M/G/1 Subject to an Initial Quorum of Customers

by

Martin Krakowski
Technical Report

Office of Naval Research
Contract No. N00014-86-K-0029

M/G/1 Subject to an Initial Quorum of Customers

by

Martin Krakowski

Report No. GMU/49146/105
November 1986

Department of Operations Research and Applied Statistics
School of Information Technology and Engineering
George Mason University
Fairfax, Virginia 22030

Copy No. 15

This document has been approved for public sale and release; its distribution is unlimited.
<table>
<thead>
<tr>
<th>1. REPORT NUMBER</th>
<th>2. GOVT ACCESSION NO.</th>
<th>3. RECIPIENT'S CATALOG NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE (and Subtitle)</th>
<th>5. TYPE OF REPORT &amp; PERIOD COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td>M/G/1 Subject to an Initial Quorum of Customers</td>
<td>Technical Report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. AUTHOR(s)</th>
<th>6. CONTRACT OR GRANT NUMBER(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin Krakowski</td>
<td>NOO014-86-K-0029</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. PERFORMING ORGANIZATION NAME AND ADDRESS</th>
<th>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Operations Research and Applied Statistics</td>
<td>Task K-B Project 4118150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. CONTROLLING OFFICE NAME AND ADDRESS</th>
<th>12. REPORT DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Office of Naval Research</td>
<td>October 1986</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. NUMBER OF PAGES</th>
<th>15. SECURITY CLASS. (of this report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Unclassified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. DISTRIBUTION STATEMENT (of this Report)</th>
<th>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unlimited</td>
<td>Unclassified</td>
</tr>
</tbody>
</table>

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

- computational probability
- computer networks
- local area networks
- computer performance evaluation
- queueing models
- server vacations

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
Abstract

The model M/G/1 is modified in the following manner. The server is idled when he runs out of customers and resumes serving when the (m+1)st customers arrives, where m \( \geq 0 \); m+1 is referred to as the "initial quorum" and m is the maximal queue size (and system size) while the server idles or vacations. The modified model is designated as M/G/1(m). When m=0 we have the regular M/G/1.

In Section 1 we derive the omni-equations for the backlog process \( B \) (equation (1.17)) and for the delay process \( w \) (equation (1.24)) by exploiting the simple relation between \( B \) and \( w \) in M/G/1(m), a relation which results in equation (1.12).

In Section 2 we derive several composition properties for the backlog and for the delay. In particular, equation (2.8) says that the backlog is distributed like the sum of the backlog in M/G/1 and of a random variable which depends on the service time; and equation (2.9) says that the delay is distributed like the sum of the delay in M/G/1 and of a random variable which depends on the service time and on the interarrival time.

Notation

\( A \) = interarrival interval
\( \lambda \) = arrival rate = \( 1/A \)
\( x \) = service duration
\( \mu = 1/x \)
\( \rho = \lambda/\mu \)
\( r \) = residual time ("residue") of \( x \); \( x \) and \( r \) are related through the omni-equation
\( E\phi(x) - \phi(0) = EXE\phi'(r); \) cf. Krakowski (Sept. 1984)
\( w \) = delay encountered by a true or virtual customer; in regular M/G/1 we have \( w=B \)
\( B \) = backlog (unfinished work) due to incumbent customers; in regular M/G/1 we have \( w=B \)
\( w_* \) = delay when server works
\( B_* \) = backlog when server works
\( m \) = maximal size of the queue when the server idles; m+1 is the "initial quorum"
\( N \) = size of the system
\( n \) = size of the queue
\( P_j = Pr(N=j) \)
\( p_j = \Pr(n=j) \)

\( Q_j = \Pr(n=j \text{ and server idles}) \), \( 0 \leq j \leq m \); \( Q_0 = P_0 \)

\( P_* = \Pr(\text{server works}) \)

\( X, Y, \text{ and } Z \) are defined by (1.5), (1.6) and (1.23)

\( \phi(\ ) \) is an arbitrary well-behaved function of its argument(s); cf. Krakowski (1984, 1985); polynomials and their limits are well-behaved if their expectation is finite;
so is the step function; \( E\phi(z) \), the expectation of \( \phi(z) \), is called the
omni-transform of \( z \).

The omni-convention calls for taking the expectation of each-side of an
omni-equation without explicitly showing the expectation operator; thus \( \phi(y) = \phi(z) \)
stands for \( E\phi(y) = E\phi(z) \)

A free copy of a random variable \( z \) is a random variable distributed like \( z \) but independent of \( z \) and of any other variable within the scope of the same expectation operator (the expectation may be implied by the omni-convention)

\[ j^*z \overset{d}{=} z_1 + \ldots + z_j \text{ where } j \text{ is a positive integer and the } z_1 \text{ are free copies of } z \text{ (the "generic" random variable)}; 0^*z = 0 \]
Section 1. The Model M/G/1 With Initial Quorum

The regular model M/G/1 is modified as follows. The server is idled (or goes off on vacation) each time he runs out of customers, and resumes serving at the instant of the arrival of the (m+1)-th customer, where m\(\geq 0\). Thus m+1 constitutes an *initial quorum* and m is the maximal queue size while the server idles; we can think of m as a "limbo." We will designate this model as M/G/1/(m). Thus, when m = 0 we have the regular M/G/1.

In a prior report (Krakowski, 1986) we dealt with the process "queue size" for M/G/1(m). (Note: In that prior report m stands for the current m+1. The change simplifies somewhat the typography of many formulas.) We have shown, in particular, that for each 0 \(< j \leq m\)

\[ Q_j = \frac{(1-\rho)}{(m+1)} \text{ and } P_* = \rho = \lambda/\mu \]  

(1.1)

where \( Q_j = \text{Pr}(n=j; \text{ server idles}) \) and \( P_* = \text{Pr}(\text{server works}) \);

and that

\[ \phi(n) = \frac{1-\rho}{m+1} [\phi(0) + \phi(m)] + \rho \phi(n+h) \]  

(1.2)

where h = number of arrivals during the residue of \( x \), the (generic) service time. To make this report readable without recourse to the report quoted above we rederive (1.1). (We will not need (1.2) in this report.)

Consider the transition diagram for M/G/1/(3):

```
server works: R states

N=0  N=1  N=2  N=3  N=4  N=4

server idles: Q states
```

The flow balance for the state "N=1; server idles" is \( \lambda Q_0 = \lambda Q_1 \) which says that the frequency of entering this state equals the frequency of exiting it. The flow balance for the state "N=2; server idles" is \( \lambda Q_1 = \lambda Q_2 \); and for the state "N=3; server idles" the flow balance is \( \lambda Q_2 = \lambda Q_3 \). Hence we have \( Q_0 = Q_1 = Q_2 = Q_3 \). Clearly, for arbitrary m we have \( Q_0 = Q_j \); 0 \(< j \leq m \). The value of \( P_* \) follows from the global balance of arrivals and departures: \( \lambda = \mu P_* \). Thus we have demonstrated (1.1).
In the model M/G/1/(m) the treatment of the delay, the theme of the current report, is more complex than the treatment of the queue-size. It is also more complex and more subtle than the analysis of the delay for the regular M/G/1. The instinctive approach in analyzing the delay, in M/G/1/(m) as well as in M/G/1, is to establish a balance equation for the virtual delay. This has been a simple task in M/G/1, mainly, in our opinion, because in this model the process "virtual delay" and the process "virtual backlog" are identical. In M/G/1/(m) when m \geq 1 the two processes differ. It turns out that the analysis of M/G/1/(m) is much simplified, and more insightful, if one analyzes the backlog process B along with the delay process w; the backlog B is the aggregate remaining service time for all incumbent customers. (Still, the direct "instinctive" approach of balancing the virtual delay, or rather, in the omni-fashion, an arbitrary function of this delay, is a viable approach.) And, of course, the process B may be of interest in its own right and not as a mere stepping stone to analyzing w.

Note: The actual delay experienced by a true customer has the same distribution as the virtual delay; cf. Wolff (1982). This follows from the fact that a true customer and a random observer see, stochastically speaking, the same picture when the arrival source is poissonian. Using the same symbol w for true and virtual delays should cause no ambiguity in the context of this report.

The Static Balance: A "random" observer will find, with probability \( Q_0 \), that \( B = 0 \) and \( N = 0 \); he will find, with probability \( Q_1 \), that \( N = 1 \) and the server idles and \( B = x \); and more generally he will find, with probability \( Q_j \) where \( 0 \leq j \leq m \), that \( N = j \) and the server idles \( B = j\cdot x \); and he will find with probability \( P_* \) that \( B = B_* \). Thus, since each \( Q_j = (1-p)/(m+1) \) and \( P_* = \rho \), we have

\[
\phi(B) = \frac{1-\rho}{m+1} \left[ \phi(0 \cdot x) \phi(m \cdot x) \right] + \rho \phi(B_*)
\]  

(1.3)

When, e.g., \( m = 3 \) equation (1.3) becomes

\[
\phi(B) = \frac{1-\rho}{4} \left[ \phi(0) + \phi(x) + \phi(2 \cdot x) + \phi(3 \cdot x) \right] + \rho \phi(B_*)
\]  

(1.3a)

To find the omni-equations for the delay w which correspond to (1.3) and (1.3a) we note that a virtual customer who finds, with probability \( Q_j \), \( j \) (true) customers in the limbo must await \( m - j \) additional customers to reach the quorum of \( m + 1 \) (this quorum includes the virtual customer) and then must wait until the \( j \) senior customers are served. The entire delay of our virtual customer is therefore \((m-j) \cdot A + j \cdot x\), where
0 ≤ j ≤ m; thus the delay of a customer who finds the server idle is composed of m intervals. (When m = 0 there is no delay when the server idles, this being the regular M/G/1.) Hence,

\[ \phi(w) = \frac{1-\rho}{m+1} [\phi(m^*A) + \phi((m-1)^*A+x) + \phi(m^*x)] + \rho \phi(w_*) \]  

(1.4)

When m = 3 equation (1.4) becomes

\[ \phi(w) = \frac{1-\rho}{4} [\phi(3^*A) + \phi(2^*A+x) + \phi(A+2^*x) + \phi(3^*x)] + \rho \phi(w_*) \]  

(1.4a)

where \( w_* \) is the delay conditioned upon the server working.

We simplify the typography of (1.4) and (1.4) by defining

\[ \phi(X) \overset{d}{=} \frac{1}{m+1} \sum_{j=0}^{m} \phi(j^*x) \]  

(1.5)

and

\[ \phi(Y) \overset{d}{=} \frac{1}{m+1} \sum_{j=0}^{m} \phi((m-j)^*A+j^*x) \]  

(1.6)

Then (1.3) and (1.4) become

\[ \phi(B) = (1-\rho) \phi(X) + \rho \phi(B_*) \]  

(1.7)

and

\[ \phi(w) = (1-\rho) \phi(Y) + \rho \phi(w_*) \]  

(1.8)

A key observation in our approach is that in the model M/G/1/(m) the virtual delay and the backlog observed at an instant when the server works are one and the same process; hence

\[ B_* = w_* \]  

(1.9)

from which follows the weaker statement that

\[ \phi(B_*) = \phi(w_*) \]  

(1.10)

which merely says that \( B_* \) and \( w_* \) are identically distributed. Hence, from (1.8 and (1.10) it follows that

\[ \phi(w) = (1-\rho) \phi(Y) + \rho \phi(B_*) \]  

(1.11)
It follows from (1.7) and (1.11) that

$$\phi(w) = \phi(B) + (1-\rho) [\phi(Y) - \phi(X)]$$  \hspace{1cm} (1.12)$$

Thus if we have an omni-equation for \(B\) we can transpose it with the aid of (1.12) into an omni-equation for \(w\). We now derive such an equation as balancing condition for an arbitrary function of \(B\).

The backlog \(B\) jumps to \(B + x\) with each arrival, these arrivals being of frequency \(\lambda\). (Departures of serviced customers do not affect the balance since in our model we assume that the value of a customer’s service time is revealed at the instant of his arrival.) On the other hand, while the server works, with probability \(P_*\), he keeps on reducing the backlog now denoted by \(B_*\) at the rate \(dB_* = dt\). Hence the omni-balance for \(B\) is

$$\lambda [\phi(B+x) - \phi(B)] = P_* \phi'(B_*)$$  \hspace{1cm} (1.13)$$

According to the shifted renewal equation (Krakowski, September 1984)

$$\phi(B+x) - \phi(B) = \bar{x} \phi'(B+r)$$  \hspace{1cm} (1.14)$$

where \(r\) is the residue of \(x\); the differentiation is with respect to the entire argument. From (1.13) and (1.14) we have

$$\lambda \bar{x} \phi'(B+r) = P_* \phi'(B_*)$$

which integrated typographically is

$$\rho \phi(B+r) = P_* \phi(B_*)$$  \hspace{1cm} (1.15)$$

Putting \(\phi() = 1\) we rederive \(P_* = \rho\) and (1.15) becomes

$$\phi(B_*) = \phi(B+r)$$  \hspace{1cm} (1.16)$$

From (1.17) and (1.16) we have

$$\phi(B) = (1-\rho) \phi(X) + \rho \phi(B+r)$$  \hspace{1cm} (1.17)$$

the sought for omni-equation for the backlog \(B\). From (1.17) we can get, recursively, successive moments of \(B\) and a convolution equation for the distribution of \(B\). And in turn, with the aid of (1.12), we can get the successive moments of \(w\) and a convolution equation for the distribution of \(w\). Of course, all needed moments and distributions of \(X\) and \(Y\) and \(Z\) have to be derived as a side exercise.
Example: Find $\bar{B}$ and $\bar{w}$

From (1.15) we have

$$\bar{X} = \frac{1}{m+1} [\bar{x} + m\bar{x}] = \frac{m(m+1)\bar{x}}{2(m+1)} = \frac{1}{2} m\bar{x}$$

(1.18)

and from (1.6) we have, similarly,

$$\bar{Y} = \frac{1}{2} m\bar{A} + \frac{1}{2} m\bar{x} = \frac{1}{2} m(\frac{1}{\lambda} + \frac{1}{\mu})$$

(1.19)

From (1.17) and (1.18) we have

$$\bar{B} = \frac{1}{2} m\bar{x} + \frac{\rho r}{1-\rho}$$

(1.20)

From (1.12), (1.18) and (1.19) we have

$$\bar{w} = \bar{B} + (1-\rho)(\bar{Y} - \bar{X}) = \frac{1}{2} m\bar{x} + \frac{\rho r}{1-\rho} + (1-\rho)m\bar{A}$$

$$= \frac{m}{2\lambda} + \frac{\rho r}{1-\rho}$$

(1.21)

Thus, $\bar{w}$ is composed of $m/\lambda$ and the expected waiting time for a regular M/G/1.

This is not incidental and we will take up the problem of composition properties in the model M/G/1(m) in Section 2.

To complete the current section we derive the omni-equation for $w$. From (1.12) and (1.17) we have

$$\phi(w) = (1-\rho)\{\phi(Y) - \rho[\phi(Y+r) - \phi(X+r)]\} + \rho \phi(w+r)$$

(1.22)

A comparison of equation (1.22) with the quite simpler (1.17)—even if we replace $\phi(X)$ by its explicit representation in (1.5)—will give the reader an idea about the more extensive need for algebraic maneuvering in a direct derivation of (1.22) from the omni-balance of the virtual delay.

The formal appearance of (1.21) is simplified by defining

$$\phi(Z) \triangleq \phi(Y) - \rho[\phi(Y+r) - \phi(X+r)]$$

(1.23)

From (1.22) and (1.23) we have

$$\phi(w) = (1-\rho)\phi(Z) + \rho \phi(w+r)$$

(1.24)

It is easy to verify that (1.24) leads to $\bar{w}$ as given in (1.21).
Section 2: Composition Theorems

When \( m = 0 \), i.e. when we consider the regular \( M/G/1/(0) \) alias \( M/G/1 \), we have according to (1.5), (1.6) and (1.23)

\[
\phi(X_0) = 0, \quad \phi(Y_0) = 0, \quad \phi(Z_0) = 0,
\]

(2.1)

We will use in this section, as we have just done in (2.1), a subscript serving as a reminder of the size of \( m \) in the model. This is not a necessity, just a convenience; most equations in this section relate a general \( M/G/1/(m) \) model to the regular \( M/G/1/(0) \) so that the reminder-subscripts will identify the models. Thus we will write \( X_m, Y_m, Z_m, B_m, \) and \( w_m \), in place of \( X, Y, Z, B \) and \( w \). Equations (1.17) and (1.22) for \( m = 0 \) are written now, taking into account (2.1),

\[
\phi(B_0) = (1-\rho)\phi(0) + \rho \phi(B_0+r)
\]

(2.2)

and

\[
\phi(w_0) = (1-\rho)\phi(0) + \rho \phi(w_0+r)
\]

(2.3)

Equations (2.2) and (2.3) imply, as can be shown, that

\[
\phi(B_0) = \phi(w_0)
\]

(2.4)

which we know from the structure of the regular \( M/G/1 \).

Let us now shift (2.2) by \( X_m \) thus obtaining

\[
\phi(B_0 + X_m) = (1-\rho)\phi(X_m) + \rho \phi(B_0+r+X_m)
\]

(2.5)

From (2.5) and (1.17) we have

\[
\phi(B_m) - \rho \phi(B_m+r) = \phi(B_0+X_m) - \rho \phi(B_0+r+X_m)
\]

(2.6)

Defining

\[
\Psi(B_m) \triangleq \phi(B_m) - \rho \phi(B_m+r)
\]

(2.7)

we write (2.6) as

\[
\Psi(B_m) = \Psi(B_0 + X_m)
\]
or, reverting to \( \phi(\cdot) \) in place of \( \Psi(\cdot) \), as

\[
\phi(B_m) = \phi(B_0 + X_m) \quad (2.8)
\]

Equation (2.8) states that the backlog \( B_m \) in M/G/1/(m) is distributed like the convolution of \( B_0 \), the backlog in M/G/1, and \( X_m \).

From (2.3) and (1.24) we derive in like manner

\[
\phi(w_m) = \phi(w_0 + Z_m) \quad (2.9)
\]

thus showing that the delay in M/G/1/(m) is distributed like the convolution of \( w_0 \), the delay in regular M/G/1, and \( Z_m \).

The composition equations (2.8) and (2.9) should be especially useful computationally when the distribution or moments of the delay \( w_0 \) are known and when \( m \) is treated as a parameter. Of course, the distribution, or moments, of \( Z_m \) have to be evaluated as a side exercise. Similar remarks hold for the distribution and moments of \( B \) should the backlog be a process of interest. The labor involved may be formidable but in some practical applications it might be justified economically.

Another composition property for the delay \( w_m \) is obtained as follows. Providing (1.12) with \( m \)-subscripts we have

\[
\phi(w_m) = (1-\rho)[\phi(Y_m) - \phi(X_m)] + \phi(B_m) \quad (2.10)
\]

Since in the regular M/G/1, i.e. in M/G/1/(0), \( B_0 = w_0 \) we can write (2.8) in the form

\[
\phi(B_m) = \phi(w_0 + X_m) \quad (2.11)
\]

From (2.10) and (2.11) we have

\[
\phi(w_m) = (1-\rho)[\phi(Y_m) - \phi(X_m)] + \phi(w_0 + X_m) \quad (2.12)
\]

Equation (2.12) appears simpler numerically than (2.9), especially for the computation of moments.
References


<table>
<thead>
<tr>
<th>Copy No.</th>
<th>Office/Person</th>
<th>Address</th>
<th>Attention</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Office of Naval Research</td>
<td>800 North Quincy Street, Arlington, VA 22217</td>
<td>Scientific Officer, Statistics and Probability, Mathematical Sciences Division</td>
</tr>
<tr>
<td>2</td>
<td>ONR Resident Representative</td>
<td>Joseph Henry Building, Room 623, 2100 Pennsylvania Avenue, N.W., Washington, DC 20037</td>
<td></td>
</tr>
<tr>
<td>3 - 8</td>
<td>Director, Naval Research Laboratory</td>
<td>Washington, DC 20375</td>
<td>Code 2627</td>
</tr>
<tr>
<td>9 - 20</td>
<td>Defense Technical Information Center</td>
<td>Building 5, Cameron Station, Alexandria, VA 22314</td>
<td></td>
</tr>
<tr>
<td>21 - 29</td>
<td>C. M. Harris</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>GMU Office of Research</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>