Journal of the Korean Statistical Society (1999), **28**: *4, pp 523–533*

Analysis of Unfinished Work and Queue Waiting Time for the M/G/1 Queue with D-policy

Yon-Il Park and Kyung-Chul Chae[1]

ABSTRACT

We consider the M/G/1 queueing model with D-policy. The server is turned off at the end of each busy period and is activated again only when the sum of the service times of all waiting customers exceeds a fixed value D. We obtain the distribution of unfinished work and show that the unfinished work decomposes into two random variables, one of which is the unfinished work of ordinary M/G/1 queue. We also derive the distribution of queue waiting time.

Keywords: Queues with D-policy; Dam Model; Storage Process; Unfinished Work; Queue Waiting Time.

1. INTRODUCTION

We consider the M/G/1 queue with D-policy. The server is turned off at the end of each busy period and is activated again only when the sum of the service times of all waiting customers exceeds a threshold of size D, a fixed non-negative real number. The rest are the same as the M/G/1 queue: Customers arrive according to the Poisson process with rate λ and *iid* service times with mean $E(S)$ are independent of the arrival process. We assume that $\rho = \lambda E(S) < 1$ to guarantee the stability of this queue and that customers are served exhaustively in the order of their arrival.

The M/G/1 queue with D-policy is originated by Balachandran (1973) and studied by Balachandran and Tijms (1975) and Boxma(1976). They are mostly interested in the optimal control of D based on the expected unfinished work. (See Tijms (1986), pp. 36 for numerical results of the optiaml D.) Recently, Lee and Ahn (1998) consider a dam model in which the input is formed by a compound Poisson process and output policy is P_λ^M. This model is equivalent to the M/G/1 queue with D-policy, if we set $M = 1$ and $\lambda = D$. They obtain the stationary

[1]Department of Industrial Engineering, KAIST, Taejon 305-701, Korea

distribution of the water level, which corresponds to the unfinished work in the queue. On the other hand, Li and Niu (1992) consider the GI/G/1 queue with D-policy and express the distribution of queue waiting time as transform-free style. However, it is not quite a complete expression since it includes an unknown distribution function of some random variable.

In this paper, we obtain the Laplace-Stieltjes transform(LST) $X^*(\theta)$ of the distribution function for the unfinished work in the M/G/1 queue with D-policy. The unfinished work $X(t)$ at time t is defined as the time to complete the service of all customers who are present in the system. Taking a different approach from Lee and Ahn's (1998), we successfully show that the unfinished work decomposes into two random variables, one of which is the unfinished work of the ordinary M/G/1 queue. In addition, we obtain the LST $W^*(\theta)$ of the distribution function for the queue waiting time. The queue waiting time is the time duration that a customer would wait before the commencement of his service, and is typically longer than the unfinished work if the customer arrives during the idle period.

This paper is arranged as follows. In section 2.1, we derive the system equations for the unfinished work. In section 2.2, we obtain the stationary distribution of initial workload of the busy period, which appears in the system equations as an unknown. In section 2.3, we obtain, by cycle analysis, some meaningful quantities to express the results. Putting all these together in section 3, we finally obtain the LST of the unfinished work and the LST of the queue waiting time.

2. PRELIMINARIES

2.1. System Equations for the unfinished work $X(t)$

Let us define the probabilities that $p_0(t) = \Pr\{X(t) = 0\}$, $f(x,t)dx = \Pr\{$ server is idling, $x < X(t) \leq x + dx \leq D\}$, and $g(x,t)dx = \Pr\{$ server is busy, $0 \leq x < X(t) \leq x + dx\}$, where $X(t)$ is the unfinished work at time t. Also, the density function of the service time S is denoted by $s(\cdot)$.

Considering mutually exclusive events which can occur during $(t, t+dt)$, we have the following Chapman-Kolmogorov equations. For $X(t+dt) = 0$,

$$p_0(t+dt) = p_0(t)(1 - \lambda dt) + g(0,t)dt. \tag{2.1}$$

For $0 < X(t+dt) \leq D$ wherein the idle period,

$$\begin{aligned} f(x,t+dt) &= \lambda dt\, p_0(t)s(x) + f(x,t)(1-\lambda dt) \\ &\quad + \lambda dt \int_0^x f(y,t)s(x-y)dy. \end{aligned} \tag{2.2}$$

For $X(t+dt) > 0$ wherein the busy period,

$$\begin{aligned} g(x-dt, t+dt) &= g(x,t)(1-\lambda dt) + \lambda dt \int_0^x g(y,t)s(x-y)dy \\ &\quad + I_{\{x>D\}} \lambda dt \int_0^D f(y,t)s(x-y)dy, \end{aligned} \tag{2.3}$$

where $I_{\{x>D\}}$ denotes an indacator variable such that $I_{\{x>D\}} = 1$ if $x > D$ and $I_{\{x>D\}} = 0$ if $x \leq D$.

Rearranging (2.1) through (2.3), dividing by dt, and letting $dt \to 0$, we have the integro-differential equations as follows: For $x = 0$,

$$\frac{d}{dt} p_0(t) = -\lambda p_0(t) + g(0,t). \tag{2.4}$$

For $0 < x \leq D$ wherein the idle period,

$$\frac{\partial}{\partial t} f(x,t) = \lambda p_0(t)s(x) - \lambda f(x,t) + \lambda \int_0^x f(y,t)s(x-y)dy. \tag{2.5}$$

For $x > 0$ wherein the busy period,

$$\begin{aligned} \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) g(x,t) &= -\lambda g(x,t) + \lambda \int_0^x g(y,t)s(x-y)dy \\ &\quad + I_{\{x>D\}} \lambda \int_0^D f(y,t)s(x-y)dy, \end{aligned} \tag{2.6}$$

Taking the limit $t \to \infty$ to (2.4) through (2.6), we obtain the steady-state integro-differential equations. We put $\frac{d}{dt} p_0(t) = \frac{\partial}{\partial t} f(x,t) = \frac{\partial}{\partial t} g(x,t) = 0$ and express the equations in terms of the following: $p_0 = \lim_{t\to\infty} p_0(t)$, $f(x) = \lim_{t\to\infty} f(x,t)$, and $g(x) = \lim_{t\to\infty} g(x,t)$. For $x = 0$,

$$p_0 = \lambda^{-1} g(0). \tag{2.7}$$

For $0 < x \leq D$ wherein the idle period,

$$f(x) = p_0 s(x) + \int_0^x f(y)s(x-y)dy. \tag{2.8}$$

For $x > 0$ wherein the busy period,

$$\begin{aligned} -\frac{d}{dx} g(x) &= -\lambda g(x) + \lambda \int_0^x g(y)s(x-y)dy \\ &\quad + I_{\{x>D\}} \lambda \int_0^D f(y)s(x-y)dy. \end{aligned} \tag{2.9}$$

The unknowns $g(0)$ in (2.7) and $\lambda \int_0^D f(y)s(x-y)dy$, $x > D$, in (2.9) will be found in section 2.3.

2.2. Initial Workload of the Busy Period

Let U denote the initial workload of the busy period, and let $u(\cdot)$ and $U^*(\theta)$ respectively denote the density function of U and its LST. Also, let $S^{(n)}$ denote the n-convolution of the service time distribution, and let $s^{(n)}(x) = \frac{d}{dx}S^{(n)}(x)$. The LST of $s(x)$ (or $s^{(1)}(x)$) is denoted by $S^*(\theta)$.

Theorem 2.1. *$u(x)$ and $U^*(\theta)$ are given by*

$$u(x) = s(x) + \int_0^D s(x-y)dm(y), \qquad x > D, \tag{2.10}$$

$$U^*(\theta) = (1 - S^*(\theta)) \int_D^\infty e^{-\theta x} dm(x), \tag{2.11}$$

where

$$m(y) = \sum_{n=1}^\infty S^{(n)}(y), \ \ \frac{d}{dy}m(y) = \sum_{n=1}^\infty s^{(n)}(y).$$

Proof: Let $N = Min\{n | \sum_{i=1}^n S_i > D\}$ denote the number of customers arriving during an idle period. The distribution and the expectation of N are obtained from renewal arguments as follows: For $n \geq 1$,

$$\begin{aligned} \Pr\{N = n\} &= \Pr\{\sum_{i=1}^{n-1} S_i \leq D, \sum_{i=1}^{n} S_i > D\} \\ &= S^{(n-1)}(D) - S^{(n)}(D), \end{aligned} \tag{2.12}$$

$$E(N) = \sum_{n=1}^\infty \Pr\{N \geq n\} = 1 + m(D), \tag{2.13}$$

where $S^0(D) = 1$ and $m(\cdot)$ is a renewal function. Therefore, U is represented by $\sum_{i=1}^N S_i$. Conditioning on N, we have that, for $x > D$,

$$\begin{aligned} u(x)dx &= \Pr\{x < U \leq x + dx\} \\ &= \Pr\{x < U \leq x + dx, N = 1\} + \sum_{n=2}^\infty \Pr\{x < U \leq x + dx, N = n\} \\ &= \Pr\{x < S_1 \leq x + dx\} + \sum_{n=2}^\infty \Pr\{\sum_{i=1}^{n-1} S_i \leq D, x < \sum_{i=1}^{n} S_i \leq x + dx\}. \end{aligned}$$

Let us consider the second term on the r.h.s. of above equation. By conditioning on $\sum_{i=1}^{n-1} S_i$, we have that

$$\sum_{n=2}^\infty \int_{y=0}^D \Pr\{x < \sum_{i=1}^{n} S_i \leq x + dx | \sum_{i=1}^{n-1} S_i = y\} d\Pr\{\sum_{i=1}^{n-1} S_i \leq y\}$$

$$= \int_{y=0}^{D} s(x-y)dxdm(y),$$

which completes (2.10).

Replacing $\int_{y=0}^{D}$ in (2.10) with $\int_{y=0}^{x} - \int_{y=D}^{x}$ yields

$$\Pr\{x < U \leq x + dx\} = dm(x) - \int_{D}^{x} s(x-y)dxdm(y).$$

Taking LST for the above equation, we have that

$$\begin{aligned} U^*(\theta) &= \int_{x=D}^{\infty} e^{-\theta x} \Pr\{x < U \leq x + dx\} \\ &= \int_{x=D}^{\infty} e^{-\theta x} dm(x) - \int_{y=D}^{\infty} e^{-\theta y} \int_{x=y}^{\infty} e^{-\theta(x-y)} s(x-y)dxdm(y) \\ &= (1 - S^*(\theta)) \int_{D}^{\infty} e^{-\theta x} dm(x), \end{aligned}$$

which completes (2.11). □

Remark 2.1. Since N is a stopping time, we have Wald's equation

$$E(U) = E(N)E(S), \tag{2.14}$$

where $E(N)$ is given in (2.13).

Remark 2.2. From (2.11) and (2.14), we have the following relation:

$$\frac{1 - U^*(\theta)}{\theta E(U)} = \frac{1 - S^*(\theta)}{\theta E(S)} \frac{1 + \int_0^D e^{-\theta x} dm(x)}{E(N)}. \tag{2.15}$$

(2.15) will be used in section 3 for demonstrating the decomposition property of the unfinished work.

2.3. Cycle Analysis

The length of an idle period can be expressed as the random sum of N inter-arrival times. Thus, its expected value is given by

$$E(I) = \lambda^{-1} E(N). \tag{2.16}$$

On the other hand, the length of a busy period is a delay cycle (see e.g. Takagi (1991)) generated by initial workload U. Thus, its expected value is given by

$$E(B) = E(U)/(1 - \rho). \tag{2.17}$$

Let $E(C)$ denote the expected regeneration cycle. Then, from (2.14), (2.16), and (2.17), we have

$$E(C) = E(I) + E(B) = E(N)/\lambda(1-\rho). \tag{2.18}$$

Remark 2.3. The probability that the server is busy is identified as $\rho = E(B)/E(C)$ by (2.14), (2.17), and (2.18).

Theorem 2.2. *$g(0)$ in (2.7) is given by*

$$g(0) = 1/E(C) = \lambda(1-\rho)/E(N). \tag{2.19}$$

Proof: By the renewal reward theorem, p_0 in (2.7) is given by $\lambda^{-1}/E(C)$, where λ^{-1} is the expected duration since the beginning of an idle period until the first customer arrives. $E(C)$ is as given in (2.18). □

Remark 2.4. Based on the sample path analysis (El-taha and Stidham (1999)), (2.19) is directly interpreted as follows: The joint event that the server is busy and that the unfinished work becomes zero occurs once every regeneration cycle.

Theorem 2.3. *$\lambda\int_0^D f(y)s(x-y)dy$ in (2.9) is given by*

$$\lambda \int_0^D f(y)s(x-y)dy = u(x)/E(C), \quad x > D. \tag{2.20}$$

Proof: $\lambda\int_0^D f(y)s(x-y)dy$, $x > D$, is the rate of transitions from an idle period to the busy period having initial workload x (see El-taha and Stidham (1999)). Thus, it is the product of $E(C)^{-1}$ and $u(x)$, where $E(C)^{-1}$ is the transition rate from an idle period to a busy period and $u(x)$, $x > D$, is the rate that the workload right after a transition to a busy period is x. □

Remark 2.5. An alternative proof for (2.20), based on the renewal reward theorem, is as follows: For $x > D$,

$$\begin{aligned}
&\int_0^D \lambda dt\, s(x-y)dx\, f(y)dy \\
&= \Pr\{\text{a transition from an idle period to a busy period occurs during } dt, \\
&\qquad x < U \le x + dx\} \\
&= \{dt/E(C)\}u(x)dx.
\end{aligned}$$

Corollary 2.1. *p_0 and $f(x)$ are given by*

$$p_0 = (1-\rho)/E(N) \tag{2.21}$$

$$f(x) = (1-\rho)\frac{d}{dx}m(x)\Big/E(N), \quad 0 < x \le D. \tag{2.22}$$

Proof: (2.21) is obtained directly from (2.7) and (2.19). Since (2.8) is a renewal-type equation, its solution is given by

$$f(x) = p_0 s(x) + \int_0^x p_0 s(y) dm(x-y).$$

Substituting (2.21) and solving based on the definition of the renewal function $m(x)$, yields (2.22). □

3. MAIN RESULTS

Theorem 3.1. *The LST of the distribution of unfinished work, $X^*(\theta)$, in the M/G/1 queue with D-policy is given by*

$$X^*(\theta) = \frac{\theta(1-\rho)}{\theta - \lambda + \lambda S^*(\theta)} \frac{1 + \int_0^D e^{-\theta x} dm(x)}{E(N)}. \tag{3.1}$$

Proof: We decompose $X^*(\theta)$ into $X^*_{idle}(\theta)$ and $X^*_{busy}(\theta)$:

$$X^*(\theta) = X^*_{idle}(\theta) + X^*_{busy}(\theta), \tag{3.2}$$

where $X^*(\theta) = E(e^{-\theta X})$, $X^*_{idle}(\theta) = E(e^{-\theta X} |\text{server is idling}) \cdot \Pr\{\text{server is idling}\}$, and $X^*_{busy}(\theta) = E(e^{-\theta X}|\text{ server is busy}) \cdot \Pr\{\text{server is busy}\}$.
$X^*_{idle}(\theta)$ is obtained by taking LST for (2.21) and (2.22).

$$\begin{aligned} X^*_{idle}(\theta) &= e^{-\theta 0} p_0 + \int_0^D e^{-\theta x} f(x) dx \\ &= (1-\rho)\left[1 + \int_0^D e^{-\theta x} dm(x)\right] \Big/ E(N). \end{aligned} \tag{3.3}$$

Substituting (2.20) into (2.9) and then taking LST for (2.9), we have that

$$-\theta X^*_{busy}(\theta) + g(0) = -\lambda X^*_{busy}(\theta) + \lambda X^*_{busy}(\theta) S^*(\theta) + U^*(\theta)/E(C).$$

Solving this equation for $X^*_{busy}(\theta)$ and substituting (2.14), (2.18) and (2.19), we obtain that

$$X^*_{busy}(\theta) = \rho \frac{\theta(1-\rho)}{\theta - \lambda + \lambda S^*(\theta)} \frac{1 - U^*(\theta)}{\theta E(U)}. \tag{3.4}$$

Substituting (3.3) and (3.4) (together with (2.15)) into (3.2), we obtain (3.1). □

Remark 3.1. It is noteworthy that the unfinished work LST given in (3.1) is decomposed into two LSTs; one is the unfinished work LST of the ordinary M/G/1 and the other is the unfinished work LST at an arbitrary time during the idle period.

Remark 3.2. (3.4) is the product of the probability that the server is busy and the queue waiting time LST for the customer who arrives during the delay cycle generated by the initial delay U (see Takagi (1991)).

Corollary 3.1. *The expected unfinished work, E(X), is given by*

$$E(X) = \frac{\lambda E(S^2)}{2(1-\rho)} + \frac{\int_0^D x dm(x)}{E(N)}.$$

Let A and $A^*(\theta)$ respectively denote the inter-arrival time and the LST of its distribution. Note that $A^*(\theta) = \lambda/(\theta + \lambda)$.

Theorem 3.2. *The LST of the distribution of queue waiting time, $W^*(\theta)$, in the M/G/1 queue with D-policy is given by*

$$\begin{aligned} W^*(\theta) &= \frac{\theta(1-\rho)}{\theta - \lambda + \lambda S^*(\theta)} \frac{1 + \int_0^D e^{-\theta x} dm(x)}{E(N)} + (1-\rho)(A^*(\theta) - 1) \\ & \Big(\sum_{i=1}^{\infty} A^*(\theta)^{i-1} \frac{S^{(i)}(D)}{E(N)} + \int_0^D e^{-\theta x} \sum_{i=1}^{\infty} A^*(\theta)^{i-1} \frac{S^{(i)}(D-x) dm(x)}{E(N)} \Big). \end{aligned} \tag{3.5}$$

Proof: As (3.2), we decompose $W^*(\theta)$ into $W^*_{idle}(\theta)$ and $W^*_{busy}(\theta)$:

$$W^*(\theta) = W^*_{idle}(\theta) + W^*_{busy}(\theta), \tag{3.6}$$

where $W^*(\theta) = E(e^{-\theta W})$, $W^*_{idle}(\theta) = E(e^{-\theta W} |\text{ idle}) \cdot \Pr\{\text{idle}\}$, and $W^*_{busy}(\theta) = E(e^{-\theta W} |\text{ busy }) \cdot \Pr\{\text{busy}\}$.
Due to PASTA (Wolff (1982)) property, $W^*_{busy}(\theta)$ is identical to $X^*_{busy}(\theta)$ given in (3.4). On the other hand, the procedure for $W^*_{idle}(\theta)$ is relatively long.
Let $E_{n,k}$ denote the joint event that $N = n$ and that an arriving customer finds the server idle and is the k^{th} among n customers, $1 \leq k \leq n$. Then

$$\Pr\{E_{n,k}\} = (1-\rho) \frac{n \Pr\{N = n\}}{E(N)} \frac{1}{n}, \tag{3.7}$$

where $1-\rho$ is due to PASTA property and the rest are based on the so-called inspection paradox.

Let $W_{n,k}$ denote the queue waiting time conditional on $E_{n,k}$. Then, in terms of LSTs, we have that

$$W^*_{idle}(\theta) = \sum_{n=1}^{\infty}\sum_{k=1}^{n} \Pr\{E_{n,k}\} W^*_{n,k}(\theta). \tag{3.8}$$

$W_{n,k}$ consists of two random variables; one is the total service time of $k-1$ customers who are already present in the system and the other is the sum of $n-k$ inter-arrival times yet to occur. That is,

$$W_{n,k} = \sum_{i=0}^{k-1} S_i|_{N=n} + \sum_{j=0}^{n-k} A_j, \qquad 1 \le k \le n, \tag{3.9}$$

where $S_0 = A_0 = 0$. Note that $\sum_{j=1}^{n-k} A_j$ is independent of both N and $\sum_{i=1}^{k-1} S_i|_{N=n}$. Now suppose our policy is not D-policy but $(D-x)$-policy, $x < D$. Then the number of arrivals during an idle period, denoted by N', has the following distribution:

$$\Pr\{N' = n-k+1\} = S^{(n-k)}(D-x) - S^{(n-k+1)}(D-x), \qquad 1 \le k \le n,$$

which is analogous to (2.12). Note that we have the following relation:

$$\begin{aligned} & \Pr\{x < \sum_{i=1}^{k-1} S_i \le x+dx,\ N=n\} \\ = \; & \Pr\{x < \sum_{i=1}^{k-1} S_i \le x+dx\} \Pr\{N' = n-k+1\} \\ = \; & s^{(k-1)}(x)dx \left[S^{(n-k)}(D-x) - S^{(n-k+1)}(D-x)\right]. \end{aligned} \tag{3.10}$$

Dividing (3.10) by $\Pr\{N=n\}dx$ yields the density function of $\sum_{i=1}^{k-1} S_i|_{N=n}$. Then taking LST on (3.9), we have that

$$\begin{aligned} & W^*_{n,k}(\theta) \\ = \; & \frac{A^*(\theta)^{n-k} \int_0^D e^{-\theta x} s^{(k-1)}(x) \left[S^{(n-k)}(D-x) - S^{(n-k+1)}(D-x)\right] dx}{\Pr\{N=n\}}. \end{aligned} \tag{3.11}$$

We finally obtain (3.5) by substituting (3.7) and (3.11) into (3.8), and then (3.8) and (3.4) into (3.6). □

Corollary 3.2. *The expected queue waiting time, E(W), is given by*

$$\begin{aligned} E(W) &= \frac{\lambda E(S^2)}{2(1-\rho)} + \frac{\int_0^D x dm(x)}{E(N)} + \frac{1-\rho}{\lambda E(N)} \int_0^D (1 + m(D-x)) dm(x) \\ &= E(X) + (1-\rho) \frac{1}{\lambda} \frac{E(N(N-1))}{2E(N)}. \end{aligned} \tag{3.12}$$

Remark 3.3. In (3.12), we notice the resemblance between the D-policy and the N-policy. The main difference, or perhaps the only difference, is that N is a stopping time in the D-policy whereas N is a constant in the N-policy.

ACKNOWLEDGEMENT

The authors are grateful to the referees for their comments which improved the presentation of this paper.

REFERENCES

Balachandran, K. R. (1973). "Control policies for a single server system," *Management Science*, **19**, 1013-1018.

Balachandran, K. R. and Tijms, H. (1975). "On the D-Policy for the M/G/1 queue," *Management Science*, **21**, 1073-1076.

Boxma, O. J. (1976). "Note on a control problem of Balachandran and Tijms," *Management Science*, **22** , 916-917.

El-taha, M. and Stidham, J. S. (1999). *Sample-Path Analysis of Queueing Systems*, Kluwer Academic, Massachusetts.

Lee, E. Y. and Ahn, S. K. (1998). "P_λ^M-policy for a Dam with Input Formed by a Compound Poisson Process," *Journal of Applied Probability*, **35**, 482-488.

Li, J. and Niu, S. C. (1992). "The Waiting-Time Distribution for the GI/G/1 Queue under the D-policy," *Probability in the Engineering and Informational Sciences*, **6**, 287-308.

Takagi, H. (1991). *Queueing Analysis: a foundation of performance evaluation*, North-Holland, Amsterdam.

Tijms, H. (1986). *Stochastic Modelling and Analysis: A Computational Approach*, John Wiley & Sons, New York.

Wolff, R. W. (1982). "Poisson Arrivals See Time Averages," *Operations Research*, **30**, 223-231.