An Approximate Performance Model for a Multislot Integrated Services System

KE-QIANG LIAO AND LORNE G. MASON

Abstract—**In this paper, we present an approximate model (with finite or infinite waiting room) for an integrated service system with three types of traffic: a first offered narrow-band traffic, an overflow narrow-band traffic and a wide-band traffic. A narrow-band call requires a single server, while the number of servers required to serve a wide-band call is *N*. The blocked narrow-band calls are lost while the blocked wide-band calls are delayed in a finite or infinite waiting room. Based on two assumptions with regard to the characteristics of the system, we resolve the system by decomposition. The corresponding improvements in numerical efficiency as well as in computational storage requirements are significant enough to enable use of the model within network optimization algorithms. The model provides a very good approximation for the system performance, that is the blocking probabilities of the two narrow-band traffics, the loss probability (in the case of finite waiting room), the probability of nonwaiting and the average waiting time of wide-band traffic.**

I. Introduction

AMONG the many architectures proposed for the ISDN (Integrated Services Digital Network), those based on variable rate circuit switching ($N \times 64$ kbits/s channels or slots) are attractive because they evolve naturally from the present telephone network. The centrally controlled digital cross-connect devices currently being deployed by most administrations provide an opportunity to offer new slow-switched services such as multimedia, multipoint teleconferencing and customer-controlled private networks. The similarities in functional requirements and traffic characteristics of these services suggests a potential for integration with existing circuit-switched services.

Before introducing the contribution of this paper, we shall briefly review previous work in this area. A more extensive review of the pertinent queueing literature can be found in [2]. In [8], an architecture was proposed for integrating wide-band (WB) slow-switched calls, such as video teleconferences, with narrow-band (NB) calls such as telephone traffic. A method was described for computing end-to-end set-up delay, for multipoint WB calls, and loss probability, for point-to-point NB calls, in terms of link level models for these performance measures.

An M-server queueing model was proposed for the transmission links. The demand consisted of two classes, WB and NB, having Poisson arrival rates λ_2 and λ_1, respectively. WB and NB holding times are exponentially distributed with means μ_2^{-1} and μ_1^{-1}, respectively. Each WB call requires N 64 kbits/s channels which are seized and released simultaneously. If sufficient idle servers are not available at the arrival instant of a WB call, it is placed in an infinite queue until it can be served in order of arrival. A cutoff parameter r_0 specifies the maximum number of WB calls which are allowed to be simultaneously connected. NB calls require a single 64 kbits/s server, and blocked calls are cleared from the system.

A link performance model was derived, under the approximating assumption of independent channel release for WB calls, using a technique originally due to Green [5]. Unfortunately, this model suffered from inaccuracy in some quantities of interest. In [2], exact results were derived, using Z-transform and matrix geometric techniques, and compared numerically. While these exact models provided insight into system behavior they involved considerable computation and were only applicable to trunk groups smaller than those typically found in practice. Problems of numerical stability were encountered for systems having a large number of servers. Fast accurate link performance models for a large range of server capacities are essential for network design purposes as these subroutines are typically called many times in the course of a routing optimization or network dimensioning procedure.

In our quest for accurate efficient performance models, a very successful approximation approach has been discovered and introduced in [10]. It is this approach which is pursued and significantly extended in this paper. Stated informally, the approximation involves a decomposition on the number of WB calls in service and in queue. For each state of the WB process, the NB process is assumed to be in stochastic equilibrium, while the WB process is assumed to be Markov, with some transition rates dependent on the distribution and certain first passage times of the NB process. In [10], a finite buffer system was analyzed where the stationary distribution of the WB Markov process was obtained by recursively solving the finite set of balance equations. An approximation to the infinite buffer system usig Welch's results [12] for the special case where $r_0 = 1$, at most one WB call allowed in service, was also described. The analysis described in this paper differs in two important respects. First, we employ a generating function approach which permits the solution of both finite and infinite buffer systems. Moreover, the new method is more efficient in terms of computation time and storage for finite buffer systems as well. More importantly, we have extended the model to include two types of NB traffic, and have added two additional control parameters. This was found to be necessary when using the model in networks employing adaptive routing for NB traffic. We let λ_{11} and λ_{12} denote the arrival rates of first offered and overflow NB calls, respectively. While overflow traffic is peaked in conventional automatic alternate routing networks, the Poisson assumption is reasonable in the context of adaptive routing as state-dependent routing tends to smooth the overflow traffic offered to each trunk group [4]. First offered and overflow NB traffic must be handled differently, in nonhierarchical routing schemes, in order to avoid network instability as demonstrated in [1], [7], [9], [11]. The control parameter introduced for this purpose r_1 is a static reservation parameter for first offered NB traffic. More precisely, an overflow NB call is lost if upon arrival, the number of occupied servers is $\geq r_1$. The other control parameters r_2 also introduced here for the first time,

Paper approved by the Editor for Communication Switching of the IEEE Communications Society. Manuscript received August 21, 1987; revised January 8, 1988 and March 23, 1988.

The authors are with INRS Telecommunications, Verdun, P.Q. H3E 1H6, Canada.

IEEE Log Number 8825980.

protects the WB traffic from overload of the narrow-band traffic. More specifically, the system blocks NB calls when there are some WB calls waiting if the number of WB calls connected or in service is $\leq r_2$. It is shown that this flow control parameter improves link performance as quantified by a power measure. The model also includes the parameter r_0 of previous models, which protects NB traffic from overload of WB traffic and prevents bimodal instability with the attendant long blocking periods for NB traffic.

To state succintly, the integrated service system considered in this paper supports three types of traffic: a first offered NB traffic, an overflow NB traffic and a WB traffic. A narrow-band call requests a single server while each wide-band call requires concurrent service from N of the M servers. Servers allocated to a WB call are seized and released simultaneously. Blocked NB calls are lost but blocked WB calls are delayed in an infinite or finite waiting room. Eventually, in the finite waiting room case, WB calls are lost if they arrive when the queue is full. The three different traffics are generated according to Poisson process with rates λ_{11}, λ_{12}, and λ_2. The service times are exponentially distributed and the mean service times of NB and WB calls are, respectively, μ_1^{-1} and μ_2^{-1}. WB calls enter service in their order of arrival.

The paper is organized as follows. In Section II, the assumptions made, and the decomposition approach employed, are stated precisely. The coupling between the narrow-band and wide-band processes is made explicit in terms of certain coupling parameters. In Section III, distributions and blocking probabilities are derived, for the two types of NB traffic, along with the coupling parameters which quantify NB interference effects on the WB process. In Section IV, the WB process is analyzed by the method of generating functions for both finite and infinite buffer systems. Expressions are given for the average number of WB calls in queue, the mean waiting time and the probability of nonwaiting for wide-band calls. The buffer overflow or loss probability for WB calls, in the case of a finite buffer, is also given. Numerical results are presented in Section V, which show the accuracy of the approximation relative to an exact solution. The effect of various system parameters on performance is also displayed. Section VI concludes.

II. The Approximate Model

For the WB services considered, it is expected that service times will be significantly longer than that of NB calls. Consequently, the NB process will have sufficient time to approach equilibrium conditions for each state of the WB process. It is, therefore, natural to make the assumption that the NB process achieves steady state instantaneously for each state of the WB process. Clearly, the NB distribution depends upon the number of WB calls in service as these servers are unavailable to the NB traffic. If the number of WB calls in service is les than r_0, the NB distribution also depends upon whether the WB queue is empty or not. If this queue is empty, then all servers not occupied by WB calls are available to the NB traffic. On the other hand, if there are WB calls queued, this implies that the number of free servers must be less than N, for otherwise, the WB call at the head of the queue would enter service. Accordingly, the distribution of NB calls in service must reflect this fact.

The NB traffic also has an influence on the WB traffic. In particular, when a WB call arrives, and the WB queue is empty, it can only enter service immediately if there are at least N servers idle. Assuming that there are i WB calls in service when such a WB call arrives, it will be placed in the queue with probability $B(i)$, which depends upon the current distribution of NB calls. Also if there are $j > 0$ WB calls queued and i WB calls being served, the WB call at the head of the queue can only enter service when N servers become available. This can occur with a WB departure or when

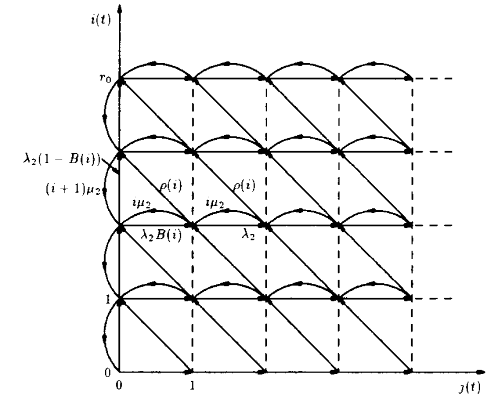

Fig. 1. State diagram and transition rates.

sufficient NB calls depart to free up N servers. The latter event will occur at an average rate $\rho(i)$ which is the inverse of the average first passage time of the NB call occupancy process, from states with fewer than N free servers to the state with N free servers.

This interference time, also called the exceptional service time or exceptional delay, is distributed in a complex fashion. In order to make the WB process tractable, we approximate this complex distribution with an exponential distribution having the same mean. In cases tested, the approximation appears to be a good two moment approximation. This assumption, together with the previous assumptions of Poisson arrivals and exponential service times, enables us to model the WB process $(i(t), j(t))$ as a two-dimensional Markov process where $i(t)$ denotes the number of WB calls in service, and $j(t)$ is the number in the queue at time t. The state transition diagram for this WB process is shown in Fig. 1 where the influence of the NB coupling parameters $B(i)$ and $\rho(i)$ determine certain transition rates. In the following sections, we analyze the NB and WB process in detail and derive performance measures in terms of these processes.

For the reader's convenience we now summarize the two assumptions and the parameters of the system considered. In effect, the approximate model is based on the following two assumptions.

1) The narrow-band occupancy achieves the stationary distribution while there are a fixed number of wide-band calls in service except when the WB service protection control operates, since in that case, the NB process is a pure death process.

2) The exceptional service times are exponentially distributed.

The system parameters of the model are as follows:

1) total number of servers: M
2) number of servers allocated to a WB call: N
3) cutoff parameter, maximum number of WB calls that can be in service simultaneously: r_0
4) trunk reservation parameter: r_1
5) wide-band service protection parameter r_2
6) arrival rates of NB and WB calls: λ_{11}, λ_{12}, λ_2
7) service rates of NB and WB calls: μ_1, μ_2.

The integrated service system is considered to be in the stationary state. For the finite waiting room case, no conditions are required for the existence of steady state for the joint process. As for the infinite waiting room case, the steady state will exist under the stability condition $\lambda_2 < r_0\mu_2$. No formal proof is provided here for this stability condition, but it can be justified as follows. If there are some WB calls in queue, and the number of WB calls in service is less than r_0, whenever there are N idle servers, the first WB call in queue will seize them and the seized servers are not released until the queue empties. We can also see that for any values of λ_{11} and λ_{12}, if the number of WB calls in service is less than r_0 ($r_0 \leq M/N$), there is a nonzero probability of having N idle servers. The maximum rate at which the system can serve the wide-band traffic is therefore $r_0\mu_2$.

III. Narrow-Band Process

We will first calculate the influence of NB traffic on the WB process using the two assumptions made in Section II. This influence is characterized by $\rho(i)$ and $B(i)$ where $B(i)$ is the probability that a WB arrival must wait when there are i WB calls in service and the queue is empty. Let X_1 be the number of narrow-band calls in service. Due to the two assumptions made in Section II, $\rho(i)^{-1}$, the average first passage time to the state $X_1 = M - (i+1)N$, while there are i WB calls in service and some WB calls in queue, can be written

$$\rho(i)^{-1} = \sum_{j=M-(i+1)N+1}^{M-iN} \rho(j, M-(i+1)N)^{-1} \Pr(j/i) \quad (1)$$

where $\Pr(j/i)$ is the conditional probability that there are j NB calls in the system, assuming that there are i WB calls and at least $M-(i+1)N+1$ NB calls in service, and $\rho(j, M-(i+1)N)^{-1}$ denotes the average first passage time from the state $X_1 = j$ to the state $X_1 = M-(i+1)N$.

Due to the assumption that the system is in the steady state for NB calls when the WB call arrives to an empty queue, one can calculate the conditional probabilities $\Pr(j/i)$ by using the following formulae.

If $M - r_1 \geq N$ (the overflow NB calls are blocked),

$$\Pr(j/i) = \frac{\rho_{11}^j}{j!} C_i^{-1} \quad M-(i+1)N+1 \leq j \leq M-iN \quad (2)$$

with

$$C_i = \sum_{n=M-(i+1)N+1}^{M-iN} \frac{\rho_{11}^n}{n!}$$

otherwise

$$\Pr(j/i) = \begin{cases} \dfrac{(\rho_{11}+\rho_{12})^j}{j!} C_i^{-1} \\ M-(i+1)N+1 \leq j \leq r_1 - iN \\ (\rho_{11}+\rho_{12})^{r_1-iN} \dfrac{\rho_{11}^{j-r_1+iN}}{j!} C_i^{-1} \\ r_1 - iN + 1 \leq j \leq M - iN \end{cases} \quad (3)$$

with

$$C_i = \sum_{n=M-(i+1)N+1}^{r_1-iN} \frac{(\rho_{11}+\rho_{12})^n}{n!} + (\rho_{11}+\rho_{12})^{r_1-iN} \sum_{n=r_1-iN+1}^{M-iN} \frac{\rho_{11}^{n-r_1+iN}}{n!}$$

where

$$\rho_{11} = \frac{\lambda_{11}}{\mu_1}$$

$$\rho_{12} = \frac{\lambda_{12}}{\mu_1}.$$

The average first passage time to the state $X_1 = M - (i+1)N$, while there are i WB calls in service and some WB calls in queue, can be calculated as follows. If we denote the first passage time from the state $X_1 = j$ to the state $X_1 = j - k$ by $\theta_{j,j-k}$ and define

$$m_j = E\{\theta_{j,j-1}\} \quad (4)$$

then

$$\rho(j, j-k)^{-1} = E\{\theta_{j,j-k}\} = m_j + \cdots + m_{j-k+1}. \quad (5)$$

In the case where $0 \leq i \leq r_2$, since the narrow-band arrivals are blocked, we have the simple formula

$$m_j = \frac{1}{j\mu_1} \quad M-(i+1)N+1 \leq j \leq M-iN. \quad (6)$$

We note that during such a period, since all NB arrivals are blocked, the NB process is a pure death process and is not in stochastic equilibrium. We do not assume such stationarity.

If $i > r_2$, we can calculate the m_j by using the following recurrence relation:

$$\begin{aligned} m_j &= \frac{1}{j\mu_1} + \frac{\lambda_{i,j}^*}{j\mu_1} m_{j+1} \quad M-(i+1)N+1 \leq j \leq M-iN-1 \\ m_{M-iN} &= \frac{1}{(M-iN)\mu_1} \end{aligned} \quad (7)$$

where

$$\lambda_{i,j}^* = \begin{cases} \lambda_{11}+\lambda_{12} & \text{for } M-(i+1)N+1 \leq j \leq r_1 - iN - 1 \\ \lambda_{11} & \text{for } r_1 - iN \leq j \leq M - iN. \end{cases} \quad (8)$$

Here, $\lambda_{i,j}^*$ denotes the arrival rate of NB calls when there are i WB calls and j NB calls in service. During this period, the NB process is assumed to be in stochastic equilibrium.

$B(i)$, is the probability that a WB arrival must wait, when there are i WB calls in service and there is no WB call in queue. Noting that $B(i)$ is the probability that $M-(i+1)N+1 \leq X_1 \leq M - iN$, we can therefore write

$$B(i) = \sum_{n=M-(i+1)N+1}^{M-iN} \frac{\prod_{j=0}^{n-1} \lambda_{i,j}^*}{n!(\mu_1)^n} \times \left\{ 1 + \sum_{n=1}^{M-iN} \frac{\prod_{j=0}^{n-1} \lambda_{i,j}^*}{n!(\mu_1)^n} \right\}^{-1} \quad 0 \leq i < r_0 \quad (9)$$

where $\lambda_{i,j}^*$ are given by (8).

The calculation of conditional blocking probability of the two types of NB traffic can be decomposed into two regions as follows. In the first region, the NB process is assumed to be stationary and all NB blocking occurs because of no available servers where we take into account that $(M - r_1)$ servers are reserved for first offered NB traffic. The states of the wide-band process defining this region are given by $\{i(t) > r_2\}$ and $\{i(t) \leq r_2, j(t) = 0\}$ where $i(t)$ denotes the number of WB calls in service and $j(t)$ the number in queue. In this region, the conditional blocking probability of the two NB traffic types, given that there are i WB calls in service, can be written (see [1]) if $r_1 > iN$,

$$\begin{aligned} PB_{11}(i) &= \frac{(\rho_{11}+\rho_{12})^{r_1-iN} \rho_{11}^{M-r_1}}{(M-iN)!} P_0(i)^{-1} \\ PB_{12}(i) &= 1 - \sum_{n=0}^{r_1-iN-1} \frac{(\rho_{11}+\rho_{12})^n}{n!} P_0(i)^{-1} \end{aligned} \quad (10)$$

where

$$P_0(i) = \sum_{n=0}^{r_1-iN} \frac{(\rho_{11}+\rho_{12})^n}{n!} + (\rho_{11}+\rho_{12})^{r_1-iN} \sum_{n=r_1-iN+1}^{M-iN} \frac{\rho_{11}^{n-r_1+iN}}{n!}.$$

If $r_1 \le iN$,

$$PB_{11}(i) = \frac{\rho_{11}^{M-iN}}{(M-iN)!} P_0(i)^{-1}$$

$$PB_{12}(i) = 1 \tag{11}$$

where

$$P_0(i) = \sum_{n=0}^{M-iN} \frac{\rho_{11}^n}{n!}.$$

In the second region, all the NB calls are blocked due to the WB service protection control associated with the parameter r_2. The set of the states of the WB process is the complement of that in the first region and is given by $\{i(t) \le r_2, j(t) > 0\}$.

Summing the conditional probabilities for these two complementary sets gives

$$PB_{11} = \sum_{i=0}^{r_2} \left\{ PB_{11}(i)P(i, 0) + \sum_{j=1}^{\infty} P(i, j) \right\} + \sum_{i=r_2+1}^{r_0} PB_{11}(i) \sum_{j=0}^{\infty} P(i, j)$$

$$PB_{12} = \sum_{i=0}^{r_2} \left\{ PB_{12}(i)P(i, 0) + \sum_{j=1}^{\infty} P(i, j) \right\} + \sum_{i=r_2+1}^{r_0} PB_{12}(i) \sum_{j=0}^{\infty} P(i, j) \tag{12}$$

where $P(i, j)$ denotes the probability that there are i wide-band calls in service and j in queue. Evidently, if the number of waiting places is limited, for example, K, we have $P(i, j) = 0$ for $j > K$.

IV. Wide-Band Process

Now, we consider the wide-band process. With the two assumptions, exponentially distributed exceptional service times and stationary distribution for NB calls, the system characterizing the WB process can be modeled as a two dimensional Markov process $(i(t), j(t))$ where $i(t)$ is the number of WB calls in service and $j(t)$ is the number in queue. The state space is then the set $\{(i, j)/0 \le i \le r_0, 0 \le j \le K\}$ in the case of finite waiting room (K is the number of waiting places) and $\{(i, j)/0 \le i \le r_0, 0 \le j \le \infty\}$ if the queue is infinite. The steady-state probability that the system is in state (i, j) is denoted by $P(i, j)$.

Welch [12] analyzed a generalized $M/G/1$ queue system in which the first customer of each busy period receives an exceptional service. The model studied here can be considered as a generalization of his model with regard to the number of servers in the special case where the exceptional service times and service times are exponentially distributed.

We will now develop the analysis for the two cases of infinite and finite waiting room.

A. Infinite Waiting Room Case

Fig. 1 gives the state diagram in the case of infinite waiting room. We can obtain the global balance equations by equating the total rate at which the process leaves a state to the total rate at which it enters that state. We give here the balance equations.

In the case where $i = 0$,

$$\lambda_2 P(0, 0) = \mu_2 P(1, 0)$$

$$(\rho(0)+\lambda_2)P(0, 1) = \lambda_2 B(0)P(0, 0)$$

$$(\rho(0)+\lambda_2)P(0, j) = \lambda_2 P(0, j-1) \quad 2 \le j. \tag{13}$$

In the case where $0 < i < r_0$,

$$(\lambda_2 + i\mu_2)P(i, 0) = (i+1)\mu_2 P(i+1, 0) + i\mu_2 P(i, 1) + \lambda_2(1-B(i-1))P(i-1, 0) + \rho(i-1)P(i-1, 1)$$

$$(\lambda_2 + i\mu_2 + \rho(i))P(i, 1) = i\mu_2 P(i, 2) + \lambda_2 B(i)P(i, 0) + \rho(i-1)P(i-1, 2)$$

$$(\lambda_2 + i\mu_2 + \rho(i))P(i, j) = i\mu_2 P(i, j+1) + \lambda_2 P(i, j-1) + \rho(i-1)P(i-1, j+1) \quad 2 \le j. \tag{14}$$

In the case where $i = r_0$,

$$(r_0\mu_2 + \lambda_2)P(r_0, 0) = r_0\mu_2 P(r_0, 1) + \rho(r_0-1)P(r_0-1, 1) + \lambda_2(1-B(r_0-1))P(r_0-1, 0)$$

$$(r_0\mu_2 + \lambda_2)P(r_0, j) = r_0\mu_2 P(r_0, j+1) + \rho(r_0-1)P(r_0-1, j+1) + \lambda_2 P(r_0, j-1) \quad 1 \le j. \tag{15}$$

We define

$$Q_i(z) = \sum_{j=0}^{\infty} P(i, j)z^j. \tag{16}$$

In order to calculate the blocking probabilities of NB traffics and the probability of nonwaiting of WB traffic, we need the probability of having i WB calls in service

$$Q_i(1) = \sum_{j=0}^{\infty} P(i, j) \tag{17}$$

and the probability that there are i WB calls in service and no WB calls in queue

$$Q_i(0) = P(i, 0). \tag{18}$$

From the balance (13), we obtain

$$Q_0(z) = P(0, 0) + P(0, 0)B(0) \frac{\lambda_2 z}{\rho(0) + \lambda_2(1-z)}. \tag{19}$$

In the case where $0 < i < r_0$, we have

$$\begin{aligned}(\lambda_2 + \rho(i) + i\mu_2)Q_i(z) &= \lambda_2 z Q_i(z) - \lambda_2(1-B(i))P(i, 0)z \\ &+ \frac{i\mu_2}{z}(Q_i(z) - P(i, 0)) \\ &+ \frac{\rho(i-1)}{z}(Q_{i-1}(z) - P(i-1, 0)) \\ &+ \rho(i)P(i, 0) + (i+1)\mu_2 P(i+1, 0) \\ &+ \lambda_2(1-B(i-1))P(i-1, 0).\end{aligned} \tag{20}$$

To obtain the various performance quantities of interest, we first must evaluate the quantities $P(i, 0)$, $Q_i(1)$, $Q_i'(1)$ and $Q_i^{(2)}(1)$ [$Q_i'(1)$ and $Q_i^{(2)}(1)$ are, respectively, the first and second derivatives of $Q_i(z)$ at $z = 1$]. These quantities are computed recursively, iterating on i by using (21), (23), (24), and (25) for $i < r_0$ and (21), (26), and (28) for $i = r_0$. The recursion is initiated by using the closed-form expression for $Q_0(z)$ given by (19) and its derivatives. $P(0, 0)$ is initially set to a constant and, following the calculation of the recursion, the probabilities are normalized. For a fixed value of i, the above quantities are computed in the order listed. The quantities $P(i, 0)$, $Q_i(1)$, and $Q_i'(1)$ are used directly in the calculation of performance measures, whereas the quantities $Q_i^{(2)}(1)$ for $0 \le i \le r_0 - 1$ are not directly used for performance evaluation but are required to evaluate $Q_{r_0}'(1)$ by L'Hospital's rule.

From the state diagram given in Fig. 1, we obtain the following recurrence relations:

$$i\mu_2 P(i, 0) = \lambda_2(1 - B(i-1))P(i-1, 0) + \rho(i-1)\sum_{j=1}^{\infty} P(i-1, j) \quad (21)$$

since the rate at which the process goes from the state $(i, 0)$ to the state $(i - 1, 0)$ must equal to the rate at which it changes from the state set $\{(i - 1, j), j = 1, \cdots, \infty\}$ to the state set $\{(i, j), j = 0, \cdots, \infty\}$.

Applying (21), we can obtain

$$Q_i(z) = \frac{H_i(z)}{G_i(z)} \quad (22)$$

where

$$H_i(z) = \rho(i)Q_i(1)z + \rho(i-1)(Q_{i-1}(z) - Q_{i-1}(1)) + \lambda_2(1-z)\{(1-B(i))P(i, 0)z - (1-B(i-1))P(i-1, 0)\}$$

$$G_i(z) = -\lambda_2 z^2 + (\lambda_2 + \rho(i) + i\mu_2)z - i\mu_2.$$

As for the case $i = r_0$, we can derive $Q_{r_0}(z)$ in a similar manner and the expression is the same as (22) except that $\rho(r_0) = 0$ and $B(r_0) = 1$.

Given $Q_{i-1}(z)$, we have two unknowns $P(i, 0)$ and $Q_i(1)$ in (22). However, we can calculate $P(i, 0)$ by applying formula (21) and then it leaves us only one unknown to determine. According to the definition of a generating function, $Q_i(z)$ should be regular for $|z| < 1$, continuous for $|z| \le 1$. Hence, every zero z of the denominator of $Q_i(z)$ must be a zero of its numerator if $|z| \le 1$.

We can easily demonstrate that both the zeros of the denominator are real and that there must be at least one zero on or inside the unit circle. Therefore, $Q_i(1)$ for $1 \le i < r_0$ can be determined in a recursive manner. Following the above mentioned line of reasoning, if $-1 \le \xi_i \le 1$ and

$$-\lambda_2\xi_i^2 + (\lambda_2 + \rho(i) + i\mu_2)\xi_i - i\mu_2 = 0$$

we have the following relation:

$$Q_i(1) = -\frac{\rho(i-1)(Q_{i-1}(\xi_i) - Q_{i-1}(1)) + \lambda_2(1-\xi_i)\{(1-B(i))P(i, 0)\xi_i - (1-B(i-1))P(i-1, 0)\}}{\rho(i)\xi_i}. \quad (23)$$

Differentiating (22) and evaluating at $z = 1$ for $1 \le i < r_0$, we then have the following equation:

$$\rho(i)Q_i'(1) = (\lambda_2 - i\mu_2)Q_i(1) + \rho(i-1)Q_{i-1}'(1) - \lambda_2(1-B(i))P(i, 0) + \lambda_2(1-B(i-1))P(i-1, 0). \quad (24)$$

In the same manner, differentiating (22) twice and evaluating at $z = 1$, we obtain

$$\rho(i)Q_i^{(2)}(1) = \rho(i-1)Q_{i-1}^{(2)}(1) + 2\lambda_2(Q_i(1) - (1-B(i))P(i, 0)) - 2\{i\mu_2 + \rho(i-1) - \lambda_2\}Q_i'(1). \quad (25)$$

In the case where $i = r_0$, we have $\rho(i) = 0$. By using L'Hospital's rule, it is easy to obtain

$$Q_{r_0}(1) = \frac{\rho(r_0-1)Q_{r_0-1}'(1) + \lambda_2(1-B(r_0-1))P(r_0-1, 0)}{r_0\mu_2 - \lambda_2}. \quad (26)$$

The first derivative of $Q_{r_0}(z)$ can be written

$$Q_{r_0}'(z) = \frac{d}{dz}\left\{\frac{H_{r_0}(z)}{G_{r_0}(z)}\right\} = \frac{H_{r_0}'(z)G_{r_0}(z) - G_{r_0}'(z)H_{r_0}(z)}{G_{r_0}(z)^2} \quad (27)$$

where

$$H_{r_0}(z) = \rho(r_0-1)(Q_{r_0-1}(z) - Q_{r_0-1}(1)) - \lambda_2(1-z)(1-B(r_0-1))P(r_0-1, 0)$$

$$G_{r_0}(z) = -\lambda_2 z^2 + (\lambda_2 + r_0\mu_2)z - r_0\mu_2.$$

Using L'Hospital's rule, we therefore have

$$Q_{r_0}'(1) = \left.\frac{H_{r_0}^{(2)}(z)G_{r_0}(z) - G_{r_0}^{(2)}(z)H_{r_0}(z)}{2G_{r_0}(z)G_{r_0}'(z)}\right|_{z=1} = \frac{1}{2G_{r_0}'(1)}\{H_{r_0}^{(2)}(1) - G_{r_0}^{(2)}(1)Q_{r_0}(1)\} = \frac{\rho(r_0-1)Q_{r_0-1}^{(2)}(1) - 2\lambda_2 Q_{r_0}(1)}{2(r_0\mu_2 - \lambda_2)}. \quad (28)$$

If we denote the average number of waiting customers by AWN, the following relation is evident:

$$\text{AWN} = \sum_{i=0}^{r_0} Q_i'(1). \quad (29)$$

We can therefore calculate the average waiting time by using Little's formula

$$\text{AWT} = \frac{\text{AWN}}{\lambda_2}. \quad (30)$$

The probability of nonwaiting is given by the following equation:

$$\text{PNW} = \sum_{i=0}^{r_0-1} (1 - B(i))P(i, 0). \quad (31)$$

B. Finite Waiting Room Case

We consider now the finite waiting room model. The system is almost the same as in the case of infinite waiting room except that the queue length is limited to K. In a similar manner, we can easily write the balance equations.

In the case where $i = 0$,

$$\begin{aligned} &\lambda_2 P(0,0) = \mu_2 P(1,0) \\ &(\rho(0)+\lambda_2)P(0,1) = \lambda_2 B(0) P(0,0) \\ &(\rho(0)+\lambda_2)P(0,j) = \lambda_2 P(0,j-1) \quad 2 \le j < K \\ &\rho(0)P(0,K) = \lambda_2 P(0,K-1). \end{aligned} \tag{32}$$

In the case where $0 < i < r_0$,

$$\begin{aligned} (\lambda_2 + i\mu_2)P(i,0) &= (i+1)\mu_2 P(i+1,0) + i\mu_2 P(i,1) \\ &\quad + \lambda_2(1-B(i-1))P(i-1,0) + \rho(i-1)P(i-1,1) \\ (\lambda_2 + i\mu_2 + \rho|i|)P(i,2) &= i\mu_2 P(i,2) \\ &\quad + \lambda_2 B(i)P(i,0) + \rho(i-1)P(i-1,2) \\ (\lambda_2 + i\mu_2 + \rho(i))P(i,j) &= i\mu_2 P(i,j+1) + \lambda_2 P(i,j-1) \\ &\quad + \rho(i-1)P(i-1,j+1) \quad 2 \le j < K \\ (i\mu_2 + \rho(i))P(i,K) &= \lambda_2 P(i,K-1). \end{aligned} \tag{33}$$

In the case where $i = r_0$,

$$\begin{aligned} (r_0\mu_2 + \lambda_2)P(r_0,0) &= r_0\mu_2 P(r_0,1) \\ &\quad + \rho(r_0-1)P(r_0-1,1) + \lambda_2(1-B(r_0-1))P(r_0-1,0) \\ (r_0\mu_2 + \lambda_2)P(r_0,j) &= r_0\mu_2 P(r_0,j+1) \\ &\quad + \rho(r_0-1)P(r_0-1,j+1) + \lambda_2 P(r_0,j-1) \quad 1 \le j < K \\ r_0\mu_2 P(r_0,K) &= \lambda_2 P(r_0,K-1). \end{aligned} \tag{34}$$

As in the preceding section, we define

$$Q_i(z) = \sum_{j=0}^{K} P(i,j) z^j. \tag{35}$$

In order to calculate the various performance measures of the system, besides the $P(i, 0)$, $Q_i(1)$, $Q_i'(1)$, and $Q_i^{(2)}(1)$, we must also calculate $P(i, K)$ for obtaining the loss probability of WB traffic. We compute these quantities in a similar way as in the infinite waiting room case. The recursion equations are (38), (40), (42), and (43) for $i < r_0$ and (38), (41), (44), and (46) for $i = r_0$.

It is easy to obtain from the balance equation (32)

$$\begin{aligned} Q_0(z) = P(0,0) \Bigg\{ 1 + B(0) &\frac{\lambda_2 z}{\rho(0) + \lambda_2(1-z)} \\ &\cdot \left(1 - \left(\frac{\lambda_2 z}{\lambda_2 + \rho(0)}\right)^K\right) \\ &+ B(0) \frac{\lambda_2 + \rho(0)}{\rho(0)} \left(\frac{\lambda_2 z}{\lambda_2 + \rho(0)}\right)^K \Bigg\}. \end{aligned} \tag{36}$$

In the case where $0 < i < r_0$, we have

$$\begin{aligned} &(\lambda_2 + \rho(i) + i\mu_2)Q_i(z) \\ &= \lambda_2 z Q_i(z) - \lambda_2(1-B(i))P(i,0)z \\ &\quad + \frac{i\mu_2}{z}(Q(z) - P(0)) \\ &\quad + \frac{\rho(i-1)}{z}(Q_{i-1}(z) - P(i-1,0)) \\ &\quad + \rho(i)P(i,0) + (i+1)\mu_2 P(i+1,0) \\ &\quad + \lambda_2(1-B(i-1))P(i-1,0) + \lambda_2(1-z)z^K P(i,K). \end{aligned} \tag{37}$$

Since we have the following relation as (21) in Section IV-A

$$i\mu_2 P(i,0) = \lambda_2(1-B(i-1))P(i-1,0) + \rho(i-1)\sum_{j=1}^{K} P(i-1,j) \tag{38}$$

we can write

$$Q_i(z) = \frac{H_i(z)}{G_i(z)} \tag{39}$$

where

$$\begin{aligned} H_i(z) &= \rho(i)Q_i(1)z + \rho(i-1)(Q_{i-1}(z) - Q_{i-1}(1)) \\ &\quad + \lambda_2(1-z)\{(1-B(i))P(i,0)z - (1-B(i-1)) \\ &\quad \cdot P(i-1,0)\} + \lambda_2(1-z)z^{K+1}P(i,K) \\ G_i(z) &= -\lambda_2 z^2 + (\lambda_2 + \rho(i) + i\mu_2)z - i\mu_2. \end{aligned}$$

As for the case $i = r_0$, we can derive $Q_{r_0}(z)$ in a similar manner and the expression is the same as (39) except that $\rho(r_0) = 0$ and $B(r_0) = 1$.

We can see that given $Q_{i-1}(z)$, there are three unknowns $P(i, 0)$, $P(i, K)$ and $Q_i(1)$ in (39) instead of two in the infinite waiting room case. We can calculate $P(i, 0)$ by applying (38) and then it leaves only two unknowns to determine. According to the definition of $Q_i(z)$, $Q_i(z)$ are polynomials of degree K and, therefore, they should be bounded in the complex plane $\{z, |z| < \infty\}$. It follows that all the zeros of $G_i(z)$ must also be zeros of $H_i(z)$.

It is easy to show that the denominator $G_i(z)$ has just two zeros and that both of them are real. Then we can determine $Q_i(1)$ and $P(i, K)$ for $1 \le i < r_0$ in a recursive manner as in the preceding section. Denoting the two zeros of $G_i(z)$ by $\xi_{i,1}$ and $\xi_{i,2}$, we then have

$$\begin{aligned} H_i(\xi_{i,1}) &= 0 \\ H_i(\xi_{i,2}) &= 0. \end{aligned} \tag{40}$$

This is a linear system of two equations from which we can calculate the two unknowns $P(i, K)$ and $Q_i(1)$. In the case where $i = r_0$, one of the two roots $\xi_{r_0,1}$ is 1 and the other $\xi_{r_0,2}$ equals $r_0\mu_2\lambda_2^{-1}$. Since we should determine $P(r_0, K)$, we can not use the root 1 which nullifies the term associated with $P(r_0, K)$. Therefore, we use the following equality:

$$H_{r_0}(\xi_{r_0,2}) = 0 \tag{41}$$

to calculate $P(r_0, K)$.

Differentiating (39) and evaluating at $z = 1$ for $1 \le i < r_0$, we then have the following equation:

$$\begin{aligned} &\rho(i)Q_i'(1) \\ &= (\lambda_2 - i\mu_2)Q_i(1) + \rho(i-1)Q_{i-1}'(1) \\ &\quad - \lambda_2(1-B(i))P(i,0) + \lambda_2(1-B(i-1)) \\ &\quad \cdot P(i-1,0) - \lambda_2 P(i,K). \end{aligned} \tag{42}$$

In the same manner, differentiating (39) twice and evaluating at $z = 1$, we obtain

$$\begin{aligned} \rho(i)Q_i^{(2)}(1) &= \rho(i-1)Q_{i-1}^{(2)}(1) + 2\lambda_2(Q_i(1) \\ &\quad - (1-B(i))P(i,0) - (K+1)P(i,K)) \\ &\quad - 2\{i\mu_2 + \rho(i-1) - \lambda_2\}Q_i'(1). \end{aligned} \tag{43}$$

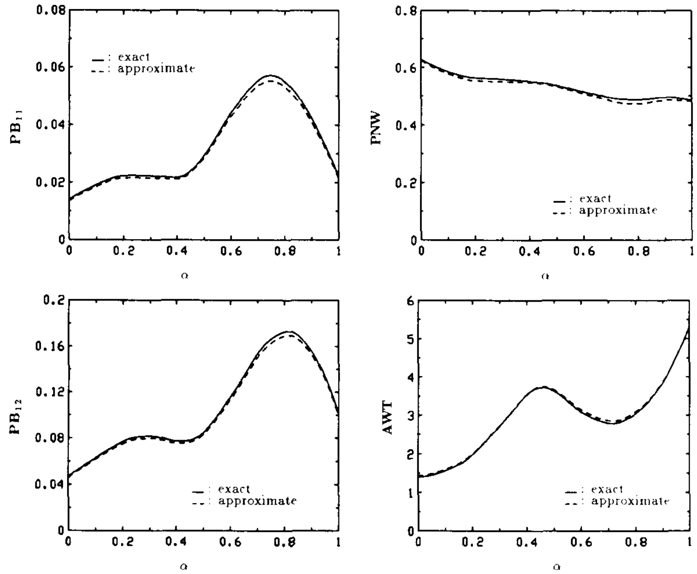

Fig. 2. Infinite waiting room case.

In the case where $i = r_0$, we have $\rho(i) = 0$. By using L'Hospital's rule, it is easy to obtain

$$Q_{r_0}(1) = \frac{\rho(r_0-1)Q'_{r_0-1}(1) + \lambda_2(1-B(r_0-1))P(r_0-1,\,0)}{r_0\mu_2 - \lambda_2}. \tag{44}$$

The first derivative of $Q_{r_0}(z)$ can be written

$$Q'_{r_0}(z) = \frac{d}{dz}\left\{\frac{H_{r_0}(z)}{G_{r_0}(z)}\right\} = \frac{H'_{r_0}(z)G_{r_0}(z) - G'_{r_0}(z)H_{r_0}(z)}{G_{r_0}(z)^2} \tag{45}$$

where

$$\begin{aligned} H_{r_0}(z) &= \rho(r_0-1)(Q_{r_0-1}(z) - Q_{r_0-1}(1)) \\ &\quad - \lambda_2(1-z)(1-B(r_0-1))P(r_0-1,\,0) \\ &\quad + (1-z)z^K P(r_0,\,K) \\ G_{r_0}(z) &= -\lambda_2 z^2 + (\lambda_2 + r_0\mu_2)z - r_0\mu_2. \end{aligned}$$

Using L'Hospital's rule, we therefore have

$$\begin{aligned} Q'_{r_0}(1) &= \left.\frac{H^{(2)}_{r_0}(z)G_{r_0}(z) - G^{(2)}_{r_0}(z)H_{r_0}(z)}{2G_{r_0}(z)G'_{r_0}(z)}\right|_{z=1} \\ &= \frac{1}{2G'_{r_0}(1)}\{H^{(2)}_{r_0}(1) - G^{(2)}_{r_0}(1)Q_{r_0}(1)\} \\ &= \frac{\rho(r_0-1)Q^{(2)}_{r_0-1}(1) - 2(K+1)\lambda_2 P(r_0,\,K) - 2\lambda_2 Q_{r_0}(1)}{2(r_0\mu_2 - \lambda_2)}. \end{aligned} \tag{46}$$

If we denote the average number of waiting customers by AWN, the following relation is evident:

$$\mathrm{AWN} = \sum_{i=0}^{r_0} Q'_i(1). \tag{47}$$

Letting PB_2 be the loss probability of WB customers, we have

$$PB_2 = \sum_{i=0}^{r_0} P(i,\,K). \tag{48}$$

We can, therefore, calculate the average waiting time by using Little's formula

$$\mathrm{AWT} = \frac{\mathrm{AWN}}{\lambda_2(1-PB_2)}. \tag{49}$$

The probability of nonwaiting is given by the following equation:

$$\mathrm{PNW} = \sum_{i=0}^{r_0-1} (1-B(i))Q_i(0). \tag{50}$$

V. Numerical Results and Discussion

In this section, the results of the approximate calculations are compared to the values determined from an exact analysis. The exact values are calculated by a method based on solving a set of balance equations for a finite waiting room system. The same technique was used by Gimpelson [3] to analyze a similar integrated services system but without any service protection mechanism. For the infinite waiting room case, in order that we can consider a finite waiting room system as approximately an infinite system, we choose the number of waiting places in such a way that the loss of wide-band traffic is negligible ($<$ 0.0001). For convenience, we define the following parameters:

1) direct narrow-band traffic: $A_{11} = \lambda_{11}/\mu_1$
2) overflow narrow-band traffic: $A_{12} = \lambda_{12}/\mu_1$
3) total narrow-band traffic: $A_1 = A_{11} + A_{12}$
4) wide-band traffic: $A_2 = N\lambda_2/\mu_2$
5) ratio of service times: $\beta = \mu_1/\mu_2$.

Figs. 2–3 show the corresponding approximate and exact

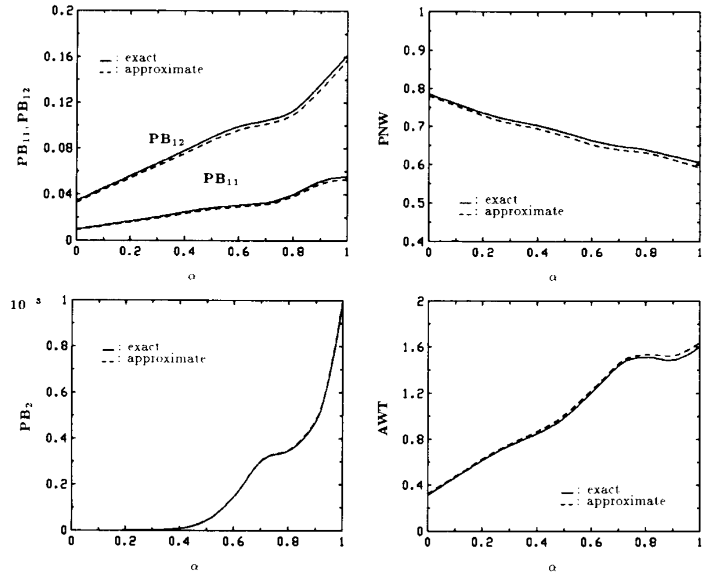

Fig. 3. Finite waiting room case.

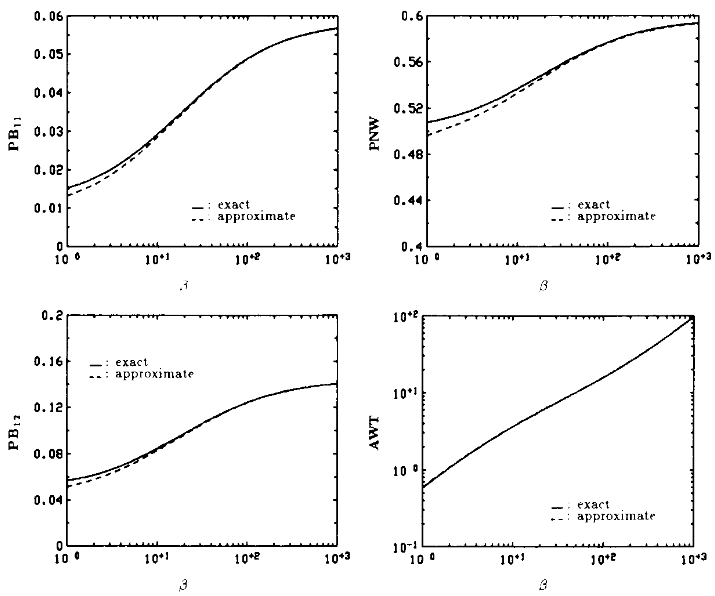

Fig. 4. Effect of β on the approximation.

curves for the blocking probabilities PB_{11}, PB_{12}, respectively, for direct and overflow NB traffics, the loss probability PB_2 (in the case of finite waiting room), the probability of nonwaiting PNW and the average waiting time AWT for WB traffic. The system parameters are 1) $M = 96$, $N = 24$, $r_0 = 3$, $r_1 = 94$, $r_2 = 0$, $\beta = 10$; 2) $M = 96$, $N = 16$, $K = 10$, $r_0 = 5$, $r_1 = 94$, $r_2 = 1$, $\beta = 10$. The offered traffics are $A_1 = 20 + 40(1 - \alpha)$, $A_{11} = 0.8A_1$, $A_{12} = 0.2A_1$, $A_2 = 10 + 40\alpha$. The approximate results are presented by dotted curves. In these figures, we calculate the system performance with the parameter μ_1 set to one. The unit of waiting time is therefore the mean service time of NB calls.

Since wide-band service times are generally longer than narrow-band service, we have assumed that narrow-band occupancy achieves the stationary distribution while the number of WB calls in service is fixed. It would be interesting to see the effect of the ratio of service times β on the proposed approximation. For the infinite waiting room case, Fig. 4 gives a comparison of the approximate and exact results as a function of β. In these graphs, we take the same system parameters as Fig. 2, but the offered traffics remain fixed: $A_{11} = 32$, $A_{12} = 8$, and $A_2 = 30$.

From these figures, we can see that the models give a very good approximation to the system performance. In the case of equal service rates $\mu_1 = \mu_2$, which is in contradiction with the assumptions, the approximation also gives extremely good results, except for PB_{11} and PB_{12} (the relative errors are, respectively, 0.134 and 0.089). From the cases we have tested, we noted that the relative errors concerning the probability of nonwaiting for WB traffic increase as the total traffic is raised (the relative error passes beyond 0.05 when the total traffic $>M$). However, throughout the useful range of

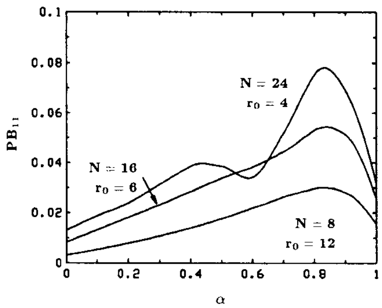

Fig. 5. Effect of N on narrow-band blocking.

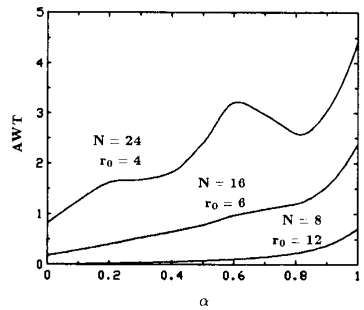

Fig. 6. Effect of N on the average waiting time.

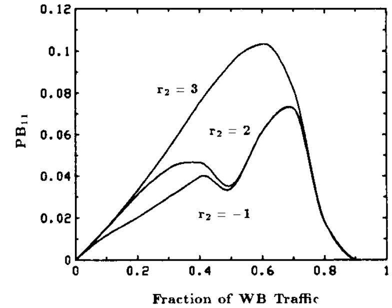

Fig. 7. Effect of r_2 on narrow-band blocking.

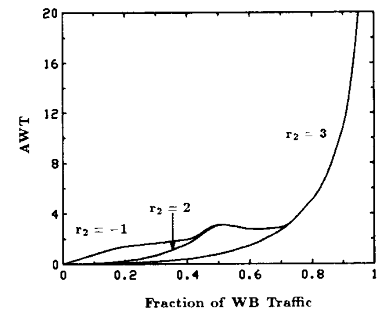

Fig. 8. Effect of r_2 on the average waiting time.

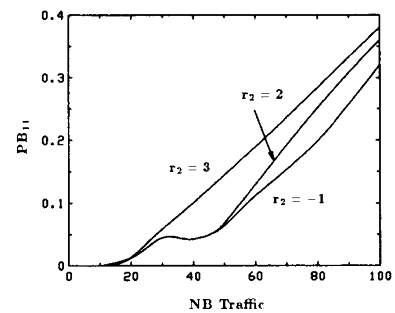

Fig. 9. Effect of r_2 on narrow-band blocking.

operation (the blocking probability of direct NB traffic is less than 0.2), the agreement is very good (relative error <0.05) for all the cases we have tested.

In the following, all the numerical results are calculated by the approximation. The systems we consider are with an infinite waiting room and $r_1 = M$. It follows that $PB_{11} = PB_{12}$ and $PB_2 = 0$.

The oscillatory variation of the blocking probability of NB traffic was discovered by Gimpelson [3]. In [2], De Serres and Mason pointed out that the oscillatory variation appears in both the blocking probability of NB traffic and the mean waiting time of WB traffic in a delay/loss model. The oscillations can be explained by the fact that a single NB call can keep a WB call out of service by effectively reserving a number of servers for narrow-band traffic. The degree of oscillation depends on the number of servers required by a WB call as shown by Figs. 5–6. In these two graphs, the total number of servers M is 120 and there is no protection of the wide-band traffic ($r_2 = -1$). (Here we use $r_2 = -1$ to indicate the case without wide-band protection; in fact, -1 has no meaning.) The offered traffics are $A_1 = 20 + 60(1 - \alpha)$, $A_2 = 10 + 60\alpha$ and the total traffic is 90. This characteristic feature of systems carrying mixtures of traffic with different bandwidth requirements can result in dramatic consequences since the relative percentage of traffic generally varies with time. It can also complicate dimensioning procedures.

At the cost of degradation of the grade of service of narrow-band traffic, we can remove this undesirable feature by using a large r_2. Figs. 7–8 present the curves of the blocking probability of the NB traffic and the average waiting time of WB traffic for different r_2. In the two graphs, we choose the system parameters: $M = 120$, $N = 24$, $r_0 = 4$ and the total offered traffic is 90. It is easy to understand that the protection of wide-band traffic prevents the case where a small number of NB calls keep a WB call out of service from lasting. As mentioned in Section II, the parameter r_2 protects the wide-band traffic against overload of the narrow-band traffic. Augmenting this parameter results in smaller wide-band delay and larger narrow-band blocking. In order to choose a suitable r_2, it is necessary to characterize the combined effect of blocking and delay. As in [2] and [6], the combined effect is measured by

$$PR = \frac{\lambda_{11}(1 - PB_{11}) + \lambda_{12}(1 - PB_{12})}{(\lambda_{11} + \lambda_{12})(1 + \mu_2 A\,WT)} \tag{51}$$

called power factor. Figs. 9–11 present, for different r_2, the curves of blocking probability of NB traffic, the average waiting time of WB traffic and the power as a function of the narrow-band traffic. The system parameters are $M = 120$, $N = 24$, $r_0 = 4$ and the wide-band traffic remains fixed $A_2 = 50$. From these figures, we can see that the parameter r_2 does protect WB traffic against NB traffic. If we measure the combined effect of blocking and delay by the power factor, without constraint concerning the narrow-band blocking, it seems to us that $r_2 = r_0 - 1$ is the best solution.

Fig. 12 presents the effect of the narrow-band protection parameter r_0 on the power factor as a function of wide-band traffic A_2. The NB traffic remains fixed $A_1 = 40$ and the system parameters are $M = 120$, $N = 24$, $r_2 = -1$, $\beta = 10$. We can see from this figure that reducing r_0 results in smaller power factor, which is the price paid to protect the narrow-

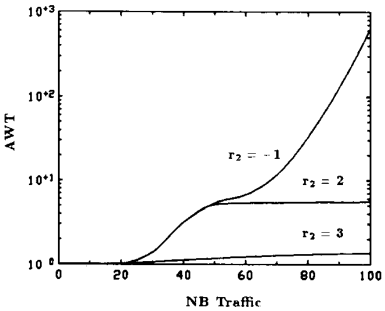

Fig. 10. Effect of r_2 on the average waiting time.

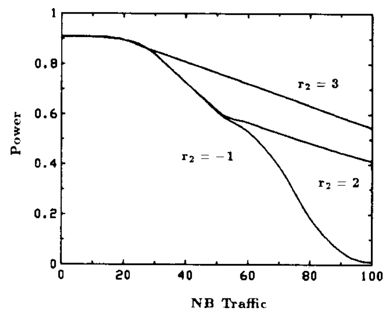

Fig. 11. Effect of r_2 on the power factor.

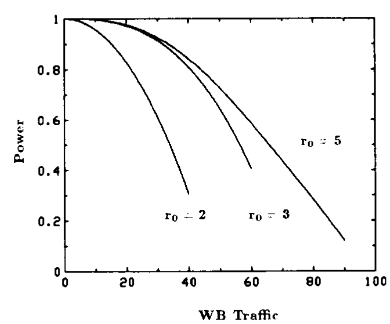

Fig. 12. Effect of r_0 on the power factor.

band traffic. We note that the protection mechanism of narrow-band traffic has always an effect on the wide-band traffic process since r_0 operates irrespective of the state of system. This is not the case for wide-band protection which operates as a function of the state of system (the queue is empty or not). This is why the protection of wide-band traffic results in larger power factor.

VI. Conclusion

An approximate performance model for an integrated service system in which some calls require concurrent service by more than one server has been presented. The model is based on two assumptions: 1) the narrow-band occupancy attains the stationary distribution for a fixed number of WB calls in service and 2) the exceptional service times are exponentially distributed. From the numerical results, we can see that the approximation gives extremely good results throughout the useful range of operation. Since there is no linear system to solve, this model is efficient in calculation speed and the memory space is no longer a limitation. These important improvements enable its use within network optimization algorithms.

The cutoff priority parameter r_0, which protects narrow-band traffic results in smaller power factor. The feasible range of choice for r_0 is limited by the requirement for system stability, $\lambda_2 < r_0\mu_2$. However, it is required to prevent long blocking periods and bistable behavior for some values of the system parameters [10]. The trunk reservation parameter r_1 will result in a smaller power factor for a single link, but it significantly improves the overall performance of an integrated service network employing alternative routing for NB traffic. The introduction of the wide-band traffic protection parameter r_2 protects WB traffic against overload of NB traffic and improves the overall system performance even in the case where NB traffic is not unusually heavy. We have also shown that using a large r_2 can remove the undesirable feature of oscillation.

The results to date indicate the importance of properly selecting the control parameters r_0, r_1, and r_2. We are currently investigating the problem of optimal control of this and other multiservice networks.

Acknowledgment

The authors wish to express their appreciation to Prof. M. Kaplan for his helpful suggestions and valuable comments on the model and to L. Fortier who wrote the program for calculating the exact results which were used to validate the approximate model. We also wish to thank the anonymous reviewers whose comments led to an improvement of our paper.

References

[1] J. M. Akinpelu, "The overload performance of engineered networks with nonhierarchical and hierarchical routing," ITC-10, pp. 3.2.4.1–3.2.4.7, June 1983.

[2] Y. De Serres and L. G. Mason, "A multi-server queue, with narrow- and wide-band customers and WB restricted access," *IEEE Trans. Commun.*, vol. 36, June 1988.

[3] L. A. Gimpelson, "Analysis of mixtures of wide- and narrow-band traffic," *IEEE Trans. Commun. Technol.*, vol. COM-30, pp. 258–266, Sept. 1965.

[4] A. Girard, "Blocking probability of noninteger groups with reservation," *IEEE Trans. Commun.*, vol. COM-33, pp. 113–120, Feb. 1985.

[5] L. Green, "A queuing system in which customers require a random number of servers," *Oper. Res.*, vol. 28, no. 6, pp. 1335–1346, Nov.–Dec. 1980.

[6] B. Kraimeche and M. Schwartz, "Analysis of traffic access control strategies in integrated service networks," *IEEE Trans. Commun.*, vol. COM-33, pp. 1085–1092, Oct. 1985.

[7] R. S. Krupp, "Stabilization of alternate routing networks," in *Proc. ICC'82*, 1982, pp. 31.2.1–31.2.5.

[8] L. Mason, Y. De Serres, and C. Meubus, "Circuit-switched multipoint service performance models," *Int. Teletraffic Congr.*, vol. ITC-11, pp. 2.1A-5-1–2.1A-5-6, Sept. 1985.

[9] L. G. Mason, "On the stability of circuit-switched networks with non-hierarchitcal routing," presented at Proc. 25th Conf. Decision Cont., Athens, Greece, Dec. 1986.

[10] L. G. Mason, K.-Q. Liao, L. Fortier, and Y. De Serres, "Performance models for a multi-slot system," *Proc. Int. Traffic Congr.*, vol. ITC-12, June 1988.

[11] Y. Nakagome and H. Mori, "Flexible routing in global communications networks," *Int. Teletraffic. Congr.*, vol. ITC-7, pp. 426.1–426.8, 1973.

[12] P. D. Welch, "On a generalized $M/G/1$ queuing process in which the first customer of each busy period receives exceptional service," *Oper. Res.*, vol. 12, no. 5, pp. 736–752, Sept.-Oct. 1964.

★

Ke-Qiang Liao received the B.Sc. degree in radio and electronic engineering from the South China Institute of Technology, Canton, China, in 1982 and a postgraduate diploma (DEA) in control theory and signal processing from the Université Paris-Sud, France, in 1983. In 1986, he received the Ph.D. degree from the Ecole Nationale Supérieure des Télécommunications, Paris, France, for his work done at the Centre National d'Etudes des Télécommunications on the modeling of prebooked telecommunications services.

He has been with the INRS-Télécommunications, Verdun, P.Q., Canada, since 1986. His current interests concern the performance evaluation and routing problems arising in integrated services networks.

Lorne G. Mason received the B.Sc. and Ph.D. degrees in mechanical engineering, control division, from the University of Saskatchewan, Saskatoon, Canada, in 1963 and 1973, respectively.

From 1963 to 1965 he was a Rocket Design Engineer with Bristol Aerospace, Winnipeg, Man., Canada. Since that time he has been involved in the areas of telecommunications and automatic control in both industrial and academic environments. He was employed with the B.C. Telephone Company, Vancouver, B.C., Canada, in the traffic studies and network planning areas. Following the completion of his postgraduate studies, he served as a consultant to Yale University, New Haven, CT, concerning the application of learning techniques to network management and control. From 1974 to 1977, he was employed by Bell-Northern Research, Ottawa, Ont., and Montreal, P.Q., Canada, where he was instrumental in developing a planning system for the introduction of digital transmission and switching equipment into telephone networks. In 1977 he joined INRS-Telecommunications, Montreal, where he pioneered a course on network planning which has subsequently been offered also at McGill University, Montreal. He currently holds the position of Full Professor of Telecommunications. His research interests involve the application of control and optimization techniques to network design and management.