be arrivals during breakdown (i.e. $\lambda_s > 0$). This comparison shows that mean queue length ($L_q$) in latter figures increases or decreases gradually than in former cases. When $\lambda_s = 0$,

![Figure 4. Fast arrival rate vs. average queue length](image)

![Figure 5. Service rate vs. average queue length](image)

![Figure 6. Service rate vs. average queue length](image)

5. DISCUSSION

By using partial generating function method we have obtained various performance measures for a single server, finite capacity queueing system with the provision of vacations and server breakdowns. The imposition of various arrival rates to the system may have the applicability in the real life queueing system in which arrival rates can be varied so as to reduce sufficient queue length. Proposed model may have potential application in the telecommunication system, computer communication systems, machining system, etc.

6. ACKNOWLEDGEMENT

Second author is thankful to both University Grant Commission (UGC) Nepal, for providing financial support and also to the School of Science, Kathmandu University (Nepal).

7. REFERENCES

10. Zhang, Y., Yue, D. and Yue, W., “Analysis of an M/M/1/N queue with balkng, reneging and server vacations”, *Proceedings of the Fifth International...*


