1984

Models and techniques for the analysis of integrated voice/data transmission schemes /

Nick R. Mikroudis
Lehigh University

MODELS AND TECHNIQUES FOR THE ANALYSIS OF INTEGRATED VOICE/DATA TRANSMISSION SCHEMES

By

Nick K. Mikroudis

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Electrical Engineering

Lehigh University
December 1984

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

12/17/84

(date)

Professor in Charge

Chairman of Department

AKNOWLEDGEMENT

The valuable guidance and suggestions given by Professor Kenneth K. Tzeng, the supervisor of this thesis, are gratefully acknowledged. His encouragement and the instructive discussions with him were of invaluable assistance to the author.

The Computer Science and Electrical Engineering Department of Lehigh University is also gratefully acknowledged for providing a teaching assistanship to the author in supporting his pursuing of graduate studies.

Table of Contents

List of Figures

ABSTRACT

The main variations of an integrated voice/data multiplex structure are considered, namely, the Fixed Frame - Fixed boundary, the Fixed Frame - Movable boundary, and the Variable Frame - Movable boundary schemes, where voice traffic is circuit switched, and data traffic is packet switched.

The performance measures of the above schemes are the blocking probability for voice calls and the expected delay for data packets. A general method is formulated for the analytical evaluation of the above measures. The method is based on the use of Z-transform and is applied to a number of different queueing models which are used to describe the system. A discrete time model, a simplified discrete time model, and a continuous time model are used for the analysis of the fixed frame (with fixed or movable boundary) scheme while a discrete time model and a round robin processor sharing scheme are used for the analysis of the variable frame (movable boundary) scheme.

The results of the above analyses are presented and justified, and based on them, the performance of the

various integrated schemes is compared. The Fixed Frame - Movable boundary scheme provides lower delays for the data customers and performs better the Fixed Frame - Fixed Boundary scheme. The Variable Frame - Movable Boundary scheme provides even lower delays for the data customers. Therefore it exhibits the best performance, but at the same time it increases the hardware and software complexity of the system.

Finally, the presented analytical models and techniques are compared and criticized in order to decide which is the more appropriate and more accurate one tor the analysis of each integrated transmission scheme.

1. INTRODUCTION

1.1 Switching techniques in computer networks

In computer networks, the users and computers communicate with each other by interchanging messages. There are two major techniques for interconnecting computers and interchanging messages: a) Circuit switching and b) Packet switching.

A. Circuit Switching

In circuit switching, a complete circuit or path of connected lines is established from the origin to the destination of the "call". After the path is set up, a return signal informs the source that data transmission may proceed and then communication begins. The circuits or channels which are used by this path cannot be accessed by other users until the termination of the connection. Circuit switching was developed in connection with telephone systems and therefore, in principle, it is nearly 100 years old.

In computer networks, the user dials a sequence of digits to obtain access to another user or computer system. In some networks the required path is established automatically, according to information in the data stream. Circuit switching systems are subject to a path setup delay because some of the lines in the path may be busy and because the switches that connect the lines are speed limited. However, once a path is set up, the transmission is synchronous and "transparent" and the message is not interrupted or buffered anywhere along its path towards the destination node. Actually, because of this property, circuit switching is a favorable technique for the transmission of voice.

Although circuit switching has been considered inefficient and economically not very practical in computer networks (because it requires dedicated use of the links), it is extensively used in some networks like TYMNET and DATRAN.

B. Packet Switching.

In packet switching, each message is subdivided into blocks, called packets. Each packet is directed across the network independently of the others. Each one of them includes information about the message where it belongs, about its destination, about the total number of packets in the message, and also a serial number which distinguishes it from other packets and allows the destination node to "reassemble" the message correcly (even if the packets are not received in the right order).

Through a complex evaluation of individual circuit and switch loads, packet switching computers dynamically alter routing of packets to the circuit of minimum loading. Each packet is switched by means of a store and forward mechanism, and therefore may experience some delay along its route towards the destination node. If a packet is not received correctly by some node, then a request for retransmission is sent back to the source node. When the destination node receives all the packets of the message without errors, then it sends a special message (one packet long) back, indicating that the message was received correctly. If the destination node does not receive all the packets within a certain period of time, then it requests the retransmission of the missing packets (assuming that they were lost during the transmission, because of noise or malfunction).

Since it allows independent routing of each packet,

packet switching offers the capability of adapting the function of the network dynamically according to the variation of traffic conditions. For the same reason, the availability of the network is very high (that is the probability that the network functions correctly for a certain period of time T). Furthermore, since many packets of the same message may be in transmission simultaneously, the transmission delay is reduced by a factor which is proportional to the number of packets into which the message is broken. Because of these advantages, the use of packet switching may reduce the cost of a network. However at the same time it has the disadvantage of adding an "overhead" of extra bits on every transmitted packet. The technique of packet switching was conceived less than 10 years ago and is used in most of the existing networks (like ARPA and MERIT).

The advantages and disadvantages of using packet switching versus circuit switching have been extensively studied in recent years [1]. However, the choice of the most effective switching technique is not an easy one, and depends on the following factors: The average length of the messages, the average time period between the transmitted messages, and the average duration of the connection between the communicating nodes. In general there is a practical roule, that circuit switching is more cost-effective for traffic characterized by lengthy messages (like digitized voice), while packet switching is more cost-effective for traffic with short messages (like interactive data). For usual networks which carry both short and long messages, it seems that packet switching performs slightly better than circuit switching.

An alternate solution to the problem of choosing the right switching technique is to allow and combine both switching techniques in the same, "integrated" network. In such a network every class of traffic is served in the most effective way: circuit switching is used for the transmission of digitized voice and other lengthy data messages, while packet switching is used for the transmission of interactive data and other short data messages. This approach brings up the concept of integrated communications which is discussed in the following section.

1.2 The concept of integrated communications

Voice and data are two major classes of traffic involved in digital communications. So far, each one of them used to be accomodated by separate transmission and switching facilities, within the communications network. The desire of providing common switching and transmission facilities, that would be shared by both traffic classes, gave rise to the concept of integrated voice/data communications. Therefore, an integrated voice/data communication structure includes two schemes of integration: a) Transmission, and b) switching.

a) In an integrated transmission scheme, the transmission link is used alternatively by each class of traffic. By means of Time Division Multiplexing, the link is dedicated part of the time to voice traffic, and another part of the time to data traffic. The exact assignment of this time sharing depends on the specific strategy of integration which is used.

b) In an integrated switching scheme, a single switching facility provides both circuit, and packet switched modes of operation: voice traffic is circuit switched (synchronous transmission), while data traffic is packet switched (asynchronous transmission). This arrangement combines the transparency of circuit switching for voice calls, with the efficient utilization of line capacity for data traffic. From now on, we will be using the terms integrated or hybrid interchangeably, without discrimination.

There are several reasons that justify the realization of an integrated communications system, instead of two separate voice, and data systems: The above integration will improve the transmission efficiency, by increasing the utilization of the link, and by reducing the number of required transmission channels. Furthermore, the transmission bandwidth will be used more effectively, since each class of traffic will be using the switching concept which is best suited to it.

So far, no true integrated transmission facilities are produced for commercial use. However, many hybrid arrangements are under extensive study and experiment. In the following sections we will present some of the most important hybrid transmission schemes.

1.3 The structure for integrated transmission.

A hybrid transmission link is a digital Time Division Multiplex structure, which enables dynamic sharing of the channel bandwidth between voice, and data traffic. Voice traffic is using the circuit switched mode, and data traffic the packet switched mode of operation (Figure 1.1).

The channel is synchronously clocked, and therefore partitioned into "frames" of fixed duration (say b seconds). Within each frame there is a boundary, dividing it into two distinct regions: One dedicated to voice traffic, the other to data traffic (Figure 1.2). The frame is further decomposed into time slots of equal duration (τ seconds each); the data region of the frame consists of N slots of equal duration (τ' seconds each). Of course, we have $b=S\tau+N\tau'$. If the transmission rate over the hybrid link is r bits/sec, then each voice slot contains n bits ($n=\tau.r$), and each data slot contains n' bits ($n'=\tau'.r$). That is, each voice customer may transmit n bits per frame, and each data customer n' bits per frame. There will be totally $n_0=b.r=Sn+Nn'$ bits, in each frame.

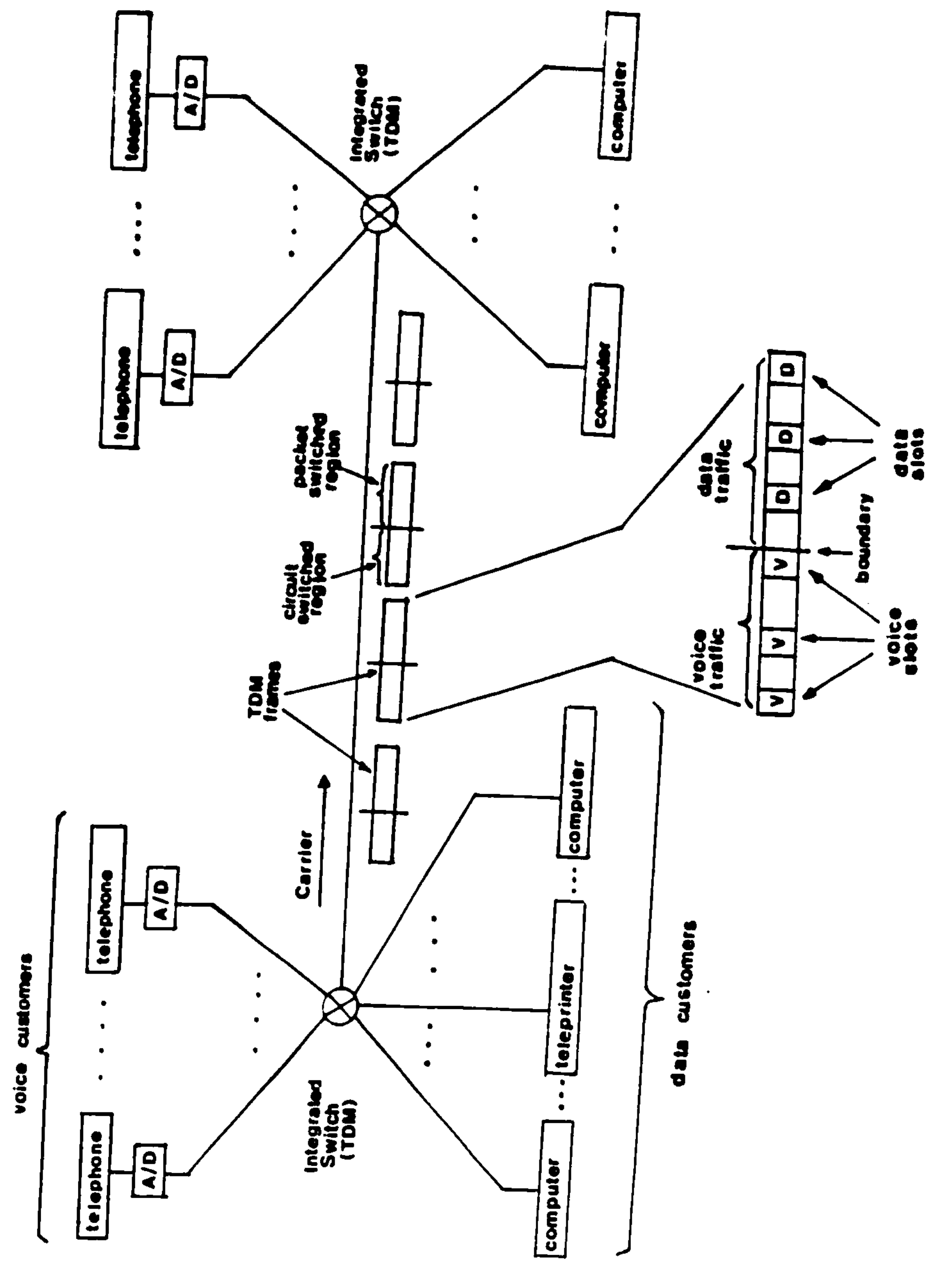

Figure 1-1: The structure of an integrated transmission link.

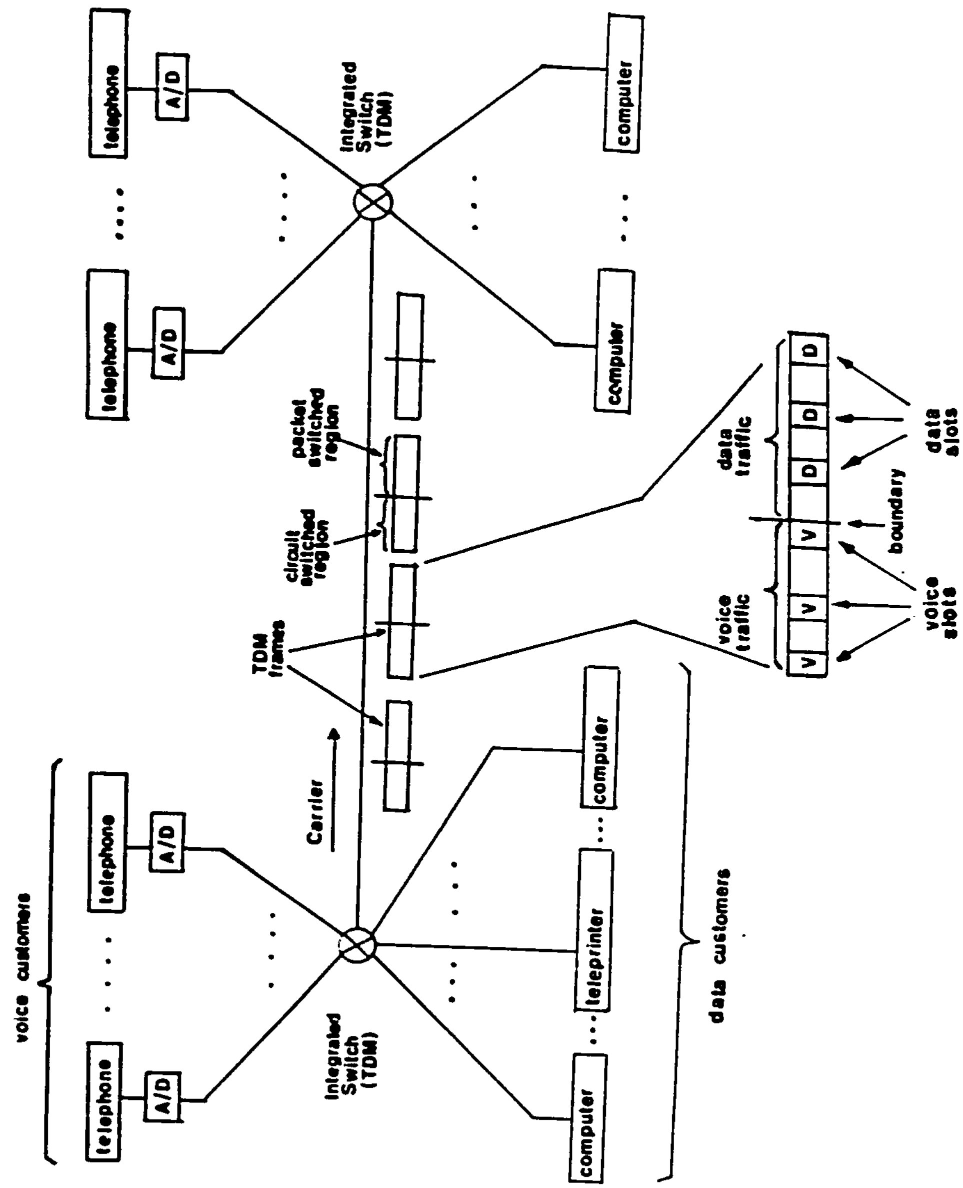

Figure 1-1: The structure of an integrated transmission link.

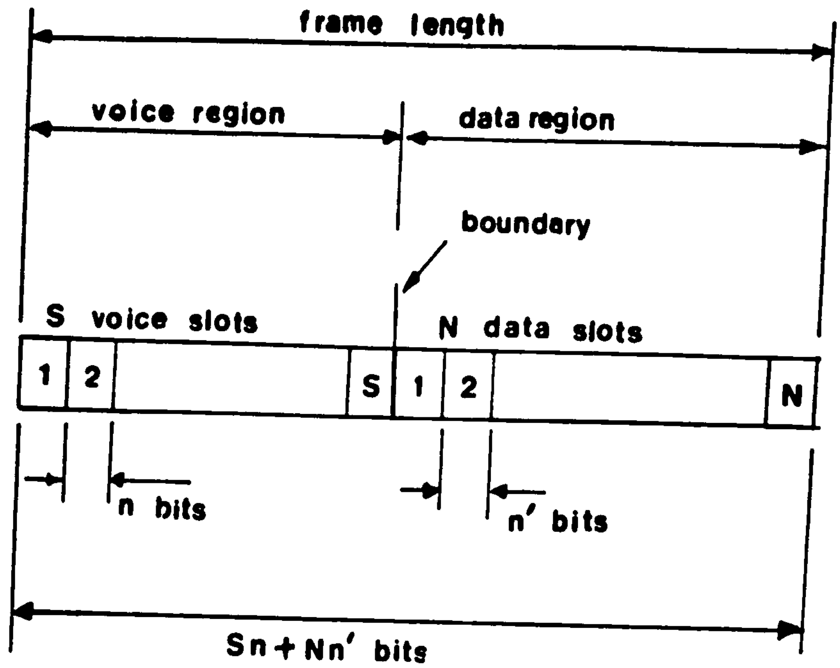

Figure 1-2:
Illustration of the frame structure

The choice of the slot size depends on the digitization rate of the voice traffic (say R bits/sec), and the frame duration (b sec). This happens because of our requirement to preserve the transparency of voice traffic over the hybrid link. The bits of digitized voice which are generated during one frame period must be accomodated in one slot; therefore the size of a voice slot must be n=R.b bits.

The data customers transmit data in the form of packets of n' bits. Each customer can occupy one data slot (equivalent to one packet) per frame period. Since data traffic is packet switched, the packet format contains information necessary for routing, security, identity, and precedence of the transmitted packets.

The structure that we described above is the so called SENET concept for integrated voice/data transmission, which was introduced in [2], and is illustrated in the following example.

Example:

Suppose that b=10 ms, and r=1.544 Mbits/sec (as for a T1 carrier). Therefore each frame contains n_0=15440 bits, and 100 such frames are transmitted per second. Now if the rate of the digitized voice is R=8 Kbits/sec, it would require an assignement of n=bR=80 bits per frame, and that would be the size of a voice slot. Therefore if a certain voice call lasts 5 min, it requires a reservation of one voice slot per frame, over 30000 consecutive frames (300 sec X 100 frames/sec). Now suppose that the frame boundary is located at the bit position 4240 within each frame, and let the data slot size be n'=800 (one data slot is equivalent to ten voice slots). Therefore each frame will contain S=4240/80=53 voice slots, and N=(15440-4240)/800=14 data slots.

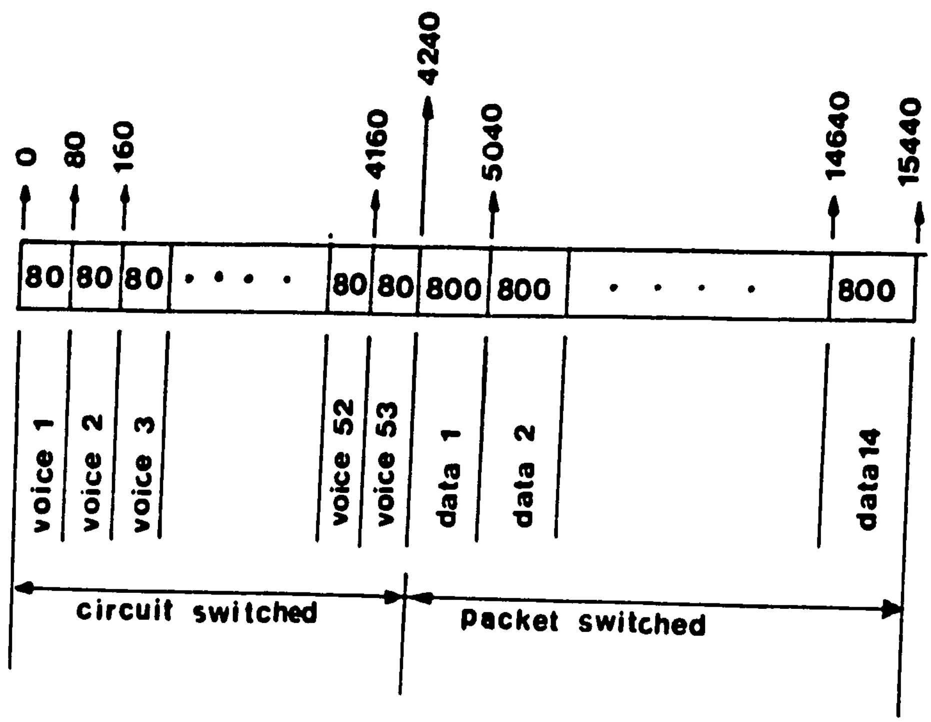

Figure 1-3:
Example of an integrated TDM frame.

At this point we must mention that some of the first voice slots in each frame (for example the first three slots) are always reserved for the Common Channel Interconnect Signaling (CCIS); CCIS actually takes care of the assignment and delivery of the subsequent voice slots. From now on, we will be neglecting the CCIS bits, for reasons of simplicity.

is equivalent to ten voice slots). Therefore each frame will contain S=4240/80=53 voice slots, and N=(15440-4240)/800=14 data slots.

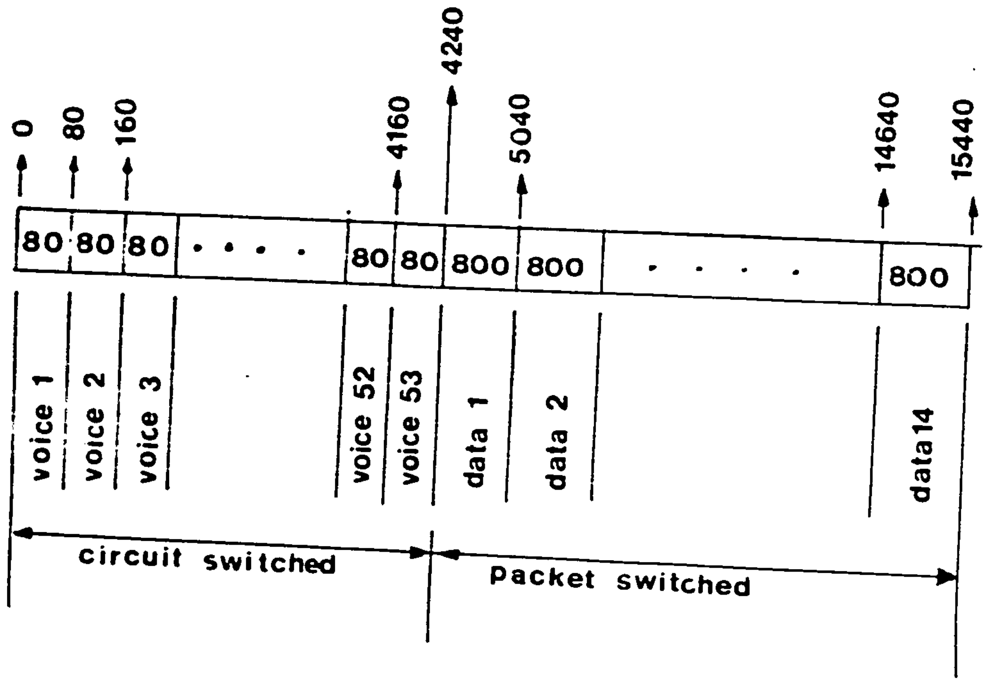

Figure 1-3:
Example of an integrated TDM frame.

At this point we must mention that some of the first voice slots in each frame (for example the first three slots) are always reserved for the Common Channel Interconnect Signaling (CCIS); CCIS actually takes care of the assignment and delivery of the subsequent voice slots. From now on, we will be neglecting the CCIS bits, for reasons of simplicity.

As we said, in a hybrid transmission scheme, each frame can accomodate up to S voice customers, and up to N data customers. Therefore, each voice or data customer arriving at the transmition link, may or may not find an available slot for transmission. In the latter case, data packets are buffered, and wait until they find some available slot within the subsequent frames, and then they are transmitted. Voice calls however are not buffered, and are simply "blocked", and "lost" if they don"t find any available slots for transmission.

More specifically, if a voice call request arrives at an arbitrary instant in the current frame, it is buffered until the start of the next frame (that is at most for b secs). In the meantime, the switch CPU determines if the occupied slots are less than S. If so, one voice slot is reserved (on a first come - first served basis) for the call, and it is transmitted within the next frame. The slot is reserved indefinitely for this call during succeeding frames, until the termination of the connection. If however all of the S voice slots are occupied, then the call is "blocked", and "lost". Let us call PL the probability that a voice customer is "lost". In the same way, an arriving data packet is buffered until the start of the next frame. If the occupied data slots are less than N, then one data slot is assigned to the packet, and it is transmitted within the next frame. If however all of the N data slots are occupied, then the packet is buffered indefinitely, until a slot becomes available for transmission. Therefore, data packets may wait for several frames in a queue, before they are transmitted. Let us call ED the mean delay of a data packet, until it receives service. In Figure 1.4 we summarize the most important properties of a hybrid transmission scheme, for the two traffic classes.

	slots per frame	switching mode	bits per slot	holding time	if not transmitted	prob. of loss	mean delay
VOICE	S	circuit switching	$n = \tau r$	until call termin.	lost	PL	0
DATA	N	packet switching	$n' = \tau' r$	b	buffered	0	ED

Figure 1-4:
Properties of an integrated voice/data link.

The capacity of the voice and data regions of each frame depends on two factors: a) The frame length, and b) the boundary position:

a) The length of the frame (b) determines the total number of bits in it, and therefore the total number of available slots per frame, for voice, and data traffic together (S+N). If b is fixed, then (S+N) will also be fixed, otherwise it will be variable.

b) The position of the boundary within the frame determines the values of S and N, that is the way of sharing the frame capacity between the two traffic classes. If the boundary is fixed, then S and N will keep the same constant values, from frame to frame. If it is movable, then their values may vary from frame to frame.

In the following sections, we will present, and analyze the following hybrid switching arrangements:

- The Fixed Frame - Fixed Boundary (FFFB) scheme.
- The Fixed Frame - Movable Boundary (FFMB) scheme.
- The Variable Frame - Movable Boundary (VFMB) scheme.

2. THE SCHEMES FOR INTEGRATED TRANSMISSION

2.1 The Fixed Frame scheme

A. Fixed Boundary

This structure utilizes a master frame of fixed length (N+S slots), with a fixed boundary. Therefore, fixed portions of the transmission capacity are assigned to the voice, and data traffic: The first S slots of each frame are reserved for voice customers, and the rest N slots for data customers. Each traffic class is not allowed to use any capacity that has not been reserved for it, even if the other class is not using it; actually, this is the chief disadvantage of this structure.

Let's call v the number of voice customers which are present just before the frame transmission, and d the number of data customers (in the frame, and in the buffer). If $v<S$, then the transmitted frame will contain $S-v$ idle voice slots. Furthermore, if $d<N$, it will also contain $N-d$ idle data slots. In Figure 2.1 we illustrate this arrangement, and we use the following notation:

$a^{+}=\max(a,0)$

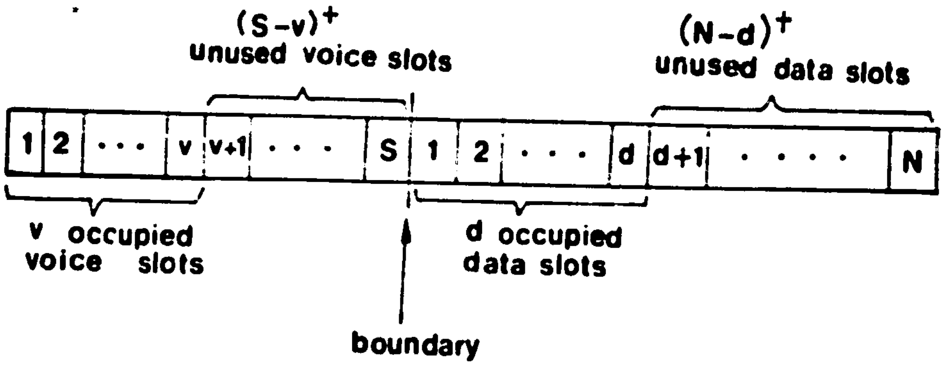

Figure 2-1:
Frame transmission, using the FFFB scheme.

B. Movable Boundary

Here, again the first S slots of the fixed - length frame are assigned to the voice traffic, and the remaining N slots to the data traffic. However, if less than S voice customers are present (say v), then they occupy the first v slots of the frame, and the boundary is positioned right after the v-th slot, assigning the rest of the frame to data customers. Therefore, data customers are allowed to occupy the unused voice slots, and N+(S-v) slots are available for them (instead of N). Although the above arrangement is more complicated, it requires less overall capacity, compared to the Fixed Boundary scheme. However, both schemes will perform in the same way in the case of heavy voice traffic (where all voice slots are occupied), or in the case of low data traffic (when there are not enough data customers to utilize the available capacity). In general, a transmitted frame will contain no iddle voice slots, and $[N+(S-v)^{+}-d]$ idle data slots. Figure 2.2 is an illustration of the above arrangement. In the above discussion we have assumed equal voice and data slot sizes (¶=¶'). However, if ¶'=a¶, then instead of (S-v) we will have [(S-v)/a] extra data slots (where [x] denotes the greatest integer less than or equal to x).

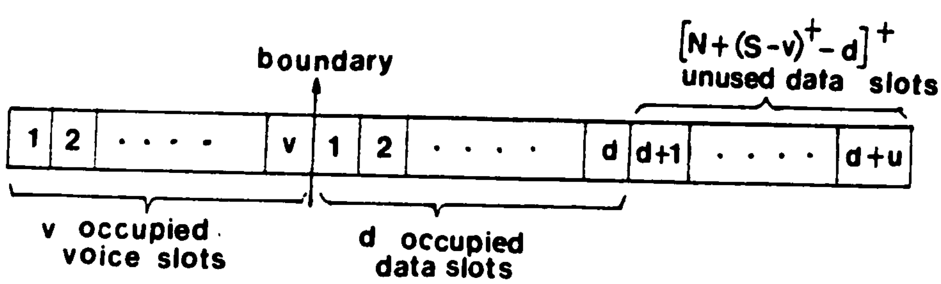

Figure 2-2:
Frame transmission, using the FFMB scheme.

The criteria for the avaluation of the performance of the Fixed Frame structure (for fixed or movable boundary), will be:

a) The probability of "loss", PL, for the voice customers.

b) The expected waiting time, EW, for the data customers. It will be:

EW=ES+ED

where ES is the expected service time of data packets, due to packet processing, and ED is the expected delay time of data packets, due to buffering.

2.2 The Variable Frame scheme

In this scheme the frame length is not fixed, but varies between some minimum and maximum value, according to the variations of the existing traffic. The maximum number of voice customers that can be accomodated within a frame is S, and the maximum number of data customers is N.

Let's assume that v voice customers, and d data customers are present just before the frame transmission. The v voice customers are accomodated in the first v slots of the frame. If v>S, then S of them are accomodated in the frame, and the rest (S-v) are lost. Now, the boundary is positioned right after the v-th slot, and the data packets are accomodated in the next d slots of the frame; then the frame is terminated right after the d-th data slot. If d>N, then only the first N packets are transmitted, and the rest d-N are buffered. If d<N, then all of the packets are transmitted, but still no part of the frame capacity remains idle, because the frame ends right after the last data packet (see Figure 2.3).

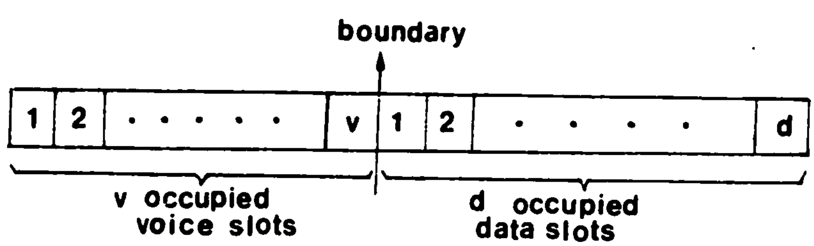

Figure 2-3:
Frame transmission, using the VFMB scheme.

Although the above arrangement involves considerable hardware complexity, it increases the transmission efficiency of the link, by reducing the idle frame capacity, and improves the packet transmission delay. However, since the frame period is not constant any more, the voice customers will have a problem of synchronization from frame to frame, with risk of losing the voice transparency through the transmission. However, we may take care of this problem, by choosing a certain maximum value for the frame size, and even by allowing some buffering of the voice for limited time intevals, without affecting the quality, and intelligibility of the connection. In Chapter 5, we will discuss more about how to preserve the transparency of the voice transnission.

The criteria for the performance evaluation of the Variable Frame structure will be:

a) The loss probability, PL, for the voice customers; besides this, the expected transmission time for an average voice message, ETR, which depends on the average frame duration.

b) The expected waiting time, EW, for the data customers, which will be again:

EW=ES+ED

as in the Fixed Frame scheme.

In the Figure 2.4, we summarize the properties of the described integrated transmission schemes, and in the next chapter we will derive analytical expressions for the evaluation of their performance characteristics.

	frame duration	available slots		unused slots		performance criteria	
		VOICE	DATA	VOICE	DATA	VOICE	DATA
FFFB	$(S+N)\tau$	S	N	$(S-v)^+$	$(N-d)^+$	PL	EW
FFMB	$(S+N)\tau$	S	$N+(S-v)$	0	$[N+(S-v)^+-d]^+$	PL	EW
VFMB	$(v+d)\tau$	S	N	0	0	PL,ETR	EW

Figure 2-4:
Properties of the various integrated transmission schemes.

3. A GENERAL TECHNIQUE FOR THE PERFORMANCE ANALYSIS OF THE SCHEMES

3.1 Description of Models

In order to evaluate the performance of a hybrid transmission scheme, it is always necessary to introduce a specific model that describes the function of the hybrid link; then, according to that model, we try to evaluate the performance characteristics of the system.

The multiplex structure of a hybrid switching scheme can be modelled in general, as a so called "queueing gate" (see Figure 3.1). The gate is a queueing system with two types of arrivals, and an operating rule, that allows the customers access to the system only at fixed intervals of time. There is some flexibility about the detailed characteristics of the above model. More specifically, within the modell definition, the following characteristics must be specified:

- The period of the gate opening, τ_g. For example, τ_g may be equal to the frame period.
- The probability distribution f_A^V (f_A) that describes the voice (data) arrival process. For

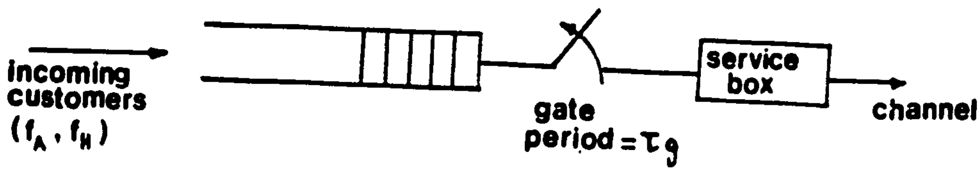

Figure 3-1:
Illustration of a queueing gate.

example, we may assume independent Poisson arrivals, for voice and data customers.

- The probability distribution f_H^V (f_H), that describes the holding times for voice (data) customers, that is, for how long a certain customer occupies a slot that has been assigned to him. For example, we may assume exponential holding times, for voice and data customers.

- The voice, and data population characteristics. For example, we may assume infinite voice and data population.

- The capacity of the existing buffer (usually for data packets). For example, we may assume infinite buffer capacity.

- The correlation assumptions, about the considered random variables. For example, we may assume that the number of data customers in the system is independent of the number of occupied voice slots.

Now we will present a general method, which can be used for the evaluation of the performance characteristics of a hybrid link. The main technique involved in that method, is the use of Z-transform.

3.2 Analysis for Voice Traffic

First we define the model that describes the switching of voice traffic. By using this model, we will evaluate the probability (PL) that a new arriving customer will be "blocked", just after the next gate opening. Now let's call

p^v_i: Probability of i busy voice slots, just after the gate opens. p^v_i, with i=0,1,...,S are the state probabilities of a finite state Markov chain, with probability transition matrix P_{ij}^v, where

P_{ij}^v: Probability of j busy voice slots just after the gate opens, under the condition that i voice slots were busy just after the last opening. The following will hold:

$$\sum_{i=0}^{S} p_i^v = 1$$

$$p_j^v = \sum_{i=0}^{S} p_i^v P_{ij}^v \quad j=0,1,\ldots,S$$

Notice that if the frame duration b is finite, the average number of voice customers arriving between two successive gate openings is finite, and those customers who do not find an empty slot are rejected from the system. Therefore, it is intuitively obvious that $p^V{}_i$ for i=0,1,...,S exists. Now, let's call

$n^V{}_i$: Probability of i voice arrivals between two successive gate openings. Of course $n^V{}_i$ depends on the voice arrivals distribution.

The next step is the evaluation of the transition probabilities $P^V{}_{ij}$, by using the probabilities $n^V{}_i$, and the holding time distribution of the voice customers.

Then we use $P^V{}_{ij}$, to evaluate $p^V{}_i$ for i=0,1,...,S. We can to this by solving the linear system of equations (E).

Now we can use the probabilities $p^V{}_i$ and $n^V{}_i$, in order to determine the blocking probability PL.

We summarize the steps involved in the performance evaluation of the voice traffic:

1. Define the model for the voice traffic.
2. Evaluate the transition probabilities $P^V{}_{ij}$, by using the distribution of the voice arrivals, and voice holding times.
3. Evaluate the state probabilities $p^V{}_i$, by using the transition probabilities.
4. Evaluate the blocking probability PL, by using the state probabilities, and the arrival distribution.

3.3 Analysis for Data Traffic

First we define the model that describes the data traffic transmission it can be the same or different than the voice traffic model. By using this model, we will evaluate the expected waiting time EW of a data customer (packet), until its transmisson. Now let's call

$p_{i,n}$: Probability of i data packets in the system (on the channels and inthe buffer), just after the n-th gate opening

Since the data queue achieves equilibrum, we have asymptotically (for large n):

$$p_{i,n}=p_{i,n-1}=p_i$$

If the frame duration b is finite, and the expected number of data arrivals during one frame is less than the system capacity available to the data customers, then p_i for i=0,1,2,.... should exist. Now let's call:

n_i: Probability of i data arrivals, between two successive openings. Of course, n_i depends on the data arrivals distribution.

The next step is the evaluation of the probabilities $p_{i,n}$ in terms of $p_{i,n-1}$, of n_1 and in terms of the holding times distribution of the data packets.

Then we use $p_{i,n}$ to determine the steady - state moment generating function P(z) for the number of data packets in the system. It will be:

$$P(z)=\sum_{i=0}^{\infty} p_i z^i \qquad (3\text{-}1)$$

with $p_i=p_{i,n}=p_{i,n-1}$ (asymptotically, for large n)

Now that we know the generating function, it is easy to evaluate the mean number of data packets in the system. It will be:

$$L=\left.\frac{dP(z)}{dz}\right|_{z=1} \qquad (3\text{-}2)$$

Then, by using Little's formula, we derive the expected delay of data packets, ED:

$$ED=●^{-1}.L \qquad (3\text{-}3)$$

where ● is the average data customer arrival rate (and is determined by the arrivals distribution)

Finally, we evaluate the expected waiting time, EW, from the expression

$$EW=ED+ES \qquad (3\text{-}4)$$

where ES is the expected service time, which is considered to be equal to half of the frame duration: ES=b/2 .

It would be useful to explain a little more about the derivation of tne generating function, from the probabilities $p_{i,n}$: In general, these probabilities will have the form:

$p_{i,n}=f(i,n)$ for $i=0,1,2,\ldots$ and asymptotically, for large n,

$$p_i=f(i)$$

If we multiply the i-th equation by z^i, and then sum all these equations for $i=0,1,2,\ldots$ then according to (3-1) we will get an expression for the generating function:

$$P(z)=\sum_{i=0}^{\infty} p_i z^i=\sum_{i=0}^{\infty} f(i) z^i$$

Now in the right hand side of this equation we attempt to identify P(z), and the Z-transform of n_i, which is $n(z)=\sum_{i=0}^{\infty} n_i z^i$

If all but a finite number of terms for these transforms are present, then we add the missing terms, to get the desired forms for P(z), N(z), and then explicitly substract them out in the equation. So we end up with an equation of the form:

$$A(z)P(z)=B(z)$$

Now we can solve for P(z):

$$P(z)=B(z)/A(z) \qquad (3-5)$$

This final form of P(z) contains a number of unknowns, which can be eliminated by the following ways:

- Using the condition p(1)=1 ; actually, this is normalizing condition _ p_i=p(1)=1). Notice that at this point we might need to use L' Hospital's rule. Using the probabilities p_i , with i=0,1,...,S-1

- Using the analyticity of the Z-transform, we observe that in the region of analyticity, the transform must have a zero cancel each pole (singularity), in order to remain bounded. Therefore, in the expression (3-5) the roots of B(z) must be equal to the roots of A(Z), for z <1.

Now we summarize the steps involved in the performance evaluation for data traffic:

1. Define the model for the data traffic.

2. Evaluate the probabilities $p_{i,n}=f(i,n)$. Set $p_{i,n}=p_{i,n-1}=p_i$

3. Multiply the i-th equation by z^i

4. Sum all these equations, for $i=0,1,2,...$

5. In the resulting single equation, attempt to identify the Z-transforms P(z) and N(z). If needed, add and substract the appropriate missing terms.

6. Solve for P(z), in the resulting algebraic equation.

7. Use the condition $p(1)=1$, to eliminate the unknown terms. Use the pole - zero cancellation of P(z) to remove any remaining unknowns.

8. Evaluate the expected number of data packets in the system:

$$L=\left.\frac{dP(z)}{dz}\right|_{z=1}$$

9. Evaluate the expected delay of data packets

   ED=[illegible]$^{-1}$L

10. Evaluate the expected waiting time: EW=ES+ED, where ES is half of the (average) frame duration.

In the following sections, different models are used to describe the hybrid link; for each specific model, we apply the above metnod in order to evaluate the performance measures of the link.

4. ANALYSIS OF THE FIXED FRAME SCHEME

Here we consider an integrated voice/data multiplex structure, which utilizes the fixed frame transmission scheme, with fixed or movable boundary. We assume the following:

- FIXED BOUNDARY CASE:

  Duration of one voice slot: τ sec

  Duration of one data slot: τ sec

  Available voice slots per frame: S

  Available data slots per frame: N

  Frame duration: $(N+S)\tau=b$ sec (constant)

- MOVABLE BOUNDARY CASE

  Duration of one voice slot: τ sec

  Duration of one data slot: τ sec

  Available voice slots per frame: S

  Available data slots per frame: $N+(S-v)^{+}$

  Frame duration: $(N+S)\tau=b$ sec (constant)

where v is the number of occupied voice slots, in the transmitted frame.

4.1 Discrete Time Model

In this section we analyze the performance of the fixed frame scheme, for the cases of fixed, and movable boundary. Both cases are studied by using the model that was introduced by Fischer et al [3].

A. Model for Voice and Data Traffic

The switching process for voice, and data traffic is modelled as a discrete time queueing gate, with the following characteristcs (see Figure 4.1):

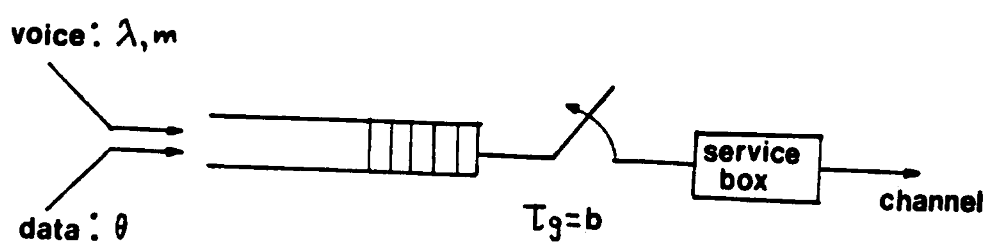

Figure 4-1:
Discrete time model, for voice, and data traffic.

- Period of gate opening: One frame period (τ_g=b sec)

- Voice arrivals distribution: Poisson, with arrival rate λ.

  Probability of i voice arrivals, between two successive openings:

$$n^{v}_{i}=\frac{e^{-\lambda b}(\lambda b)^{i}}{i!} \quad \text{for } i=0,1,2,...$$

$$n^{v}_{i}=0 \quad \text{otherwise}$$

- Data arrivals distribution: Poisson, with arrival rate $\bullet$.

  Probability of i data customers between two successive openings:

$$n_{i}=\frac{e^{-\bullet b}(\bullet b)^{i}}{i!} \quad \text{for } i=0,1,...$$

$$n_{i}=0 \quad \text{otherwise}$$

- Voice customers holding time distribution: Exponential, with mean m^{-1}.
  Probability that an occupied voice slot will still be occupied, after b secs:
  $p^{v}_{H}=e^{-mb}$

- Data customers holding time distribution: Deterministic.
  Probability that an occupied data slot will still be occupied, after x sec:

$p_H=1$ for $x<b$

$p_H=0$ for $x \geq b$

- Voice population: Infinite

- Data population: Infinite

- Data buffer capacity: Infinite

- Correlation assumptions: We assume that the number of data customers in the system is independent of the number of occupied voice slots. This is true, for the case of fixed boundary. For the movable boundary case however, although this assumption simplifies the performance analysis, it is somehow optimistic, especially under heavy traffic conditions.

B. Fixed Boundary

VOICE TRAFFIC

As we know, we have:

n^V_i: Probability of i voice arrivals, between two openings.

p^V_i: Probability of i busy voice slots, just before the gate opens.

P^V_{ij}: Probability of j busy voice slots just before the gate opens, given that i slots were busy just before the last opening. Now let's call:

q^V_{ij}: Probability of j busy voice slots just before the gate opens, given that i slots were busy, just after the last opening (or probability of i-j voice connection terminations, between two successive openings). It will be:

$$q^V_{ij} = (p^V_H)^j (1 - p^v_H)^{i-j} \text{ or}$$

$$q^V_{ij} = (1 - e^{-mb})^{i-j} (e^{-mb})^j \text{ for } i \geqslant j$$

$q^V_{ij}=0$ otherwise (4-1)

Now, the transition probabilities will be:

$$P^V_{ij}=\sum_{f=0}^{\infty} n^V_t P^V_{ij/f} \text{ for } i=0,1,..,S-1 \; j=0,1,...,S \tag{4-2}$$

where $P^V_{ij/f}$ is the same probability as P^V_{ij}, under the condition of f voice arrivals between two successive openings.

Let's consider two successive gate openings (Figure 4.2), and call:

t_n: Time moment of the n-th gate opening.

j: Number of busy voice slots, just before the n-th opening.

i: Number of busy voice slots, just before the (n-1)-th opening.

x: Number of voice connection terminations between t_n, t_{n+1}.

f: Number of voice arrivals, between t_{n-1}, t_n.

It will be: $j=(i+f)-x$

or $x=k-j$

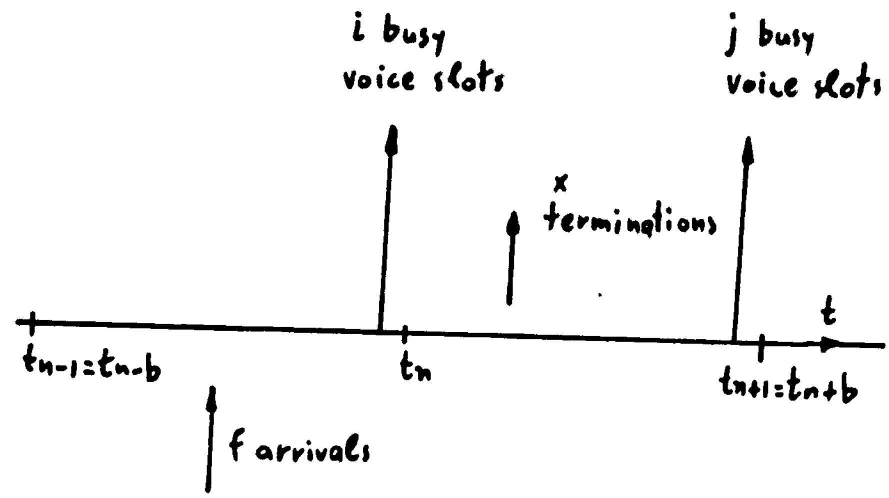

Figure 4-2:
The function of the gate, for the voice traffic

where $k=i+f$, is the number of busy voice slots just <u>after</u> the n-th opening.

However, it is always $x \geq 0$, $f \geq 0$, therefore $k > \max(i,j)$. Now we have

$$n^v_f p^v_{ij/f} = n^v_f q^v_{kj} = n^v_{k-i} q^v_{kj}$$

Now we can rewrite (4-2) as:

$$p^v_{ij} = \sum_{k=\max(i,j)}^{S-1} n^v_{k-i} q^v_{kj}$$

(because $0 \leq f \leq \infty$ $\max(i,j) \leq k \leq \infty$), or

$$P^{V}_{ij}=\sum_{k=\max(i,j)}^{S-1} n^{V}_{k-i}q^{V}_{kj}+\sum_{k=S}^{\infty} n^{V}_{k-i}q^{V}_{kj}$$

But for k>S we have $q^{V}_{kj}=q^{V}_{Sj}$ (since we can have at most S busy voice slots). Therefore

$$P^{V}_{ij}=\sum_{k=\max(i,j)}^{S-1} n^{V}_{k-i}q^{V}_{kj}+q^{V}_{Sj}\sum_{k=S}^{\infty} n^{V}_{k-i}$$

and finally

$$P^{V}_{ij}=\sum_{k=\max(i,j)}^{S-1} n^{V}_{k-i}q^{V}_{kj}+q^{V}_{Sj}\sum_{f=S-i}^{\infty} n^{V}_{f}$$

for i=0,1,...,S-1 j=0,1,...,S (4-3)

Of course, for $i\geq S+1$, $j\geq S+1$, it is $P^{V}_{ij}=0$ (4-4)

For i=S, j=0,1,...,S all voice slots are occupied just before the opening, and so we don't care about any new voice arrivals. Therefore

$P^{V}_{ij}=q^{V}_{Sj}$ for i=S j=0,1,...,S (4-5)

So the transition probabilities are given by (4-3)-(4-5). Noe we can write

$$p^{V}_{j}=\sum_{i=0}^{S} p^{V}_{i}P^{V}_{ij} \quad (4\text{-}6)$$

with j=0,1,...,S

and $p^{V}_{j}>0$ for all j. Furthermore we have

$$\sum_{j=0}^{S} p^{V}_{j}=1 \quad (4-7)$$

So finally, the state probabilities p^{V}_{i} are the solution of the linear system of equations (4-3)-(4-7) Now we can evaluate the loss probability PL: If PS is the probability that an arriving customer will receive service, just after the next opening, then PL will be:

PL=1-PS

However PS=ES/E, where

ES: Expected number of voice customers that arrive between two successive openings, and receive service, just after the gate opens.

E: Expected total number of voice customers that arrive between two openings. Of course $E=\lambda b$ (Poisson arrivals). Therefore

$$PL=1-(ES/\lambda b) \quad (4-8)$$

Now let's call

ns^V_i: Probability that i voice customers arrive between two successive openings, and receive service just after the gate opens. Then it will be

$$ES=\sum_{i=0}^{S} i\,ns^V_i=ns^V_1+\sum_{i=2}^{S} i\,ns^V_i \qquad (4\text{-}9)$$

But $ns^V_1=p^V_{S-1}\cdot(1-n^V_0)$ (4-10)

where $1-n^V_0$ is the probability of at least one arrival. Furthermore

$$\sum_{i=2}^{S} i\,ns^V_i=\sum_{i=2}^{S} i\left[\sum p^V_j\,ns^V_{i/j}\right]$$

where $ns^V_{i/j}$ are the same probabilities as ns^V_i, under the condition that j voice slots were busy just before the last opening. Notice that the index j=S-1 and j=S, because in that case, one or zero new customers will receive service (and those cases are described by ns^V_1, ns^V_0). Therefore

$$\sum_{i=2}^{S} i\,ns^V_i=\sum_{j=0}^{S-2} p^V_j\sum_{i=2}^{S} i\,ns^V_{i/j}$$

However we have

$ns^V_{i/j}=n^V_i$ for $(i+j)<S$ (some slots are still unused after opening, and so all the new customers received service).

$ns^V{}_{i/j}=1-\sum_{k=0}^{S-j-1} n^V{}_k$ for $(i+j)=S$ (all slots are occupied after opening, and so at least S-j customers arrived).

Now, (4-10) becomes:

$$\sum_{i=2}^{S} ins^V{}_i=\sum_{j=0}^{S-2} p^V{}_j[\sum_{j=2}^{S-j-1} in^V{}_i+ \\ +(S-j)(1-\sum_{k=0}^{S-j-1} n^V{}_k)] \qquad (4-11)$$

By combining (4-9)-(4-11), we get

$$ES=p^V{}_{S-1}(1-n^V{}_0)+ \\ +\sum_{j=0}^{S-2} p^V{}_j[\sum_{k=1}^{S-j-1} kn^V{}_k+(S-j)(1-\sum_{k=0}^{S-j-1} n^V{}_k)]$$

Now, if we substitute the value of ES into (4-8), we will get the desired value of PL.

DATA TRAFFIC

As we know, we have:

n_i: Probability of i data arrivals, between two successive gate openings.

$p_{j,n}$: Probability of j data packets in the systen (in the slots and in the buffer), just after the n-th opening.

Now let's consider two successive gate openings (Figure 4.3), and call:

t_n: Time moment of the n-th opening.

Q_n: Number of data packets in the system, just after t_n.

f: Number of data arrivals, between t_{n-1}, t_n.

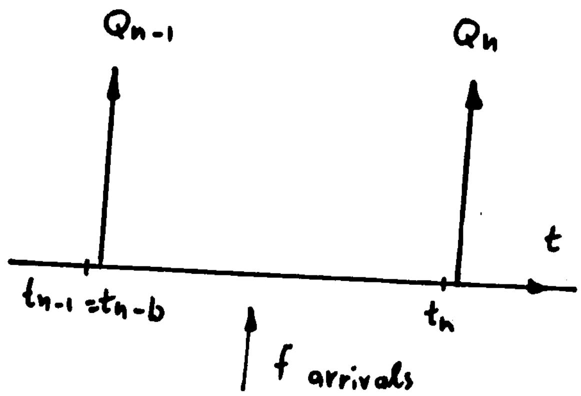

Figure 4-3:
The function of the gate, for data traffic.

The number of transmitted packets between t_{n-1}, t_n will be Q_{n-1}, if $Q_{n-1}<N$, or N if $Q_{n-1}\geq N$. Therefore $(Q_{n-1}-N)^+$ packets will remain in the system, and if we also count the new packet arrivals, we will get:

$$Q_n=(Q_{n-1}-N)^{+}+f \qquad (4\text{-}12)$$

And so, since f>0, it will be:

$f=Q_n$ if $Q_{n-1}<N$

$f=Q_n-Q_{n-1}+N$ if $N\leq Q_{n-1}\leq Q_n+N$ (4-13)

Now we can evaluate the state probabilities $p_{j,n}$:

$$p_{j,n} = Pr[Q_n = j] = \sum_{k=0}^{\infty} Pr[Q_{n-1} = k] Pr[Q_n = j / Q_{n-1} = k] \tag{4-14}$$

In order to have $Q_n = j$, under the condition that $Q_{n-1} = k$ all that we need is f data arrivals between t_{n-1}, t_n, with f satisfying (4-13). Therefore

$$Pr[Q_n = j / Q_{n-1} = k] = n_f = \begin{cases} n_j & \text{for } 0 \leqslant k < N \\ n_{j-k+N} & \text{for } N \leqslant k < N+j \end{cases}$$

Now (4-13) becomes

$$p_{j,n} = \sum_{k=0}^{N-1} p_{k,n-1} n_j + \sum_{k=N}^{N+j} p_{k,n-1} n_{j-k+N} \tag{4-15}$$

and asymptotically (for large n)

$$p_j = \sum_{k=0}^{N-1} p_k n_j + \sum_{k=N}^{N+j} p_k n_{j-k+N} \quad \text{or}$$

$$p_j = (p_0 + p_1 + \ldots + p_{N-1}) n_j + \sum_{j=N}^{j+N} p_i n_{j+N-i}$$

$$\text{with } j = 0, 1, 2, \ldots \tag{4-16}$$

At this point, from (4-16) we can see that p_j satisfies the same system of equations, as the state probabilities of a standard M/D/N queueing system.

Now we can evaluate the generating function P(z). From (4-16) we get:

$$p_j z^j = n_j z^j \sum_{i=0}^{N-1} p_i + \sum_{i=N}^{j+N} p_i n_{j+N-i} z^j$$

$$p_j z^j = n_j z^j \sum_{i=0}^{N-1} p_i + z^{-N} \sum_{i=N}^{j+N} p_i z^i n_{j+N-i} z^{j+N-i}$$

$$\sum_{j=0}^{\infty} p_j z^j = \sum_{j=0}^{\infty} n_j z^j \sum_{i=0}^{N-1} p_i +$$

$$+ z^{-N} \sum_{j=0}^{\infty} \sum_{i=N}^{j+N} p_i z^i n_{j-(i-N)} z^{j-(i-N)}$$

$$P(z) = N(z) \sum_{i=0}^{N-1} p_i + z^{-N} \sum_{i=N}^{\infty} p_i z^i N(z)$$

$$P(z)[z^N - N(z)] = N(z) \sum_{i=0}^{N-1} (z^N - z^i) p_i$$

$$P(z) = \frac{N(z) \sum_{i=0}^{N-1} (z^N - z^i) p_i}{[z^N - N(z)]}$$

But

$$N(z) = \sum_{i=0}^{\infty} n_i z^i = \sum_{i=0}^{\infty} e^{-\bullet b} (\bullet b)^i z^i =$$

$$= e^{-\bullet b} \sum_{i=0}^{\infty} (\bullet b z)^i = e^{-\bullet b(1-z)}$$ and so

$$P(z) = \frac{\sum_{i=0}^{N-1} (z^N - z^i) p_i}{z^N e^{\bullet b(1-z)} - 1} \qquad (4\text{-}17)$$

Now we need to elminate the N unknowns p_i $(i=0,1,\ldots,N-1)$. Let z_r $(r=0,1,\ldots,N-1)$ be the roots of the equation

$$z^N = e^{-\bullet b(1-z)}$$

It is obvious that these roots are unique, and all lie within the unique circle, except for $z_0=1$. Since $P(z)$ is an analytic function, the numerator of (4-17) must also have the same roots z_r $(r=0,1,\ldots,N-1)$, and since it is a polynomial of degree N, by using the fundamental theorem of algebra, we get:

$$P(z)=\frac{K(z-1)\prod_{i=1}^{N-1}(z-z_r)}{z^N e^{\bullet b(1-z)}-1}$$

Now we remove the unknown k, by using the condition $P(1)=1$. From L' Hospital's rule we get:

$$K=\frac{N-\bullet b}{\prod_{i=1}^{N-1}(1-z_r)}$$

So finally

$$P(z)=\frac{(N-\bullet b)(1-z)}{1-z^N e^{\bullet b(1-z)}} + \prod_{r=1}^{N-1}\frac{z-z_r}{1-z_r} \qquad (4\text{-}18)$$

Now we evaluate the expected number of data packets in the system (using L' Hospital's rule):

$$L=\left.\frac{dP(z)}{dz}\right|_{z=1}=\sum_{r=1}^{N-1}\frac{1}{1-z_r}+\frac{N-(N-\bullet b)^2}{2(N-\bullet b)}$$

The expected delay for data packets will be:

$$ED=\bullet^{-1}L$$

Finally, the expected waiting time will be:

$$EW=ED+ES$$

But ES=b/2 therefore

$$EW= \quad + \bullet^{-1}\left[\sum_{r=1}^{N-1}\frac{1}{1-z_r}+\frac{N-(N-\rho_2)^2}{2(N-\rho_2)}\right] \qquad (4\text{-}19)$$

where $\rho_2=\bullet b$ is the data traffic intensity.

C. Movable Boundary

VOICE TRAFFIC

The voice traffic analysis is the same as it was in the fixed boundary case. This happens because the number of available voice slots per frame is still S, and voice traffic is treated as a high - priority class. Therefore its performance will be the same as it was in the FFFB case.

DATA TRAFFIC

In the data traffic performance for the FFMB case was evaluated by using the technique of "unconditioning". Instead of that, here we will use the general method that we described before, and get the same results as those in [a].

Let's consider two successive gate openings (see Figure 4.4), and call:

$Q^V{}_n$: Number of occupied voice slots, just after t_n.

Q_n: Number of data customers in the system, just after t_n.

The state probabilities $p_{j,n}$ will be:

$$p_{j,n}=Pr[Q_n=j]=\sum_{K=0}^{S} Pr[Q^V{}_{n-1}=S-k].$$

$$\cdot\left(\sum_{i=0}^{\infty} \Pr[Q_{n-1}=i]\Pr[Q_n=j/Q_{n-1}=i,\ Q^V_{n-1}=S-k]\right) \tag{4-20}$$

The expression (4-20) is written in this form, because we assume that Q_n and Q^V_n are statistically independent (uncorrelated). Notice however, that large values of Q^V_n tend to be correlated with large values of Q_n, and this fact will affect our finals results (as we will see). If we had not assumed statistical independence, the expression for $p_{j,n}$ would be:

$$p_{j,n}=\Pr[Q_n=j]=\sum_{k,i} \Pr[Q^V_{n-1}=S-k, Q_{n-1}=i].$$
$$.\Pr[Q_n=j/Q_{n-1}=i,\ Q^V_{n-1}=S-k]$$

The difference between the fixed and the movadle boundary scheme, is actually in the number of available data slots per frame: In the FFMB, if a frame contains S-k occupied voice slots, then we have N+k available data slots. In the FFFB we always have S voice slots (occupied or not), and N available data slots. Therefore we may consider the FFFB as a special case of the FFMB wit S-0=S occupied voice slots, and N+0 available data slots. However, we recall that the state probabilities for the FFFB case ($p_{j,n}=p^F_{j,n}$), were given by (4-14), (4-17). Therefore we can rewrite (4-14) in the following form:

$$p^F_{j,n} = Pr[Q_n = j] = \sum_{i=0}^{\infty} Pr[Q_{n-1} = i] \cdot Pr[Q_n = j / Q_{n-1} = i,\ Q^V_{n-1} = S]$$

But from (4-17)

$$p^F_{j,n}(N) = \sum_{i=0}^{N-1} p_{i,n-1} n_j + \sum_{i=N}^{N+j} p_{i,n-1} n_{j-i+N}$$

Therefore, for the general case of S-k occupied voice slots, and N+k available data slots, we can write:

$$p^F_{j,n}(N+k) = \sum_{i=0}^{\infty} Pr[Q_{n-1} = i] Pr[Q_n = j / Q_{n-1} = i,\ Q^V_{n-1} = S-k] =$$
$$= \sum_{i=0}^{N+k-1} p_{i,n-1} n_j + \sum_{i=N+k}^{N+k+j} p_{i,n-1} n_{j-i+N+k} \qquad (4\text{-}21)$$

Now by comparing (4-20), (4-21) we get:

$$p_{j,n} = \sum_{k=0}^{S} pr[Q^V_{n-1} = S-k] p^F_{j,n}(N+k) \quad \text{or}$$

$$p_{j,n} = \sum_{k=0}^{S} Pr[Q^V_{n-1} = S-k] [\sum_{i=0}^{N+k-1} p_{i,n-1} n_j + \sum_{i=N+k}^{N+k+j} p_{i,n-1} n_{j-i+N+k}]$$

and asymptotically, for large n

$$p_j = \sum_{k=0}^{S} Pr[Q^V = S-k] [\sum_{i=0}^{N+k-1} p_i n_j + \sum_{i=N+k}^{N+k+j} p_i n_{j-i+N+k}] \Rightarrow$$

$$p_j = \sum_{k=0}^{S} p^V_{S-k} [n_j \sum_{i=0}^{N+k-1} p_i + \sum_{i=N+k}^{N+k+j} p_i n_{j-i+N+k}]$$

with j=0,1,2,... Now we can evaluate the generating function P(z):

$$\sum_{j=0}^{\infty} p_j z^j = \sum_{k=0}^{S} p^V{}_{S-k} [\sum_{j=0}^{\infty} n_j z^j \sum_{i=0}^{N+k-1} p_i + \sum_{j=0}^{\infty} z^j \sum_{i=N+k}^{j+N+k} p_i n_{j+N+k-i}] \Rightarrow$$

$$P(z) = \sum_{k=0}^{S} p^V{}_{S-k} [N(z) \sum_{i=0}^{N+k-1} p_i + z^{-(N+k)} \sum_{j=0}^{\infty} \sum_{i=N+k}^{j+N+k} p_i z^i n_{j+N+k-i} z^{j+N+k-i}] \Rightarrow$$

$$P(z) = \sum_{k=0}^{S} p^V{}_{S-k} [N(z) \sum_{i=0}^{N+k-1} p_i + z^{-(N+k)} N(z) \sum_{i=N+k}^{\infty} p_i z^i]$$

$$P(z) = \sum_{k=0}^{S} p^V{}_{S-k} [N(z) \sum_{i=0}^{N+k-1} p_i + z^{-(N+k)} N(z) [P(z) - \sum_{i=0}^{N+k-1} p^i z^i)] \Rightarrow$$

$$P(z)[1 - N(z) \sum_{k=0}^{S} p^V{}_{S-k} z^{-(N+k)}] = N(z) \sum_{k=0}^{S} \sum_{i=0}^{N+k-1} p^V{}_{S-k} p_i [1 - z^{-(N+k)-i}] \Rightarrow$$

$$P(z)[z^{N+S} - N(z) \sum_{k=0}^{S} p^V{}_{S-k} z^{S-k}] = N(z) \sum_{k=0}^{S} \sum_{i=0}^{N+k-1} p^V{}_{S-k} p_i [z^{N+S} - z^{S-k+i}]$$

and finally

$$P(z) = \frac{N(z) \sum_{k=0}^{S} \sum_{i=0}^{N+k-1} p^V{}_{S-k} p_i [z^{N+S} - z^{S-k+i}]}{z^{N+S} - N(z) P^V(z)}$$

where $P^V(z) = \sum p^V{}_i z^i$

However $N(z) = e^{-\lambda b(1-z)}$ and so

$$P(z)=\frac{\sum_{k=0}^{S}\sum_{i=0}^{N+k-1} p^{V}_{S-k}p_i\,(z^{N+S}-z^{S-k+i})}{z^{N+S}e^{\lambda b(1-z)}-P^{V}(z)} \qquad (4\text{-}22)$$

As we see, in (4-22) we must eliminate the (N+S) unknowns $(p_0,p_1,\ldots,p_{N+S-1})$. Let z_r $(r=0,1,\ldots,N+S-1)$ be the N+S roots of the equation

$$z^{N+S}e^{\lambda b(1-z)}=P^{V}(Z)$$

It is obvious that these roots are unique and all lie within the unit circle, except for $z_0=1$. Since $P(z)$ is an analytic function,the numerator of (4-22) must also be zero at z_r $(r=0,1,\ldots,N+S-1)$. And since it is a polynomial of degree S+N, by using the fundamental theorem of algebra we get:

$$P(z)=\frac{K(z-1)\prod_{r=1}^{N+S-1}(z-z_r)}{z^{N+S}e^{\lambda b(1-z)}-P^{V}(z)}$$

Now we can eliminate the constant K, by using the normalizing condition $P(1)=1$. By using L' Hospital's rule, we get:

$$K=\frac{N+S-\lambda b-\sum_{k=1}^{S}kp^{V}_{k}}{\prod_{r=1}^{N+S-1}(1-z_r)}$$

where we have used the fact that $\left.\frac{dP^{V}(z)}{dz}\right|_{z=1}=\sum_{k=1}^{S}kp^{V}_{k}$

So finally

$$P(z)=\frac{(N+S-\bullet b-\sum_{k=1}^{S} kp^V{}_k)(z-1)}{z^{N+S}e^{\bullet b(1-z)}-p^V(z)}+\prod_{r=1}^{N+S-1}\left(\frac{z-z_r}{1-z_r}\right)$$

Now the expected number of data customers in the system will be

$$L=\left.\frac{dP}{dz}\right|_{z=1}=$$

$$=\frac{\sum_{k=2}^{S} k(k-1)p^V{}_k-(N+S)(N+S-1)-(\bullet b)^2+2\bullet b(N+S)}{2(N+S-\bullet b-\sum_{k=1}^{S} kp^V{}_k)}+\sum_{r=1}^{N+S-1}\frac{1}{1-z_r} \qquad (4\text{-}23)$$

The expected delay for data customers will be

$$ED=\bullet^{-1}L=\bullet^{-1}\left\{\left.\frac{dP(z)}{dz}\right|_{z=1}\right\}$$

and finally, the expected waiting time will be

EW=ES+ED or

$$EW=\frac{b}{2}+\bullet^{-1}\cdot\left\{\frac{dP(z)}{dz}\right\} \qquad (4\text{-}24)$$

where $\left.\frac{dP(z)}{dz}\right|_{z=1}$ is given by (4-23)

Notice that here again the quantity $\bullet b$ represents the data traffic intensity, $\rho_2=\lambda b$. If we consider the simple case S=1, N=0 as an example, then (4-23), (4-24) will give us:

$$EW=\frac{b}{2}+\frac{b(2-\rho_2)(1+\rho_1)}{(1-\rho_2-\rho_1\rho_2)} \qquad (4-25)$$

4.2 Simplified discrete time model

In this section we analyze once more the performance of the fixed frame scheme, for the cases of fixed, and movable boundary. Both cases are studied by using the models which were introduced by Gitman et al [4].

A. Voice Traffic Model

The switching process for the voice traffic (for both the FFFB, and FFMB cases) is modelled as discrete time queueing gate, with the following characteristics:

- Period of gate opening: One frame period (τ_g=b).
- Voice arrivals distribution: Poisson, with arrival rate λ. Probability of i voice arrivals, between two successive openings:

$n^V_i = e^{-\lambda b}[(\lambda b)^i / i!]$ for $i=0,1,2,\ldots$

$n^V_i = 0$ otherwise

- Voice customers holding time distribution: Exponential, with mean m^{-1}. Probability that an occupied voice slot will still be occupied, after b secs:
  $p^V_H = e^{-mb}$

- Voice population: Infinite.

- Furthermore we assume the following: $\lambda b \ll 1$, which means that we may neglect the probability of more than one voice arrivals during one frame period (b secs) and $Smb \ll 1$, which means that we may neglect the probability of more than one voice call terminations during one frame period, without making considerable error. These approximations allow us to treat the system as a birth - death queuing system. In this way, we finally come up with a closed form solution for the voice traffic blocking probability.

A. Voice Traffic Analysis

The voice traffic analysis will be the same for the FFFB, and and FFMB case. Let's call again:

p^V_i: Probability of i busy voice slots, just before the gate opens.

$p^V_{i,j}$: Probability of j busy voice slots just before the gate opens, given that i slots were busy, just before the last opening.

Now let's consider two successive openings (time moments t_{n-1}, t_n), and call Q^V_n the number of occupied voice slots, just before t_n. We have:

$P_{i,j}=0$ for $|i-j| \geqslant 2$, according to our assumptions.

For $0<i<S$ we have:

$P_{i,i+1}=Pr[Q^V_n=i+1/Q^V_{n-1}=i]=$

$=Pr[$one arrival between $t_{n-1}, t_n].$

$.Pr[$i voice slots still busy at tn$]=$

$=n^V_1(p^V_H)^i=(\lambda be^{-\lambda b})(e^{-mb})$ or

$P_{i,i+1}=\lambda b\exp[-(\lambda+im)b]$

furthermore we have

$P_{i,i-1}=Pr[Q^V_n=1-1/Q^V_{n-1}=i]=$

$=Pr[$no arrivals between $t_{n-1}, t_n].$

$.Pr[(i-1)$ voice slots still busy at $t_n]=$

$=(e^{\lambda b})[(e^{-mb})^{i-1}(1-e^{-mb})]$

$=(e^{-\lambda b})(ie^{-(i-1)mb}mb)$, and so

$$P_{i,i-1} \cong imb\exp[-(\lambda+(i-1)mb]$$

Summarizing, we have the following transition probabilities:

$$P_{i,j}=0 \qquad \text{for } |i-j| \geqslant 2$$

$$P_{i,i+1}=\lambda b\exp[-(\lambda+im)b]=B_i/b \qquad \text{for } 0\leqslant i\leqslant S$$

$$P_{i,i-1}=i(mb)\exp[-(\lambda+(i-1)mb]=D_i/b \text{ for } 0<i<S \qquad (4-26)$$

where B_i, D_i are the "birth", and "death" coefficients of the system (they are defined to be the rates at which one customer is coming into the system or is leaving the system, per time unit).

However, in general in a birth - death process, if the values of B_i, D_i are known, then the equilibrum solution for the state probabilities $p^V{}_k$ will have the following standard form:

$$p^V{}_k=p^V{}_0\prod_{i=0}^{k-1}[B_i/D_{i+1}] \qquad \text{for } k=1,2,\dots,S$$

and $p^V_k=0$ for $k>S$ (4-27)

with a single unknown constant, p^V_0

Now, from (4-26) we get:

$$p^V_k = p^V_0 \prod_{i=0}^{k-1} [(\lambda m^{-1})/(i+1)] \text{ for } k=1,2,\ldots,S$$

$$p^V_k=0 \text{ for } k>S \qquad (4\text{-}28)$$

From (4-28) we can see that these solutions have exactly the same form as the state probabilities of a standard M/M/S queueing system. Consequently their values will be:

$$p^V_i = B(\rho_1, i) = \frac{[\rho_1^{\,i}/i!]}{\sum_{i=0}^{S} \rho_1^{\,i}/i!} \qquad (4\text{-}29)$$

with $\sum_{i=0}^{S} p^V_i = 1$

But since only one voice customer may arrive between two successive openings, the blocking probability for voice customers will be simply the probability of finding S occupied voice slots. Therefore

$$PL = p^{V}_{S} = \frac{\rho_1^{S}/S!}{\sum_{i=0}^{S} \rho_1^{i}/i!} \qquad (4\text{-}30)$$

where $\rho_1 = \lambda m^{-1}$ is the voice traffic intensity, in Erlangs.

This is the well known "Erlang-B" formula, and as we have just proved that it is a close approximation for the blocking probability of voice customers.

In order to show that our assumptions for $\lambda b \ll 1$, and $Smb \ll 1$ are true in reality, let's consider the range of parameters in practical systems: The values of the frame period are typically in the range of 5 ms$\leq b \leq$50ms. The holding time is in the order of 60 sec$\leq m^{-1} \leq$300 sec, and the number of available slots per frame (S) is a few hundred or less. Therefore $Smb \ll 1$. Furthermore, in typical, well designed, circuit switched systems, the blocking probability is low, and the number of voice channels S is in the same order of magnitude as the average traffic intensity λ/m. Therefore $\lambda b \ll 1$. For example, if $m^{-1}=300$ sec, b=30 ms, $\rho_1 \leq 100$ Erlangs, and $S<100$ slots, we obtain $Smb=0.01$ and $\lambda b=0.01$.

B. **Data Traffic Model**

The switching process for the data traffic is modelled as a queueing gate (Figure 4.5) with the following characteristics (for both the FFFB, and FFMB cases):

- Period of gate opening: Equal to the duration of one slot
  ($\tau_g = \tau$ sec).

- Data arrivals distribution: Poisson, with rate λ.
  Probability of i data arrivals between two successive openings:

$$n_i = \frac{e^{-\lambda\tau}(\lambda\tau)^i}{i!} \qquad i=0,1,2,\ldots$$

$$n_i = 0 \qquad \text{otherwise.}$$

- Data customers holding time distribution: Deterministic.
  Probability that an occupied voice slot will still be occupied after x sec:

$$p_H = 1 \quad \text{if} \quad x < \tau$$

$$p_H = 0 \quad \text{if} \quad x \geq \tau$$

- Data population: Infinite

- Data buffer capacity: Infinite

- Furthermore, we assume the following:
  Each gate opening allows one voice or data slot to be transmitted. The decision whether the transmitted slot will contain voice or data depends on the boundary allocation policy (fixed or movable). The gate is actually a channel access switch, which has two possible positions: "voice", and "data".

Now let's call δ: Probability that the channel access switch is at the "data" position (allowing the transmission of one data packet). It will be:

$$\delta = \frac{\text{average number of available data slots per frame}}{\text{total number of slots per frame}}$$

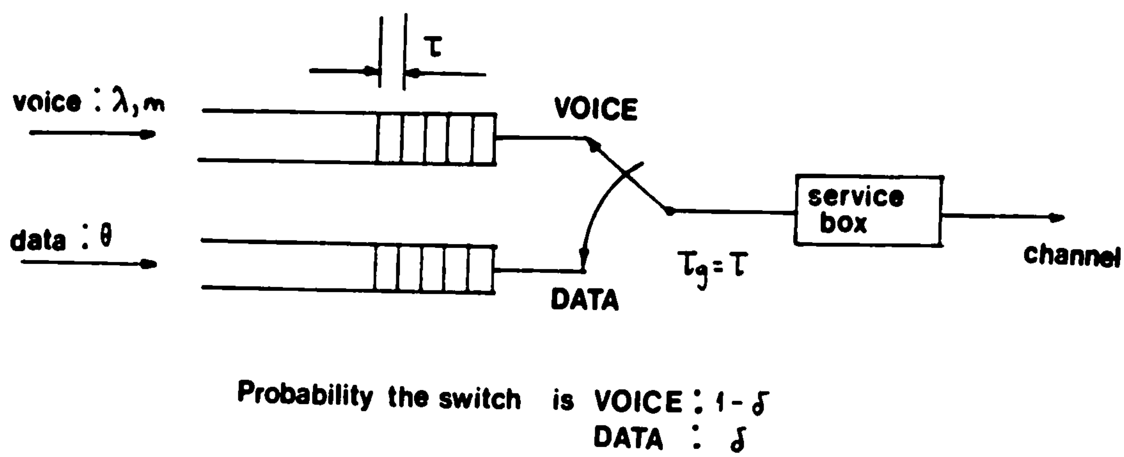

Figure 4-4:
Simplified discrete time model.

Now the values of δ will be:

For the case of FFFB: $\delta_F = N/(N+S)$

For the case of FFMB: $\delta_M = [N+(S-\bar{V})]/(N+S)$ (4-31)

where we call $\bar{V}$ the average number of occupied voice slots per frame. Of course it will be:

$V=\sum_{i=0}^{S} i\Pr[i$ occupied voice slots per frame$]$ and by using (4-29) we get:

$$V=\sum iB(\rho_1,i)=\frac{\sum_{i=0}^{S} i(\rho_1^{i}/i!)}{\sum_{i=0}^{S}\rho_1^{i}/i!}=\frac{\sum_{i=0}^{S}\rho_1^{i-1}/(i-1)!}{\sum_{i=0}^{S}\rho_1^{i}/i!}$$

$$=\frac{\sum_{k=0}^{S-1}\rho_1^{k}/k!}{\sum_{i=0}^{S}\rho_1^{i}/i!}=[\sum_{k=0}^{S-1}B(\rho_1,k)]=\rho_1[1-B(\rho_1,S)]$$

therefore $\bar{V}=\rho_1[1-B(\rho_1,S)]$ (4-32)

Notice that the switch position is a random process, described by the probability δ, which is the same for every gate opening. Therefore this process is uncorrelated from frame to frame. However, this is not true for the case of movable boundary, where the number of voice customers (and therefore the switch position) is strongly correlated, from frame to frame. As we will see later, this may introduce some error in the obtained results.

C. Data Traffic Analysis

FIXED BOUNDARY

Let's consider two consecutive gate openings, at the moments t_{n-1}, t_n. As we know, we have:

$p_{i,n}$: Probability of i data packets in the system, just after t_n.

n_i: Probability of i data arrivals, between t_{n-1}, t_n. It will be:

$$p_{i,n}=(1-\delta)p_{i,n/v}+\delta p_{i,n/d} \qquad (4\text{-}33)$$

where $\delta=\delta_F$ and

$p_{i,n/v}$: The same probability as $p_{i,n}$, under the condition that the (n-1)-th slot was used for voice transmission, and

$p_{i,n/d}$: The same probability as $p_{i,n}$, under the condition that the (n-1)-th slot was used for data transmission.

Now, suppose that the (n-1)-th slot was available for voice transmission. If j data packets were present at t_n, then we need (i-j) packet arrivals between t_{n-1}, t_n, in order to have i packets present at t_n (see Figure 4.6). Therefore, for the different values of j=0,1,...,i we get:

$$p_{i,n/v}=\sum_{j=0}^{i} n_{i-j}p_{j,n-1}=\sum_{k=0}^{i} n_k p_{i-k,n-1} \qquad (4\text{-}34)$$

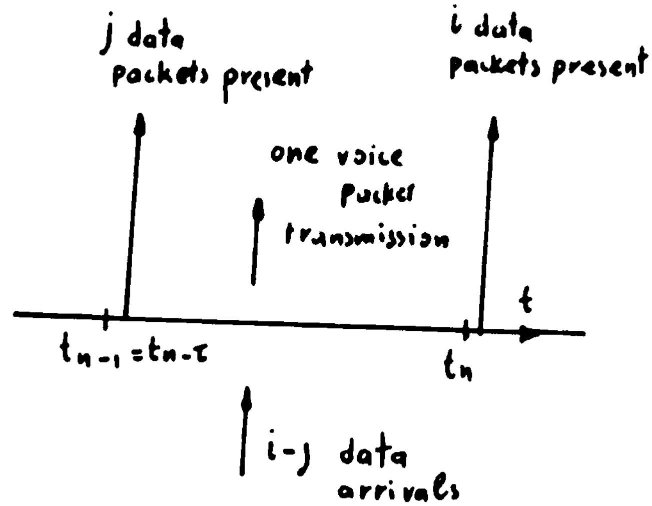

Figure 4-5: Case of (n-1)-th slot available for voice transmission

In the same way, if the (n-1)-th slot is available for data transmission, then one data packet will be transmitted between t_{n-1}, t_n. Therefore, if j data packets were present at t_n, then we need (i-j+1) data arrivals during t_{n-1}, t_n, in order to have i packets present at t_n (see Figure 4.7). Notice however, that in the case of j=0, we need only (i-j) new arrivals (since there is no packet to be transmitted, although it could). Therefore, for the different values of j=0,1,...,i+1 we get:

$$p_{i,n/d} = \sum_{j=1}^{i+1} n_{i-j+1} p_{j,n-1} + [n_{i-j} p_{j,n}]_{j=0}$$ or

$$p_{i,n/d} = \sum_{k=0}^{i} n_k p_{i-k+1,n-1} + n_i p_{0,n-1} \qquad (4\text{-}35)$$

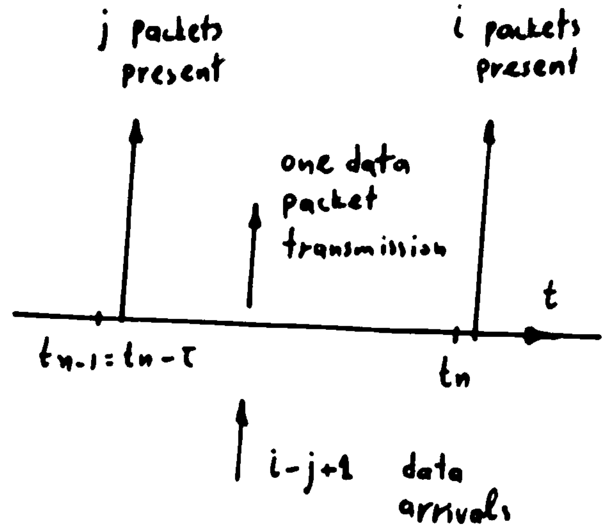

Figure 4-6: Case of (n-1)-th slot available for data transmission.

So finally, from (4-33)-(4-35) we get:

$$p_{i,n}=(1-\delta)\sum_{k=0}^{i} n_k p_{i-k,n-1}+$$

$$+\delta\Big[\sum_{k=0}^{i} n_k p_{i-k+1,n-1}+n_i p_{0,n-1}\Big] \quad \text{or}$$

$$p_{i,n}=\sum n_k\big((1-\delta)p_{i-k,n-1}+\delta p_{i+1-k,n-1}\big)$$

$$+n_i\delta p_{0,n-1}$$

and asymptotically, for large n

$$p_i=\sum_{k=0}^{i} n_k\big((1-\delta)p_{i-k}+\delta p_{i+1-k}\big)+n_i\delta p_0 \qquad (4-36)$$

Now we can derive the generating function P(z). From (4-36) we get:

$$\sum_{i=0}^{\infty} p_i z^i=$$

$$=\sum_{i=0}^{\infty} z^i [\sum_{k=0}^{i} n_k((1-\delta)p_{i-k}+\delta p_{i+1-k})+$$

$$+n_i \delta p_0] \qquad \text{or}$$

$$P(z)=(1-\delta)\sum_{i=0}^{\infty}\sum_{k=0}^{i} n_k p_{i-k} z^i+$$

$$+\delta\sum_{i=0}^{\infty}\sum_{k=0}^{i} n_k p_{i+1-k} z^i+$$

$$+\delta p_0 \sum_{i=0}^{\infty} n_i z^i \qquad \text{or}$$

$$P(z)=(1-\delta)\sum_{i=0}^{\infty}\sum_{k=0}^{i}(n_k z^k)(p_{i-k} z^{i-k})+$$

$$+F+\delta p_0 N(z) \tag{4-37}$$

where

$$F=\delta\sum_{i=0}^{\infty}\sum_{k=0}^{i} n_k p_{i+1-k} z^i=$$

$$=(\delta/z)\sum_{j=1}^{\infty}\sum_{k=0}^{j-1} n_k p_{j-k} z^j=$$

$$=(\delta/z)\sum_{j=1}^{\infty}\sum_{k=0}^{j} n_k p_{j-k}-(\delta/z)\sum_{j=1}^{\infty} n_j p_0 z^j=$$

$$=(\delta/z)P(z)N(z)-(\delta/z)p_0 N(z) \qquad \text{or}$$

$$F=(\delta/z)N(z)[P(z)-p_0] \tag{4-38}$$

From (4-37), (4-38) we get:

$$P(z)=(1-\delta)N(z)P(z)+(\delta/z)N(z)[P(z)-p_0]+\delta p_0 N(z)$$

and so

$$P(z)=\frac{\delta p_0(z-1)}{z[\delta+(1/N(z))-1]-\delta}$$

The constant p_0 is determined by the normalizing condition $P(1)=1$.

Since $N(z)=e^{-\lambda\tau(1-z)}$, using L' Hospital's rule, we get:

$$P_0=\frac{\delta-\lambda\tau}{\delta}$$

Now the average number of data packets in the system will be:

$$L=\left.\frac{dP(z)}{dz}\right|_{z=1}=\frac{\lambda\tau(2-\lambda\tau)}{2(\delta-\lambda\tau)}$$

The expected data packet delay will be:

$$ED=\lambda^{-1}L=\frac{\tau(2-\lambda\tau)}{2(\delta-\lambda\tau)}$$

and so, the expected waiting time for data packets will be:

$$EW=ES+ED$$

where $ES=b/2=[(N+S)\tau]/2$ and so

$$EW=\frac{b}{2}+\frac{\tau(2-\lambda\tau)}{2(\delta-\lambda\tau)}$$

But $\delta=\delta_F=N/(N+S)$, and $\tau=b/(N+S)$.

If we call $\rho_2=\lambda b$ the data traffic intensity we finally get:

$$EW=\frac{b}{2}+\frac{b[2(N+S)-\rho_2]}{2(N+S)[N-\rho_2]} \qquad (4\text{-}39)$$

MOVABLE BOUNDARY

The data traffic analysis for this case is exactly the same, as for the FFFB case. The only difference is that the value of ϕ is now $\phi = \phi_M$. Therefore we have:

$$EW = \frac{b}{2} + \frac{b[2(N+S) - \rho_2]}{2(N+S)[\phi_M(N+S) - \rho_2]} \tag{4-40}$$

where ϕ_M is given by (4-31).

As an example, let's consider the case of S=1, N=0.
It will be:

$\phi_M = 1 - \bar{V}$ with

$\bar{V} = \rho_1[1 - B(\rho_1, S)] = \rho_1/(1+\rho_1)$

Therefore $\phi_M = 1/(1+\rho_1)$ and finally

$$EW = \frac{b}{2} + \frac{b(2-\rho_2)(1+\rho_1)}{2(1-\rho_2-\rho_1)} \tag{4-41}$$

Now if we compare (4-41) with (4-25) we can see that both the discrete time model and the simplified discrete time model give us the same results for N=0, S=1.

4.3 Continuous Time Model

In this section we analyze the performance of the fixed frame scheme, only for the movable boundary case. Mainly we describe an exact analysis of the data traffic performance for the fixed frame movable boundary case, which was not analyzed very accurately in the last two sections (because we had assumed independence between the used random variables). The following is used for the voice, and data traffic, and was introduced by Weinstein et al [5].

A. Voice, and Data Traffic Model.

In this section, in order to take care of the correlation between the number of occupied voice slots, and the number of data customers in the system, we use a two - dimensional continuous time model to describe the system, with the following assumptions:

- We ignore the effect of time quantization due to the frame structure, and assume that voice and data customers may enter the system at any moment in time (not only at the moments which are integer multiples of the frame period).
- We describe the ~~state of the~~ system in terms of the two - dimensional continuous time state probabilities, $p_{i,j}(t)$. Let's call

$Q^V(t)$: Number of occupied voice slots, at the moment t.

$Q(t)$: Number of data customers in the system (in the slots, and in the buffer), at the moment t. It will be:

$$p_{i,j}(t) \triangleq Pr[Q^V(t)=i,\ Q(t)=j]$$

As $t \to \infty$, the system reaches its steady state, with $[dp_{i,j}(t)/dt]=0$. If $Q^V=\lim_{t\to\infty} Q^V(t)$ and $Q=\lim_{t\to\infty} Q(t)$ are the steady state numbers of customers in the system, the steady state probabilities will be:

$$p_{i,j} \triangleq Pr[Q^V=i,\ Q=j] \qquad (4\text{-}42)$$

- we assume that the voice arrivals isdistribution is Poisson, with rate λ.
  Probability of i voice arrivals in t secs:

$$n^V_i = \frac{e^{-\lambda t}(\lambda t)^i}{i!} \qquad \text{for } i=0,1,2,\ldots$$

$$n^V_i = 0 \qquad \text{otherwise}$$

- The data arrivals distribution is Poisson, with rate δ.
  Probability of i data arrivals, in t secs:

$$n_i = \frac{e^{-\delta t}(\delta t)^i}{i!} \qquad \text{for } i=0,1,2,\ldots$$

$$n_i = 0 \qquad \text{otherwise}$$

- The voice customers holding time distribution is exponential, with mean m_1^{-1}.
  Probability that an occupied voice slot will still be occupied after t sec: $p^V_H=e^{-m_1 t}$.

- Furthermore we assume that the data packet lengths are exponentially distributed, with mean m_2^{-1} (instead of being constant, and equal to the duration of one data slot, τ). Usually we consider the value $m_2^{-1}=\tau$.

  Probability that an occupied data slot will still be occupied after t sec: $p_H=e^{-m_2 t}$.

- Under these assumptions, the system can be modelled as two - dimensional continuous time Markov chain,, with state probabilities $p_{i,j}$.

- In the FFMB scheme, there is a total of S+N available slots, where S are reserved for the voice calls, with priority. In other words, if a voice call arrives and there are less than S voice calls occupying the slots, the call will seize a free slot (channel) if there is one, otherwise it will preempt a data customer, who might be using one of them. The preempted data customer returns to the buffer, and is serviced on a first - come first - served basis.

- We assume infinite voice, and data populations.

- We assume finite data buffer capacity: M buffer spaces are available for data customers.

Now, using the above model, we can analyze the data traffic performance (by evaluating the state probabilities $p_{i,j}$ and applying the described general method of analysis).

DATA TRAFFIC

First we evaluate the steady state probabilities $p_{i,j}$:
Since the voice (data) arrivals distribution is Poisson

with rate λ ($\bullet$), therefore the rate at which voice (data) arrivals occur will be λ ($\bullet$). Since the voice (data) holding times distribution is exponential with mean m_1^{-1} (m_2^{-1}), threfore the rate at which voice (data) "deaths" occur will be m_1 (m_2). The "death" of a customer happens when his transmission is completed (and he stops occupying the slot that was assigned to him).

Now let's consider the equilibrium state $E_{i,j}$ of the system, and its neighbouring states $E_{i-1,j}$, $E_{i,j-1}$, $E_{i+1,j}$, $E_{i,j+1}$, in the state transition diagram of the Figure (4.7). Since the system is in eqilibrium, the flow into and out of the state $E_{i,j}$ must be conserved. This means that the input flow must be equal to the output flow from the state $E_{i,j}$. In general, it will be:

Flow rate into $E_{i,j}=\lambda p_{i-1,j}+\bullet p_{i,j-1}+$

$$+(i+1)m_1 p_{i+1,j}+(j+1)m_2 p_{i,j+1}$$

Flow rate out of $E_{i,j}=(\lambda+\bullet+im_1+jm_2)p_{i,j}$

Therefore in equilibrium we will have:

$$(\lambda+\bullet+im_1+jm_2)p_{i,j}=\lambda p_{i-1,j}+\bullet p_{i,j-1}+$$

$$+(i+1)m_1 p_{i+1,j}+(j+1)m_2 p_{i,j+1} \quad (4\text{-}43)$$

However the equation (4-43) is very general, and

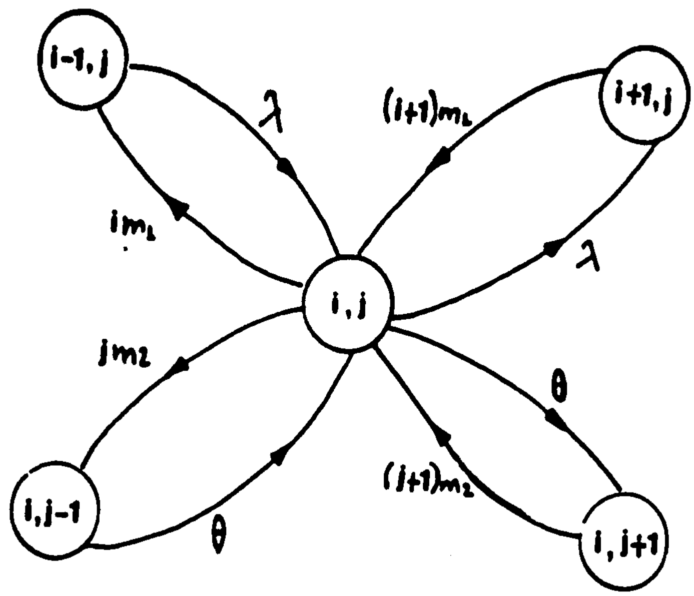

Figure 4-7:
State transition diagram: State $(i,j)=E_{i,j}$

takes different forms, depending on the different values of i,j. So we have the following cases:

a) If $i=0,1,\ldots,S-1$ and $j=0,1,\ldots,N+S-i-1$, then all voice customers (i) are transmitted, and all data customers (j) are transmitted. Therefore (4-43) takes the form:

$$(\lambda+\theta+im_1+jm_2)p_{i,j}=\lambda p_{i-1,j}+\theta p_{i,j-1}+$$
$$+(i+1)m_1p_{i+1,j}+(j+1)m_2p_{i,j+1}$$

b) If $i=0,1,\ldots,S-1$ and $N+S-i\leqslant j\leqslant M+N+S-i-1$, then all voice customers (i) are transmitted;

(N+S-i) of the data customers are transmitted, and the rest of them are buffered; the buffer is not full. Therefore

$$(\lambda+\bullet+im_1+(N+S-i)m_2)p_{i,j}=\lambda p_{i-1,j}+\bullet p_{i,j-1}+$$
$$+(i+1)m_1p_{i+1,j}+(N+S-i)m_2p_{i,j+1}$$

c) If $i=0,1,\ldots,S-1$ and $j=M+N+S-i$ then all voice customers (i) are transmitted; (N+S-i) of the data customers are transmitted, M of them are buffered, and the rest of them are "lost"; the buffer is full, and no new data customers are accepted in the system ($p_{i,j+1}=0$). Therefore

$$(\lambda+im_1+(N+S-i)m_2)p_{i,j}=\lambda p_{i-1,j}+$$
$$+\bullet p_{i,j-1}+(i+1)m_1p_{i+1,j}$$

d) If $i=S$ and $j=0,1,\ldots,N-1$ then S voice customers are transmitted, and the rest of them are lost; no new voice customers are accepted in the system, and $p_{i+1,j}=0$.
all data customers (j) are transmitted. Therfore

$$(\bullet+Sm_1+jm_2)p_{S,j}=\lambda p_{s-1,j}+$$
$$+\bullet p_{S,j-1}+(j+1)m_2p_{S,j+1}$$

e) If $i=S$ and $N \leq j \leq M+N-1$ then

S voice customers are transmitted, the rest of them are lost, and $p_{i+1,j}=0$.

N of the data customers are transmitted, and the rest of them are buffered (buffer is not full). Therefore

$$(\bullet+Sm_1+Nm_2)p_{S,j}=\lambda p_{S-1,j}+\bullet p_{S,j-1}+Nm_2 p_{S,j+1}$$

f) If $i=S$ and $j=M+N$ then

S voice customers are transmitted, and the rest of them are lost; $p_{i+1,j}=0$.

N of the data customers are transmitted, and the rest of them are lost; the buffer is full, and $p_{i,j+1}=0$. Therefore

$$(Sm_1+Nm_2)p_{S,j}=\lambda p_{S-1,j}+\bullet p_{S,j-1}$$

Notice that we assume:

$$p_{-1,j}=p_{S+1,j}=p_{i,-1}=p_{i,M+N+S+1-i}=0$$

and when N=0 the equation (4-43,d) is meaningless.

Now we can see that the state probabilities $p_{i,j}$ satisfy the system of equations (4-43,a)-(4-43,f). In the case M=oo , the above system can be solved formally, using the general technique that we described. Since we have two - dimensional state probabilities, we define the generating function:

$$P_i(z)=\sum_{j=0}^{\infty} p_{i,j} z^j \qquad i=i^v=0,1,\dots,S$$

Furthermore we have the normalizing condition

$$\sum_{i=0}^{S} P_i(1)=1$$

Now we can solve for $P_i(z)$ by developing a set of linear equations of the form

$$A(z)P(z)=B(z) \tag{4-44}$$

where $A(z)$ is an $(S+1)x(S+1)$ matrix (which is a function of the parameters $\lambda,\bullet,m_1,m_2,N,S$) and $P(z)$ is an $(S+1)x1$ vector, with components $P_0(z)$, $P_1(z),\dots,$ $P_S(z)$. $B(z)$ is an $(S+1)x1$ vector, depending on the unknown probabilities $p_{i,j}$ for $i=0,1,\dots,S$ and $j=0,1,\dots,N+S$. The solution of (4-44) for $P(z)$ can be carried out, using the general tachnique that we have described. Finally, when $P_i(z)$ has been found, the expected number of data customers in the system will be:

$$L=E\{Q\}=\sum_{i=0}^{S} (P_i(z))' \Big|_{z=1}$$

and the mean delay will be

$ED=\bullet^{-1}L$

Therefore, the expected waiting time will be

$$EW=(b/2)+\bullet^{-1}L \qquad (4\text{-}45)$$

although the above solution can be realized, it is practically difficult to evaluate, for reasonably large values of N and S. Some examples, and approximation techniques have been developed by Chang [9].

However, for the special case of S=1, N=0, we can easily evaluate the performance of the system, and get some useful information about its behaviour. For this case, the system of difference equations for $p_{i,j}$ takes the following form:

$$(\lambda+\bullet)p_{0,0}=m_1p_{1,0}+m_2p_{0,1} \qquad (a)$$

$$(\lambda+\bullet+m_2)p_{0,j}=\bullet p_{0,j-1}+m_1p_{1,j}+m_2p_{0,j+1} \qquad (b)$$

for j=1,2,...

(4-46)

and

$$(\bullet+m_1)p_{1,0}=\lambda p_{0,0} \qquad (c)$$

$$(\bullet+m_1)p_{1,j}=\lambda p_{0,j}+\bullet p_{1,j-1} \qquad (d)$$

for j=1,2,...

Now, using (4-46,b) we get:

$$(\lambda+\theta+m_2)\sum_{j=1}^{\infty} p_{0,j}z^{j-1}=\theta\sum_{j=1}^{\infty} p_{0,j-1}z^{j-1}+$$

$$+m_1\sum_{j=1}^{\infty} p_{1,j}z^{j-1}+m_2\sum_{j=1}^{\infty} p_{0,j+1}z^{j-1}$$

or

$$z^{-1}(\lambda+\theta+m_2)[P_0(z)-p_{0,0}]=\theta P_0(z)+$$

$$+m_1z^{-1}[P_1(z)-p_{1,0}]+m_2z^{-2}[P_0(z)-(p_{0,0}+zp_{0,1}]$$

and after multiplying by z^2

$$(\lambda+\theta+m_2)P_0(z)-m_2P_0(z)-\theta z^2P_0(z)=m_1zP_1(z)+$$

$$+(\lambda+\theta+m_2)zp_{0,0}-m_2p_{0,0}-(m_1p_{1,0}+m_2p_{0,1})z$$

After using (4-46,a) get

$$(-\theta z^2+(\lambda+\theta+m_2)z-m_2)P_0(z)=$$
$$=m_1zp_1(z)+m_2(z-1)P_{0,0} \qquad (4-47)$$

In the same way, from (4-46,d) we get:

$$(\theta+m_1)\sum_{j=1}^{\infty} p_{1,j}z^{j-1}=$$

$$=\lambda\sum_{j=1}^{\infty} p_{0,j}z^{j-1}+\theta\sum_{j=1}^{\infty} p_{1,j-1}$$

or

$$z^{-1}(\theta+m_1)[P_1(z)-p_{1,0}]=\lambda z^{-1}[P_0(z)-p_{0,0}]+$$

$$+\theta P_1(z)$$

and finally

$$\lambda P_0(z)=[\theta(1-z)+m_1]P_1(z) \qquad (4-48)$$

Now from the system of equations (4-47), (4-48) we get:

$$P_1(z)=\frac{\lambda m_2(z-1)p_{0,0}}{[-\theta z^2+(\lambda+\theta+m_2)z-m_2][\theta(1-z)+m_1]-\lambda m_1} \qquad (4\text{-}49)$$

and

$$P_0(z)=\frac{[\theta(1-z)+m_1]m_2(z-1)p_{0,0}}{[-\theta z^2+(\lambda+\theta+m_2)z-m_2][\theta(1-z)+m_1]-\lambda m_1} \qquad (4\text{-}50)$$

Furthermore, we have the normalizing condition

$$P_0(1)+P_1(1)=1 \qquad (4\text{-}51)$$

From (4-47) we get

$$\lambda P_0(1)=m_1 P_1(1) \qquad (4\text{-}52)$$

Furthermore, from (4-49), using L' Hospital's rule, we get

$$P_1(1)=\frac{m_2\lambda p_{0,0}}{m_1 m_2-\theta m_1-\lambda\theta} \qquad (4\text{-}52)$$

Now, from (4-51)-(4-53) we get

$$p_{0,0}=\frac{m_1 m_2-\theta m_1-\lambda\theta}{m_2(\lambda+m_1)} \qquad \text{or}$$

$$p_{0,0}=\frac{(1-\rho_2-\rho_1\rho_2)}{(1+\rho_1)} \qquad (4\text{-}54)$$

where $\rho_1=\lambda/m_1$, $\rho_2=\bullet/m_2$ are the voice, and data traffic intensities.

Now using (4-49), (4-50), combined with(4-54), we get (using L' Hospital's rule) the expected number of data customers in th system:

$$L=E\{Q\}=\left.\frac{dP_0(z)}{dz}\right|_{z=1}+\left.\frac{dP_1(z)}{dz}\right|_{z=1}$$

or

$$L=\frac{\rho_2(1+\rho_1)^2+(\rho_1\bullet)/m_1}{(1+\rho_1)(1-\rho_2-\rho_1\rho_2)} \qquad (4\text{-}55)$$

and finally

$$EW=(b/2)+\bullet^{-1}L \qquad (4\text{-}56)$$

Notice that in this section we did not analyze the voice traffic performance, because it has been accurately analyzed in previously in this chapter, using the discrete time models.

5. ANALYSIS FOR THE VARIABLE FRAME SCHEME

5.1 Frame Length Constraints

Here we consider an integrated voice/data multiplex structure that utilizes the variable frame (movable boundary) transmission scheme. We assume the following:

- Duration of one voice-slot: τ sec
- Duration of one data-slot: τ sec
- Available voice slots per frame: S
- Available data slots per frame: N
- Duration of the i-the frame: $[(v)_i+(d)_i]\tau=b_i$ sec] where (v_i), (d_i) are respectively the numbers of occupied voice and data slots in the i-th transmitted frame.

As we have mentioned before, in the variable frame scheme, the variation of the frame length affects the synchronization of the voice traffic, and may create a loss of its transmission transparency. In order to avoid that, we assume that the frame size is upper - bounded to a certain maximum value, which allows us to regain the synchronization at the receiving node. The voice digitization rate and the transmission rate over the link are designed so that they allow us to buffer the voice bits which are generated during the average frame duration; for the same reason the frame length is also lower-bounded to an appropriate minimum value. Finally with the help of some (limited time) buffering of voice bits before their transmission and some equalizing buffering at the receiving node, it is possible to realize the variable frame structure, and still preserve the transparency of the voice transmission.

In order to illustrate this, suppose that voice calls are digitized at a rate of R bits/sec and that the duration of an average generated voice message is t_v secs; in other words, the length of the average voice message is (Rt_v) bits. Since each voice slot accomodates $n=\tau r$ bits (r is the transmission rate over the link), then the average voice message requires

$$m^{-1}= Rt_v/n \tag{5-1}$$

voice slots for its transmission. These m^{-1} voice slots will be contained in consecutive frames of variable length. If all the frames had the same duration of b_c secs, then the transmission of the voice message would be completed after $(m^{-1}b_c)$ secs. So if we want the transmission to last as long as the original message duration, we must have:

$m^{-1}b_c=t_v$ or $b_c=mt_v$ or

$b_C=n/R$ (5-2)

Therefore we may consider that every b_C seconds an active voice customer "v" generates packets of bits (equivalent voice slots), which must be accomodated periodically in consecutive frames. If the duration of these frames is constant and equal to b_C seconds , the transmission of course will be transparent (see Figure).

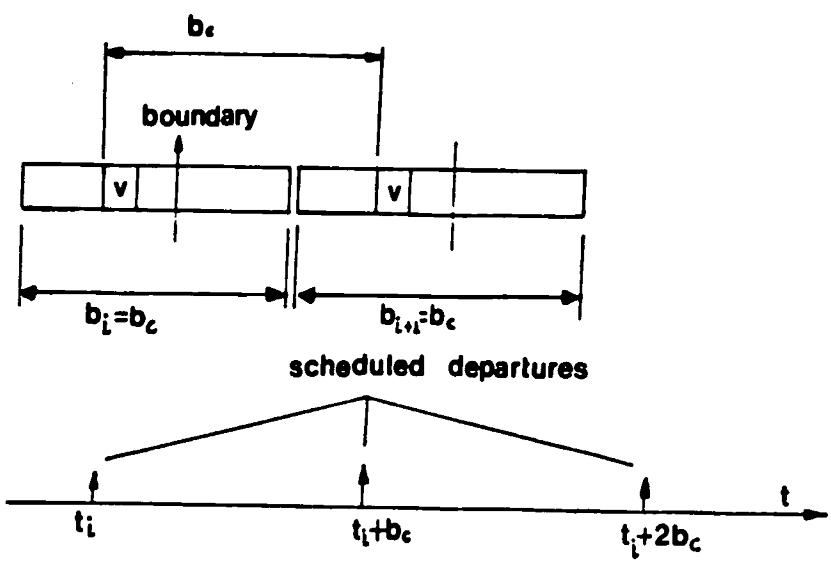

Figure 5-1: Consecutive frames of constant length, and scheduled departure times

However, in the variable frame scheme, the frames are not transmitted periodically, since their length varies. So, even if we have designed the average frame duration b to be equal to b_C secs, some of the transmitted frames may last longer than b_C secs (say $b'>b_C$). In that case, the

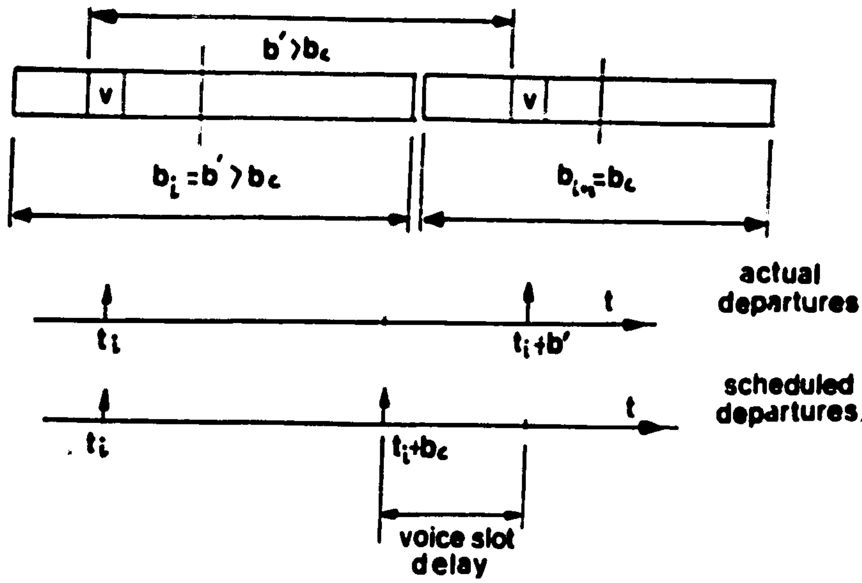

Figure 5-2: Consecutive frames of variable length, and actual departure times

voice slot will have to be buffered and be transmitted $(b'-b_c)$ secs later than its "scheduled" transmission (see figure). Even if the subsequent frames are shorter than b_c secs, the above situation may finally result in a loss of the transmission transparency (when the introduced extra delay of $(b'-b_c)$ secs exceeds a certain value). This can be avoided if the frame duration is bounded by some maximum value bmax=(N+S)τ which is appropriately chosen to allow the preservation of the transmission transparency. Furthermore, we consider the minimum frame length constraint that $b_{min}=S\tau$, and so the transmission of a new frame starts at least S slots after the start of the last frame. (If the last frame contains less than S slots, then same "idle" slots are transmitted until the completion of S slots).

So finally the length of any transmitted frame b_i satisfies the constraint:

$$S\tau \leq b_i \leq (N+S)\tau \qquad (5\text{-}3)$$

In the following sections we analyze the performance of the variable frame (movable boundary) scheme, using the models which were introduced by Maglaris et al [6].

5.2 Round Robin processor sharing, for voice traffic.

A. Voice traffic model

Here, in order to simplify the analysis of the system, we consider the case of a finite population of S voice customers. Each of them is assigned a dedicated slot, in the voice portion of the frame. If a certain customer is not active just before the start of the new frame, his slot will not be included in that frame. Since there are S available voice slots per frame, the generated voice calls will always receive service, and so the blocking probability will be zero:

$PL=0$ (5-4)

However, the transmission of a voice message may last longer than the original message duration, because of the variations of the frame length. So, it is important to evaluate the average transmission time (ETR) of voice messages, in order to know if it affects the voice quality and intelligibility.

Voice slots are assigned to active voice customers, according to the variable - frame slot allocation mechanism. This mechanism has the same properties as a processor - sharing algorithm ("round robin") where a processor serves sequentially each active customer for a short period of time (equal to one voice slot) until his transmission is completed.

So, we can model the voice switching process by a group of S voice customers permanently connected to a service box (figure 5.3); the service box serves the active customers according to the round robin service - sharing algorithm (which is represented by a loop inside the service box). the number of voice messages (S_i) in the service box is equal to the number of active voice customers: $S_i = v_i$. Let's call $\bar{V}$ and $\bar{D}$ respectively the average number of voice and data customers per frame. Now we assume the following:

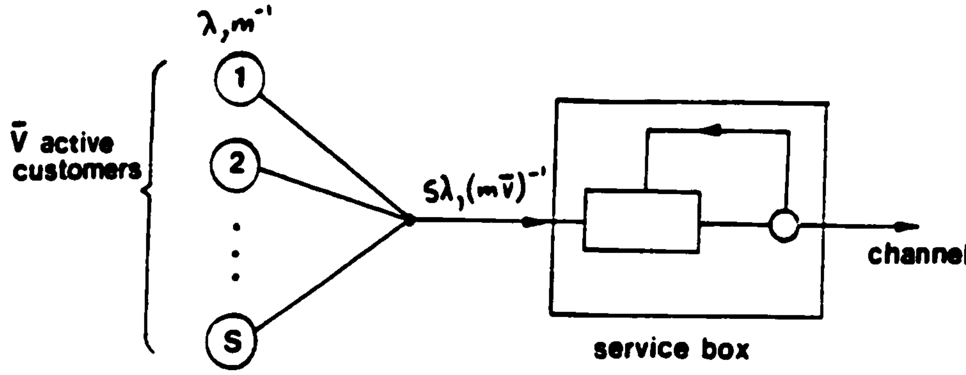

Figure 5-3:
Model for the switching of voice traffic

- Each of the S voice customers generates new calls, according to a Poisson distribution of rate λ (on the average, λ new messages are generated per second, per customer). So totally, the service box receives on the average $S\lambda$ new messages per second. This is equivalent to Poisson arrivals at the service box, with rate $S\lambda$.

- The length of the voice messages which are generated by each customer is exponentially distributed, with mean m^{-1} slots (so on the average m messages are generate per generated slot, per customer). Since the average number of active voice customers is $\bar{V}$, then totally $\bar{V}m$ messages are terminated per generated slot, among the messages which totally arrive at the service box. Equivalently, the length of the messages which totally arrive at the service box is exponentially distributed, with mean $(m\bar{V})^{-1}$ slots.

- Voice population: S customers.

- Voice buffer capacity: S customers (for buffering the generated voice packets, until the start of a new frame).

B. Voice traffic analysis

In the following analysis, we consider two parts:

1. Evaluate the probability distribution for the number of (active) voice customers in the system (that is also needed for the data traffic analysis), and

2. evaluate the average transmission time, ETR.

PART 1.

It is known that a finite population (S) round robin service sharing algorithm(Figure 5.4) has the following property:

- If the processor serves the arriving messages sequentially, in infitisimal time slices, and

- new messages arrive according to a Poisson process of rate λ_R, and

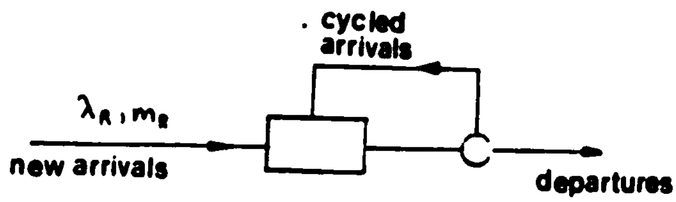

Figure 5-4:
Round Robin processor

- the duration of the new arriving messages is exponentially distributed with mean m_R^{-1} seconds,

then the number of messages (M) in the processor follows the stationary probability distribution:

$$P(M=k)=\frac{S!}{(S-k)!}\left(\frac{\lambda_R}{m_R}\right)^k \Bigg/ \left[\sum_{i=0}^{S}\frac{S!}{(S-i)!}\left(\frac{\lambda_R}{m_R}\right)^i\right] \qquad (5\text{-}5)$$

Now let's apply (5-5) to the voice traffidc model that we described before, assuming that τ is very short, compared to $1/\lambda$:

According to our model, new messages arrive at the processor, at a Poisson process of rate $S\lambda$, and so $\lambda_R=S\lambda$

(5-6) Furthermore, new messsages that arrive at the processor have exponentially distributed lengths, with mean $(mV)^{-1}$ slots. But in (5-5) we need the mean duration of arriving messages, expressed in secs (let's call it $(m')^{-1}$) - that is, the actual time needed for their transmission.

As we know, each message may occupy only one slot per frame. Therefore, an average message $-(m\bar{V})^{-1}$ slots long- will be completely transmitted after $(m\bar{V})^{-1}$ frames. Therefore the expected transmission duration will be:

$$(m')^{-1}=(m\bar{V})^{-1}b$$

where b is the average duration of a transmitted frame (in secs). Since the average frame contains $\bar{V}$ voice customers, and $\bar{D}$ data customers, it will be:

$$b=(\bar{V}+\bar{D})\tau \text{ secs} \qquad (5-7)$$

So finally

$$(m')^{-1}=(m\bar{V})^{-1}(\bar{V}+\bar{D})\tau$$

and so

$$(m_R)^{-1}=(m')^{-1}=m^{-1}\cdot[(\bar{V}+\bar{D})/\bar{V}] \qquad (5\text{-}8)$$

Therefore, according to (5-5), the number of (active) voice customers in the i-th transmitted frame , v_i, follows thw distribution:

$$P[v=k]=\frac{S!}{(S-k)!}\left(\frac{S\lambda}{m'}\right)^k \Bigg/ \left[\sum_{i=0}^{S}\frac{S!}{(S-i)!}\left(\frac{S\lambda}{m'}\right)^i\right] \qquad (5\text{-}9)$$

where m_R is given by (5-8).

Now the average number of voice customers per frame will be:

$$V=\sum_{k=0}^{S} kP[v=k]$$

and by using (5-9), and the normalizing condition

$$\sum_{k=0}^{S} P[v=k]=1$$

we get

$$V=S-(m'/\lambda)(1-P[v=0]) \qquad (5\text{-}10)$$

However, if the voice traffic is heavy, it will be very unlikely to have no customers in the system, and so we may consider that $P[v=0]-0$.

So finally , from (5-10), (5-8) we get:

$$\bar{V} = S - \frac{m'\bar{V}}{\lambda \P (\bar{V}+\bar{D})} \tag{5-11}$$

PART 2.

The length of an average message generated by an active customer is equal to m^{-1} slots. This message needs m^{-1} frames for its transmission. Therefore the transmission duration will be $m^{-1}b$ secs. So using (5-7), we find that the average transmission time of a generated voice message will be:

$$ETR = m^{-1}b = m^{-1}\P(\bar{V}+\bar{D}) \text{ secs} \tag{5-12}$$

5.3 Discrete Time Model, for data traffic.

A. Data traffic model

The switching process for the data traffic is modelled as a discrete time queueing gate, with the following characteristics (Figure 5.5):

- Times of gate opening: Just before the start of a new frame.
  So the time between the i-th and the (i+1)-th opening is equal to the duration of the i-th frame:

$$(\P_g)_i = b_i = [v_i + d_i]\P \tag{5-13}$$

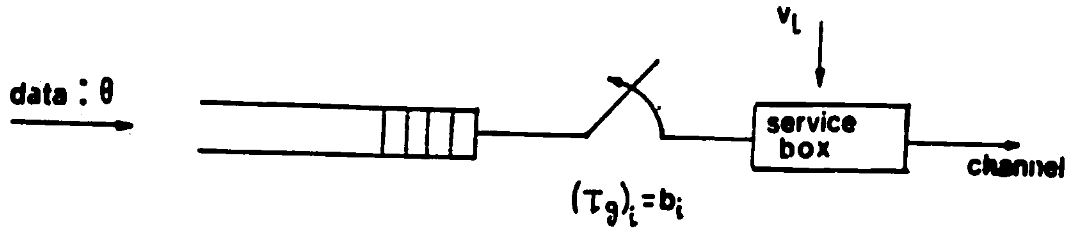

Figure 5-5:
Discrete time model for the data traffic

- Data arrivals distribution: Poisson, with arrival rate θ.
  Therefore, the probability of k data arrivals between the i-th, (i+1)-th opening will be:

$$(n_k)_i = e^{-\theta b_i}[(\theta b_i)^k/k!] \quad \text{for} \quad k=0,1,2,\ldots$$

$$(n_k)_i = 0 \quad \text{otherwise.}$$

Now let's call

$n_{i/k}$: The probability of i data arrivals between two openings, under the condition that the transmitted frame (between these openings) contained k data customers.
In general, such a frame will contain j voice customers, with $0 \leqslant j \leqslant S$. Therefore, its duration will be $(k+j)\tau$ secs, with $k\tau \leqslant (j+k)\tau \leqslant (S+k)\tau$ secs and so (unconditioning over j) we get:

$$n_{i/k} = \sum \Pr[v=j].e^{-\theta\tau(j+k)}[\theta\tau(j+k)]^i \qquad (5\text{-}13)$$

- Data customers holding time distribution: Deterministic.

- Probability that an occupied slot will still be occupied after x secs:

$p_H=1$ for $x<b$ and

$p_H=0$ for $x\geq b$

- Data population: Infinite

- Data buffer capacity: Infinite.
  Furthermore, according to the results for the voice traffic, it will be:

- Distribution of the number of active voice customers: $P[v=k]$
  (as given by (5-9).

- Voice population: Finite (S)

- Voice buffer capacity: S

- Correlation assumptions: we assume that the number of data customers in the system, and the number of active voice customers are statistically independent (uncorrelated). As we have mentioned in the previous sections, this assumption is not realistic, under heavy traffic conditions. However, if we consider slowly varying voice traffic, we can neglect the error which is introduced by the above assumption.

B. Data traffic analysis

When the system reaches its equilibrium state, the average number of data packets (slots) per frame, D, will be equal to the average number of arriving packets, during the duration of an average frame b. Therefore using (5-7), and because the average data arrival rate is ●, we get:

$\bar{D}$=●b=●¶($\bar{V}+\bar{D}$) (5-14)

Now we can solve the system of equations (5-11)-(5-14), and find:

$\bar{V}$=S-[m(1-●¶)]/(λ¶)

and (5-15)

$\bar{D}$=(●¶$\bar{V}$)/(1-●¶), with $D<N$

Let's call:

ρ_1=(λ¶S/m) the voice traffic intensity, and

ρ_2=●(N+S)¶ the data traffic intensity.

Now, we can write (5-15) as follows:

$$V=S-S\left[\frac{(N+S-\rho_2)}{(N+S)\rho_1}\right]$$

and (5-16)

$$D=\frac{V.\rho_2}{N+S-\rho_2}$$

So finally we can determine the values of b, ETR from (5-7)-(5-12):

$$b=¶S.\left[\frac{N+S}{N+S-\rho_2}-\frac{1}{\rho_1}\right] \quad (5\text{-}17)$$

and

$$\bar{s}TR = m^{-1} q S \left[\frac{N+S}{N+S-\rho_2} - \frac{1}{\rho_1} \right] \qquad (5\text{-}18)$$

The state probabilities of the system are:

$p_{j,n} = Pr[Q_n = j]$: Probability of j data packets in the system, just before the n-th opening.

Now, let's consider two successive gate openings (Figure 5.6) and call:

t_n: Time of the n-th opening.

Q_n: Number of data packets in the system, just after t_n.

d_n: Number of data packets in the n-th frame.

f_n: Number of data arrivals, between t_{n-1}, t_n.

The number of transmitted packets between t_{n-1}, t_n will be Q_{n-1} if $Q_{n-1} < N$, or N if $Q_{n-1} \geq N$. Therefore, $(Q_{n-1} - N)^+$ packets will remain in the system, and if we also count the new data arrivals, we get:

$$Q_n = (Q_{n-1} - N)^+ + f_n \qquad (5\text{-}19)$$

We can easily see that (5-19) is the same as the expression (4-12), which describes the FFFB scheme. The only difference is that here the number of data arrivals

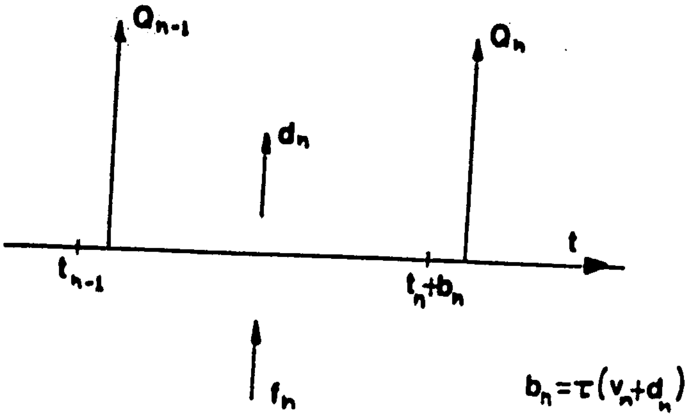

Figure 5-6:
The function of the gate, for data traffic

f_n is not constant (like f), but depends on the duration of the n-th frame, b_n. Consequently, the equation (4-15) (which describes an M/D/N queueing system) will also be true for the VFMB scheme. However, instead of the probabilities of k new data arrivals, n_k, (which do not depend on b_n, now we must use the probabilities n_{k/d_n} (which depend on b_n). Therefore, we have:

$$p_{j,n}=\sum_{k=0}^{N-1} p_{k,n-1}n_{j/d_n}+\sum_{k=N}^{N+j} p_{k,n-1}n_{j-k+N/d_n}$$

(5-20)

But the number of transmitted packets d_n will be: $d_n = Q_n$ if $Q_n < N-1$, or $d_n = N$ if $Q_n \geq N$.

Therefore, (5-20) takes the form:

$$p_{j,n} = \sum_{K=0}^{N-1} p_{k,n-1} n_{j/k} + \sum_{K=N}^{N+j} p_{k,n-1} n_{j-k+N/N} \quad (5\text{-}21)$$

Now we can apply the general method that we have described, in order to get the generating function of $p_{j,n}$, $P_n(z)$. Finally, for n->oo we find:

$$P(z) = \frac{\sum_{K=0}^{N-1} [z^N N_k(z) - z^k N_N(z)] p_k}{z^N - N_N(z)} \quad (5\text{-}22)$$

Here, the only unknown is $N_k(z)$, which is the generating function of tne probabilities $n_{j/k} = Pr[f_n = j/d_n = K]$.

But the number of arrivals (f_n) between two openings consists of two terms: The number of arrivals f_v during the "voice" portion of the n-th frame, and the number ot arrivals f_d during the "data" portion of the frame. So

$f_n = f_v + f_d$.

Since the random variable f_n is the sum of the random variables f_v, f_u, consequently, its generating function will be the product of the generating functions of these two variables. Therefore we can write:

$$N_k(z)=N_{kv}(z).N_{kd}(z)$$

But $N_{kd}(z)=e^{\lambda\tau(z-1)k}$, and

$$N_{kv}(z)=P^V(z')$$

with $z'=e^{\lambda\tau(z-1)}$, and $P^V(z)=\sum_{i=0}^{S} Pr[v=i]z^i$

So finally it will be:

$$N_k(z)=e^{\lambda\tau(z-1)}P^V[e^{\lambda\tau(z-1)}] \qquad (5\text{-}23)$$

Now that we have determined $N_k(z)$, the expression (5-22) will give us the generating function P(z). Notice however that the N probabilities p_k $k=0,1,...,N-1$ are still unknown. But P(z) must be analytic for $|z|\leq 1$. Furthermore, the equation

$$z^N-N_N(z)=0$$

has N-1 roots: z_r $r=1,2,...,N-1$ with $|z_r|<1$.

Therefore, the numerator of P(z) must also vanish at the points $z=z_r$, and the following N-1 equations must be satisfied:

$$\sum_{k=0}^{N-1} [N_k(z)-z_r^{\,i}]p_k=0$$
with r=1,2,...,N-1 (5-24)

Furthermore, we have the normalizing condition

$$P(1)=1 \qquad (5\text{-}25)$$

So the solution of the N equations (5-24),(5-25) determines the desired N unknowns.

Now that P(z) is completely determined by (5-22), the average number of data packets in the system will be:

$$L=\left.\frac{dP(z)}{dz}\right|_{z=1} \qquad (5\text{-}26)$$

Finally, the expected waiting time for the data packets will be:

$$EW=(b/2)+\lambda^{-1}L \quad (5\text{-}27)$$

where b is given by (5-8), and L is given by (5-26)

6. THE RESULTS OF THE PERFORMANCE ANALYSIS

Here we present the numerical results for the performance evaluation of the various integrated transmission schemes, according to the analysis of the models that we described in the last chapters.

6.1 Voice traffic

As we have mentioned before, the probability of loss for voice customers, PL, is the same for FFFB, and the FFMB schemes because in both of them, voice is treated as a high priority class. Furthermore, as we said, for small values of the voice traffic intensity ρ_1, PL is given by the Erlang-B formula (4-30), which is a very good approximation to the actual state probabilities of the voice traffic , which satisfy the system of equations (4-3)-(4-7)

In Figure 6.1 we can see the values of PL for the different values of ρ_1, according to the Erlang-B formula. We consider a fixed frame structure where S=10 N=5. (The data traffic intensity, ρ_2, may have any value, since it does not affect PL - because voice has priority over data.

As we expected, the values of PL are increasing, as ρ_1 increases This happens because more customers are lost when more customers arrive at the system (and occupy more of the S available voice slots.

In Figure 6.2 we can see the values of PL for the different values of N, that is for the different positions of the boundary, within the frame. We assume that the total length of the frame is always the same N+S=25 slots. As we expected, PL is increasing as N increases. This happens because of the decrease in the number of available voice slots (S=25-N)

6.2 Data Traffic

A. Fixed Frame scheme

In the Figures (6.3)-(6.5) we can see the expected waiting times EW for data customers, as a function of ρ_2. We consider the cases of fixed and movable boundary, with S=10, N=5. The duration of the frame is b=.01 secs, and the values of EW are normalized over b. The voice traffic intensity is fixed, $\rho_1=5$ (and corresponds to PL=.018).

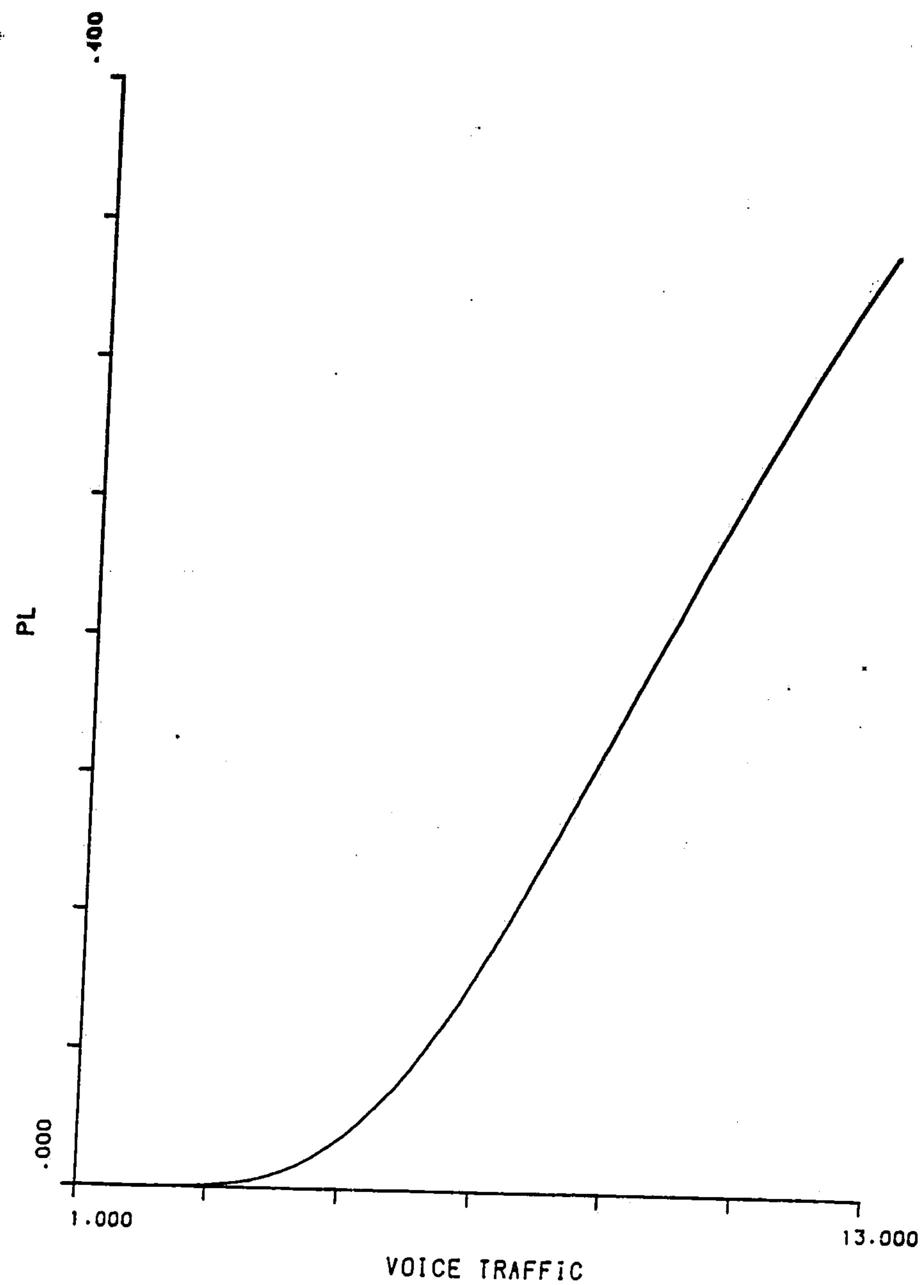

Figure 6-1:
Probability of loss, for the FFFB, FFMB schemes.
S=10, N=5, ρ_2=5.

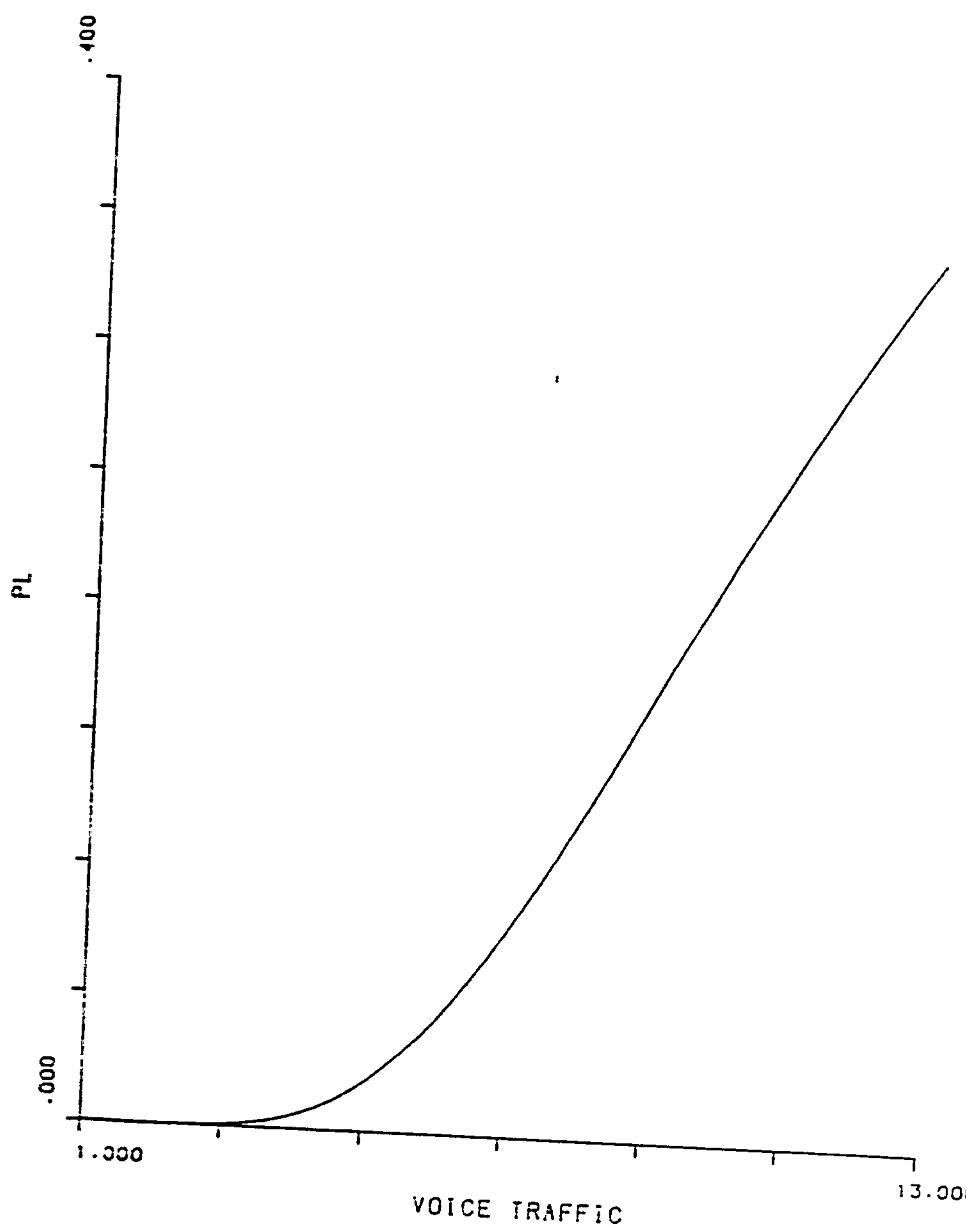

Figure 6-1:
Probability of loss, for the FFFB, FFMB schemes.
S=10, N=5, ρ_2=5.

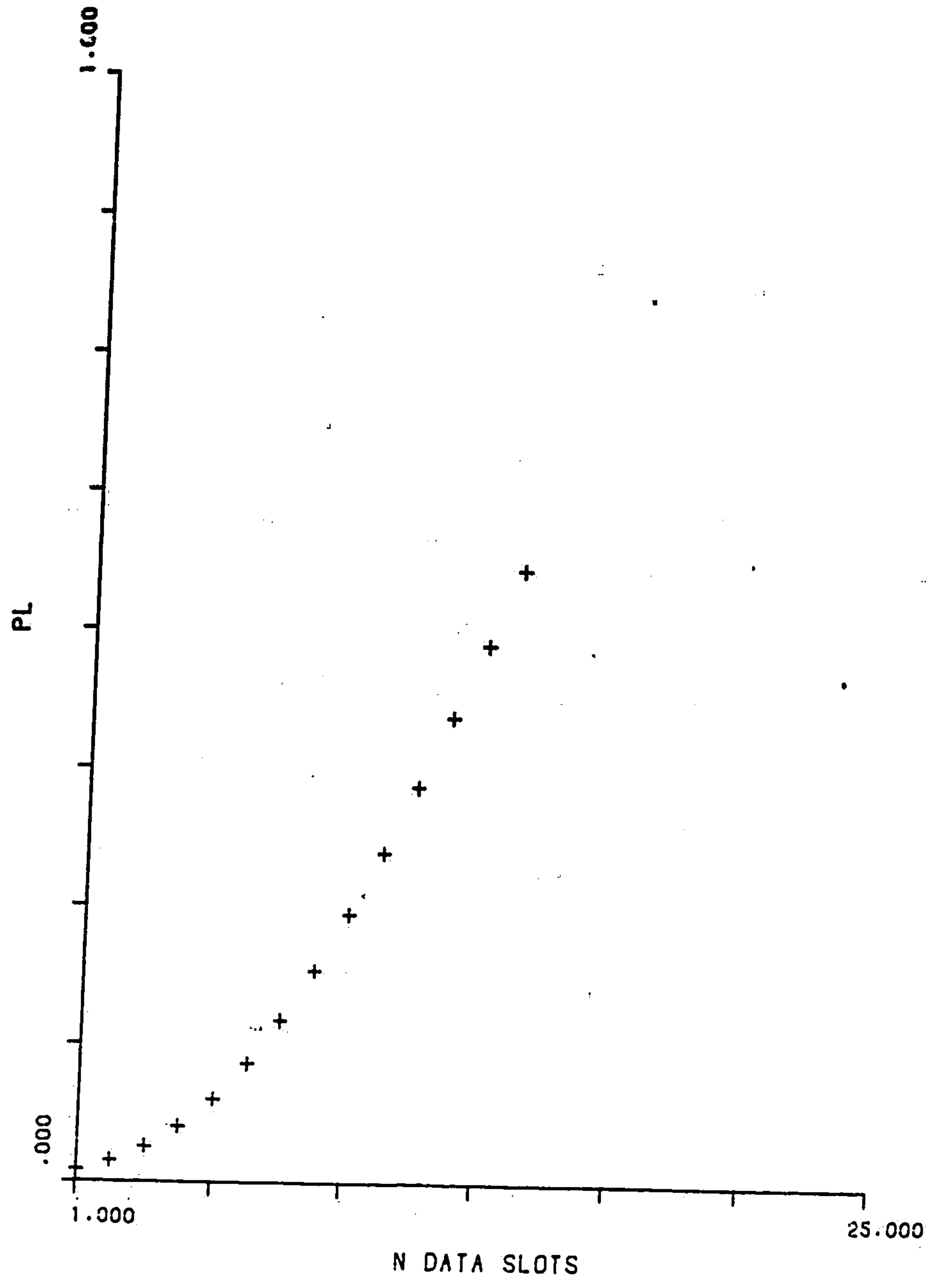

Figure 6-2:
Probability of loss, versus boundary position.
N+S=25, ρ_1=5, ρ_2=7.

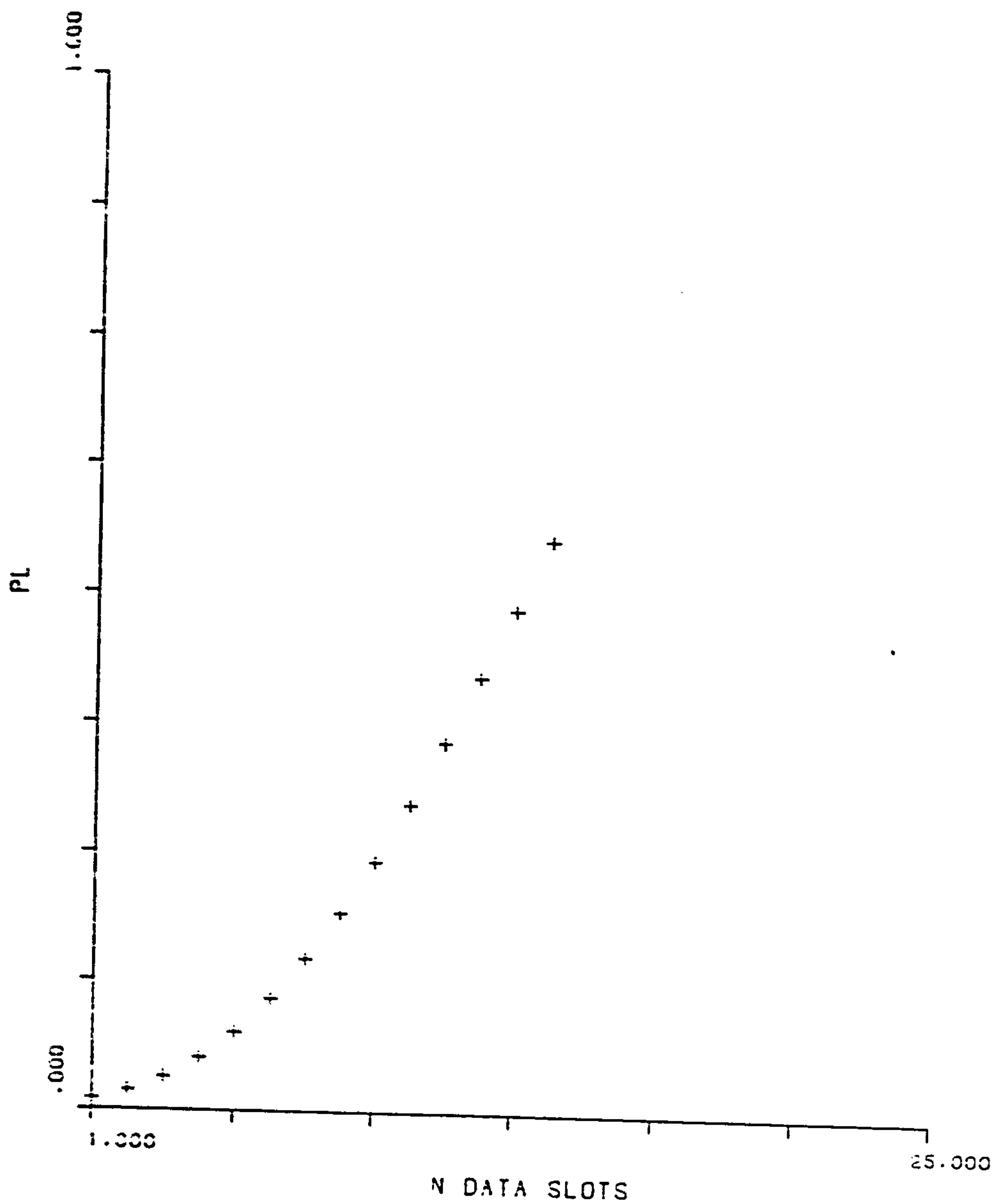

Figure 6-2:
Probability of loss, versus boundary position.
N+S=25, ρ_1=5, ρ_2=7.

In Figure 6.3, we show the results from the analysis of the discrete time model. These are according to the expressions (4-19)-(4-24). As we see, in the FFFB scheme EW goes to infinity, as ρ_2 approaches 5 (=N), which is the portion of the link capacity that is available to data customers. In the FFMB scheme however, EW remains low untill ρ_2 approaches 10 (=N+S), which is the maximum possible available capacity for data customers. The delay in the FFMB scheme is lower than the delay in the FFFB scheme. This happens because in the FFMB scheme there are more available slots for the data customers. Nevertheless, the improvement of the delay is not actually that big, as we will see in a while.

In the Figures 6.3, 6.4 we show the results from the analysis of the simplified discrete time model, under the same conditions as before. Here again we can see a similar performance of the system, with $EW \rightarrow \infty$ as $\rho_2 \rightarrow 5$ for the FFFB case, and $EW \rightarrow \infty$ as $\rho_2 \rightarrow 10$, for the FFMB case. We notice that the values of EW for a certain value of ρ_2 in the Figures 6.4, 6.5, are slightly lower than the corresponding of EW in Figure 6.3. This happens because of the approximations involved in the consideration of the simplified discrete time model. Actually, in that model

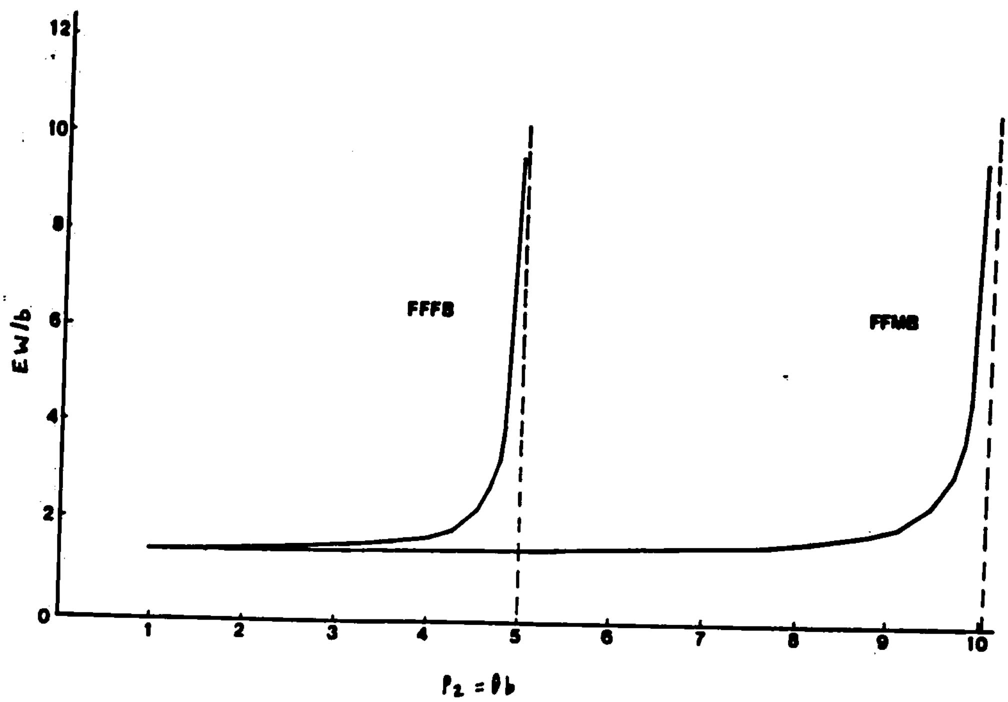

Figure 6-3:
Waiting time for the FFFB, FFMB. S=5, N=10, ρ_1=5.
(Discrete time model).

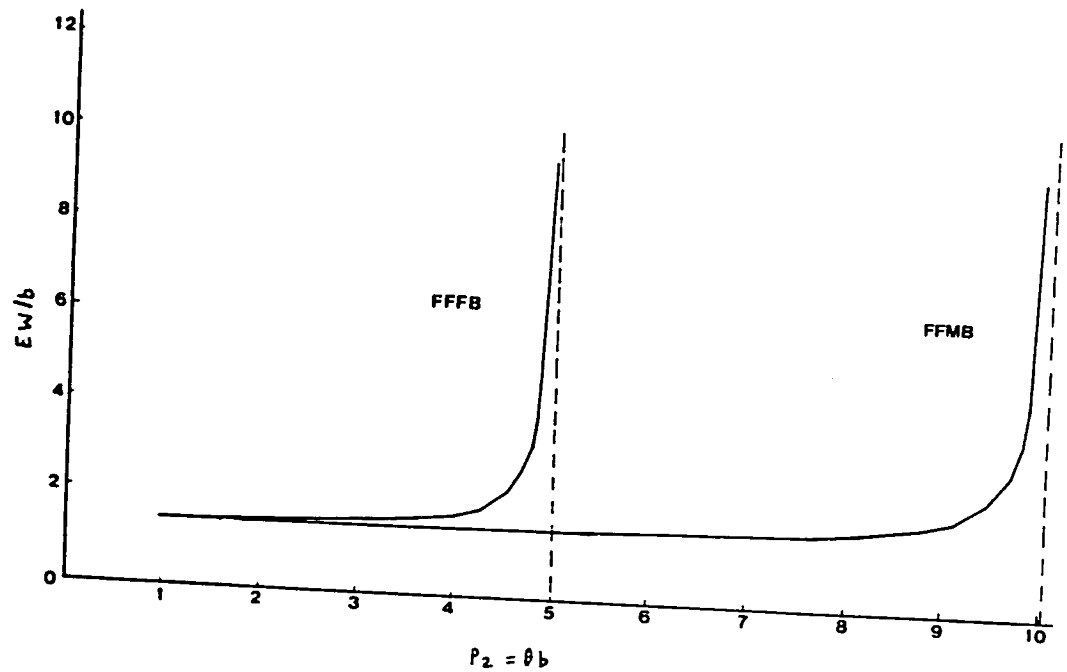

Figure 6-3:
Waiting time for the FFFB, FFMB. S=5, N=10, ρ_1=5.
(Discrete time model).

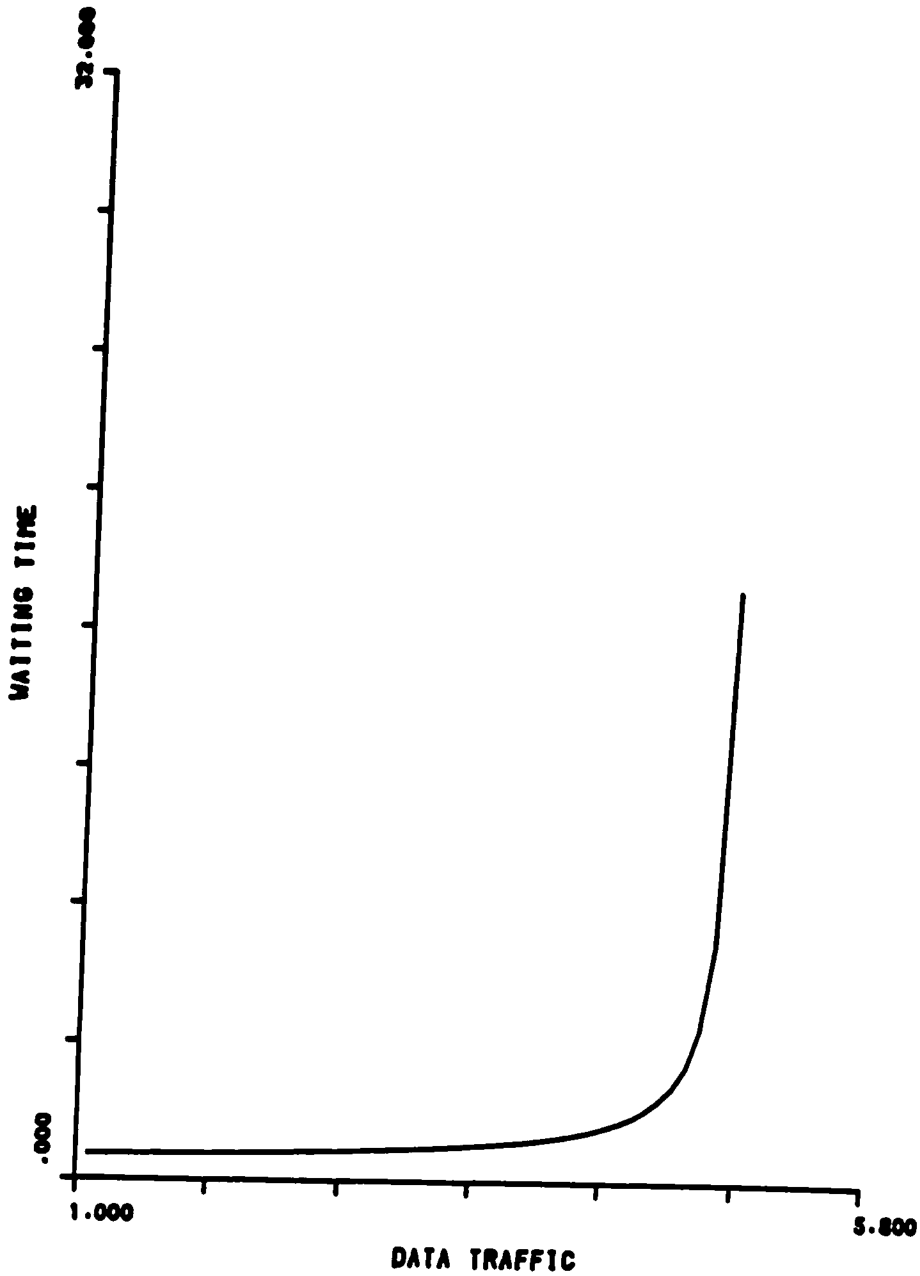

Figure 6-4:
Waiting time for the FFFB. S=10, N=5, ρ_1=5.
(Simplified discrete time model).

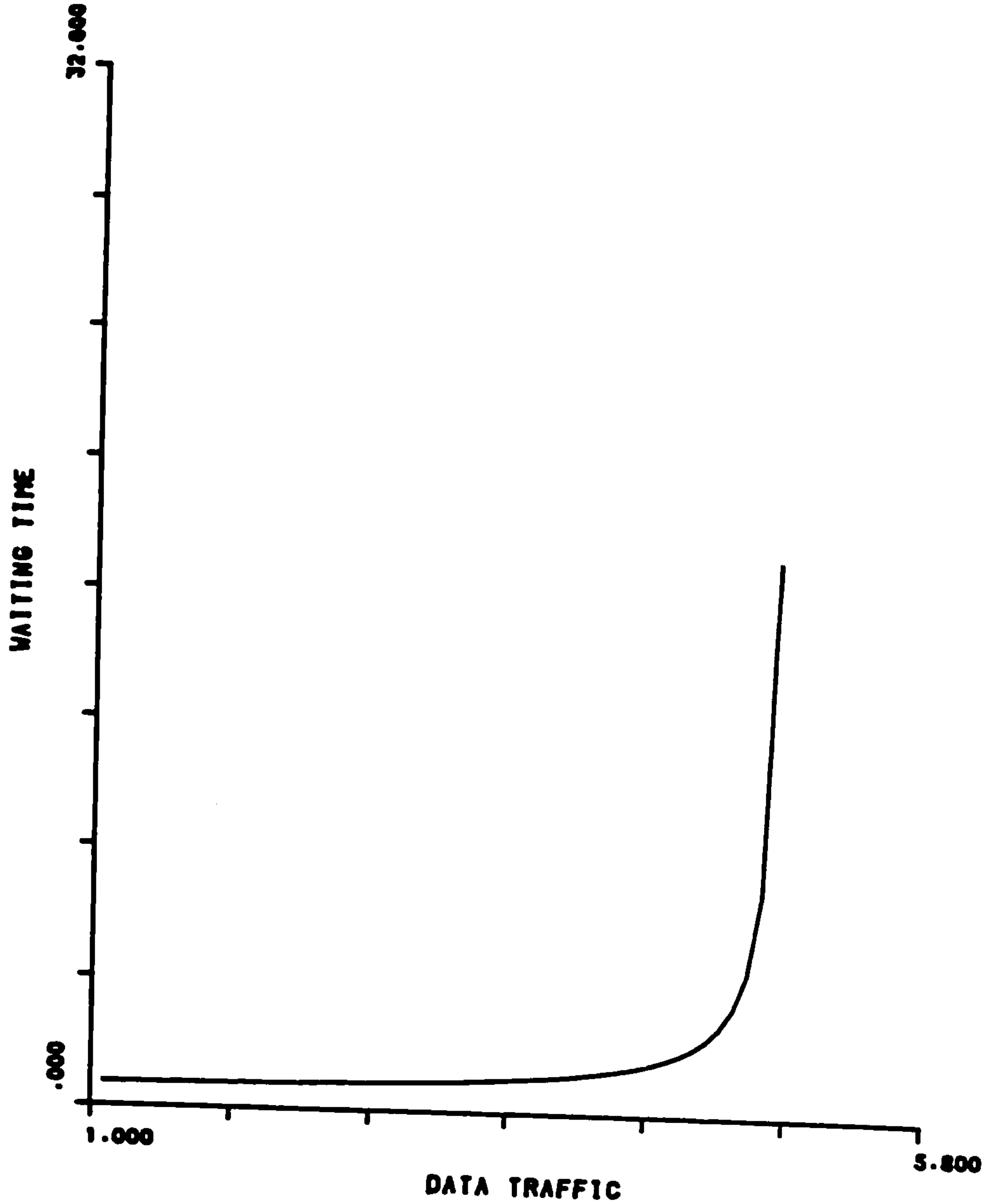

Figure 6-4:
Waiting time for the FFFB. S=10, N=5, ρ_1=5.
(Simplified discrete time model).

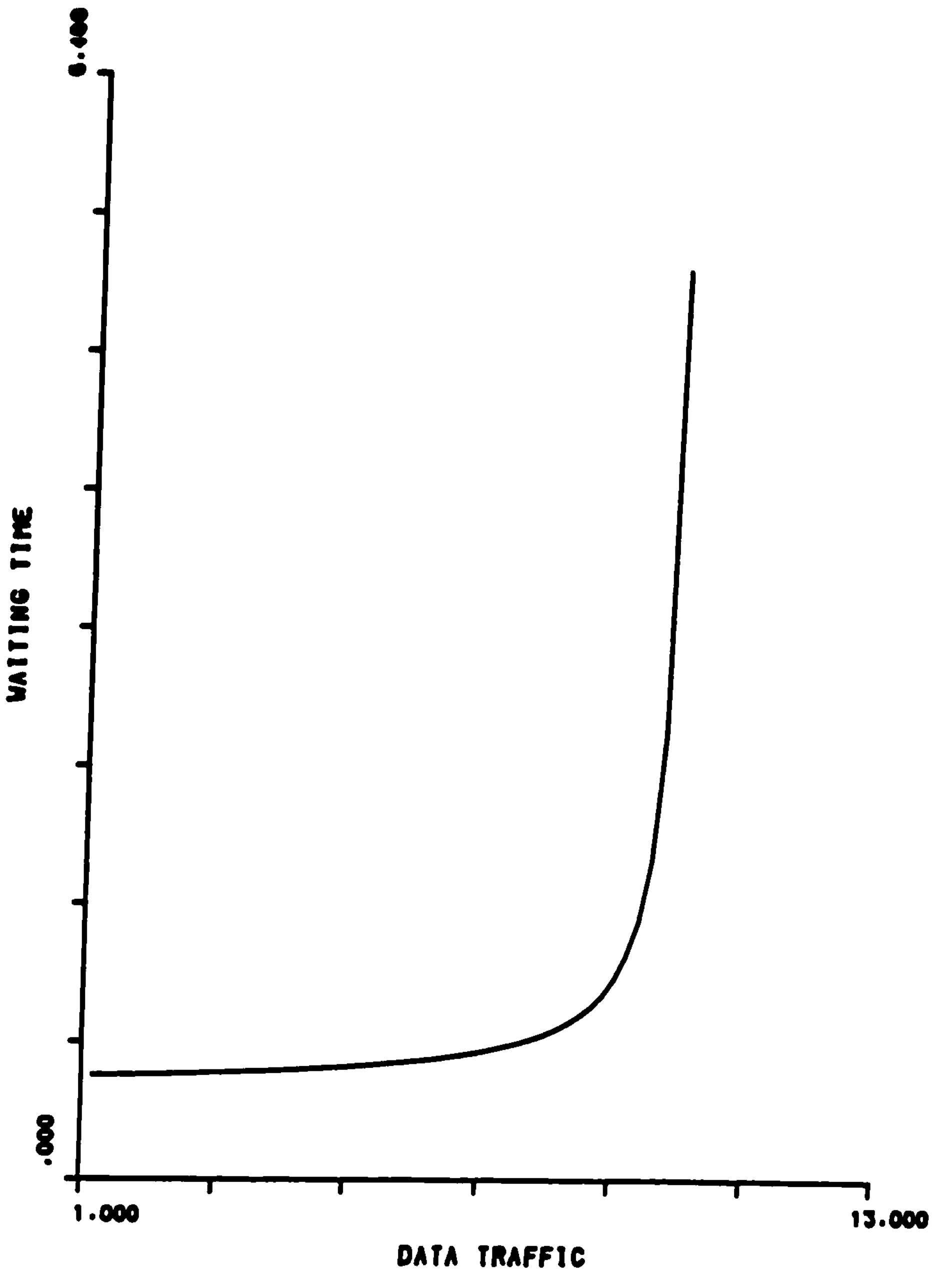

Figure 6-5:
Waiting time for the FFMB. S=10, N=5, ρ_1=5.
(Simplified discrete time model).

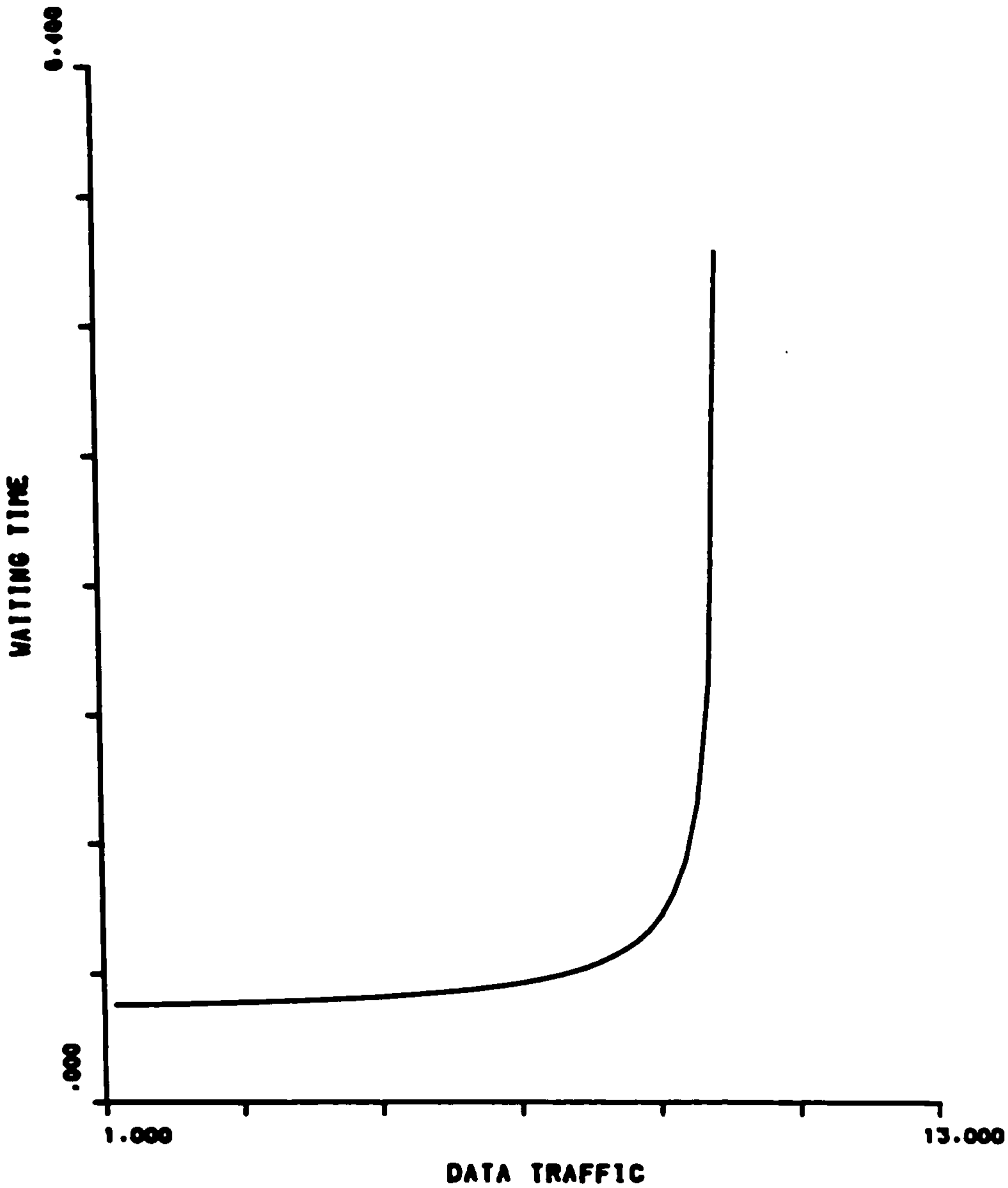

Figure 6-5:
Waiting time for the FFMB. S=10, N=5, ρ_1=5.
(Simplified discrete time model).

we consider that the system is available to the arriving customers at the start of every new slot. However actually the system is available only at the start of every new frame: Even if the N data slots are not occupied, an arriving customer will have to wait until the next frame in order to be transmitted. This waiting time is not taken into account in the above model ,(where the customer waits only until the next slot) and so the result is slightly lower values for EW.

In both Figures 6.3, 6.4, the FFFB case is described accurately. However, this is not true for the FFMB case: The improvement of the delay of the FFMB over FFFB is not actually as big as indicated by these figures. This happens because in the discrete time model, and the simplified discrete time model we have assumed that the numbers of voice and data customers are uncorrelated. This is true for the case of fixed boundary, but it is not a realistic assumption for the case of movable boundary, especially under heavy traffic conditions.

This fact is illustrated very clearly by the simulation results for the FFMB, under the same conditions. These results are shown in Figure 6.6, as

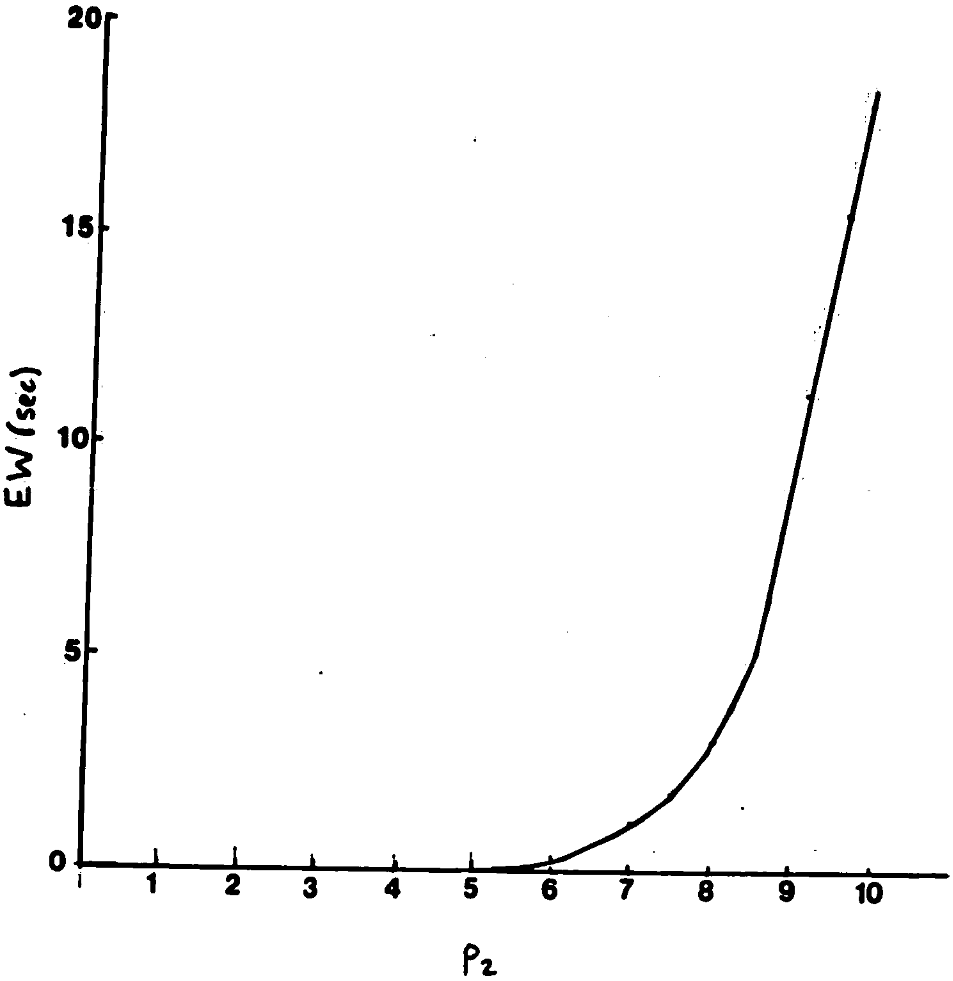

Figure 6-6:
Waiting time for the FFMB. S=5, N=10, ρ_1=5, ml^{-1}=100 secs
(Simulation results)

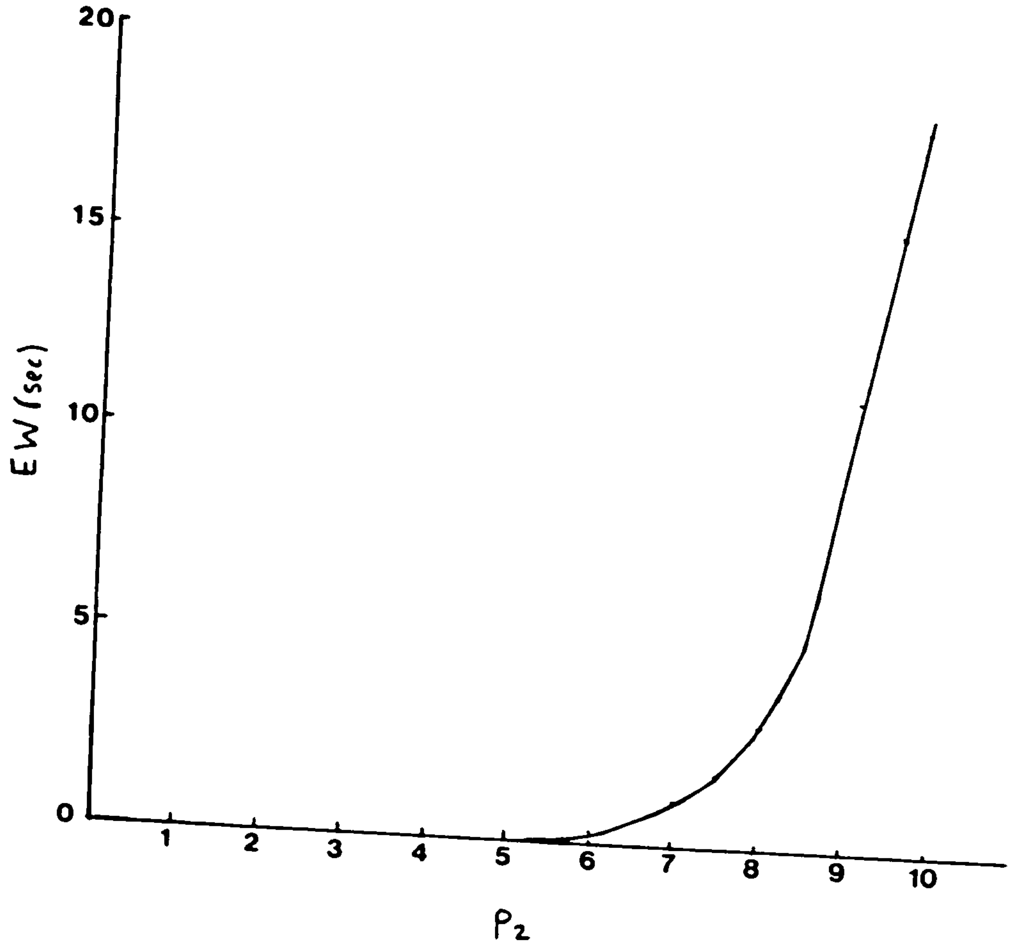

Figure 6-6:
Waiting time for the FFMB. S=5, N=10, ρ_1=5, $m1^{-1}$=100 secs
(Simulation results)

they were presented in [5],[6]. Specifically, we can see that the actual value of EW for ρ_2=9 is 5 secs, according to the simulation. But the corresponding value of EW according to the above two models is in the order of 2 msecs, which is 250 times lower than the actual value. Therefore these two models cannot describe the FFMB under heavy traffic conditions.

However, the correlation of voice and data customers is taken into account in the continuous time model, which gives results in agreement with the above simulation [5],[6]. Therefore, the continuous time model provides an exact description of the FFMB case. As we have mentioned before, the analytical solution of that model is not easy to evaluate (because of the complexity of the system of equations (4-43a)-(4-43f). However, for the simple case of N=0, S=1, the expression (4-5) provides a closed form solution for EW. So we can compare it with the corresponding solutions of the discrete time models, which is given by (4-41), for both of them. With this comparison, we want to find out whether under certain conditions the discrete time models can still describe the system correctly. So in the Figures 6.6-6.10 we present a comparison of the discrete time models versus the continuous time model (which gives the right solution): We consider the FFMB scheme, with S=1, N=0, ρ_1=.6. Since ρ_1=.6=λ/m_1, we consider three different values of m_1: m=100, 1000, 10000.

We notice that the values of EW=EW_D for the discrete time models are not affected by the variations of the parameter ρ_1 (EW_D depends only on ρ_1, ρ_2). For the continuous model however, the values of EW=EW_C change drastically, for the different values of m_1: For example, for ρ_2=1, it is EW_C=EW_D=.8 secs when m=10000, but EW_C=42 secs (=125EW_D) when m=10. Therefore we conclude that when the traffic consists of short messages (m^{-1})=10 then the discrete time models do not describe the system accurately, and we have to use the continuous time model instead. Based on this conclusion we are convinced that the use of a discrete time model for the analysis of the VFMB scheme, with m=10000 (as we will see), will be accurate enough.

B. Variable Frame scheme

In Figure 6.11 we can see the expected transmission time ETR, versus ρ_1, for the VFMB scheme, where S=10,

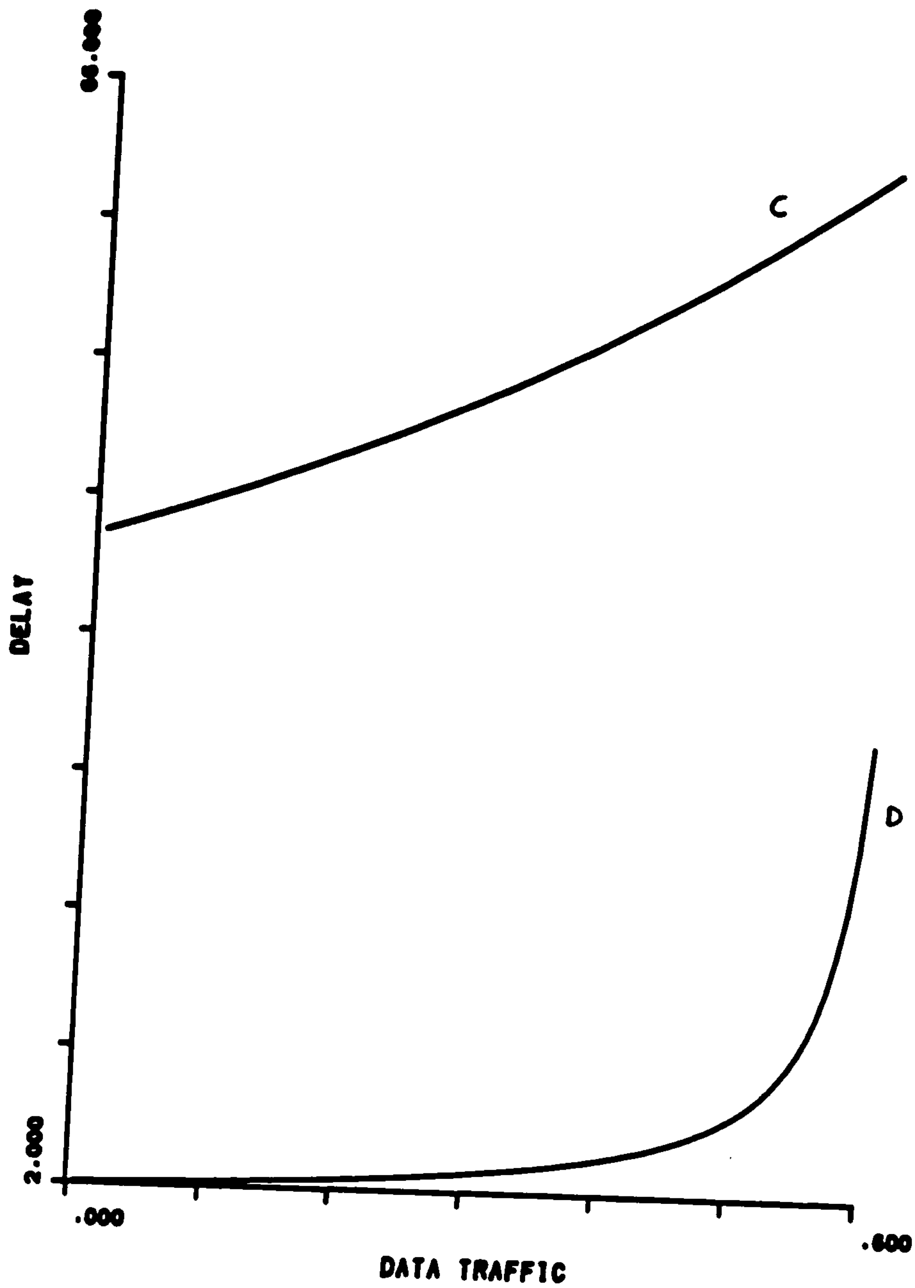

Figure 6-7:
FFMB scheme, continuous - discrete time models.
S=1, N=0, $\rho_{\perp}$=.6, m=100

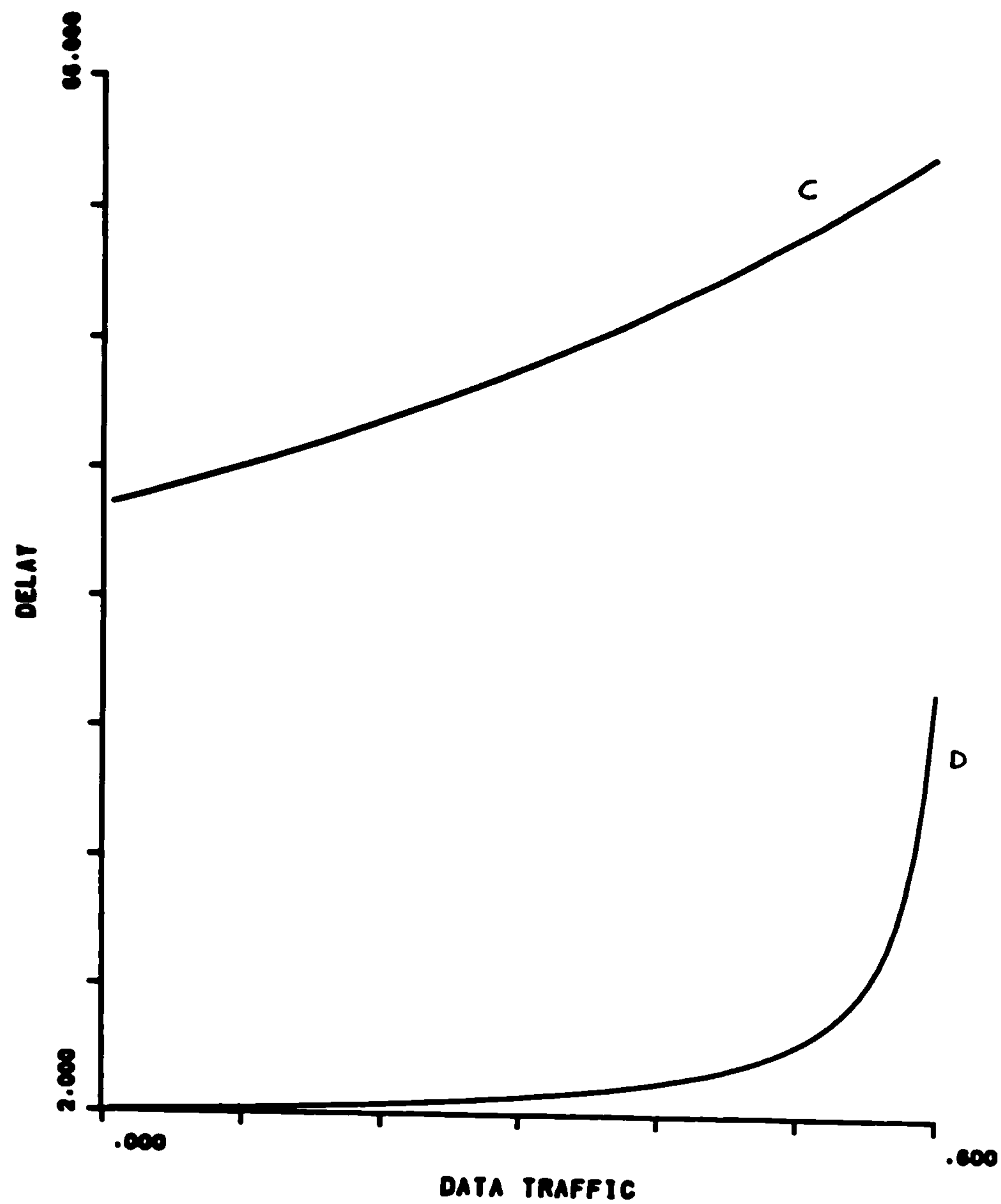

Figure 6-7:
FFMB scheme, continuous - discrete time models.
S=1, N=0, ρ_1=.6, m=100

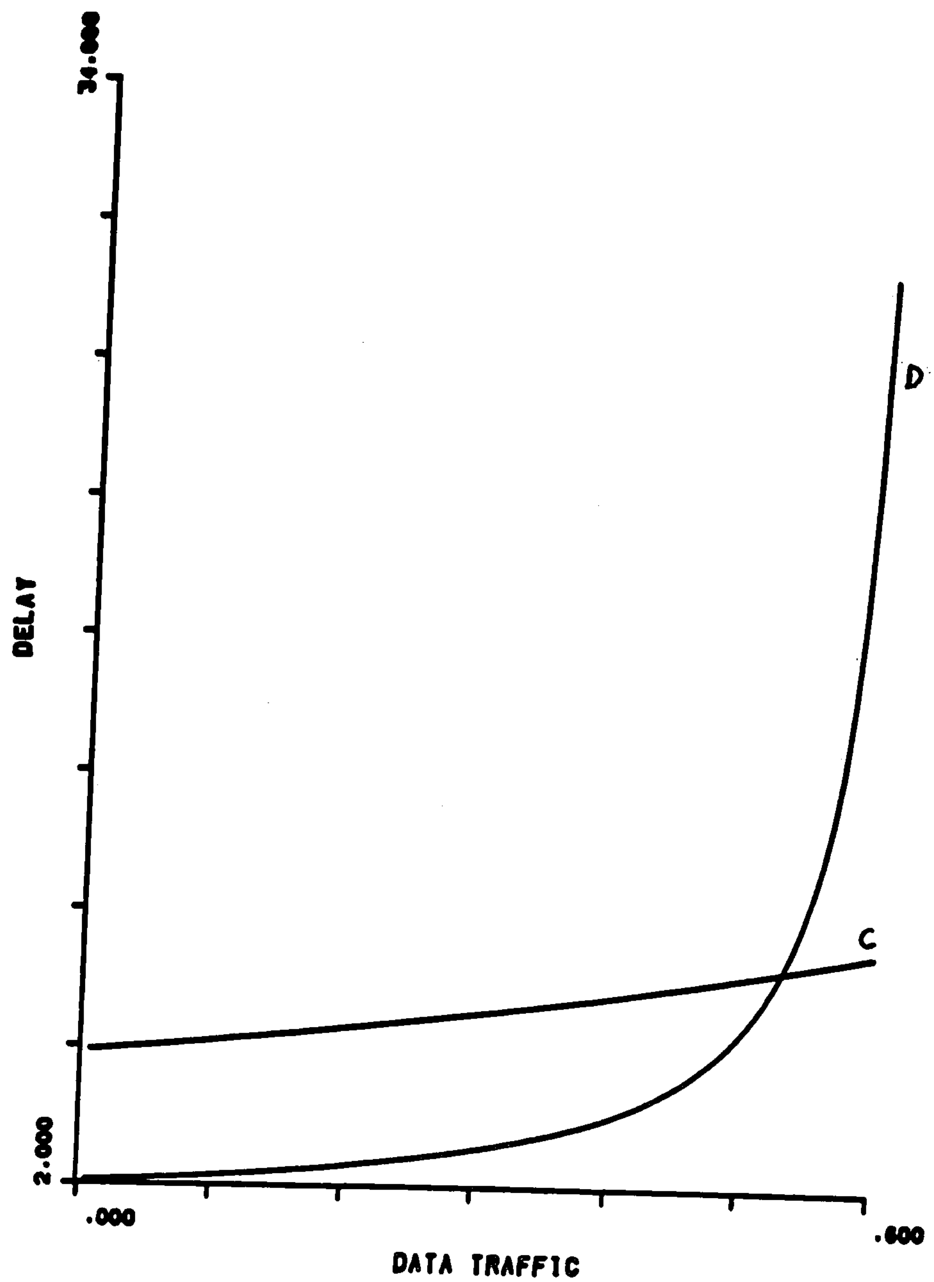

Figure 6-8:
FFMB, continuous - discrete time models.
S=1, N=0, $\rho_\perp$=.6, m=1000

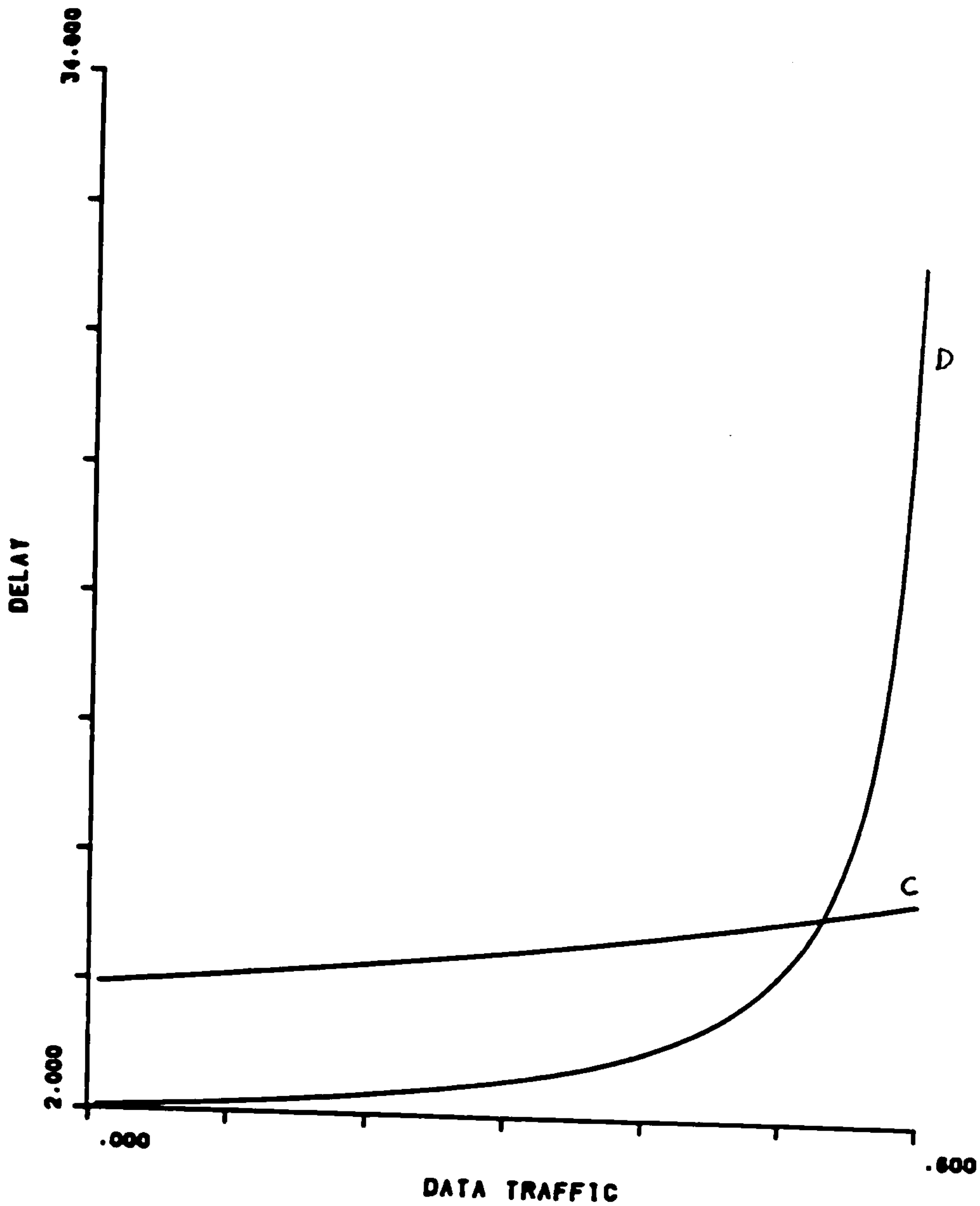

Figure 6-8:
FFMB, continuous - discrete time models.
S=1, N=0, ρ_1=.6, m=1000

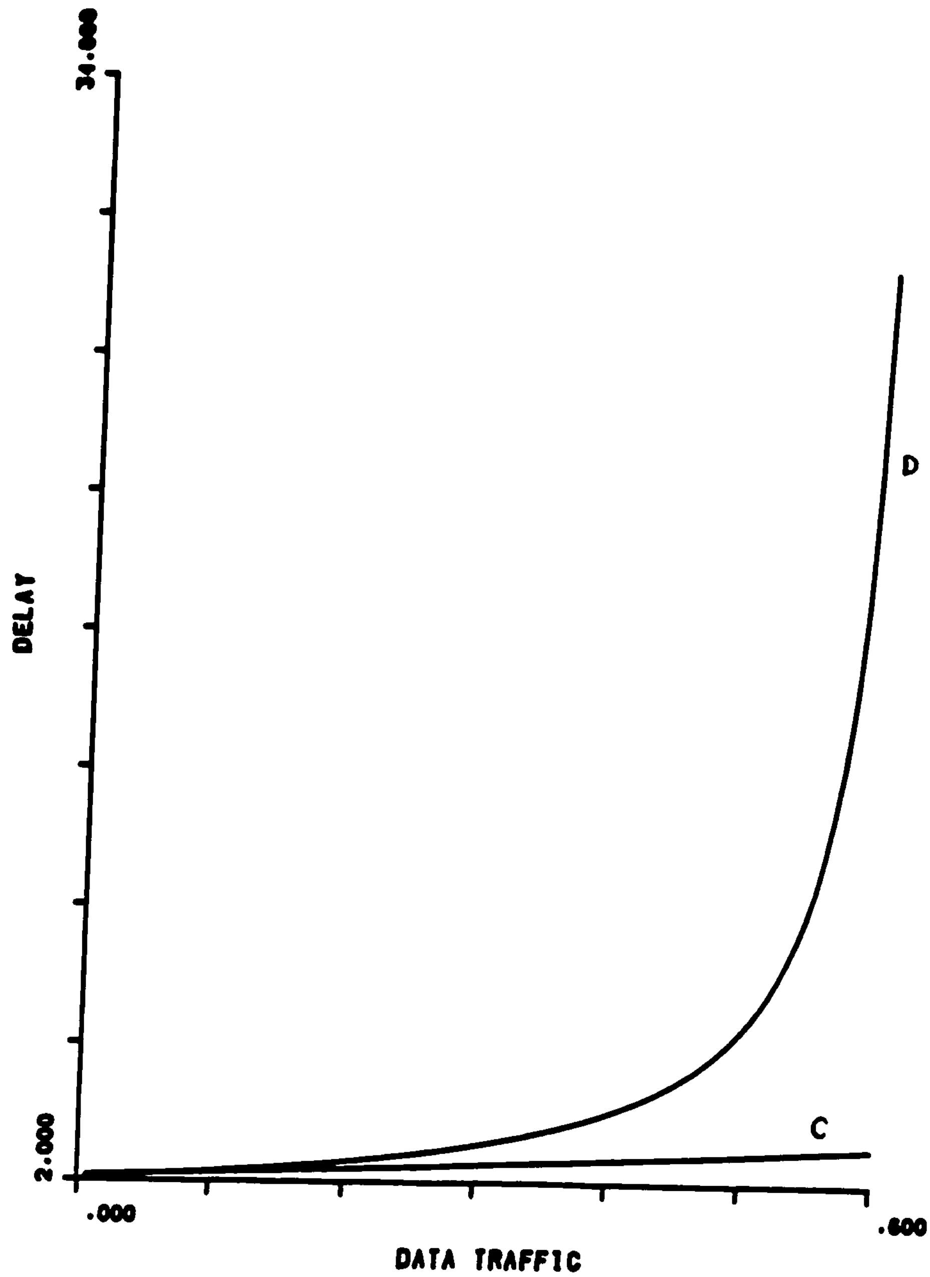

Figure 6-9:
FFMB, continuous - discrete time models,
S=1, N=0, ρ_1=.6, m=10000

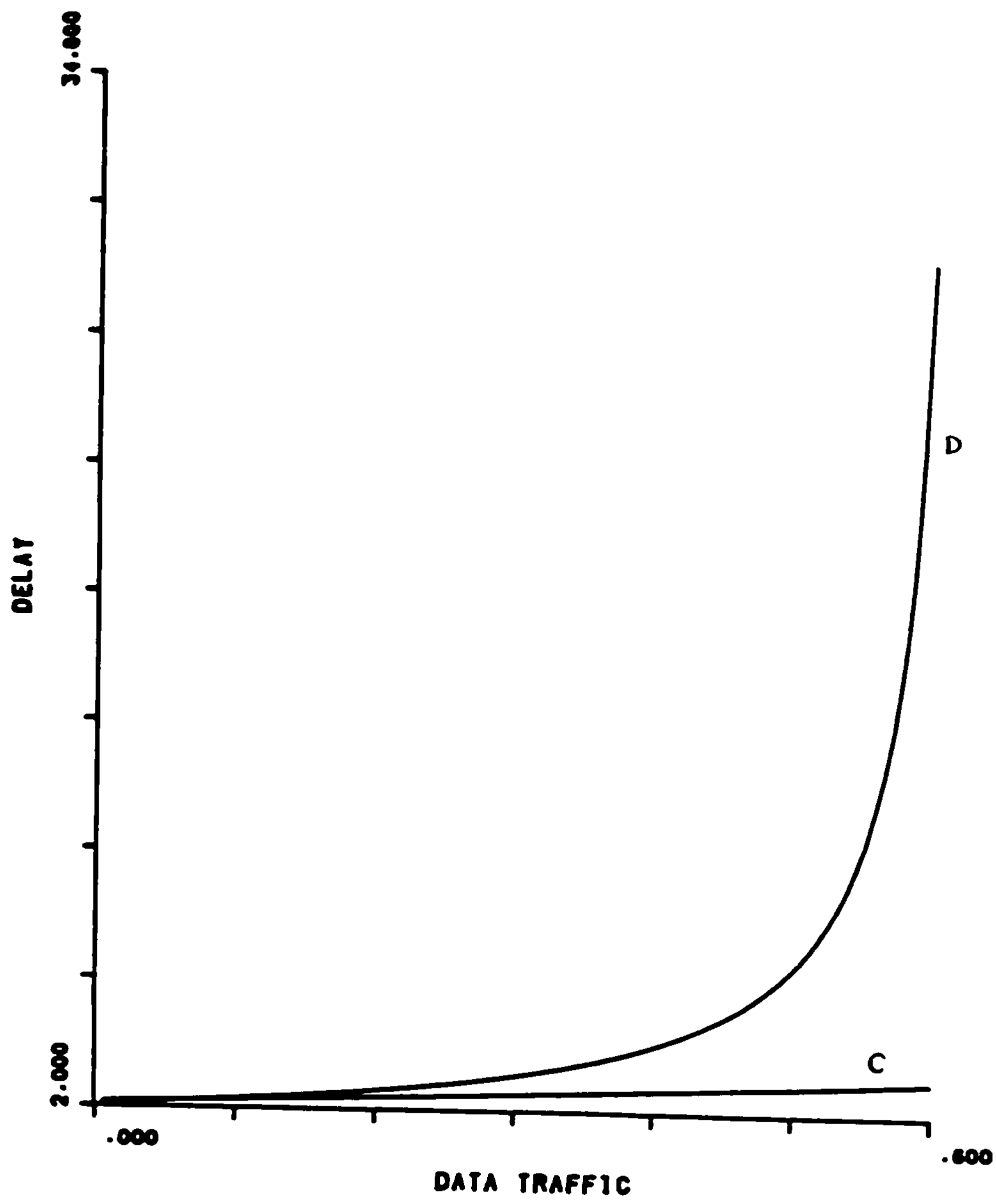

Figure 6-9:
FFMB, continuous - discrete time models,
S=1, N=0, ρ_1=.6, m=10000

N=10, ρ_2=7. The values of ETR are normalized over the service time of the voice customers, tm^{-1}. As we see, the transmission time increases as ρ_1 increases, because at the same time the transmitted frames are longer.

In Figure 6.12 we can see the values of EW for the variable frame scheme, where S=10, N=10, ρ_1=6, m=10000. The results are according to the expression (5-27), and they are compared with the corresponding values for the FFMB scheme, with S=10, N=10, ρ_1-.6. The values of EW are nirmalized over the duration of the fixed length frame: $(N+S)\tau=20\tau$ secs. As we see, $EW \rightarrow \infty$ as $\rho_2 \rightarrow 10=N$, for both of these schemes. Furthermore, the delay in the VFMB scheme is lower than the delay in the FFMB scheme, for most part of the range of ρ_2. However, for large values of ρ_2, the performance of the VFMB deteriorates faster than the FFMB scheme, which provides lower delays in that case.

This happens because when the data traffic is very high, then the transmitted frames of the VFMB tend to have the maximum length of (N+S) slots. Therefore the VFMB scheme tends to perform like a FFMB scheme with the same characteristics (which provides higher delays than the FFMB scheme).

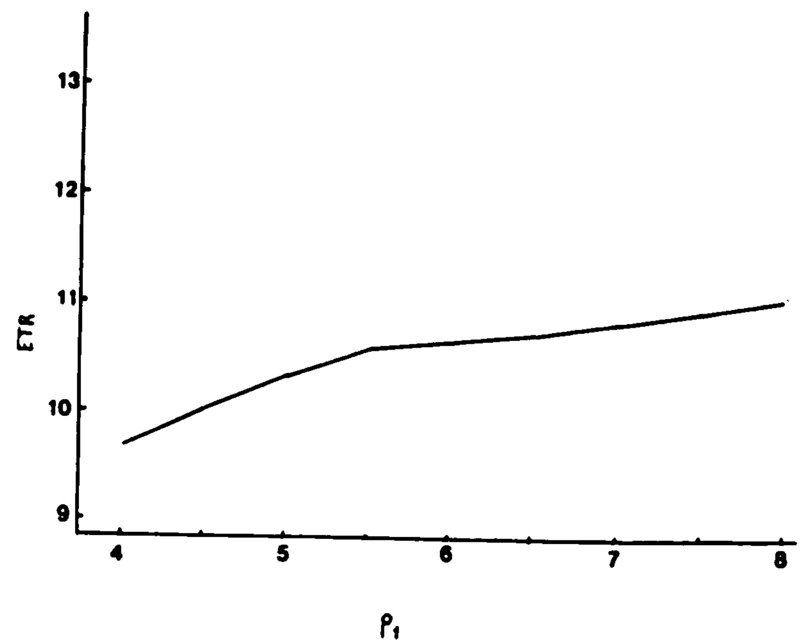

Figure 6-10:
Normalized average transmission time for the VFMB scheme.
S=10, N=10, ρ_2=7.

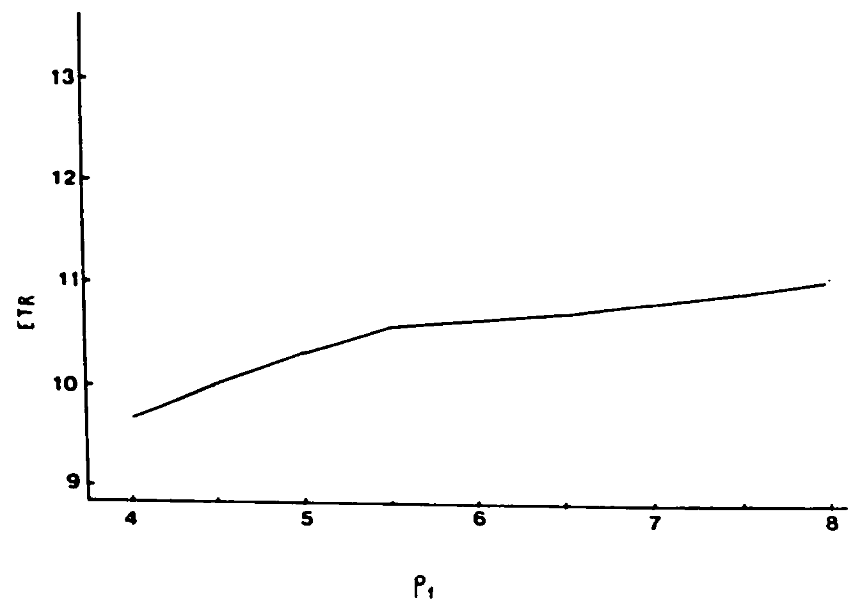

Figure 6-10:
Normalized average transmission time for the VFMB scheme.
S=10, N=10, ρ_2=7.

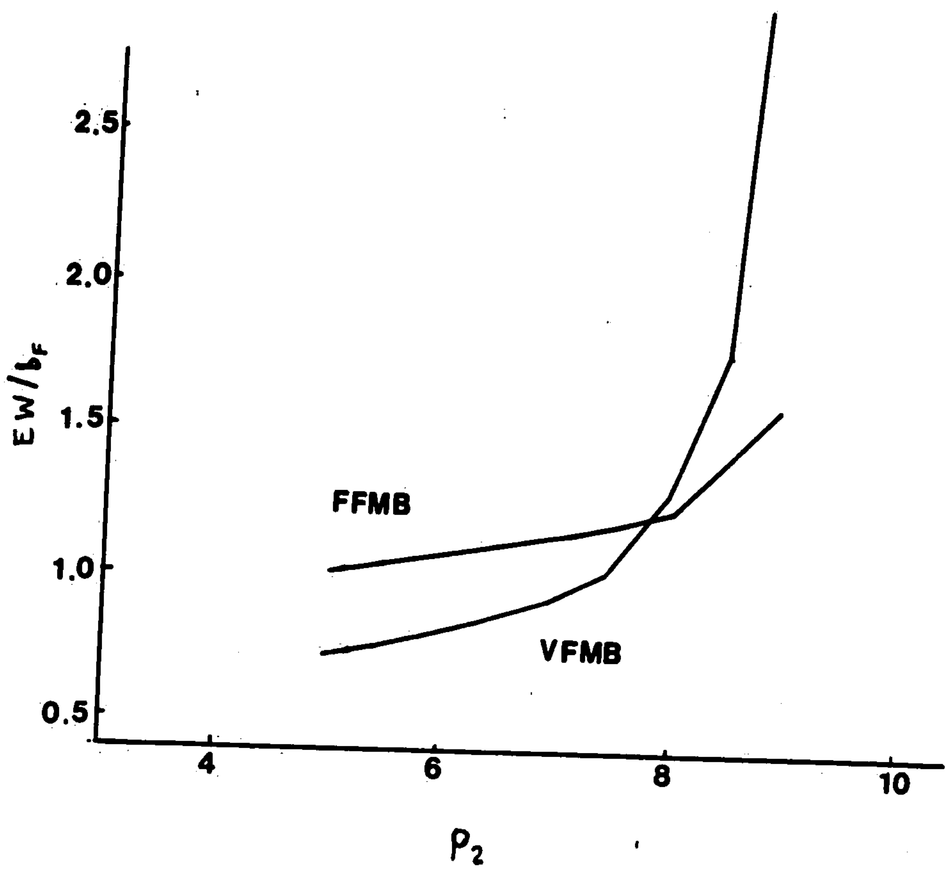

Figure 6-11:
Waiting time for the VFMB versus FFMB.
S=10, N=10, ρ_1=6, m=10000.

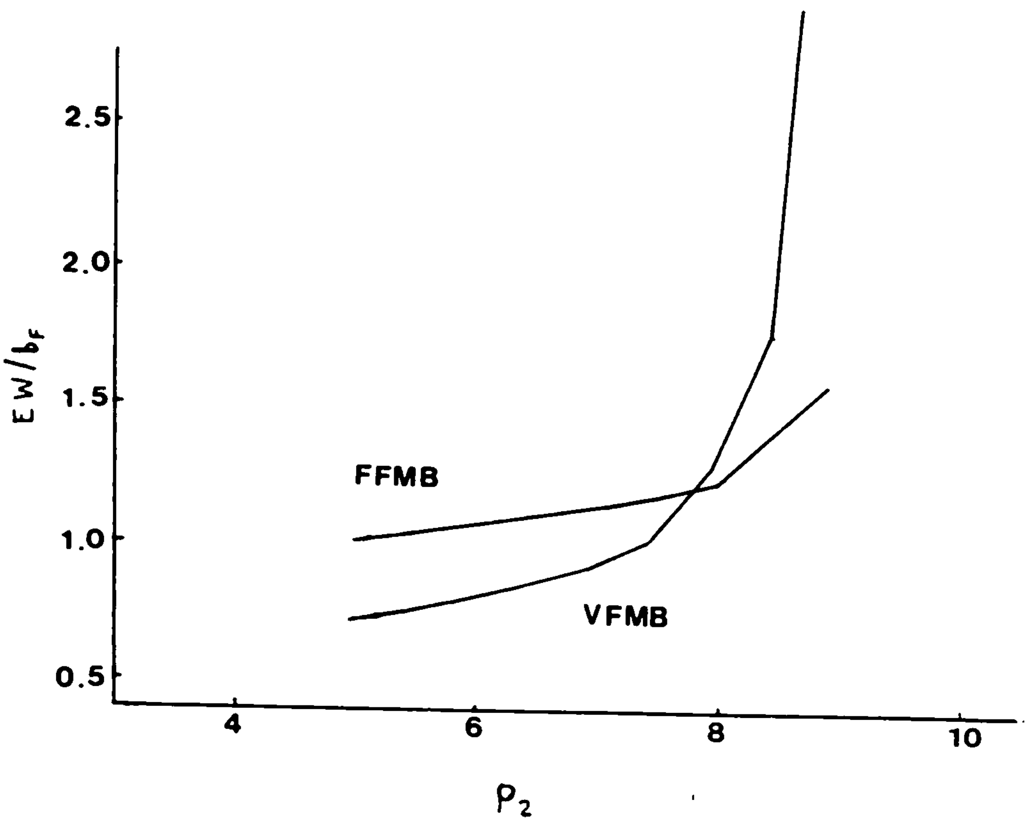

Figure 6-11:
Waiting time for the VFMB versus FFMB.
S=10, N=10, ρ_1=6, m=10000.

7. CONCLUSIONS

The main variations of an integrated voice/data transmission scheme have been considered and their performance has been analyzed by using a general technique applied on various analytical models. The results of these analyses have been presented and justified. Based on the above results, we are able to compare the performance and efficiency of the various integrated schemes. Furthermore, we can compare the presented analytical models and decide which is more appropriate and more accurate for the analysis of each scheme:

7.1 On the integrated transmission schemes

- The performance of the voice traffic transmission is the same for all of the described schemes, because voice always has priority over data. For the variable frame scheme however, care must be taken so that the average time is kept low enough to preserve the voice transmission transparency.

- The FFMB scheme performs better than the FFFB scheme. It provides lower delays for the data customers and improves the channel utilization of the system by allowing data customers to utilize the unused capacity of the voice customers. However, the improvement of the delay does not take place when the voice traffic varies fast and the overall traffic is very high. But even in that case, the delays of the FFMB scheme may be kept low by applying two special mechanisms as suggested in [5,6].

* a) A flow control mechanism (limiting the data butter capacity).

* b)A voice rate control mechanism (using lower voice digitization rate, when the traffic is high).

Notice finally, that the realization of the FFMB scheme involves more hardware complexity than thE FFFB scheme.

- The VFMB scheme performs better than the FFMB scheme: It provides even lower delays for the data customers and even higher channel utilization because it does not allow any idle transmission capacity. Therefore, the VFMB scheme exhibits the best performance and may be considered as the most efficient scheme for integrated voice data transmission. However, the realization of this scheme involves even more hardware and software complexity than the VFMB scheme.

In other words, there is always a trade-off between the impovement of the performance characteristics and the increase of the hardware and software complexity of the system.

7.2 On the presented analytical models

- For the fixed frame scheme, (fixed or movable boundary), the analysis of the voice traffic is carried out simply and accurately by using the standard Erlang-B formula.

- For the FFFB scheme, the data traffic is analyzed accurately by the discrete time model with a reasonable amount of comlexity. This simplified discrete time model does not provide a very accurate analysis but it has the advantage of a simple closed form solution. Actually, it provides a simple way of getting an approximation for the performance of the system.

- For the FFMB scheme, the continuous time model provides the exact analysis of the system performance. However, the numerical evaluation of the results is really complicated, except for the very simple cases with small number of slots per frame. In the case of slowly varying voice traffic and not very high overall traffic, the FFMB scheme can be analyzed correctly and more simply, by using the discrete time model, or approximately by using the simplified discrete time model.

- Finally, for the VFMB scheme, the discrete time model provides an exact analysis of the system performance, with a reasonable amount of complexity. But this model was applied only for the case of slowly varying voice traffic. In the case of fast varying voice traffic, the above modell might not be accurate any more, and probably we would have to introduce a different model for the analysis of the VFMB scheme: That could possibly be a continuous time model, properly modified for the description of this scheme. We might even need to use a different technique of analysis, when is not based on the use of Z-transform. Such a technique could be the fluid approximation method which has been efficiently used for the analysis of the fixed frame scheme [11]. But these special cases could be the subject of a further study, and are beyond the scope of this thesis

Other subjects of further study could be the detailed analysis of the tradeoffs in the realization of the integrated transmission schemes, like the buffer storage requirements, the complexity of the designed integrated multiplexers, or even the quality of the voice transmission.

REFERENCES

1. G. Coviello, "Comparative discussion of circuit - vs. packet switched voice", IEEE Transactions on Communications, vol. COM-27, August 1979.

2. G. Coviello and P.A. Vena, "Integration of circuit/packet switching by a SENET (slotted envelope network) concept", in the Proceedings of 2nd International Conference in Computer Communications, Stockholm, Sweden, August 1974, pp 517-523

3. M. Fisher and T. Harris, "A model for evaluation of the performance of integrated circuit and packet switched multiplex structure", IEEE Transactions on Communications, vol. COM-24, February 1976, pp. 195-202.

4. I. Gitman, W. Hsieh, and B. Occhiogrosso, "Analysis and design of Hybrid Switching Networks", IEEE Transactions on Communications vol. COM-29, September 1981, pp 1290-1300.

5. G.J. Weinstein, M.K. Malpass, and M.J. Fischer, "Data traffic performance of an integrated, circuit- and packet- switched multiplex structure", IEEE Transactions on Communications vol. COM-28, June 1980, pp 873-877.

6. C.J. Weinstein, M.K. Malpass, and M.J. Fischer, "Data traffic performance of an integrated circuit- and packet- switched structure", MIT Lincoln Lab, Lexington, MA, Technical Note 1978-41, October 1978.

7. B. Maglaris, and M. Schwartz, "Performance Evaluation of a variable frame multiplexer for integrated switched networks", IEEE Transactions on Communications, vol. COM-29, June 1981, pp 800-807.

8. M.J. Ross, and O.A. Mowafi, "Performance analysis of hybrid switching concepts for integrated voice/data communications", IEEE

Transactions on Communications, vol. COM-30, May 1982, pp 1073-1087.

9. Lih-Hsing Chang, "Analysis of integrated voice and data communication network", PhD Thesis, Department of Electrical Engineering, Carnegie-Melon University, November 1977.

10. L.K. Kleinrock, Queueing Systems, vols I, II, New York: Wiley 1976.

11. A. L. Garcia, R. H. Kwong, and G. Williams, "Performance evaluation methods for an integrated voice/data link", IEEE Transactions on Communications, vol. COM-30, pp. 1848-1857, August 1982.

12. N. Janakiraman, B. Pagurek, and J. E. Neilson, "Performance analysis of an integrated switch, with fixed or variable frame rate and movable voice/data boundary", IEEE Transactions on Communications, vol. COM-32, pp. 34-39, January 1984.

13. H. Miyahara, and T. Hasegawa, "Integrated switching with variable frame and pakcet", in Procceedings of International Communications Conference, Toronto, pp. 20.2.1-20.3.5, Toronto, June 1978.

14. K. Kummerle, "Multiplex performance for integrated line-and packet switched traffic", in Procceedings of 2nd International Conference in Computer Communications, Stockholm, Sweden, August 1974, pp. 517-523.

VITA

Nick Mikroudis was born in Athens, Greece, in February, 1960. He received the Diploma in Electrical / Electronics Engineering from the National Technical University of Athens, in June 1983. During his stay at the University of Athens he received honorary scholarship from the National Scholarship Foundation of Greece.

In August 1983, he joined the Lehigh University to pursue his graduate studies towards a Master's of Science degree in Computer Science and Electrical Engineering. He was a Teaching assistant from August 1983 to December 1984.