

APPROXIMATIONS IN FINITE CAPACITY

MULTI-SERVER QUEUES WITH POISSON ARRIVALS

Operations Research Center Research Report No. 77-34

Shirley A. Nozaki and Sheldon M. /Ross

Dec. 1977 / /// 1.111

U. S. Army Research Office - Research Triangle Park

5 N 17714-71-C 7290

Operations Research Center University of California, Berkeley

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

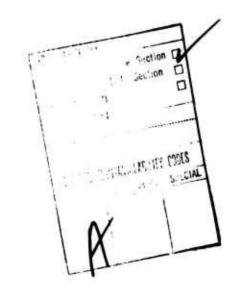
<sup>&</sup>lt;sup> $\dagger$ </sup>Partially supported by the Office of Naval Research under Contract N00014-77-C-0299 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, UNLESS SO DESIGNALED BY OTHER AUTHORIZED DOCUMENTS.

PARTNERS       1 GOVY ACCESSION NO.       1 ARECHIENT CATAGONUMENT         PRESENTED       I TYPE OF ARDON AND IN FINITE CAPACITY MULTI-SERVER       1 TYPE OF ARDON AT SETUD COVENED         PUPENSINGTONS IN FINITE CAPACITY MULTI-SERVER       I TYPE OF ARDON AT SUBJECT       Research Report         PUPENSINGTONS IN FINITE CAPACITY MULTI-SERVER       I TYPE OF ARDON AT SUBJECT       PERFORMING ORD. REPORT NUMBER         PUPENSINGTON SIN FAME AND ADDRESS       I TYPE OF ARDON ADDRESS       PERFORMING ORD. REPORT NUMBER         Proming GRAND FATCH AND ADDRESS       I AREONT ACCESSION FOR CONTRELS       I AREONT AND ADDRESS         Proming GRAND FATCH AND ADDRESS       P-13179-M       IIII ARDON TO ATTER         Proming GRAND FATCH AND ADDRESS       I AREONT CARE AND ADDRESS       P-13179-M         Proming GRAND FATCH AND ADDRESS       III AREONT CARE AND ADDRESS       III AREONT CARE AND ADDRESS         Proming GRAND FATCH AND ADDRESS       III AREONT CARE AND ADDRESS       III AREONT CARE AND ADDRESS         Proming GRAND FATCH AND ADDRESS       IIII AREONT CARE AND ADDRESS       IIIII AREONT CARE AND ADDRESS         Proming GRAND FATCH AND ADDRESS       IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
Tet and Andrew         PPRCXIMATIONS IN FINITE CAPACITY MULTI-SERVER         PPRCXIMATIONS IN FINITE CAPACITY MULTI-SERVER         PERFORMING ONG. REPORT & PERFORMENT ON COVERED         Research Report         Provide         hirley A. Nozaki and Sheldon M. Ross         ANDUEST OF Capacity HUBER()         DAAG29-76-6-0042         Performing Onderstrong Coverage         Performing Coverage         Performin	LPOAT NUMBER	
PPROXIMATIONS IN FINITE CAPACITY MULTI-SERVER QUELES WITH POISSON ARRIVALS	)RC 11- (4	
QUELES WITH POISSON ARRIVALS <ul> <li>PERFORMING ORG.REPORT NUMBER</li> <li>Intervention and Sheldon M. Ross</li> <li>CONTRACT OR GRANT NUMBER;</li> <li>DAAG29-76-G-0042</li> <li>DECENT, PARAGET, TARK</li> <li>P-13179-M</li> <li>P-13179-M</li> <li>December 1977</li> <li>NewERACT ADAGES</li> <li>Securption and PARAGENERS</li> <li>DAAG2010</li> <li>Securption and PARAGENERS</li> <li>DECEMPTICE NAME AND ADAGESES</li> <li>DECEMPTICE NAME AND ADA</li></ul>	TLE and Subirity)	S TYPE OF REPORT & PERIOD COVERED
************************************	PPROXIMATIONS IN FINITE CAPACITY MULTI-SERVER	Research Report
hirley A. Nozaki and Sheldon M. Ross ##domus Jadawization take Ab Address perations Research Center niversity of California erkeley, California 94720 Divroluws Office Name And Address . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Under Carling Office . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Under Contract . Not Adds . S. Army Research Under Contract . Novel for public release; distribution unlimited. 	QUEUES WITH POISSON ARRIVALS	6 PERFORMING ORG. REPORT NUMBER
hirley A. Nozaki and Sheldon M. Ross ##domus Jadawization take Ab Address perations Research Center niversity of California erkeley, California 94720 Divroluws Office Name And Address . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . December 1977 . Numer of Adds . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Under Carling Office . S. Army Research Office . December 1977 . Numer of Adds . S. Army Research Under Contract . Not Adds . S. Army Research Under Contract . Novel for public release; distribution unlimited. 	UTHOR(a)	A CONTRACT OR GRANT NUMBER(A)
AFTORING GARGANIZATION HARE AND ADDRESS Perations Research Center niversity of California erkeley, California 94720 NHTOOLUNG OFFICE HARE AND ADDRESS 0. S. Artwo Research Office 1. REPORT DATE December 1977 1. NUMBER OF PAGES 1. REPORT DATE 1. RE		
perations Research Center         niversity of California         erkeley, California 94720         N=Tocume office wate and abortss         .5. Army Research Office         .0. Box 12211         esearch Triangle Park, North Carolina 27709         17         N=Tocume office wate a abortsikif different free Centraling Office         18       MECONTY CLASS. (et this reperi)         Description Statewert (et this Reperi)         pproved for public release; distribution unlimited.         Statewert (et the exerted in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved by the Office of Naval Research under Contract N00014-77-C-0299.         verge Delay         pproximations         title capacity         verge Delay         pproximations         title capacity         statewer informations         title capacity         statewer informations         title capacity         statewer         title capacity <td>hirley A. Nozaki and Sheidon M. Ross</td> <td></td>	hirley A. Nozaki and Sheidon M. Ross	
perations Research Center         niversity of California         erkeley, California 94720         N=Tocume office wate and abortss         .5. Army Research Office         .0. Box 12211         esearch Triangle Park, North Carolina 27709         17         N=Tocume office wate a abortsikif different free Centraling Office         18       MECONTY CLASS. (et this reperi)         Description Statewert (et this Reperi)         pproved for public release; distribution unlimited.         Statewert (et the exerted in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved for public release; missed in Black 20, if different free Reperi)         pproved by the Office of Naval Research under Contract N00014-77-C-0299.         verge Delay         pproximations         title capacity         verge Delay         pproximations         title capacity         statewer informations         title capacity         statewer informations         title capacity         statewer         title capacity <td>ERFORMING ORGANIZATION NAME AND ADDRESS</td> <td>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</td>	ERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
erkeley, California 94720 DermoLumo Griter Mame and address S. Active Research Office J. Box 12211 esearch Triangle Park, North Carolina 27709 DertoRhing addacev Wame's Address Mild Millerent from Controlling Office DirioShing addacev Wame's Address Mild Millerent from Controlling Office Statewert (of the sector) pproved for public release; distribution unlimited. Statewert (of the address millered in Block 10, 11 different from Report) proceed by the Office of Naval Research under Contract N00014-77-C-0299. V GORD (Continue on reverse side if necessary and identify by Neck number) ueueing uiti Server inite Capacity verage Delay pproximations State I necessary and identify by Neck number) SEE ABSTRACT)		
DN TAOLLING OFFICE NAME AND ADDRESS       11 AEPORT DATE         D. Son 12211       December 1977         Search Triangle Park, North Carolina 27709       17         DNITONING ADENCY NAME & ADDRESS(II different from Controlling Office)       18 SECURITY CLASS. (at this report)         Unclassified       18-DECLASSIFICATION DOWNGRADING         STRIBUTION STATEMENT (at this Report)       Unclassified         pproved for public release; distribution unlimited.         STRIBUTION STATEMENT (at the essence in Stored in Block 30, if different from Report)         ueueing         NO0014-77-C-0299.         Y wondb (Continue on reverse side if necessary and identify by block number)         ueueing         uit Capacity         verage Delay         proximations         STRICT (Continue on reverse side if necessary and identify by block number)         SEE ABSTRACT)		2-131/9-31
<ul> <li>.0. Box 12211 <ul> <li>search Triangle Park, North Carolina 27709</li> <li>It substrates a contrasting of a acts 17</li> <li>It substrates 18</li> <li>It substrates</li></ul></li></ul>		12 REPORT DATE
Image Park, North Carolina 27709       17         Introduce Adduce Ad		
DNITORING ADENCY WAME & ADDRESSUIT different from Controlling Office) 18 SECURITY CLASS for the report Unclassified 18 SECURITY CLASS for the Development STRIBUTION STATEMENT (of the Report) pproved for public release; distribution unlimited. STRIBUTION STATEMENT (of the electric entered in Bloch 30, if different from Report) (POLIMENTARY HOTES Iso supported by the Office of Naval Research under Contract N00014-77-C-0299. V WORDS (Continue on records olds if necessary and identify by block number) until String Delay proximations STRICT (Continue on records olds if necessary and identify by block number) SEE ABSTRACT) SEE ABSTRACT)		
The DECLASSIFICATION DOWNGRADING         STRIBUTION STATEMENT (of the Report)         pproved for public release; distribution unlimited.         STRIBUTION STATEMENT (of the executer microd in Block 20, if different free Report)         PPLEMENTARY NOTES         Iso supported by the Office of Naval Research under Contract         N00014-77-C-0299.         Y WORDS (Continue on reverse side if necessary and identify by block number)         ueueing         uiti Server         inite Capacity         yerage Delay         pproximations         STRACT (Continue on reverse side if necessary and identify by block number)         SEE ABSTRACT)	DNITORING AGENCY NAME & ADDRESS(I different from Controlling Office)	18 SECURITY CLASS. (of this report)
STRIBUTION STATEMENT (of this Report) pproved for public release; distribution unlimited. STRIBUTION STATEMENT (of the observat microal in Block 20, if different from Report) PPLEMENTARY HOTES Iso supported by the Office of Naval Research under Contract N00014-77-C-0299. V UORDS (Continue on reverse side if necessary and femility by block number) ueueing ulti Server inite Capacity verage Delay pproximations STRICT (Continue on reverse side if necessary and identify by block number) SEE ABSTRACT)		Unclassified
STRIBUTION STATEMENT (a) this Report) pproved for public release; distribution unlimited. STRIBUTION STATEMENT (a) the abstract microd in Block 20, if different from Report) PPL EMENTARY HOTES Iso supported by the Office of Naval Research under Contract N00014-77-C-0299. (v WORDS (Continue on reverse size if necessary and identify by block number) ueueing Ulti Server inite Capacity verage Delay pproximations STRIFUCT (Continue on reverse size if necessary and identify by block number) SEE ABSTRACT) FORM 1473		15. DECLASSIFICATION DOWNGRADING
pproved for public release; distribution unlimited. STRIBUTION STATEMENT (of the electron original in Block 20, if different from Report) PPLEMENTARY NOTES Iso supported by the Office of Naval Research under Contract N00014-77-C-0299. Y MORDS (Continue on reverse side if necessary and identify by block number) ueueing ulti Server inite Capacity verage Delay pproximations STRACT (Continue on reverse side if necessary and identify by block number) SEE ABSTRACT)		
<pre>lso supported by the Office of Naval Research under Contract N00014-77-C-0299. vwoROS (Continue on reverse side if necessary and identify by block number) ueueing ulti Server inite Capacity verage Delay pproximations stract (Continue on reverse side if necessary and identify by block number) SEE ABSTRACT) SEE ABSTRACT)</pre>	pptoved for public release, distribution and	
<pre>iv woROS (Centinue en reverse side il necessery end identify by bleck number) ueueing ulti Server inite Capacity verage Delay pproximations si TACT (Centinue en reverse side il necessery end identify by bleck number) SEE ABSTRACT)</pre>		
<pre>ueueing ulti Server inite Capacity verage Delay pproximations 3) RACT (Continue on reverse side if necessary and identify by block number) SEE ABSTRACT)</pre>	DISTRIBUTION STATEMENT (of the observed on Block 20, if different to SUPPLEMENTARY NOTES Also supported by the Office of Naval Research	'een Ropert)
	DISTRIBUTION STATEMENT (of the observed on Block 20, if different to supplementary notes Also supported by the Office of Naval Research	'een Report)
	DISTRIBUTION STATEMENT (of the exertact entered in Block 20, if different is SUPPLEMENTARY NOTES Also supported by the Office of Naval Research NO0014-77-C-0299. TEV WORDS (Continue on reverse elde if necessory and identify by block number Queueing Multi Server Finite Capacity Average Delay Approximations	under Contract
	DISTRIBUTION STATEMENT (of the observed on Block 20, if different to SUPPLEMENTARY NOTES Also supported by the Office of Naval Research	under Contract
	TRIBUTION STATEMENT (of the observed in Block 20, if different in PPLEMENTARY NOTES so supported by the Office of Naval Research NO0014-77-C-0299. TRORDS (Continue on reverse side if necessary and identify by block number sever in inte Capacity rerage Delay proximations TRACT (Continue on reverse side if necessary and identify by block number SEE ABSTRACT) Data 1473 EDITION OF 1 NOV 68 IS DESOLETE	under Contract

# ABSTRACT

In this paper, we consider an M/G/k queueing model having finite capacity N. That is, a model in which customers, arriving in accordance with a Poisson process having rate  $\lambda$ , enter the system if there are less than N others present when they arrive, and are then serviced by one of k servers, each of whom has service distribution G. Upon entering, a customer will either immediately enter service if at least one server is free or else join the queue if all servers are busy. Our results will be independent of the order of service of those waiting in queue as long as it is supposed that a server will never remain idle if customers are waiting. To facilitate the analysis, however, we will suppose a service discipline of first come first to enter service."



N. Barr

### APPROXIMATIONS IN FINITE CAPACITY MULTI-SERVER QUEUES WITH POISSON ARRIVALS

by

Shirley A. Nozaki and Sheldon M. Ross

### 0. INTRODUCTION

In this paper, we consider an M/G/k queueing model having finite capacity N. That is, a model in which customers, arriving in accordance with a Poisson process having rate  $\lambda$ , enter the system if there are less than N others present when they arrive, and are then serviced by one of k servers, each of whom has service distribution G. Upon entering, a customer will either immediately enter service if at least one server is free or else join the queue if all servers are busy. Our results will be independent of the order of service of those waiting in queue as long as it is supposed that a server will never remain idle if customers are waiting. To facilitate the analysis, however, we will suppose a service discipline of "first come first to enter service."

Our objective is to obtain an approximation for the average time spent waiting in queue by an entering customer. This is done mainly by means of an approximation assumption, presented in Section 2, and used in Section 3 to derive the approximation. In Section 4, we let  $N = \infty$  and relate the approximation to the existing literature.

#### 1. BASIC DEFINITIONS AND FUNDAMENTAL EQUATION

We shall need the following notation:

P<sub>i</sub>: the steady state probability that there are i people in the system.

C. P. Tan Con The Control of the

2

- S : a service time random variable, i.e.,  $P\{S \le x\} = G(x)$ .
- $W_Q$ : the average amount of time that an entering customer spends waiting in queue (does not include service time).
- L<sub>Q</sub>: the (time) average number of customers waiting in queue.
  V: the (time) average amount of work in the system, where the work in the system at any time is defined to be the total (of all servers) amount of service time necessary to empty the system of all those presently either being served or waiting in queue.
- $V_{a}$  : the average amount of work as seen by an entering arrival.

We will make use of the following idea (previously exploited in such papers as [1], [2] and [8]) that if a (possibly fictitional) cost structure is imposed, so that entering customers are forced to pay money (according to some rule) to the system, then the following identity holds--namely,

time average rate at which the system earns = average arrival rate
(1)
of entering customers × average amount paid by an entering customer.

A heuristic proof of the above is that both sides of (1) times T is approximately equal to the total amount of money paid to the system by time T, and the result follows by dividing by T and then letting  $T + \infty$ .<sup>†</sup>

A rigorous proof along these lines can easily be established in the models we consider since all have regeneration points. More general conditions under which it is true are presented in [1].

Ev choosing appropriate cost rules, many useful formulae can be obtained as special cases of (1). For instance, by supposing that each customer pays \$1 per unit time while in service, Equation (1) yields that

average number in service = 
$$\lambda(1 - P_N)E[S]$$
.

Similarly, by supposing that each customer pays \$1 per unit time while waiting in queue, we obtain from (1) that

$$L_{Q} = \lambda (1 - P_{N})W_{Q} .$$

Also, if we suppose that each customer in the system pays x per unit time whenever its remaining service times is x, then (1) yields that

(2) 
$$V = \lambda (1 - P_N) E \left[ SW_Q^* + \int_0^S (S - x) dx \right] = \lambda (1 - P_N) \left[ E[S]W_Q + E[S^2]/2 \right]$$

where  $W_Q^{\dagger}$  is a random variable representing the (limiting) amount of time that the n<sup>th</sup> entering customer spends waiting in queue.

Another important fact which we shall use is that, since our arrival stream of customers is a Poisson process, the probability structure of what an arrival observes is identical to the steady state probability structure of the system.

### 2. THE APPROXIMATION ASSUMPTION

Let  $G_{e}$  denote the equilibrium distribution of G. That is,

$$G_{e}(x) = \int_{0}^{x} \frac{(1 - G(y))}{E[S]} dy$$
,

also let

$$\delta(\mathbf{x},\mathbf{y}) = \begin{cases} 1 & \text{if } \mathbf{x} = \mathbf{y} \\ 0 & \text{if } \mathbf{x} \neq \mathbf{y} \end{cases}.$$

We assume throughout that  $\int x dG_e(x) = E[S^2]/2E[S]$  is finite. We make the following approximation assumption.

## Approximation Assumption:

Given that a customer arrives to find i busy servers,  $i \ge 0$ , then at the time that he enters service, the remaining service times of the other  $i = \delta(i,k)$  customers being served has a joint distribution that is approximately that of independent random variables each having distribution  $G_p$ .

#### Heuristic Remarks Concerning the A.A.:

1. In the infinite capacity case, the A.A. appears to be approximately true either in heavy traffic (that is, as  $\lambda E[S] \neq k$ ) or in light traffic (that is, as  $\lambda E[S] \neq 0$ ). This is so in heavy traffic since the great majority of arrivals will encounter a large queue and as a result the k departure processes (one for each server) they observe will be approximately independent delayed renewal processes. Hence, considering those customers served by server 1, it follows that when they enter service they would have been observing k - 1 independent delayed renewal processes for a large time, and the A.A. follows since the limiting distribution of excess in a renewal process is just  $G_e$ .

.

In extremely light traffic, the great majority of arrivals will find either 0 or 1 busy servers. Now, since Poisson arrivals see the system as it is (averaged over all time), it follows that arrivals finding 1 server busy would encounter the same additional service time (for the busy server) as would random (and uniform) time sampling of the excess of a renewal process. Hence, the A.A. follows in light traffic from the renewal process (excess) result.

2. Additional heuristics for the A.A. follows from the fact that it is known to be (exactly) true when no queue is allowed (see [9]).

### 3. THE APPROXIMATION

Since V is the average amount of work as seen by an arrival, it follows by conditioning upon whether or not an arrival enters the system that

 $V = (1 - P_N)V_P + P_N \times$  (average work as seen by a lost customer).

In accordance with our basic A.A., it seems reasonable to approximate the average work as seen by a lost customer by  $k \frac{E[S^2]}{2E[S]} + (N - k)E[S]$ . Hence, we have that

(3) 
$$V = (1 - P_N)V_e + P_N\left(k \frac{E[S^2]}{2E[S]} + (N - k)E[S]\right)$$
.

Now for any arbitrary customer that enters the system we have the following identity

work as seen by the entering customer = k × time entering customer spends waiting in queue + R

where R is defined to be the sum of the remaining service times of those being served when the customer enters service. Taking expectations yields that

(4) 
$$V_{0} = kW_{0} + E[R]$$
.

To obtain E[R], we condition on  $B_e$ , the number of servers that are busy when the customer enters the system:

$$E[R] = E(E[R - B])$$

(5)

$$= E[B_e - B_e, k)] \frac{E[S^2]}{2E[S]}$$
 by the A.A.

Now,

(6)  $\frac{(1 - P_N)E[S]}{(1 - P_N)E[B_P] + kP_N}$ 

Also,

(7) 
$$E[\delta(B_{e},k)] = \left(1 - P_{N} - \frac{k-1}{j=0}P_{j}\right) / (1 - P_{N})$$

and so from (3)-(7) and Equation (2) we obtain that

(8) 
$$W_{Q} = \frac{\frac{E[S^{2}]}{2E[S]} \left(1 - P_{N} - \frac{k-1}{V}P_{j}\right) - (N-k)P_{N}E[S]}{(1 - P_{N})(k - VE[S])}$$

Therefore, it remains to obtain  $P_N$  and  $P_j$ ,  $0 \le j \le k - 1$ . To do so, we impose the following fictitional cost structure--namely, that the i oldest customers in the system pay \$1 per unit time, i = 1, 2, ..., k, where the age of a customer is measured from the moment it enters the system. Hence, letting  $S_1^e, S_2^e, ..., S_{k-1}^e$  denote k - 1 independent random variables each having distribution  $G_e$ , we obtain from Equation (1) that

$$P_{1} + 2P_{2} + \dots + (i - 1)P_{i-1} + i(1 - P_{0} - \dots - P_{i-1})$$

$$= \lambda(P_{0} + \dots + P_{i-1})E[S] + \lambda P_{i}E\left[\left(S - \min(S_{1}^{e}, S_{2}^{e}, \dots, S_{i}^{e})\right)^{+}\right]$$

$$+ \lambda P_{i+1}E\left[\left(S - 2nd \text{ smallest of } (S_{1}^{e}, \dots, S_{i+1}^{e})\right)^{+}\right]$$

$$+ \lambda P_{k-2}E\left[\left(S - (k - 1 - i)th \text{ smallest of } (S_{1}^{e}, \dots, S_{k-2}^{e})\right)^{+}\right]$$

$$+ \lambda(1 - P_{N} - P_{0} - \dots - P_{k-2})E\left[\left(S - (k - i)th \text{ smallest of } (S_{1}^{e}, \dots, S_{k-2}^{e})\right)^{+}\right]$$

$$i = 1, \dots, k - 1$$

$$P_{1} + 2P_{2} + \dots + (k - 1)P_{k-1} + k(1 - P_{0} - \dots - P_{k-1}) = \lambda(1 - P_{N})E[S]$$

 $\begin{pmatrix} \text{where } x^+ = \begin{cases} x & \text{if } x \ge 0 \\ 0 & \text{if } x < 0 \end{cases} \end{pmatrix}. \text{ To understand the above equations, suppose } \\ \text{first that } i < k . Now, as only the i oldest pay, it follows that when } \\ \text{j customers are present the system earns at a rate j when } j \le i \text{ and at } \\ \text{a rate } i & \text{when } j > i \text{ . Hence, the left side of Equation (9) represents } \\ \text{the average rate at which the system earns. On the other hand, an arrival } \\ \text{finding fewer than i customers already in the system will immediately go } \\ \text{into service and will pay a total amount equal to his service time; while } \\ \text{an arrival finding j present, } k - 1 \ge j \ge i \text{ will also go immediately } \\ \text{into service but will only begin paying when } j - i + 1 & \text{of the } j & \text{others } \\ \text{in service leave. Thus, in this latter case, under the A.A., the arrival } \\ \text{would expect to pay a total of } E \left[ \left( S - (j + 1 - i) \text{th smallest of } \\ \left( S_1^e, S_2^e, \ldots, S_j^e \right) \right)^+ \right]. \\ \text{Finally, if the arrival found more than } k - 2 & \text{busy} \end{cases}$ 

servers, then he will begin paying after k = 1 of those customers in service when he enters service leave the system. This explains the first k = 1 of the set of Equation (9). The last equation (when i = k) easily follows since in this case each customer will pay a total equal to his time in service.

To simplify the set of Equation (9), we will need the following lemma.

### Lemma 1:

If  $S_1S_1^e$ , ...,  $S_r^e$  are independent random variables such that S has distribution G and the others  $G_e$ , then

$$E\left[\left(S - jth \text{ smallest of } \left(S_1^e, \ldots, S_r^e\right)\right)^+\right] = \frac{r+1-j}{r+1} E[S] .$$

Proof:

Using the identity  $(x - y)^+ = x - \min(x, y)$ , we have that

$$E\left[\left(S - jth \text{ smallest of } S_1^e, \dots, S_r^e\right)^+\right] = E\left[S\right] - E\left[\min\left(S, jth \text{ smallest of } S_1^e, \dots, S_r^e\right)\right]$$

Now,

$$E\left[\min\left(S, jth \text{ smallest of } S_{1}^{e}, \dots, S_{r}^{e}\right)\right]$$

$$= \int_{0}^{\infty} P\{S \ge a\}P\left\{jth \text{ smallest of } \left(S_{1}^{e}, \dots, S_{r}^{e}\right) \ge a\right\}da$$

$$= \int_{0}^{\infty} (1 - G(a)) \frac{j-1}{\frac{V}{i=0}} {r \choose i} (G_{e}(a))^{i} (1 - G_{e}(a))^{r-i} da$$

$$= E[S] \frac{j-1}{\sum_{i=0}^{r} {r \choose i}} \int_{0}^{1} y^{i} (1-y)^{r-i} dy$$
$$= E[S] \frac{j-1}{\sum_{i=0}^{r} {r \choose i}} \frac{1!(r-i)!}{(r+1)!}$$
$$= E[S] \frac{j}{r+1}$$

which proves the lemma.

It follows from Lemma 1 that the equations for  $P_N$  and  $P_j$ ,  $0 \le j \le k - 1$  depend on G only through E[S]. Hence, as the equations are exactly true when G is exponential, it follows (since for fixed  $P_N$ , it can be shown that the set of equations has at most one solution) that the  $P_j$ ,  $0 \ge j \le k - 1$  have the same relationship to  $P_N$  as when G is exponential. Thus, we are left to determine  $P_N$ , which we will approximate by the answer in the exponential case. In other words, we shall use the exact result for  $P_j$ ,  $0 \ge j \ge k - 1$ ,  $P_N$  when G is exponential as our approximation. This yields, from Equation (8), that

(10) 
$$W_{Q} = \frac{\frac{E[S^{2}]}{2E[S]} \frac{N-1}{j=k} \frac{(VE[S])^{j}}{k!k^{j-k}} - (N-k) \frac{E[S](VE[S])^{N}}{k!k^{N-k}}}{\begin{bmatrix} k-1 \\ j=0 \end{bmatrix} \frac{(VE[S])^{j}}{j!} + \sum_{j=k}^{N-1} \frac{(VE[S])^{j}}{k!k^{j-k}} (k-VE[S])}{(k-VE[S])}$$

## 4. THE INFINITE CAPACITY CASE

In the infinite capacity case  $N = \infty$ , the approximation (10) reduces, when N = N = N = N = 0.

(11) 
$$W_{Q} \stackrel{:}{=} \frac{\sqrt{k_{E[S^{2}](E[S])}^{k-1}}}{2(k-1)!(k-\sqrt{E[S]})^{2} \left[ \frac{k-1}{j!} \frac{(\sqrt{E[S]})^{j}}{j!} + \frac{(\sqrt{E[S]})^{k}}{(k-1)!(k-\sqrt{E[S]})} \right]}$$

Some remarks are in order:

1. In [4], Kingman obtained bounds on  $W_Q$  for the general queueing system GI/G/k. When adapted to the M/G/k case of Poisson arrivals, his inequalities are

$$\frac{E[S^2]}{2E[S](k - \sqrt{E[S]})} = \frac{[E[S^2] + k/\sqrt{2} - (E[S])^2/k]}{2E[S]} = \frac{W_Q}{2E[S]} = \frac{\sqrt{[E[S^2] - (E[S])^2] + k/\sqrt{2}}}{2(k - \sqrt{E[S]})}.$$

It is easily verified that our approximation for  $W_Q$  is consistent with Kingman's upper and lower bounds.

2. In [5], Kingman conjectured a heavy traffic approximation for  $W_Q$  in GI/G/k models. In the special case of Poisson arrivals, his conjecture is that

$$W_Q \approx \frac{\sqrt{2}E[S^2] - \sqrt{2}(E[S])^2 + k^2}{2\sqrt{k(k - \sqrt{E[S]})}}$$
 when  $\sqrt{E[S]} \approx k$ .

Calling the right side of the above K and our approximation, as given by (8), N - R, we have

$$\frac{K}{N-R} = \frac{\lambda E[S]}{k\overline{P}_{k}} + \frac{E[S]}{\overline{P}_{k}E[S^{2}]} \left(\frac{k^{2} - (\lambda E[S])^{2}}{\lambda k}\right)$$

where  $\bar{P}_{k} = 1 - \sum_{0}^{k-1} P_{j}$ . Hence, since in heavy traffic,  $E[S] \approx k/\lambda$ ,  $\bar{P}_{k} \approx 1$ , we see that our approximation is consistent with Kingman's heavy traffic conjecture.

- 3. Numerical tables for  $L_Q$  have been published by Hillier and Lo in the special case  $M/E_r/k$ , where  $E_r$  represents an Erlang distribution with r phases. Table 1 compares our approximate formula for  $L_Q$  (=  $\lambda W_Q$ ) with the Hillier-Lo tables.
- 4. Another heavy traffic conjecture was given by Maaloe who in [6] conjectured that for the model  $M/E_r/k$

$$W_Q \approx \frac{\lambda E[S^2]}{2k(k - \lambda E[S])}$$
 when  $\lambda E[S] \approx k$ .

As the ratio between our approximation and the above approaches unity in heavy traffic, we see that our approximation is also consistent with this conjecture.

TABLE 1

L <sub>i</sub> for $M/E_r/k$ , $L_{ij} = \frac{(o_k)^{k+1}}{2^{(k+1)k(1-c)} 2^{\frac{k-1}{2}} (.k)^{n/n!} + \frac{(o_k)^{k}}{k!(1-c)}}$ , $-\frac{E[S]}{k}$

1		1			t	•	•	1	×	5	2
	<u>r=2</u> .000309	<u>r-1</u> 2002.74	.000257	<u>1=2</u> .000066	[-] .000059	<u>1-2</u> .000015	.00000		<u>[</u> ] e	21 -	[] =
	176000.	611000.	010000.	.000075	1/0000.	.000016	000000.	1000001	e	-	:
	.022509	.02008	.018758	606110"	.010586	627900.	E82E00.	\$00200.	261100.	.000648	.000372
	166120.	.022024	.0210483	.0129386	6066110.	.607142	100100	.002269	.001297	\$71000	1(-000-
	.177632	157895	.148026	20,0001.	.115942	811140.	126.210.	671250.	.044283	115760.	.0220.
	.184420	.16697	.158227	2130761.	124807	.103778	.079616	.0616708	.048126	97760.	961950.
	.861603	. 765869	.718002	. 750145	.666796	.661217	.587961	. 526293	\$\$\$674.	116/27	. 366030
	.877853	762187.	. 7420435	J68936	.6916764	.681288	.608581	.547007	.494070	10877.	07627
	5.515162	4.902366	4.595968	3.11735	4.726520	5.146829	878566.7	4.859712	876561.4	4.620612	76[ [5
-	2.54407	4.9405	4.638435	5.354183	4.775027	5.189889	5.04 3990	4.912132	4.79145	4.6799	57607
\$	12.924871	11.488775 10.770726	10.770726	12.702721	11.291307	12.508636	12.334640	12.175944	12.029403	11.892822	11.74-59
	12.95731	11.53154	10.81833	12.74479	11.34665	12.55858	12.39130	12.23848	12.09718	11.96535	11.6-1-7
35	72.851736	660121.43	60.709780	72.609459	272175.29	72. 395734	12.202451	12.024744	71.859-17	11.704239	71.55-5
	72.88707	64.80367	60.76161	72.65587	64.60278	72.45146	72.26631	72.09590	71.93724	71.7662	71.64728

bottom number in box - exact value as given by Hillfer-Lo

13

...

#### REFERENCES

- Brumelle, S. L., "On the Relation Between Customer and Time Averages in Queues," <u>Journal of Applied Probability</u>, Vol. 8, pp. 508-520, (1971).
- [2] Brumelle, S. L., "A Generalization of L = W to Moments of Queue Length and Waiting Times," <u>Operations Research</u>, Vol. 20, No. 6, pp. 1127-1136, (December 1972).
- [3] Hillier, F. S. and F. D. Lo, "Tables for Multiple-Server Queueing Systems Involving Erlang Distributions," Technical Report No. 31, Department of Operations Research, Stanford University, (December 1971).
- [4] Kingman, J. F. C., "Inequalities in the Theory of Queues," <u>Journal of</u> the Royal Statistical Society, Series B, Vol. 32, pp. 102-110,
- [5] Kingman, J. F. C., "The Heavy Traffic Approximation in the Theory of Queues," <u>Proceedings of the Symposium on Congestion Theory</u>, pp. 137-170, University of North Carolina, (1965).
- [6] Maaloe, E., "Approximation Formulae for Estimation of Waiting-Time in Multiple-Channel Queueing System," <u>Management Science</u>, Vol. 19, No. 6, pp. 703-710, (February 1973).
- [7] Newell, G., "Approximate Stochastic Behavior of n-Server Service Systems with Large n," Lecture Notes in Economics and Mathematical Systems, M. Beckmann, G. Goos and H. P. Künzi, eds., Springer-Verlag, (1970).
- [8] Stidham, S., "Static Decision Models for Queueing Systems with Non-Linear Waiting Costs," Technical Report No. 9, Stanford University, (1968).
- [9] Takács, L., "On Erlang's Formula," <u>Annals of Mathematical Statistics</u>, Vol. 40, No. 1, pp. 71-78, (1969).

the second s