SERIE RESEARCH MEMORANDA

PRACTICAL APPROXIMATIONS FOR FINITE-
BUFFER QUEUEING   MODELS   WITH BATCH-
ARRIVALS

R.D. Nobel

VRIJE UNIVERSITEIT
FACULTEIT DER ECONOMISCHE WETENSCHAPPEN
'EN ECONOMETRIE
AMSTERDAM

Practical approximations for finite-buffer queueing models with batch-arrivals.

By
R. D. Nobel,
Operations Research Group,
Dept. of Econometrics,
Vrije Universiteit,
Postbox 7161,
1007 MC Amsterdam, The Netherlands.

May 1987.

ABSTRACT

In this paper we discuss the finite capacity queueing model $G^X/G/1/N$ with batch-arrivals, a single server and having room for only N customers. For this model two different rejection strategies are conceivable: a batch finding upon arrival not enough space in the buffer is rejected completely or the buffer is filled up and only a part of the batch is rejected. For either strategy we are interested in the rejection-probabilities both for a batch and for an individual customer. Also we want to investigate the waiting-time distribution for an accepted customer. In general we cannot find analytical solutions for this model. However by specifying the service-time distribution to be an Erlang-k distribution, a Markov-chain approach is possible and exact results can be obtained. The next step is to get approximate results for the general case via interpolation with respect to the squared coefficient of variation of the service-time. We give approximations for the waiting-time percentiles and for the minimal bufferspace such that the rejection-probability is below a prespecified level. Also numerical results are given to illustrate the quality of the approximations.

0. INTRODUCTION

In modern telecommunication-technology and computer networks we are
confronted with the phenomenon that messages are sent in batches over a
communication-line and must be buffered at their destination-node before
they can be handled. The messages are handled one at a time in order of
arrival and within a batch in random order. Batches which upon arrival find
not enough space in the buffer are, at least partly, lost. So a natural
design problem arises with respect to the buffersize required to assure a
reasonable behaviour of the system, given the probability-distribution of
the batchsize, the interarrival-time of the batches and the service-time of
a message. The service level can be measured in terms of an upperbound for
the rejection-probability of a message or batch, or in terms of the
waiting-probability for an accepted message. This design problem motivates
the study of the $G^X/G/1/N$ model for which we present approximative results
using the exact solutions of the $G^X/E_k/1/N$ model, where E_k stands for an
Erlang-k distributed service-time. In this special case we can interpret
the service-time as the sum of k independently identically distributed
exponential phases. Thus the total number of uncompleted phases seen by an
arriving batch forms a sufficient state-description to enable an embedded
Markov-chain approach.

A simpler analysis is possible in case of Poisson arrivals ($M^X/E_k/1/N$)
because under this specific arrival process the total number of uncompleted
phases at an arbitrary epoch forms a continuous Markov-chain. Then we can
find a recursive solution for the steady-state probabilities. In section 1
we shall discuss the $M^X/E_k/1/N$ model and derive the formulae for the
rejection-probabilities for a batch and an individual customer. Also the
waiting-time distribution for an accepted customer will be given.

In section 2 we give the exact solution for the above mentioned $G^X/E_k/1/N$
model. In section 3 we discuss the general $G^X/G/1/N$ case and use the exact
results for the special case of the $G^X/E_k/1/N$ model to get approximative
results for both the waiting-time percentiles and the minimal buffersize
for which the rejection-probability does not exceed a given value. The
approximations involve interpolation with respect to the squared
coefficient of variation of the service-time.

In section 4 at last we give some comparisons between approximate values and exact values for models which also allow an analytical approach.

1. The $M^X/E_k/1/N$ queueing model

For the $M^X/E_k/1/N$ model we can give a recursive solution for the steady-state probabilities of the total number of uncompleted phases at an arbitrary epoch via a continuous Markov-chain analysis. We describe the model and this analysis in subsection 1.1. Once these probabilities are known we can derive the rejection-probabilities for a batch and an individual customer for each of the following two rejection/acceptance strategies. If an arriving batch does not fit completely in the remaining capacity of the buffer, then

a. under the whole batch acceptance strategy (WBAS) the whole batch is rejected.

b. under the partial batch acceptance strategy (PBAS) the remaining places in the buffer are filled up and only the customers for whom there is no place left are rejected.

Note that under the PBAS strategy every customer of a partially rejected batch has the same probability to be rejected because of the random order of customers within a batch. Under the WBAS strategy this probability equals one independent of the service order.

In the second subsection we discuss the waiting-time distribution of an accepted customer.

1.1. Description of the model and its Markov-chain analysis.

Batches of customers arrive at a service-facility according to a Poisson process with rate λ. The batch size X has a general discrete probability distribution: $P\{X = k\} = \alpha_k$ ($k = 1, 2, ..$). The service-time S for an individual customer has an Erlang-r distribution with parameter μ, so $E(S) = r/\mu$. The service-facility can handle one customer at a time and works at unity rate. The batch size distribution is independent of the arrival process and the service-times.

For customers who cannot be taken into service immediately there is a

buffer of N-1 waiting-places. As soon as the server becomes free a new customer is taken into service. Here we assume a first-come-first-served (FCFS) queue discipline for customers from different batches and a random selection for service (RSS) queue discipline for customers from the same batch. If no new customer is present the server becomes idle.

If an arriving batch contains too many customers to fit into the remaining places of the buffer we make a distinction between the two already mentioned rejection/acceptance strategies: WBAS and PBAS. So we have to deal with two different models, which we will discuss separately. Using that the Erlang-k distributed service-time can be interpreted as a sum of k identically, independently exponentially distributed phases, we can describe for both strategies the state of the system at an arbitrary epoch t as

$$X(t) = \text{the total number of uncompleted phases present}$$
$$\text{in the system at time } t.$$

We now have that $\{X(t), t>0\}$ is a continuous Markov-chain, where the steady-state probabilities can be calculated via the standard technique of equating the rate at which the process $\{X(t)\}$ leaves any state to the rate at which the process enters that state. The state-space of $\{X(t)\}$ is $\{0, 1, \ldots, Nr\}$.

Let f_i denote the time-average probability of having i uncompleted phases in the system at an arbitrary time i.e.

$$f_i = \lim_{t \to \infty} P\{X(t) = i\}$$

By the property 'Poisson arrivals see time-averages' (PASTA) we have that f_i can also be interpreted as the probability that a batch sees i uncompleted phases in the system upon its arrival.

Next we define

$$p_i = \lim_{t \to \infty} P\{\text{at epoch t there are i customers in the system}\}.$$

So $\{p_i\}$ is the time-average distribution for the number of customers in the system at an arbitrary time. Clearly, we have

$$p_j = \sum_{k=(j-1)r+1}^{jr} f_k$$

By the PASTA property p_j is also the probability that a batch finds upon arrival j other customers already present in the system.

Our first task is now to give the balance equations and the rejection-

probabilities for the two different strategies WBAS and PBAS.

A recursive equation for the state probabilities is obtained by applying the balance principle,

> the rate at which the process leaves a set A of states =

> the rate at which the process enters that set of states

to an appropriately chosen set A of states rather than to a single state.
If we take $A = \{j, \ldots, Nr\}$ ($j = 1, 2, ..$) we get in case of a WBAS strategy the following equations:

$$\mu * f_j = \lambda * \sum_{i=0}^{j-1} f_i * \sum_{k \geq (j-1)/r}^{[(N-i)/r]} \alpha_k$$

Here $[x]$ stands for the integer part of x. Under a PBAS strategy the balance equations show only a minor difference:

$$\mu * f_j = \lambda * \sum_{i=0}^{j-1} f_i * \sum_{k \geq (j-1)/r}^{\infty} \alpha_k$$

In either case we can solve the given system of equations recursively by starting with $f_0 = 1$. The steady-state probabilities f_j can subsequently be found by using the normalizing equation

$$\sum_{j=0}^{Nr} f_j = 1$$

We now consider the rejection-probabilities under the WBAS strategy.
First we look at the probability that a batch is rejected. This can be done easily by conditioning on the number of customers present upon arrival of the batch:

P(batch rejected) =

$$\sum_{k=0}^{N} P(\text{batch rejected} \mid \text{batch sees k customers upon arrival}) * p_k =$$

$$\sum_{k=0}^{N} P(\text{batch consists of more than N-k customers}) * p_k =$$

$$\sum_{k=0}^{N} p_k * \left(1 - \sum_{i=1}^{N-k} \alpha_i \right) \tag{1}$$

For the calculation of the probability that a customer is rejected we need the

following renewal-theoretic result [BURKE]:

P{an arbitrary customer belongs to a batch of size i} $= i\alpha_i/E(X)$. (2)

Now we can deduce

$$P\{\text{customer rejected}\} = \sum_{k=0}^{N} P\left\{ \begin{array}{l|l} \text{customer} & \text{customer sees k customers} \\ \text{rejected} & \text{already present upon arrival} \end{array} \right\} * p_k$$

$$= \sum_{k=0}^{N} P\{\text{customer belongs to a batch of size more than N-k}\} * p_k =$$

$$= \sum_{k=0}^{N} p_k * \left(1 - \sum_{i=1}^{N-k} i\alpha_i/E(X) \right). \tag{3}$$

Next we look at the rejection-probabilities under PBAS. Under this strategy we prefer to speak of 'batch overflow' when an arriving batch does not find space enough for all its customers, because under PBAS we have only partial batch-rejection. Apart from this change in name the formula for the probability of batch-overflow is exactly the same as formula (1) for the probability of batch-rejection under WBAS, where of course the probabilities p_k are different. Then

$$P\{\text{batch overflow}\} = \sum_{k=0}^{N} p_k * \left(1 - \sum_{i=1}^{N-k} \alpha_i \right)$$

To find the rejection-probability for an individual customer we define:

$\eta_j = P\{\text{an arbitrary customer takes the j-th position in his batch}\}$. (4)

From the formula (2) we get:

P{an arbitrary customer takes the j-th position in his batch} =

$$\sum_{i=j}^{\infty} \alpha_i/E(X), \quad j \geq 1$$

Now we can deduce

$$P\{\text{customer rejected}\} = \sum_{k=0}^{N} P\left\{ \begin{array}{l|l} \text{customer rejected} & \text{customer belongs to a batch} \\ & \text{which sees k customers} \\ & \text{already present upon arrival} \end{array} \right\} * p_k =$$

$$\sum_{k=0}^{N} P\{\text{customer takes a position greater than N-k in his batch}\} * p_k =$$

$$\sum_{k=0}^{N} p_k * \left(1 - \sum_{i=1}^{N-k} \eta_i \right). \tag{5}$$

1.2. The waiting-time distribution for an accepted customer

Let W_q be defined as the waiting-time of an accepted customer.
To find the probability distribution of W_q, we define the random variable

F = the total number of uncompleted phases in front of an arbitrarily
accepted customer just after his entrance into the system.

Using the fact that the waiting-time of a customer having j phases in front of
him has an Erlang-j distribution, we have for either strategy:

$$P\{W_q > x\} = \sum_{j=1}^{(N-1)r} P\{W_q > x \mid F=j\} * P\{F=j\} =$$

$$\sum_{j=1}^{(N-1)r} P\{F=j\} * \sum_{i=0}^{j-1} e^{-\mu x} (\mu x)^i / i! \qquad (6)$$

So it remains to calculate the distribution of F for either of the two
strategies. To compute the probability distribution of F, let us define the
following events:

A = the event that an arbitrary customer is accepted

A_j = the event that an arbitrary customer is accepted and has j
uncompleted phases in front of him just after his entrance
into the system.

Then, by the defintion of conditional probability,

$$P\{F = j\} = P\{A_j\} / P\{A\}. \qquad (7)$$

For the two strategies WBAS and PBAS the rejection probability 1-P{A} is
given by the respective formulae (3) and (5). To find $P\{A_j\}$, let B_{ki} denote
the joint event that an arbitrary customer belongs to a batch of size k and
that i uncompleted phases are in the system just prior to the arrival of his
batch. Using (2) we have,

$$P\{B_{ki}\} = (k\alpha_k / E(X)) * f_i. \qquad (8)$$

By the law of total probability, we have for any fixed j

$$P\{A_j\} = \sum_{k,i} P\{A_j \mid B_{ki}\} * P\{B_{ki}\}. \qquad (9)$$

Noting that each newly arriving customer represents exactly r phases, it
follows that for fixed j the probability $P\{A_j \mid B_{ki}\}$ cannot be positive
unless i and k satisfy $i \leq j$ and $j-i \in \{0, r, \ldots, kr\}$ and $k \leq u_i$, where

$$u_i = \begin{cases} [(Nr-i)/r] & \text{for the WBAS strategy} \\ \infty & \text{for the PBAS strategy} \end{cases} \qquad (10)$$

In case i and k satisfy the above characteristics we have $P\{A_j \mid B_{ki}\} = 1/k$,

otherwise $P\{A_j | B_{ki}\} = 0$. Together the formulae (8) - (10) yield for both strategies the result

$$P\{A_j\} = \sum_{i=0}^{j} \sum_{k=[(j-i)/r]+1}^{u_i} P\{A_j | B_{ki}\} * f_i * \alpha_k / E(X). \tag{11}$$

Substituting (11) in (7) gives the desired probability $P\{F = j\}$.

2. The $G^X/E_k/1/N$ queueing model

For the $G^X/E_k/1/N$ model we can also give a Markov-chain analysis if we look only at the arrival epochs. This so called embedded Markov-chain approach leads to a system of equilibrium equations for the steady-state probabilities that an arriving batch sees j uncompleted phases in the system ($j = 0, 1, ..$). Once these probabilities are known we can use the same formulae as deduced in section 1 for the rejection-probabilities and the waiting-time distribution. Thus we can confine ourselves to the description of the model and its Markov-chain analysis. This will be done in subsection 2.1, where we will also give the results for some specific interarrival-time distributions.

2.1. The description of the model and its Markov-chain approach

The only difference with the model discussed in section 1 lies in the arrival process: here we consider a generally distributed interarrival-time with density-function $g(t)$. For the other characteristics of the model we refer to the previous section.

Now define

X_n = the total number of uncompleted phases in the system just prior to the arrival of the n-th batch.

Then $\{X_n, n = 1, 2, ..\}$ is a discrete aperiodic Markov-chain with a finite state-space $\{0, 1, ..., Nr\}$. Let

$$z_i = \lim_{n \to \infty} P\{X_n = i\} \quad (i=0,1,..)$$

be the steady-state probability that an arriving batch sees i uncompleted phases in the system. Then we know that the z_i can be determined as the

unique solution of the lineair equations

$$z_j = \sum_{i=0}^{Nr} z_i * p_{ij}, \quad j = 0,1,.. \tag{12}$$

together with the normalization equation

$$\sum_{j=0}^{Nr} z_j = 1 .$$

Here the p_{ij} stand for the one-step transition probabilities of the Markov-chain $\{X_n, n = 1, 2, ..\}$. So it remains to calculate the $p_{ij} = P\{X_{n+1} = j \mid X_n = i\}$. It suffices to do this for $j \neq 0$, since

$$p_{i0} = 1 - \sum_{j=1}^{Nr} p_{ij} .$$

Let D_n be the size of the n-th batch and let T_n stand for the interarrival-time between the (n-1)-th and the n-th batch. Then by conditioning on the number of customers in the n-th batch and the interarrival-time we get the following formula,

$$p_{ij} = \sum_k \int_0^{\infty} P\{X_{n+1} = j \mid X_n = i, D_n = k, T_{n+1} = t\} * g(t)dt * \alpha_k . \tag{13}$$

Next we must calculate the conditional probability

$$P\{X_{n+1} = j \mid X_n = i, D_n = k, T_{n+1} = t\} .$$

To do this put for abbreviation

$$POIS(t, h) = e^{-\mu t} (\mu t)^h / h! .$$

Note that POIS(t, h) represents the probability of h service-phase completions during a time interval of length t when the server would be continuously busy during this time. Using this observation it is not difficult to see the following results:

(i) WBAS strategy. Then for $j \neq 0$,

$$P\{X_{n+1}=j \mid X_n = i, D_n = k, T_{n+1} = t\} = \begin{cases} POIS(t, i+kr-j) & \text{if} \\ \quad \max((j-i)/r, 1) \leq k \leq [(Nr-i)/r] \\ POIS(t, i-j) & \text{if} \\ \quad i \geq j \text{ and } k > [(Nr-i)/r] \\ 0 & \text{otherwise.} \end{cases}$$

(ii) PBAS strategy. Then for $j \neq 0$,

$$P(X_{n+1}=j \mid X_n= i, \ D_n= k, \ T_{n+1}= t) = \begin{cases} POIS(t, \ i+kr-j) & \text{if} \\ \quad \max((j-i)/r, \ 1) \leq k \leq [(Nr-i)/r] \\ POIS(t, \ i+[(Nr-i)/r]r-j) & \text{if} \\ \quad k > [(Nr-i)/r] \\ 0 & \text{otherwise.} \end{cases}$$

Now we have to plug in these formulae into (13) to get the transition probabilities p_{ij}. For completeness we give the results:

1) strategy is WBAS

1a) $j > i$

$$P_{ij}= \sum_{k \geq (j-i)/r}^{[(Nr-i)/r]} \int_0^\infty POIS(t,i+kr-j)*g(t)dt * \alpha_k. \tag{14a}$$

1b) $1 \leq j \leq i$

$$P_{ij}= \sum_{k=1}^{[(Nr-i)/r]} \int_0^\infty POIS(t,i+kr-j)*g(t)dt * \alpha_k +$$

$$\int_0^\infty POIS(t, \ i-j)*g(t)dt * (1 - \sum_{k=1}^{[(Nr-i)/r]} \alpha_k). \tag{14b}$$

2) strategy is PBAS

2a) $j > i$

$$P_{ij}= \sum_{k \geq (j-1)/r}^{[(Nr-i)/r]} \int_0^\infty POIS(t,i+kr-j)*g(t)dt * \alpha_k +$$

$$\int_0^\infty POIS(t, \ i+[(Nr-i)/r]r-j)*g(t)dt * (1 - \sum_{k=1}^{[(Nr-i)/r]} \alpha_k). \tag{15a}$$

2b) $1 \leq j \leq i$

$$P_{ij}= \sum_{k=1}^{[(Nr-i)/r]} \int_0^\infty POIS(t,i+kr-j)*g(t)dt * \alpha_k +$$

$$\int_0^\infty POIS(t, \ i+[(Nr-i)/r]r-j*g(t)dt * (1 - \sum_{k=1}^{[(Nr-i)/r]} \alpha_k). \tag{15b}$$

At last, for $j = 0$ we can write $P_{i0} = 1 - \sum_{j=1}^{Nr} P_{ij}$

Now that we have calculated the one-step transition probabilities p_{ij} we can solve the system (12) to get the desired steady-state probabilities z_j.

In their turn these probabilities should replace the probabilities f_j in the formulae discussed in section 1.

Next we show how the formulae for the p_{ij} can be easily evaluated for a constant interarrival-time and an Erlang-s distributed interarrival-time.

First let the interarrival-time be a constant D. Then the integrals

$$\int_0^\infty POIS(t, n)*g(t)dt$$

become simply $POIS(D, n)$. Substitution of this result in the formulae (14a) to (15b) gives the transition probabilities for the deterministic case. Next we consider an Erlang-s distribution with scale parameter λ for the interarrival-time, i.e. the density function $g(t)$ has the form

$$g(t) = e^{\lambda t} \lambda^s t^{s-1} / (s-1)!.$$

Now we can easily derive the following equality

$$\int_0^\infty POIS(t, n)*g(t)dt = \begin{bmatrix} n+s-1 \\ s-1 \end{bmatrix} * \mu^n * \lambda^s / (\mu + \lambda)^{n+s} \tag{16}$$

Substitution of (16) in the formulae (14a) to (15b) gives the p_{ij} when the interarrival-time has an Erlang-s distribution with scale parameter λ.
In case the interarrival-time density is a mixture of Erlangian densities the integrals evaluate to a mixture of formulae of the form as in (16). So also for this type of arrival process we have simple formulae for the p_{ij}.

3. The approximative approach for the $G^X/G/1/N$ queueing model

So far we have been succesful in analyzing the $M^X/E_k/1/N$ and the $G^X/E_k/1/N$ model. Things get worse if we allow the service-time to have a general distribution because then a Markov-chain analysis is not possible anymore. Nevertheless we can approximate performance measures such as the rejection probability and the waiting-time probabilities. This cannot be achieved directly, but should be done indirectly via percentiles. Much empirical evidence is given in [TIJMS] that the percentiles of queue size and waiting-time distribution can often be accurately approximated by interpolating the corresponding percentiles for simpler queueing systems.
It is well known that a polynomial interpolation of degree n with support points (x_i, f_i), $i = 0, 1, .., n$, is given by

$$F(x) = \sum_{i=0}^n f_i * \prod_{\substack{k=0 \\ k \neq i}}^n (x - x_k) / (x_i - x_k). \tag{17}$$

We are mostly interested in linear interpolation (n = 1):

$$F(x) = f_0 * (x-x_1) / (x_0-x_1) + f_1 * (x-x_0) / (x_1-x_0).$$

In our case we approximate the percentiles for the $G_X/G/1/N$ queue by interpolating the percentiles for two models $G_X/E_k/1/N$ and $G_X/E_m/1/N$, where the interpolation is with respect to the squared coefficient of variation c_S^2 of the service-time. Denote by ξ_p the p-th conditional waiting-time percentile for the $G^X/G/1/N$ model i.e. $P\{W_q \leq \xi_p \mid W_q > 0\} = p$, $0<p<1$. It is more convenient to consider the conditional waiting-time percentiles than the unconditional percentiles, since the former ones are defined for all $0<p<1$. Then ξ_p can be approximated by

$$\xi(k,p) * (c_S^2 - 1/m)/(1/k - 1/m) + \xi(m, p) * (c_S^2 - 1/k)/(1/m - 1/k). \quad (18)$$

Here $\xi(r,p)$ stands for the p-th conditional waiting-time percentile of the $G^X/E_r/1/N$ model. Usually we can take k=1 and m=2 to get reasonable results. An analogous approach is possible with respect to the minimal buffersize needed to assure that the rejection-probability for a customer (or a batch) does not exceed a prespecified value β. Again we can obtain the exact values for the models $G^X/E_k/1/N$ and $G^X/E_m/1/N$ and use the same type of interpolation as for the waiting-time percentiles. The approximated buffersize becomes:

$$\nu(k,\beta) * (c_S^2 - 1/m)/(1/k - 1/m) + \nu(m,\beta) * (c_S^2 - 1/k)/(1/m - 1/k). \quad (19)$$

Here $\nu(r,\beta)$ stands for the minimum buffersize for which the rejection-probability in the $G^X/E_r/1/N$ model does not exceed a prespecified value β. The next question is how good these approximations are. To check this simulation is not necessary, but we can simply test the approximation for models which also allow an exact solution.
In the next section we will discuss this item.

4. Numerical results

To get a good idea of the quality of the approximation method outlined in the previous section we start with calculating the exact values of the conditional waiting-time percentiles for the following models: $M^X/M/1/N$, $M^X/E_2/1/N$ and $M^X/E_8/1/N$. In all cases we took E(S) = 1. Next we used formula (17) with k = 1, m = 2 and $c_S^2 = 1/8$ to get the proposed approximative percentiles for the $M^X/E_8/1/N$ model. In table 1a we show the results both for a constant and a geometric batch size distribution, where in either case we have taken E(X) = 3. Further, the buffersize N = 25 and the

rejection strategy is WBAS. We have included the results for exponential service-times to show that the first moment is not sufficient.

Notice that the batch size distribution has a considerable effect on the percentiles. This can also be seen from the values of the rejection probability of an arbitrary customer. For the respective values of $\rho = 0.5$, 0.8, 0.9, 1.0 and 1.4 these values are 0.00004 (0.00779), 0.00851 (0.04376), 0.02682 (0.07049), 0.06340 (0.10628) and 0.28678 (0.29306), where the values between brackets correspond to the case of a geometric batch size. To get an even better idea of the quality of the approximation we also calculated the conditional waiting-time percentiles for the $M^X/E_{20}/1/N$ model and we compared these with the approximations which result from formula (18) if we use the exact results of the $M^X/M/1/N$ and the $M^X/E_2/1/N$ model.

It turned out that better approximations are obtained if we use a three-point approximation in this case (n = 2 in formula (17)). We used the percentiles of the $M^X/E_3/1/N$ model besides the percentiles of the earlier mentioned models. To show the differences between the two approximations we list some results in table 1b. Again the buffersize N has been taken 25 and the mean batch size $E(X) = 3$. We only give the results for a geometric batch size distribution and a WBAS strategy.

table 1a.
two-point approximation for the conditional waiting-time percentiles
of the $M^X/E_8/1/N$ model; N = 25; E(X) = 3

	constant				geometric			
ρ	0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
0.5								
exa	4.612	6.353	8.109	12.16	8.223	11.35	14.20	19.30
appr	4.624	6.380	8.133	12.21	8.236	11.37	14.20	19.26
expo	5.745	8.209	10.67	16.27	9.137	12.74	16.02	22.32
0.8								
exa	11.03	14.83	17.89	22.09	13.48	16.90	19.29	22.66
appr	11.15	14.97	17.92	22.16	13.50	16.88	19.26	22.85
expo	12.83	17.03	20.40	26.32	14.42	18.31	21.39	26.96
0.9								
exa	15.14	18.61	20.77	23.62	15.46	18.48	20.49	23.37
appr	15.20	18.56	20.75	23.97	15.45	18.44	20.50	23.69
expo	16.18	20.06	23.07	28.46	16.22	19.90	22.82	28.13
1.0								
exa	18.49	20.85	22.33	24.56	17.16	19.72	21.42	23.94
appr	18.38	20.83	22.49	25.16	17.12	19.70	21.50	24.39
expo	18.84	22.27	24.99	30.01	17.80	21.25	24.02	29.13
1.4								
exa	22.30	23.47	24.38	26.02	20.82	22.33	23.43	25.31
appr	22.50	23.90	25.02	27.07	20.86	22.55	23.84	26.12
expo	23.46	26.17	28.46	32.95	21.79	24.67	27.08	31.73

table 1b.

Comparison between the two-point and the three-point approximation for the conditional waiting-time percentiles of the $M^X/E_{20}/1/N$ model; $N = 25$, $E(X) = 3$, geometric batch size distribution

ρ	0.800	0.900	0.950	0.990	0.995	0.999
0.5						
exact	8.141	11.23	14.05	19.06	20.54	22.84
appr2	8.159	11.25	14.05	19.00	20.52	23.18
appr3	8.141	11.23	14.05	19.04	20.51	22.93
0.8						
exact	13.40	16.79	19.13	22.24	23.06	24.41
appr2	13.42	16.75	19.07	22.50	23.54	25.48
appr3	13.40	16.78	19.09	22.27	23.18	24.80
0.9						
exact	15.39	18.37	20.31	22.86	23.55	24.75
appr2	15.38	18.31	20.30	23.31	24.26	26.05
appr3	15.39	18.35	20.27	22.96	23.77	25.24
1.0						
exact	17.12	19.61	21.19	23.34	23.94	25.03
appr2	17.06	19.57	21.29	23.98	24.86	26.54
appr	17.11	19.57	21.17	23.52	24.24	25.61
1.4						
exact	20.75	22.08	22.99	24.40	24.85	25.73
appr2	20.78	22.37	23.56	25.64	26.36	27.79
appr3	20.71	22.11	23.13	24.83	25.40	26.54
1.8						
exact	21.97	22.95	23.66	24.86	25.26	26.06
appr2	22.20	23.51	24.54	26.41	27.07	28.41
appr3	21.98	23.08	23.93	25.42	25.95	27.00

Let us now look at the conditional waiting-time percentiles for the $G^X/G/1/N$ model. We only considered the case of constant interarrival-times and calculated the exact values for three service-time distributions: exponential service-time, Erlang-2 service-time and Erlang-8 service-time. In all cases we took $E(S)=1$. Next we used the results for the exponential case and the Erlang-2 case to get also approximative results for the $D^X/E_8/1/N$ model. We show the results in table 1c. both for a constant and a geometric batch size distribution. Also, by comparing the results with those in table 1a., notice the significant effect of the interarrival-time distribution. This can also be seen from the values for the rejection probability of an arbitrary customer.

table 1c.
two-point approximation for the conditional waiting-time percentiles
of the $D^X/E_8/1/N$ model; $N = 25$; $E(X) = 3$

		constant				geometric		
ρ	0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
0.5								
exa	2.095	2.407	2.663	3.156	5.250	7.308	9.343	13.89
appr	2.070	2.433	2.729	3.270	5.242	7.285	9.307	13.82
expo	2.546	3.485	4.396	6.455	6.018	8.581	11.11	16.73
0.8								
exa	2.126	2.456	2.749	3.325	9.086	12.29	15.02	19.57
appr	1.790	2.045	2.278	2.818	9.148	12.37	15.07	19.58
expo	4.546	6.407	8.271	12.60	10.92	14.70	17.90	23.82
0.9								
exa	2.356	2.824	3.271	4.310	12.23	15.47	17.86	21.48
appr	2.106	2.706	3.438	5.876	12.34	15.54	17.90	21.61
expo	8.309	11.69	14.90	21.33	13.89	17.63	20.65	26.15
1.0								
exa	17.62	19.90	21.39	23.70	15.76	18.33	20.13	22.93
appr	17.62	20.00	21.66	24.36	15.78	18.37	20.23	23.26
expo	18.31	21.70	24.40	29.42	16.79	20.21	22.96	28.09
1.4								
exa	23.58	24.49	25.24	26.68	21.31	22.57	23.55	25.31
appr	24.10	25.25	26.20	28.02	21.52	22.96	24.11	26.23
expo	25.35	27.84	29.99	34.27	22.64	25.32	27.61	32.10

Now we will look at the minimal buffer size for which the rejection probability
of an individual customer is below a prespecified value β. Exact values can be
obtained for the $M^X/E_k/1/N$ model, so we can get approximations for the minimal
buffer sizes of the $M^X/G/1/N$ model via the results of the $M^X/M/1/N$ model and
the $M^X/E_2/1/N$ model. So we calculated minimal buffer sizes with respect to the
values $\beta = 0.01$, 0.001, 0.0001 and 0.00001 for either model and we used the
results to get approximative minimal buffer sizes for the $M^X/E_8/1/N$ model.
The two point interpolations were rounded upward to the nearest integer.
In table 2a we give a comparison between the approximative and the exact values
for both a geometric and a constant batch size distribution. We give only the
results for the WBAS strategy.

table 2a
two-point approximation for the minimal buffer sizes in the $M^X/E_8/1/N$ model; $E(X) = 3$; strategy is WBAS

	constant batch size				geometric batch size			
ρ	0.01	0.001	0.0001	0.00001	0.01	0.001	0.0001	0.00001
0.5								
exa	12	18	23	29	24	36	47	59
app	13	18	23	30	26	37	48	60
exp	14	21	28	35	25	38	51	63
0.7								
exa	18	29	39	50	33	52	72	91
app	19	29	40	51	33	53	72	93
exp	22	35	48	61	36	58	80	101
0.8								
exa	24	41	57	74	43	72	101	130
app	25	42	58	75	44	72	101	130
exp	30	50	71	92	47	80	114	147
0.9								
exa	38	72	107	141	65	122	180	239
app	38	73	107	142	66	122	181	239
exp	48	90	134	178	74	139	207	275
0.95								
exa	57	123	193	264	95	205	322	440
app	57	123	195	266	96	206	323	441
exp	72	156	245	335	109	237	373	510

At last we consider the minimal buffer sizes for the $G^X/G/1/N$ model. As for the waiting-time percentiles we take a constant interarrival-time and an Erlang-8 service-time distribution. Exact results have been calculated for exponential service-times and Erlang-2 service-times and these have been used to get approximations for the minimal buffer sizes of the $D^X/E_8/1/N$ model. In table 2b. we give a comparison between these approximations and the exact results for the $D^X/E_8/1/N$ model. We give only the results for constant batch sizes and a WBAS strategy. We also include the minimal buffer sizes for the case of exponential service to show that the exact results for the case of exponential service cannot be used as first-order approximations for the case of general service. Note also from table 2b. the empirical finding that the rejection probability decreases exponentially as function of the capacity of the buffer size. It is remarked that approximations of a comparable quality were obtained when a PBAS strategy was followed. Finally we remark that for practical purposes a two-point approximation is sufficient.

- 16 -

table 2b
two-point approximation for the minimal buffer sizes in the $D^X/E_8/1/N$ model; $E(X) = 3$; strategy is WBAS

constant batch size

ρ	0.01	0.001	0.0001	0.00001	0.000001	0.0000001
0.7						
exa	4	4	5	5	5	6
app	4	5	7	6	7	7
exp	7	10	13	16	19	22
0.8						
exa	4	5	5	6	7	7
app	4	6	6	7	9	8
exp	9	14	19	24	29	34
0.9						
exa	5	7	8	9	11	12
app	6	7	8	10	11	13
exp	14	24	35	46	56	67
0.95						
exa	6	9	11	14	17	20
app	7	8	11	16	18	20
exp	20	41	63	85	107	130

References.

[BURKE] Burke, P. J. (1975) 'Delays in Single-Server Queues with Batch Input', Oper. Res., 23, 830-833.

[MANF] Manfield, D. R. and Tran-Gia, P. (1982) 'Analysis of a Finite Storage System with Batch Input Arising out of Message Packetization', IEEE Transactions on Communications, vol. com-30, no. 3, 456-463.

[TAKA] Takahashi, Y. and Katayama, T. (1985) 'Multi-Server System with Batch Arrivals of Queueing and Non-Queueing Calls', in Proceedings of the 11th International Teletraffic Congress, Kyoto, Japan, pp. 3.2A-4-1/7.

[TIJMS] Tijms, H. C. (1986) Stochastic Modelling and Analysis, Wiley, New York.