Analysis of a Multiaccess/Switching Technique for Multibeam Satellites in a Prioritized ISDN Environment

Mohamed Hachicha

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science at
Concordia University
Montreal, Canada

August 1992

ISBN 0-315-84632-1

Canada

Abstract

Analysis of a Multiaccess/Switching Technique for Multibeam Satellites in a Prioritized ISDN Environment

Mohamed Hachicha

An integrated multibeam satellite system with a new multiple access technique and an on-board baseband switch with priority rule is analyzed. The performance analysis includes the computation of the various packet loss, call blocking, and packet delays of a typical user in an integrated satellite internetworking environment. The uplink technique employed is a hybrid packet/circuit switched approach of the Demand Assignment type, while the down link is a Time Division Multiplexing technique. On board the satellite, a baseband nonblocking switch is employed to route packets from input to output ports. Various amounts of input and output buffering as well as priority rules and a head of line blocking resolution algorithm are employed to improve the performance of the data services without affecting other services performance. The delay and throughput characteristics of the services are determined by examinig the relationships between the head of line resolution, buffer sizes, and internal speedup ratio of the baseband switch.

Table of Contents

Table of Contents

Table of Contents

LIST OF ABBREVIATIONS

TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
FDMA	Frequency Division Multiple Access
CDMA	Code Division Multiple Access
DAMA	Demand Assignment Multiple Access
HOL	Head Of Line
VSAT	Very Small Aperture Terminal
SREJ	Selective Reject
VCPB	Variable Channel Per Burst
ACTS	Advanced Communications Technology Satellite
LAN	Local Area Network

LIST OF SYMBOLS

L: Knockout switch speedup ratio

N: Number of uplink and downlink beams

r_f : file data user bit rate

r_v : voice user bit rate

r_w : video user bit rate

h_f : file data call duration in frames

h_v : voice call duration in frames

h_w : video call duration in frames

λ_f' : file data calls generated per frame

λ_v' : voice calls generated per frame

λ_w' : video calls generated per frame

D_{Rf} : Average call establishment delay for file data users

D_{Rv} : Average call establishment delay for voice users

D_{Rw} : Average call establishment delay for video users

D_{di} : Average access delay for interactive data users

F_{up} : Overall size of the uplink frame

N_P : Number of slots reserved for interactive data users

P_k : Probability of the number of reserved slots on the uplink beam

P_B : Blocking probability on the uplink beam

ρ_{t1} : Average throughput for connection traffic

ρ_{t2} : Average throughput for connectionless traffic

$\bar{P}_{bw}$: Packet blocking probability for video users

$\bar{P}_{bv}$: Packet blocking probability for voice users

$\bar{P}_{bdf}$:Packet blocking probability for file data users

$\bar{P}_{bdi}$: Packet blocking probability for interactive data users

P_A : Probability of arrival of a video packet

P_B : Probability of arrival of a voice packet

P_C : Probability of arrival of a file data packet

P_D : Probability of arrival of a interactive data packet

$\bar{W}_i$: Mean input queuing delay

$\bar{W}_o$: Mean output queuing delay

K_i : Input buffer size

K_o : Output buffer size

$\bar{D}_s$: System delay

λ_{max} : Maximum switch throughput

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my thesis supervisors Drs. A. K. Elhakeem and T. Le Ngoc for their invaluable assistance and constant guidance throughout this research, and for their advice and constructive criticism during the preparation of this thesis as well as to the University Mission of Tunisia in Washington D.C. for its support.

Dedicated to my family

CHAPTER 1

INTRODUCTION

Satellite technology has long been recognized as a potentially attractive alternative to terrestrial media for wide-area data networks. With the allocation and use of the 12/14 GHz Ku band, terrestrial interference is no longer a problem for the design of efficient Very Small Aperture Terminals (VSAT) networks without being constrained by the limitations of the spread spectrum technology. The recent emergence of VSAT based networks has simulated renewed interest in random access satellite channel sharing protocols. Multiple access technology which provides good throughput (bandwidth efficiency), low delay, robust operation, and minimal implementation cost is critical to the commercialization of VSAT based wide-area networks.

The move of on-board processing satellite communications to serve the internetworking needs of distant small earth terminals for personal/mobile communications and private business networks becomes more and more justified whenever terrestrial links become scarce or expensive or fault prone [1 - 3]. Multispot beam limited coverage switched satellite systems will result in the deployment of small antennas (.8 - 2.4 m) directly at customer sites thus allowing direct user to user communication systems.

1.1. Multiple Access Protocols.

The accomodation of integrated video, voice, and data services plus the traditional point to multipoint satellite channels will necessitate the integration of many satellite multi-access techniques. The following is a review of the most important channel access protocols. A multiple access protocol is a set of rules by which a number of distributed remote stations communicate efficiently over a shared channel; these protocols can be classified as slotted or unslotted depending on whether or not time slotting was used.

Satellite multi-access protocols are divided into three categories: fixed assigned multi-access, contention/random access protocols, and finally reservation/contention access protocols.

1.1.1. Fixed Assignement Multi-Access Protocols.

Fixed assigned multi-access protocols include unslotted fixed assignement and slotted fixed assignment.

1.1.1.1. Frequency Division Multiple Access (FDMA) Protocol.

In FDMA, earth stations using the satellite are assigned specific uplink and downlink carrier frequency bands within the allotted satellite bandwidth. Station seperability is therefore achieved by seperation in frequency. After retransmission through the satellite a receiving station can receive the transmitted waveform of an uplink station by simply tuning to the proper frequency band. FDMA is the simplest and most basic satellite multi-access format. If the frequencies are not seperated adequately, intersymbol interference may result which is one of the disadvantages of FDMA [4].

1.1.1.2. Code Division Multiple Access (CDMA) Protocol.

In CDMA, a specific coded address waveform is used with each carrier to achieve carrier separation. Information is transmitted by superimposing it onto the addressing waveform, and modulating the combined waveform onto the station carrier. The entire satellite bandwidth can be used by any station at any time. All carrier waveforms are overlapping whenever there are stations that transmit simultaneously. The reception of only the appropriate carrier is allowed by receiving the entire satellite transmission and demodulating it with the proper address waveform. Consequently, exact time and frequency separation are no longer needed, but in order to carry out the address selection

required, the receiving station equipment tends to be more complicated. CDMA is also called *spread spectrum multiple access* since addressing waveforms tend to produce carrier spectra over a relatively wide bandwidth [5].

1.1.1.3. Time Division Multiple Access (TDMA) Protocol.

In Time Division Multiple Access (TDMA), each uplink station is assigned a specific time slot in which it uses the satellite. Each station must carefully ensure that its waveform passes through the satellite during its prescribed interval only. Receiving stations receive an uplink station by receiving the downlink only at the proper timeperiod. Clearly, TDMA involves most complicated station operations, including some form of precise time synchronization among all users. Frequency crosstalk between users is no longer a problem, since theoretically only one station uses the satellite at one time [6].

Conventional TDMA or FDMA will neither give the delay efficiency nor the flexibility required. However Slotted Aloha and Demand Assignment Multiple Access Techniques [7], seemed to provide the delay throughput, channel error, and fault tolerance characteristics required for the uplink techniques.

1.1.2. Contention/Random Access Protocols.

Random access techniques which are characterized by the possibility of collision between data messages have attracted considerable attention since the ALOHA system was introduced by Abramson in 1970 [8]. The need for contention protocols originates from the realization that for a very large population of low average rate bursty users, fixed allocation of channel time, even for control signals is not feasible since the overhead increases linearly with the number of terminals supported. For such large population systems, the basic principle is to permit all users unrestricted access to the channel, with channel loading low enough to assure a relatively low rate of collisions between

user messages.

As it is the case in fixed assigned multi-access protocols, contention/random access protocols are divided into two categories: Unslotted contention/random access and slotted contention protocols.

1.1.2.1. Unslotted Contention/Random Access Protocols.

In the unslotted protocols we find the pure Aloha protocol, Selective/Reject (SREJ) Aloha and Random Access Code Division Multiple Access (RA-CDMA). As mentioned earlier, pure contention protocols are suitable for a large population of bursty users because the overhead incurred by contention protocols for assigning channel access to ready users is independent of the number of users and the channel propagation time, but is dependent upon the level of traffic.

1.1.2.1.1. The Pure Aloha Protocol.

Under the Aloha protocol, users are not synchronized in any way. Each user transmits a data packet whenever one is ready. In the event that two or more packets collide, i.e overlap in time, each of the users involved realizes this after R seconds (the channel propagation time) and retransmits his packet after a randomized delay. It was shown that [8] the aggregate packet " birth process " is a Poisson process at a rate of S packets per packet transmission time. Assuming that transmission and retransmission in the channel (channel traffic) is a Poisson process at a rate of G packets per packet time, the probability that a transmitted packet is successful is $P_{succ} = S/G = e^{-2G}$. From which the Maximum Channel throughput is obtained at G=0.5 and is equal to (1/2e) =0.184. Fig.1.1 summarizes the channel events of the pure Aloha protocol [8].

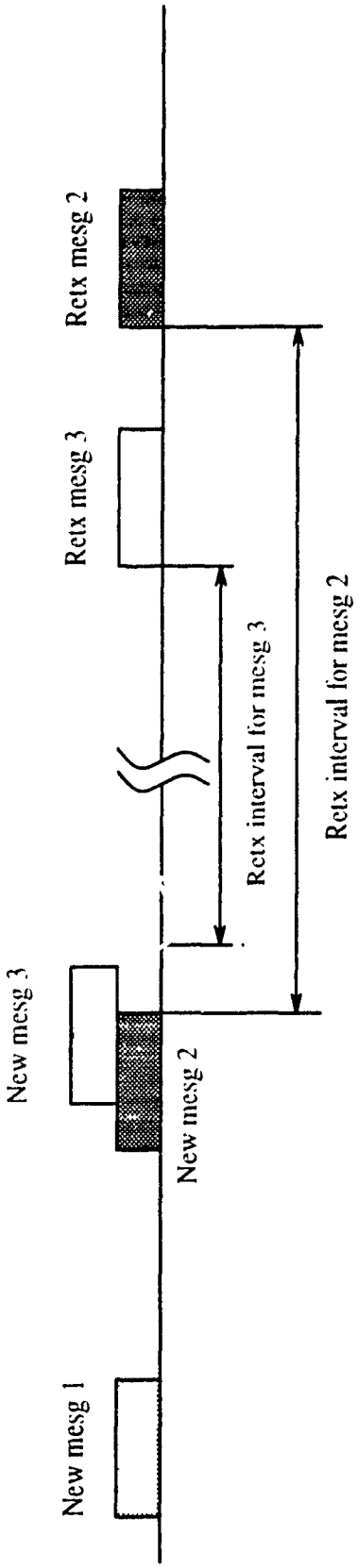


Fig. 1.1 Channel Events of Pure Aloha Protocol.

1.1.2.1.2. The Selective Reject (SREJ) Aloha Protocol.

Without Synchronization, SREJ Aloha, an unslotted random access technique, achieves a maximum throughput of 0.368 using subpacketization of messages in conjuntion with a selective reject ARQ strategy [9]. In other words, messages are formatted into a contigous sequence of independently detectable fixed length subpackets each with its own header and acquisition preamble. Since only partial overlap is the result of collisions in an asynchronous channel, retransmission is necessary only for those subpackets encountering conflict. Theoretically it has been shown that the maximum throughput approaches 0.368 but since there is a need for acquisition preamble and header for each subpacket, the Maximum throughput is limited to the range of 0.2 to 0.3. Fig.1.2 summarizes the channel events for the SREJ Aloha access protocol discussed above [9].

1.1.2.1.3. The Asynchronous Packet RA-CDMA Protocol.

The characteristics of asynchronous Aloha operation can be improved by using spread spectrum modulation. Specifically, it might be possible for packets to encounter multiple interferences without being destroyed if a spread spectrum system were used in conjunction with an Aloha random access protocol. It was shown [10] that the stability issues in random access CDMA are similar to those in Aloha channels and that unstable behaviour can be avoided by operating with sufficiently long retransmission delays

1.1.2.2. Slotted Contention/Random Access Protocols.

The best known and most analysed random access protocol is the slotted contention protocol: slotted Aloha which is based on the principle of reducing the vulnerable period of a packet in Aloha by constraining the channel transmissions to begin and end at slot boundaries.

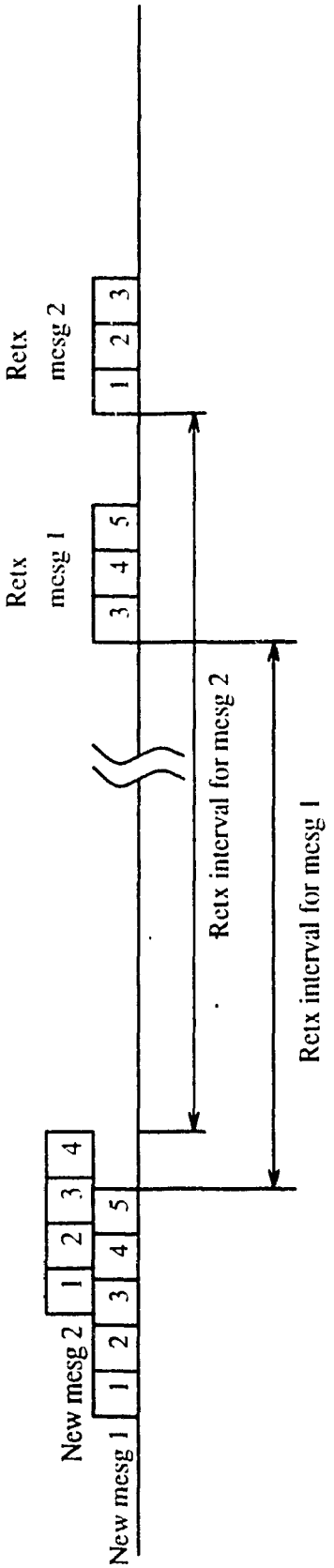


Fig. 1.2 Channel Events for SREJ Aloha Protocol

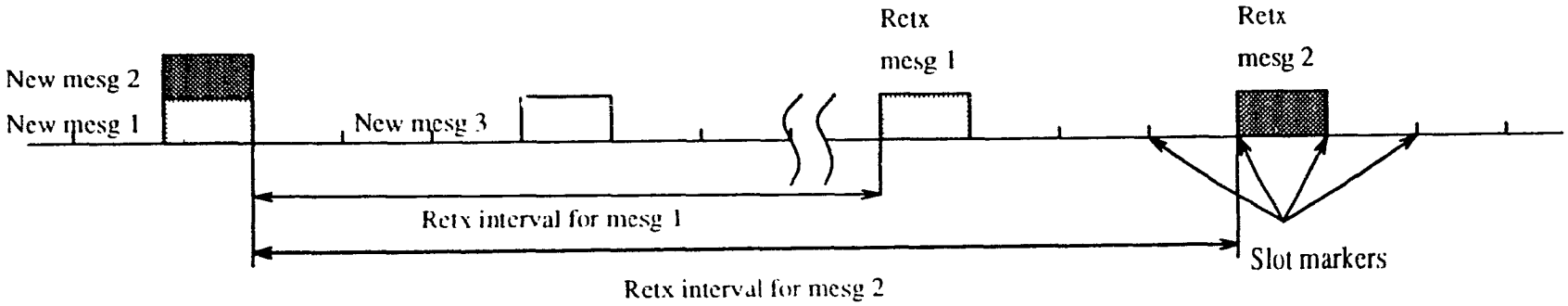


Fig.1.3 Channel Events for Slotted Aloha Protocol.

In other words, slotted Aloha protocol is like the Aloha protocol with the additional requirement that the channel is slotted in time. Under the same assumptions given in the pure Aloha protocol the successful probability $P_{succ} = S/G = e^{-G}$ and the Maximum throughput happens at G=1 and is equal to 1/e = 0.368. Fig.1.3 summarizes the channel events for the slotted Aloha random access protocol [11].

1.1.3. Reservation/Controlled Access Protocols.

Besides the contention access protocols mentioned earlier there is the demand assigned or controlled accesss strategies that consist of access allocation to time slots in TDMA-like channel. This scheme is one of the simplest reservation systems particularly used in satellite communications. The main idea behind this system is to employ a two level access procedure. The first is to place a short reservation for a message transmission using a fraction of the channel time either in TDMA or in contention (such as slotted Aloha) mode. Once the reservation message is successfully transmitted, message slots are allocated in a conflict free manner by the central controllers. Such a scheme is going to be the basic scheme for this thesis.

In the switched multispot limited coverage satellite channels there is a necessity to route messages of different kinds through the satellite on-board switch economically and efficiently [12] while respecting the various criteria for various users such as losses and delay jitter for voice users and end to end delay and throughput for data users. One of the earliest solutions appeared in [13] where switching was used on the on-board regenerative satellite. A regenerative repeater on-board of the satellite would have some important functions such as carrier recovery, coherent detection, clock recovery, decision, differential decoding and encoding, data processing, carrier generation, and modulation; since the on-board satellite repeater must have high reliability, light weight, and low power consumption, it is not an easy task to realize all the functions mentioned above with a lot of efficiency.

Recently a Variable Channel Per Burst technique (VCPB) was introduced [14] and its principle is the following: on uplinks, earth stations transmit synchronization and traffic bursts to the satellite. The beginning of each burst includes preamble for synchronizatin and identification and guard time for prevention of collisions; later preamble and guard time are extracted and the call data is forwarded to the transmitter through a baseband switch. The number of channel uplink bursts depends on the burst configuration; the fewer the number of bursts in the frame, the less is the overhead such as preamble and guard time, the greater is the number of available channels.

A recent study of the Advanced Communication Technology Satellite (ACTS) appeared in [15]. ACTS consists mainly of an on-board baseband processor that, after demodulation, switches traffic betwen 30 Ghz uplink and 20 Ghz downlink hopping beams and then modulates the traffic for retransmission. The ACTS system will also demonstrate the use of very high speed intersatellite laser links which will provide the means to connect future switching satellites in order to form a completely global satellite network. Such a satellite network will have a distinct advantage over ground based systems which are plagued by govermental regulations and equipment incompabilities.

In all the above, some issues related to the switch itself were not addressed. In [16] it was shown that queuing was unavoidable to have the best throughput-delay performance. Future satellite systems will accomodate different services each with its own requirements including allowable delays and rates. Among other issues is the use of the Banyan type of networks [17] rather than the classical nonblocking crosspoint architectures [18].

1.2. Baseband Switches.

The majority of packet switches include either a form of Banyan switch or hybrids of fully connected switches. When Banyan switches are used, the switch architecture is

minimized while the blocking probability is maximized. Fully connected switches consists of minimizing the blocking probability and maximizing the switch complexity.

1.2.1 Banyan Switch

Banyan switches were originally defined by Goke and Lipovski [19] with the property of having exactly one path from any input to any output port. Banyan switches are blocking, meaning that packets can collide with each others. Blocking can take place in the internal fabric of the switch as well as the output ports. Internal link blocking happens when two or more packets fight for a particular link inside the switch. On the other hand and when two or more packets contend for the same output port, we talk about output port blocking. Patel [20] and Kumar and Jump [21] showed that the performance of the Banyan switch under uniform traffic decreases sharply as the switch size increases. Jenq [22] proposed internal buffering at each stage to reduce internal blocking. Theimer and al [23,24] and Dias [25] showed that it is possible to increase the throughput with large internal buffers. Still to minimize the disadvantages and to avoid collisions of packets having to traverse the same element, other solutions include introducing buffers, operating the switch at higher speeds than the inputs, or both. Collisions can be avoided totally by sorting the inputs so their paths will not cross .

1.2.1.1 Batcher-Banyan Switch

A variety of functions involving reordering based on address relations can be performed by the basic switching element. For example, Batcher proposed to build switches that sort the incoming packets in descending (or ascending) address order. For example, if packets with destination addresses 6, 5, 0, 7, 3 enter an 8x8 Batcher sorting network on any set of 5 of the 8 inputs, the output would be ordered 0, 3, ,5, 6, 7, x, x, x, where x represents no outputs for the highest three ports. If there are multiple packets for the same output port, this arrangement would let only one packet pass and would drop the rest [26].

A simpler approach to use sorting for switching has been proposed by Huang and Knauer [27]. They use the idea of combining a Batcher sorting switch with a Banyan switch for what they called a Starlite switch. If the inputs to the Banyan are presorted, collisions within it are avoided. Thus, the Banyan network becomes unblocking. The exception is output blocking, in which multiple packets aim for the same output port. In this instance, the duplicates are intercepted in a trap network illustrated by Fig. 1.4. Seen differently, the sorted output from the Batcher switch (0, 3, 5, 6, 7, x , x , x , in our example) is expanded to skip over unaddressed slots (0, x , 3, x , x , 5, 6, 7). The network elements interact so unsorted packets entering on the left leave the leftmost network sorted They then enter a trapping stage in the middle that removes the duplicates; The switching function is completed when the sorted but portunique packets finally are expanded by the Banyan network on the right.

Under uniform traffic, a single Batcher-Banyan network may provide acceptable throughput for certain applications. However, and when dealing with nonuniform traffic, the throughput is severely limited due to blocking. Blocking can be minimized by increasing the number of packets destined to a certain output. This can be done either by increasing the internal speed of the switch which is impossible since the switch is already operating at the maximum speed or by using parallel Batcher-Banyan planes which increases the complexity of the switch.

Another disadvantage of the Batcher-Banyan switch is low fault tolerance due to its high complexity. In other words, if an internal (2x2) switch fails, more than one output maybe affected depending on the location of the fault.

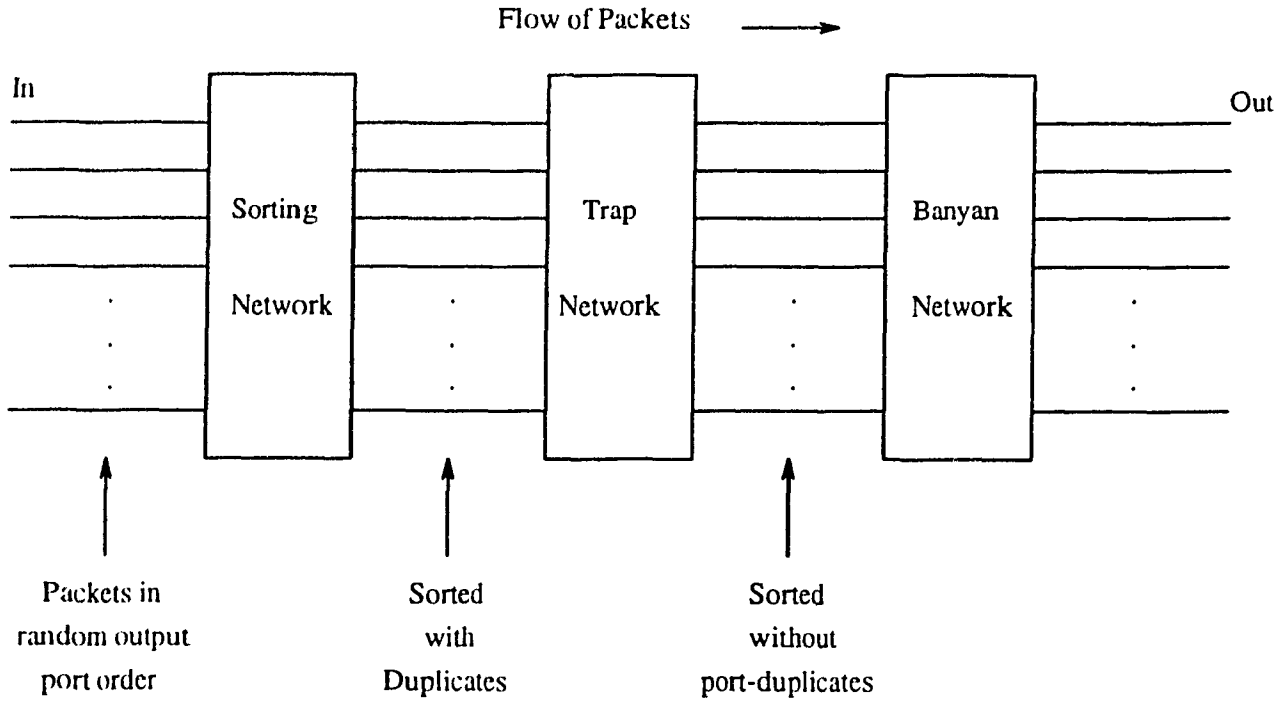


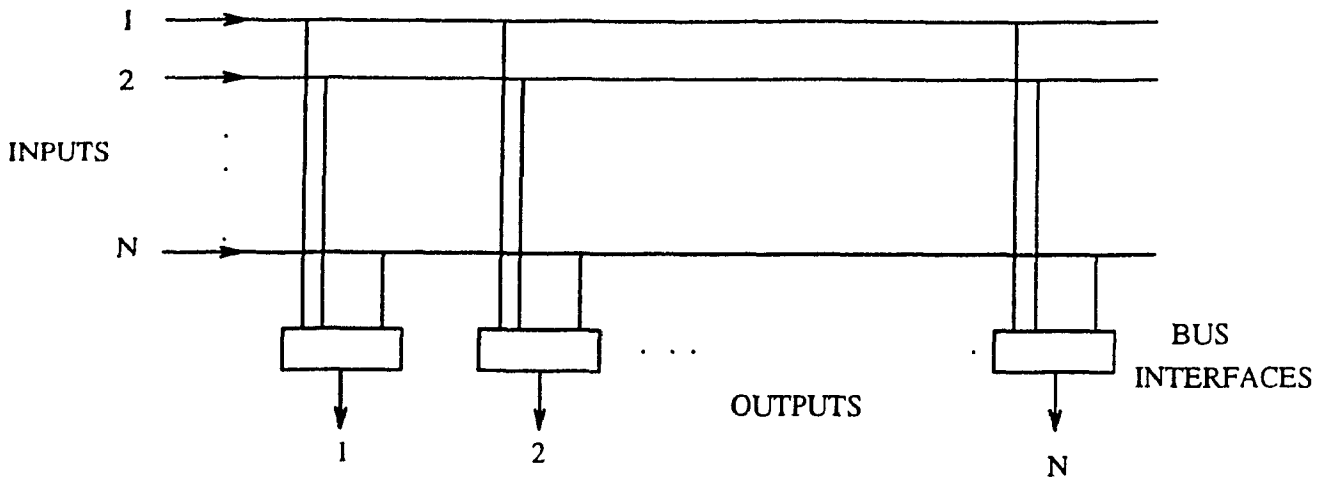
Fig. 1.4 Basic Batcher Banyan Switch

1.2.2 Knockout Switch

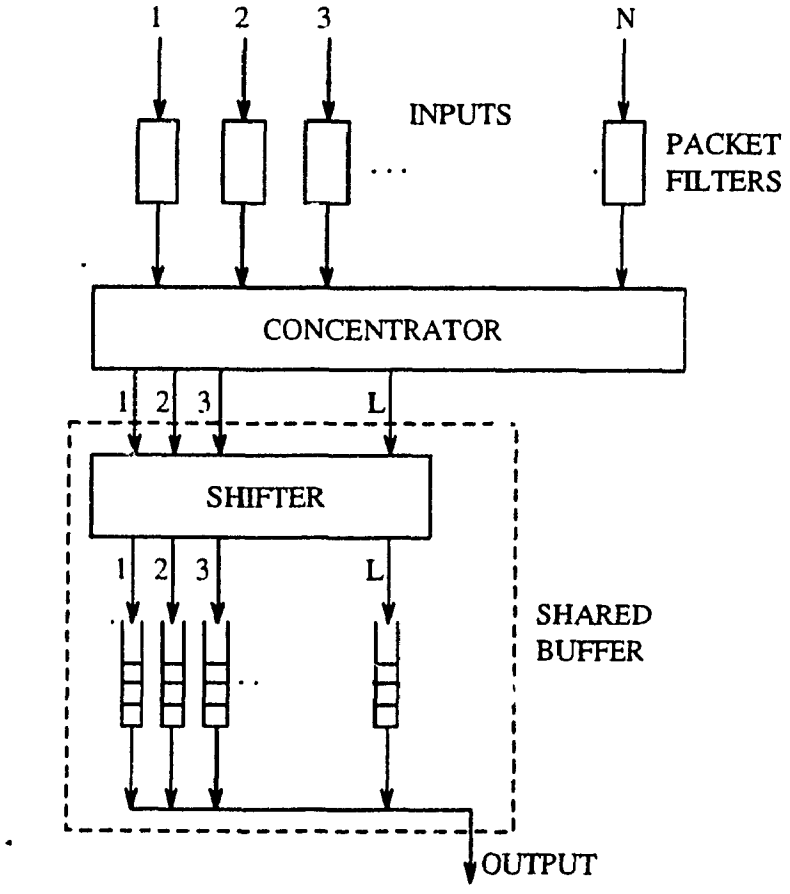
In [28], the knockout switch was presented for a pure packet switched environment that can handle either fixed length or variable length packets. The Knockout switch architecture has low latency, and is self-routing and nonblocking. Moreover, its simple interconnection topology allows for easy modular growth along with minimal disruption and easy repair for any fault. In the knockout switch, as seen in Fig. 1.5, all output ports are connected to every input port by a broadcast input bus meaning that each output port has a bus interface that enables him to receive packets from each input bus line or input port. No collision occurs for packets destined to different output ports. In the case where packets are destined to the same output port, an algorithm is used to choose L packets out of the n incoming lines ($1 \leq L \leq N$). It is a concentrator circuit shown in Fig. 1.5.b that implements the corresponding algorithm to select the L packets. The rest of the packets (n-L) will simply be dropped. It was shown that [29], with L=8 it was possible to achieve a probability of packet loss P_{loss} less than 10^{-6} at a utilization rate of 0.9.

Knockout switches not only increase the throughput but also fault tolerance since faults in a knockout switch effect only one output. The tolerance can be increased by adding an additional spare output circuit that can be used in case of fault.

In comparison to Batcher-Banyan switch, knockout switch presents a major disadvantage in that the total number of basic switching elements For each output port the Knockout switch requires L input packet filters which can be implemented using a minimum of five gates each. In addition each of the L stages of a concentrator requires about N switching elements (2x2) for a total of NL switches per output. The shifter which can be built using an omega network in addition requires a minimum of $\left\lceil \frac{L}{2} \log_2 L \right\rceil$ (2x2) switching elements [28]. Assuming that each (2x2) switch element can be implemented using a minimum of 16 gates.



A: INTERCONNECTION FABRIC.



B: BUS INTERFACE

Fig. 1.5. KNOCKOUT SWITCH

Table 2.1 [28] compares the total number of gates required for each configuration with L=1 and shows that knockout requires much more gates especially for a large N. However the implementation of a large Knockout switch is not limited by any means. Because of the modular nature of the design of the concentrator larger concentrators may be built out of smaller elements. Additional outputs can be accomodated by adding more plug-in modules into a daisy chain configuration. In addition the L+N pin requirement of the concentrator may be satisfied by building larger concentrators from smaller ones. With current VLSI technology limiting the number of pinouts to around 256 and the number of gates to approximately 150 000 per chip a concentrator for N=128 can still be easily be built. It is assumed that such a switch based on gallium arsenide will be able to operate around 150 Mbps. In addition it has been shown by Eng [30] that an all optical Knockout switch can be implemented which would allow the rate to be increased to 2-4 Gbps.

Table 2.1:

Number of gates required for an NxN switch with L=1

N	Batcher_banyan	Knockout
8	576	10048
32	5120	142336
128	35840	$2.203\ 10^6$
512	221184	$3.496\ 10^7$
1024	532480	$1.398\ 10^8$

1.2.3 Other types of switches.

A nonblocking self routing copy network with constant latency appeared in [31]: Simultaneously, the copy network replicates input packet from various sources and then copies of broadcast packets are routed through the switch to their final destination. An encoding and a decoding processes are used to accomplish packet replication. The destinations of the copies are determined by the trunk number translator.

1.3 Scope of the Thesis

Throughout this thesis, and through performance analysis, we investigate the problems involved with integrating the baseband switching technique used on-board the satellite with the multiaccess techniques used in the uplink and downlink channels of the satellite. Also to be introduced is a hybrid reservation/random technique in conjunction with a new version of a switch with speed-up ratio equal to L of the knockout type. Four classes of integrated services namely interactive data, file data, video, and voice users are handled and a priority scheme is used at the switch level in conjunction with a new head of line resolution algorithm to ease the problem of 2 or more input beams contending for the same output.

In other words the motivations of this work are mainly to analyse cases where there are different types of traffic trying to access a satellite when some types of traffic are given priority over others as well as to forward more and more traffic through a Knockout switch even when the Head Of Line of the switch is blocked. In the second Chapter, a description of the multibeam VSAT network is given. In the third Chapter, the analysis of the uplink multiaccess technique is detailed; it includes the reservation/random access technique as well as the moving frame policy. The fourth Chapter will be related to the architecture and the analysis of the prioritized satellite switch including the Head Of Line resolution algorithm and the different priority structures applied to different ser-

vices. Finally a conclusion that summarizes the whole work as well as suggestions for future research are given in Chapter V.

CHAPTER 2

THE DESCRIPTION OF THE PROPOSED MULTIBEAM VSAT NETWORK

The proposed satellite based network, shown in Fig.2.1 is actually a very efficient backbone network to connect LANs and/or just sparse users scattered in a wide region. Fig.2.2 shows the potential services that can use this network, while the part pertaining to baseband satellite switching is depicted in Fig.2.3

The coverage region is divided into Z zones served by A transmitting and receiving satellite antennas, where $Z \geq A$. In this work and for convenience of analysis, we assume Z = A.

The various video, voice, file data, or interactive data users are connected directly to a VSAT terminal or multiplexed in groups in a certain LAN configuration to the VSAT terminal with the latter possibility evidently preferred in the case of $Z \geq A$. (If a single beam is serving many geographically spaced zones, then it would be better to hop to a new zone to serve LAN traffic rather than a single data terminal). The various connection type of users (video,voice, file data) from now on we call it type C, use a (TDMA based) DAMA technique superimposed on a slotted Aloha -like random access technique for interactive users (connectionless type), from now on we call it CL type. Once a C type user becomes active he tries to issue a reservation using one of the mini-slots (N_i) of the uplink frame (Fig.2.4). If he alone used this slot of the specific frame, his request will be answered by the satellite switch manager (reservation or blocking) depending on the current load, priorities, etc. If the requesting user collides with others in the reservation field of the uplink, he will know this after R_T seconds where R_T is the round trip propagation delay for the satellite. Therefore it is necessary to number the various uplink frames up to few multiples of the number of frames included in one round trip propagation delay, so that the user will be able to get correct feedback information pertaining to

the specific frame that he used for the request. In other words, we need a superframe whose minimum duration is set to exceed the round trip delay so that the reservation slots at the beginning of one frame allocate the data slots in the next frame. Once the reservation request is processed, two possibilities are encountered: success or failure. If the request is successful, the user is assigned one or more slots of each frame depending on the kind of services required, e.g, video call may require 4 slots for certain frame sizes. This acknowledgement of reservation is embedded in the control part of the downlink frame (AR field:Fig.2.5). If a collision takes place in the reservation field of the uplink signal of the beam, the involved users will know by tracking the CR field of the downlink beam (Fig.2.5) which gives the identity of these slots as well as the number of the involved uplink frames. In case of collision of the reservation packet at the uplink, or reservation failure at the satellite, the user retransmits the reservation packet in a randomly selected reservation slot.

Retransmission of collided uplink requests will be randomly scheduled in a typical fashion and retransmission of the connection requests takes place k frames after the elapse of R_T sec. from previous transmission. k is a random variable between 1 and 10 since values of transmission span exceeding 10 were found to yield approximately the same delay and throughput as infinity [32]. It is easily seen that each request collision costs R_T sec. which is the major drawback of a reservation method and so it is imperative to have sufficient reservation minislots in the frame so that collision is minimized. It is also assumed that the frame numbering sequence is sufficiently long to span the round trip propagation delay R_T as well as the onboard satellite delays so that there is no ambiguity in which frame is being received. The user has to wait at least one round-trip propagation delay to know what happened with the reservation request. It is of course possible for the requesting user to repeat his request in several consecutive or randomly spaced frames (i.e. request flooding technique). This may work at low loads, but its apparent inefficiency at high loads discouraged further investigation.

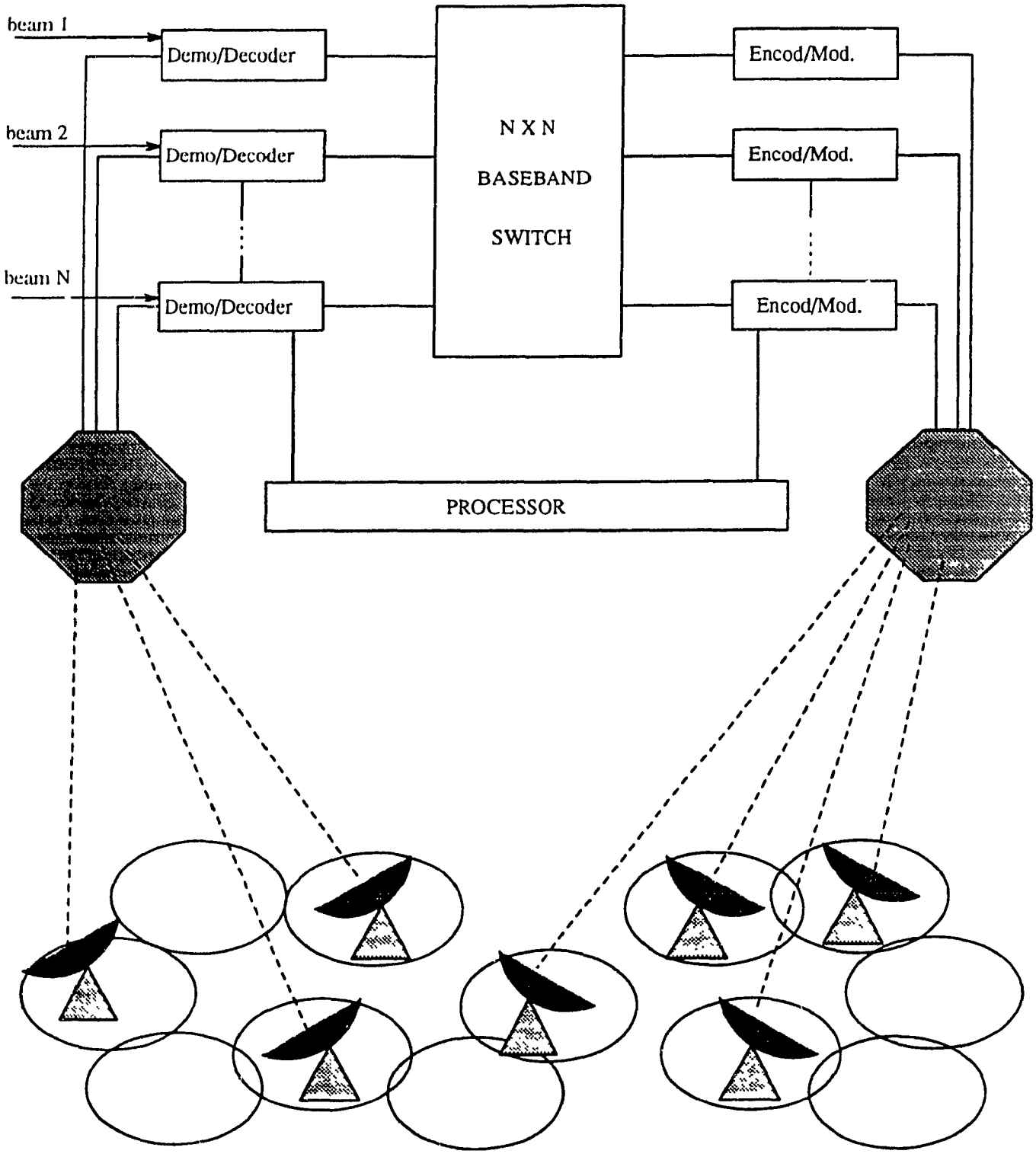


Fig.2.1 . Proposed Multibeam Switched Satellite System.

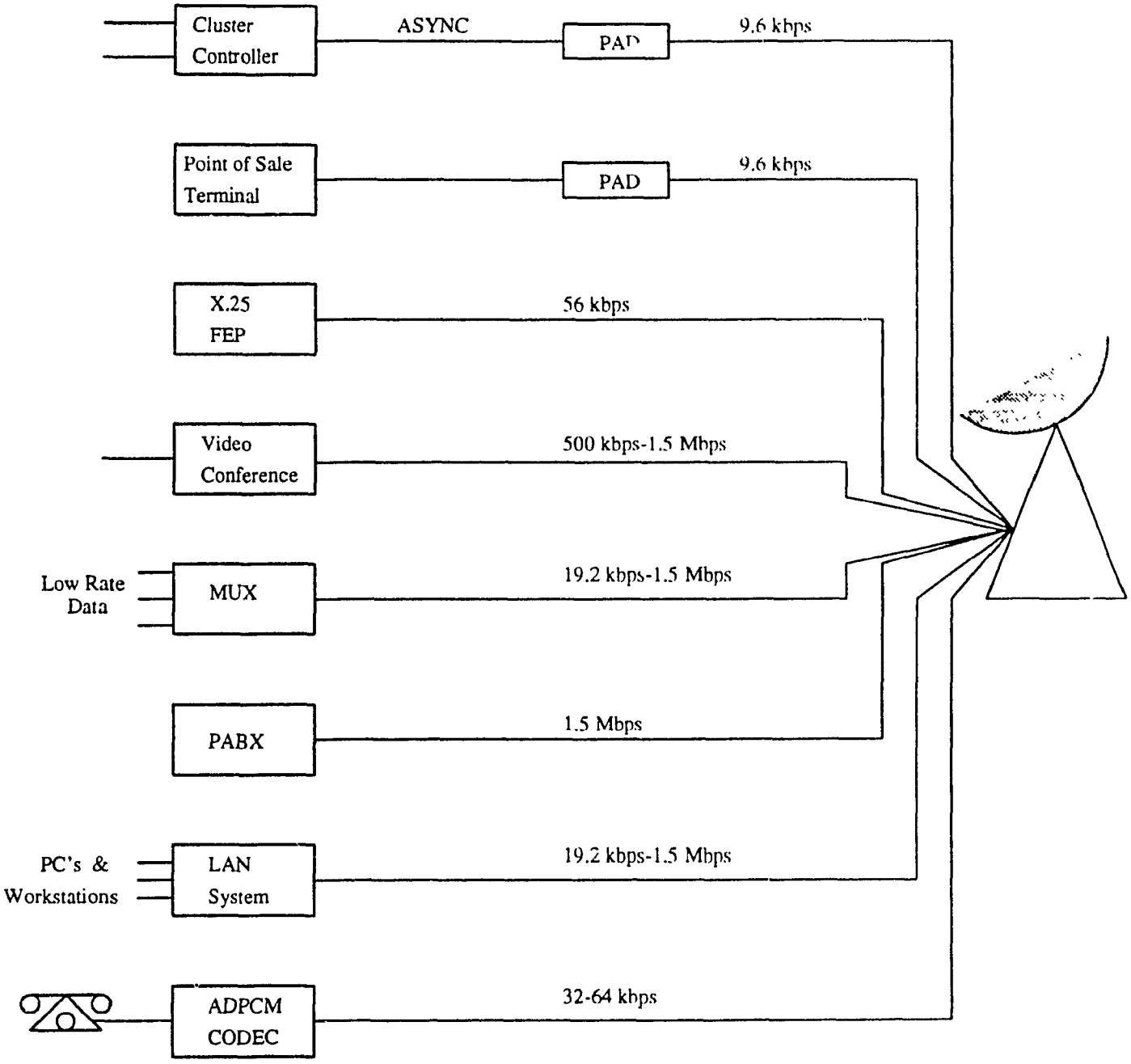


Figure 2.2: Potential services integration for the switched multibeam satellite.

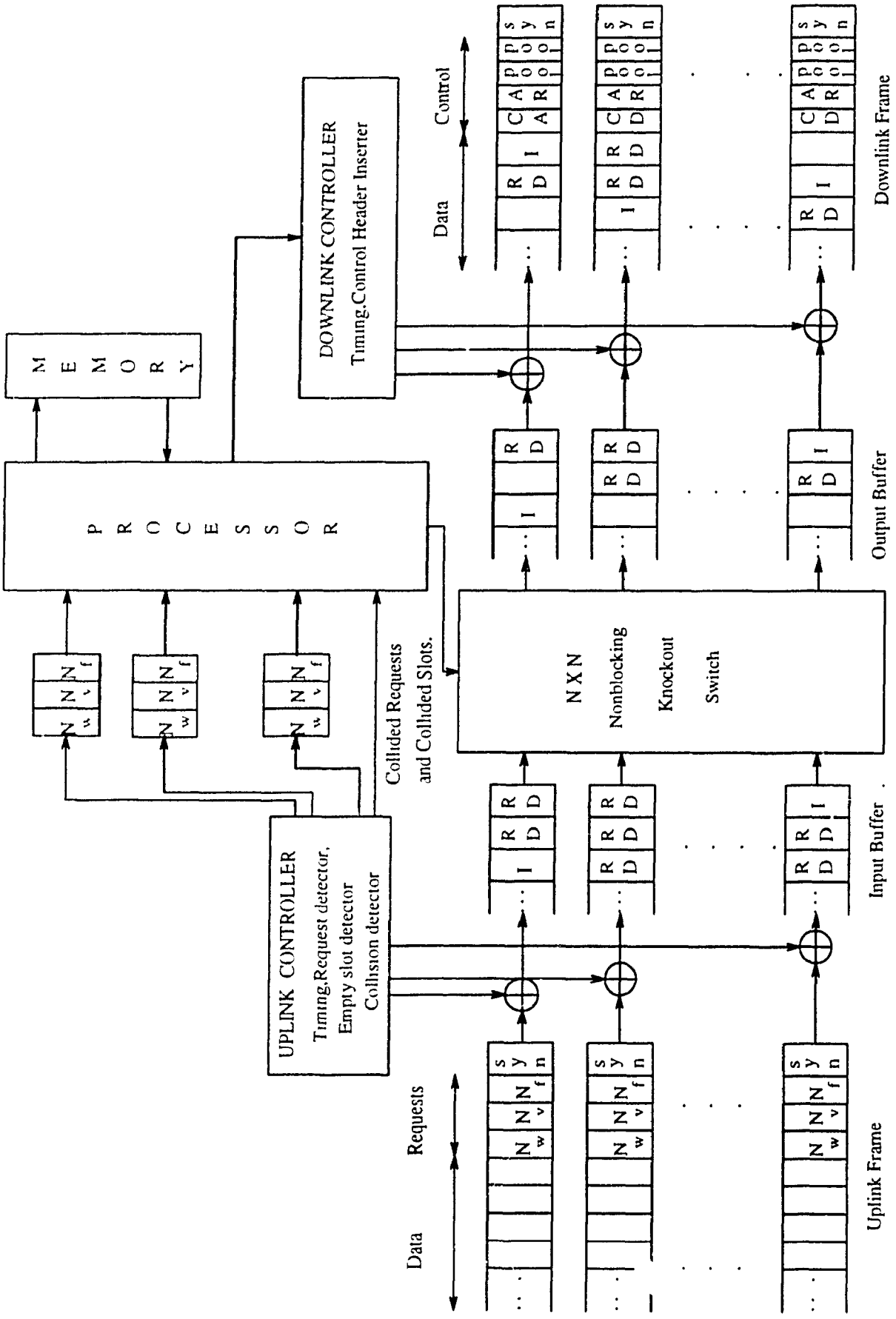
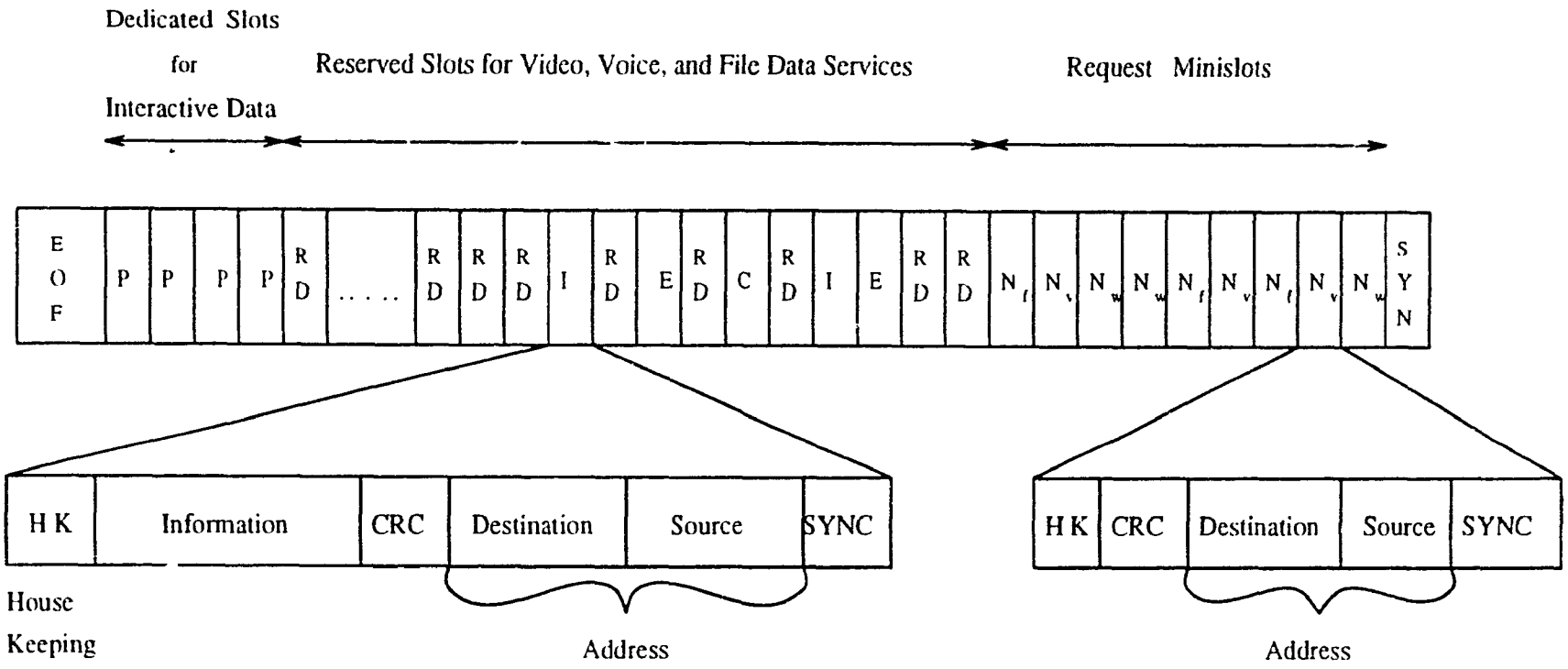


Figure 2.3: Onboard Satellite Switch Configuration.



RD -Denotes a typical reserved slot. Wideband Services such as Video or File Data may use a number of RD Slots in the same frame

C -Denotes Collided Interactive Data Users.

I -Denotes an Interactive Data Packet.

E -Denotes an Empty Slot not Used by any Service.

P -Denoted a Dedicaed Slot for Interactive User.

N_v -Denotes a Reservation Minislot for Voice Service.

N_w -Denotes a Reservation Minislot for Video Service.

N_f -Denotes a Reservation Minislot for for File Data Service.

Figure 2.4: Uplink Frame for a Typical Beam.

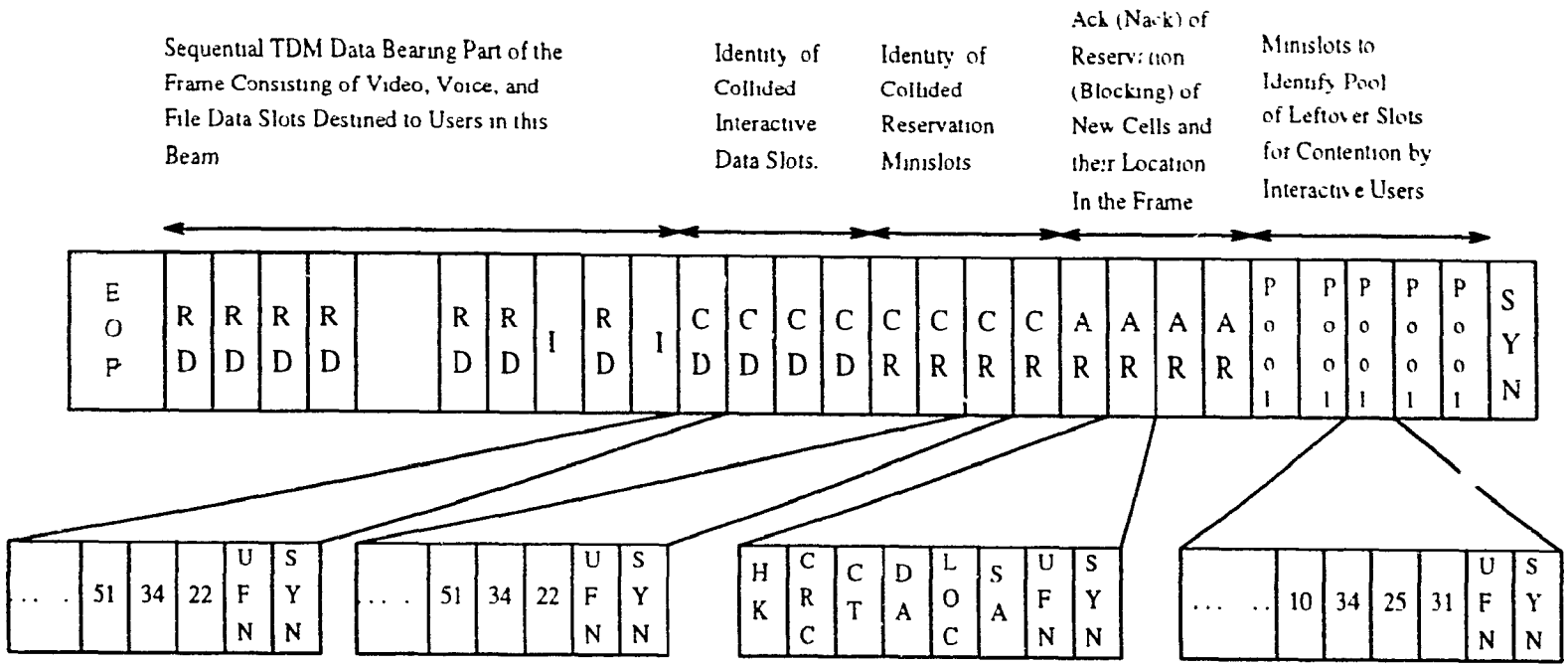


Figure 2.5: Downlink Frame Format for a Typical Beam.

The last paragraph equally applies to random transmission of interactive users. It is noted that once the C type user is connected, he now transmits one or more slots in the same location in each frame and the superframe definition thus applies only to the request stage.

CL type of users (interactive data) will pick one slot of the data part of the uplink frame (P in Fig.2.4) at random and insert his full data packet in it. To allow for flexibility in assigning the capacity for video, voice, and data users (i.e. for employing a moving frame policy), the remaining slots from the C type of data slots in the uplink frame are announced and assigned to the CL type of users. This is achieved by the field labeled POOL in the control section of the downlink frame. The interactive user continuously monitors the downlink frames and combines the capacity unknown by the POOL with that given by the P field. The satellite switch manager dynamically updates the POOL field in each frame to reflect the new call reservation granted.

It is possible that two or more CL type users collide at a certain slot. In this case they will know that they collided after an additional interval of R_T sec and then they may randomly select one out of the next data slots (among only those assigned to CL service) for data packet retransmission. It is assumed that the number of interactive users is limited with an acceptable safety margin so that Slotted Aloha is always operating in steady state at an assumed throughput of 0.32. It is also assumed that in both uplink and dowlink frames, synchronization is assured by other fields such as SYNC and slot SYNC. Similar to most TDMA techniques, users use some pilot synchronizing technique to exactly squeeze their packets in the appropriate time slot in the uplink beam. However, an individual SYNC in each packet is necessary to further aid in slot synchronization at the satellite switch.

In the downlink, transmission is sequential from the satellite to all earth-terminals. The slot SYNC is crucial so destinations pick all packets destined to them in each frame.

Now, we are ready to follow the sequence of events that take place at the satellite space switch. Also, we assume an internally nonblocking switch with L speed up ratio, i.e., it can transmit L packets to the same output within one packet time (slot).

The switch (Fig.2.3) has N uplink beams, N downlink beams, and operates on the baseband data. This requires onboard storage and baseband switching capability. Signal routing for a multiple beam satellite can be accomplished by a baseband processor which regenerates the uplink baseband signal and through storage and switching routes individual circuits to apprpriate downlink beams as shown in Fig.2.3. The satellite demodulates the uplink signal from each beam to recover each channel and stores the content of each channel in an input port. An NxN baseband switch routes the contents of the input ports to output ports so that all channels intended for a given beam can be multiplexed, remodulated, upconverted and transmitted in a single TDM burst.

Two processors are assumed to be coordinating their actions continuously: one as previously discussed is the reservation, framing, and empty slot detection manager. This reservation manager will keep tables in his RAM to store the sources and destinations of on-going calls and the identities and number of slots they are using within each frame. Other functions of the reservation manager include implementing the priority rules for call establishment at the uplink level (not at the space switch level). It will also monitor the unused POOL of slots to be given to CL type of users, increment it or decrement it depending on the case. Moreover the reservation manager prepare the control part of the downlink (to reflect the status of reservation acknowledgement, etc. Fig.2.5), amend it to the data part, and finally number the frames and reset the frame counter.

The second manager is the switch manager who will be responsible of the preparation of the downlink data part, omission of collided and empty data slots, grouping of all data packets from all input ports of the switch destined to the same output and setting the command control to configure the switch (on a slot basis). The switch manager will also

implement any switch priority rules (video packets preferred over voice, for example, if both are destined to the same output and only one is allowed), as well as the randomization necessary to lessen the Head Of Line effect (HOL) as will be shown later on. Finally the switch manager will implement an efficient L knockout strategy to resolve the output contention as will be seen in the analysis; this is amenable to priority rules accomodation at the switch level.

Basically the switch manager performs all functions of the switch while the reservation manager performs all functions related to the uplink and downlink frame circuits establishments. It is also assumed that the switch represents the space part of the T-S-T configuration of the satellite network; in other words the first time (T) stage is represented by the uplink multiaccess technique, the space stage (S) is represented by the satellite baseband switch and the second time ()T stage being the downlink multiaccess technique.

CHAPTER 3

ANALYSIS AND PERFORMANCE OF THE PROPOSED HYBRID DAMA/TDMA UPLINK MULTI-ACCESS TECHNIQUE.

3.1. Analysis

In the following, the hybrid DAMA technique that has been presented is analysed and the performance of the various classes of users is evaluated. The combined traffic of file data users including the retransmission from the slotted Aloha reservation traffic and from the on-board satellite reservation frame manager is assumed to follow a Poisson distribution. This means that the probability of having k new messages arriving during a time interval of t seconds is given by the Poisson distribution to be [33]:

$$P_k(t) = \frac{(\lambda t)^k \, exp\,(-\lambda t)}{k\ !} \tag{3.1}$$

where λ is the average message arrival rate. Another assumption is that blocked reservation traffic due to insufficiency of the frame slots (high load conditions) are nonqueued on board the satellite and that they return back to the Aloha reservation state. (Either by NoACK sent from the satellite or by a negative ACK procedure). This is intended partially to ease memory requirements on board the satellite and for analysis convenience purposes, in which case, performance (connection establishment delay) is slightly sacrificed. The same discussion above equally applies to voice and video traffic reservation mechanisms. The analysis is started by assuming a state diagram (Fig.3.1) representing the number of occupied slots k on the uplink frame (reserved file, voice, and video allowed traffics).

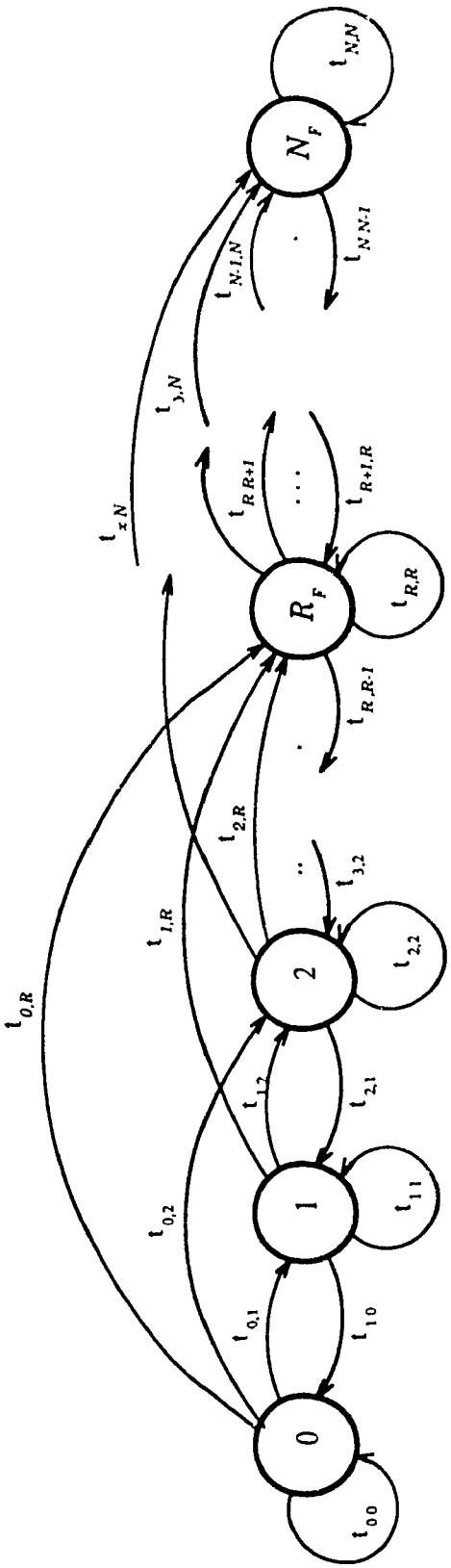


Figure 3.1: Markov Chain of the Uplink Frame

At a certain state k, the reservation success probabilities is denoted to be $P_{sdf,k}$. Once calculated, each success probability $P_{sdf,k}$ is averaged over all states of the uplink frame to yield the average request success probability P_{sdf} for the file data traffic [35]. The potential reservation request traffic at this condition including all retransmissions is given by:

$$\lambda_f'' = \frac{S_f \lambda'_f}{P_{sd2}}$$

$$= S_f \lambda_f (M_f - [KR_f])/(N_f \cdot [1 - P_{bdf}]) \tag{3.2}$$

$$\lambda'_f = \lambda_f (M_f - [KR_f])/N_f \tag{3.3}$$

$$P_{sdf} = 1 - P_{bdf} \tag{3.4}$$

$$R_f = \frac{M_f \lambda_f}{(M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{3.5}$$

where λ'_f is the generated file data calls per frame from all potential users, λ_f'' is the request file data traffic per reservation slot delivered to the satellite on board reservation manager, λ_f is the file data calls generated per uplink frame per user, P_{sdf} is the probability of reservation success (capacity available on the uplink frame), S_f is a traffic mapping factor representing the fact that reservation for the call takes place only once at the call establishment phase. M_f is the total design number of file data users, N_f is the number of reservation minislots for file data users, $(M_f - [KR_f])$ is the potential number of file data users that can generate call reservation requests , R_f is the file data traffic ratio , and $[KR_f]$ is the number of active file data calls contributing to state k.

Naturally video, voice, file data calls all don't have the same call duration or data rate, however, we assume this has been absorbed in the process of calculating λ_f , λ_v, and λ_w (as will follow shortly) such that each call has only one dedicated slot on the frame (once a call is allowed on the uplink frame). For a large number of bursty requesting users M_f, a Poisson assumption for call reservation request traffic is justified, so the throughput of this Aloha state is given by:

$$\lambda_f'' = G_{df} e^{-G_{df}} \tag{3.6}$$

From which the reservation success probability in the Aloha state becomes,

$$p_{sf,k} = (\lambda_f''/G_{df}) \tag{3.7}$$

Repeating all the discussions and assumptions leading to (1) - (4) for the case of voice reservation traffic, we obtain:

$$\lambda_v'' = \frac{S_v \lambda'_v}{P_{sv}}$$

$$= S_v \lambda_v (M_v - [KR_v])/(N_v \cdot [1 - P_{bv}]) \tag{3.8}$$

$$\lambda'_v = \lambda_v (M_v - [KR_v])/N_v \tag{3.9}$$

$$P_{sv} = 1 - P_{bv} \tag{3.10}$$

$$R_v = \frac{M_v \lambda_v}{(M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{3.11}$$

where λ'_v is the generated voice calls per frame from all potential users, λ_v'' is the request voice traffic per reservation slot delivered to the satellite on board reservation

manager, λ_v is the voice calls generated per uplink frame per user, P_{sv} is the probability of reservation success (capacity available on the uplink frame), S_v is a traffic mapping factor representing the fact that reservation for the call takes place only once at the call establishment phase. M_v is the total design number of voice users, N_v is the number of reservation minislots for voice users, $(M_v - [KR_v])$ is the potential number of voice users that can generate call reservation requests , R_v is the voice traffic ratio , and $[KR_v]$ is the number of active voice calls contributing to state k.

The throughput of this Aloha state is given by:

$$\lambda_v'' = G_v\, e^{-G_v} \tag{3.12}$$

From which the reservation success probability in the Aloha state becomes,

$$p_{sv,k} = (\lambda_v''/G_v) \tag{3.13}$$

For the case of video users, the corresponding equations are:

$$\lambda_w'' = \frac{S_w \lambda'_w}{P_{sw}}$$

$$= S_w \lambda_w (M_w - [KR_w])/(N_w \cdot [1 - P_{bw}]) \tag{3.14}$$

$$\lambda'_w = \lambda_w (M_w - [KR_w])/N_w \tag{3.15}$$

$$P_{sw} = 1 - P_{bw} \tag{3.16}$$

$$R_w = \frac{M_w \lambda_w}{(M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{3.17}$$

where λ'_w is the generated video calls per frame from all potential users, λ_w'' is the request video traffic per reservation slot delivered to the satellite on board reservation manager, λ_w is the video calls generated per uplink frame per user, P_{sw} is the probability of reservation success (capacity available on the uplink frame), S_w is a traffic mapping factor representing the fact that reservation for the call takes place only once at the call establishment phase. M_w is the total design number of video users, N_w is the number of reservation minislots for video users, $(M_w - [KR_w])$ is the potential number of video users that can generate call reservation requests , R_w is the video traffic ratio , and $[KR_w]$ is the number of active video calls contributing to state k.

The throughput of this Aloha state is given by:

$$\lambda_w'' = G_w e^{-G_w} \tag{3.18}$$

From which the reservation success probability in the Aloha state becomes,

$$p_{sw,k} = (\lambda_w''/G_w) \tag{3.19}$$

As explained in the discussion preceeding equation (3), the following mapping factors are added:

$$\lambda_v = \frac{r_v}{r_f} \times \frac{h_v}{h_f} \times \lambda_v' \tag{3.20}$$

$$\lambda_w = \frac{r_w}{r_f} \times \frac{h_w}{h_f} \times \lambda_w' \tag{3.21}$$

$$\lambda_f = \frac{r_f}{r_f} \times \frac{h_f}{h_f} \times \lambda_{df}' \tag{3.22}$$

where

r_f = file data user bit rate

r_v = voice user bit rate

r_w = video user bit rate

h_f = file data call duration in frames

h_v = voice call duration in frames

h_w = video call duration in frames

λ_f' = file data calls generated per frame

λ_v' = voice calls generated per frame

λ_w' = video calls generated per frame

$\lambda_w, \lambda_v, \lambda_f$ represent call traffic generation mapped to an equivalent file data call generation as in equations (3.20) - (3.22).

As will follow shortly, the successful Aloha reservation requests going to the uplinks, (eqn. (3.30)) represents the aggregate of all video, voice, and file data requests. The state diagram (Fig.3.1) will be solved, the call request blocking probability P_B will be found (eqn. (3.40)), P_{bdf}, P_{bv}, P_{bw} will be found (eqns. (3.42) - (3.47) where they will be compared to their starting values in eqns (3.2), (3.8), (3.14). If they are equal, iterations will stop, otherwise we go in an iterative search scheme to find the solution (steady state value of P_B and hence P_{bdf}, P_{bv}, P_{bw}). At these values, the single user reservation performance are just $P_{sf,k}$, $P_{sv,k}$, and $P_{sw,k}$. It follows because of the random nature of retransmission in the request slot that the call establishment delay of file data, voice, and video calls (at state k) is given by :

$$D_{Rf,k} = R_T \sum_{i=0}^{\infty} i P_{sf,k} (1-P_{sf,k})^{i-1} = R_T / P_{sf,k} \tag{3.23}$$

$$D_{Rv,k} = R_T \sum_{i=0}^{\infty} i P_{sf,k} (1-P_{sf,k})^{i-1} = R_T / P_{sv,k} \tag{3.24}$$

$$D_{Rw,k} = R_T \sum_{i=0}^{\infty} iP_{sw,k}(1-P_{sw,k})^{i-1} = R_T/P_{sw,k} \tag{3.25}$$

where R_T is the roundtrip propagation delay of the satellite (assuming that a requesting user waits for the result of his previous transmission request before issuing the next request slot), and a geometric distribution is fit to the request process.

The average call establishment delays for the different connection type of users are given by:

$$D_{Rf} = \sum_{k=0}^{NF} D_{Rf,k} \cdot P_k \tag{3.26}$$

$$D_{Rv} = \sum_{k=0}^{NF} D_{Rv,k} \cdot P_k \tag{3.27}$$

$$D_{Rw} = \sum_{k=0}^{NF} D_{Rw,k} \cdot P_k \tag{3.28}$$

and NF = maximum possible number of calls (slots) per frame

$$= F_{up} - \alpha(N_f + N_v + N_w) - N_p \tag{3.29}$$

F_{up} = the overall size of uplink frame in units of data slots.

N_p = the part of frame assigned solely for interactive data (in units of data slots).

α = the ratio of reservation subslot duration to data slot size

P_k = steady state distribution of the number of active calls on the uplink beam.

The dependence of all the equations (3.2) - (3.25) on the state of the number of calls on the uplink channel k is clearly seen. To find the probability density function P_k in the steady state, we refer to the state diagram of Fig(3.1) where the state dependent transition probabilities α_i^k are given by:

$$\alpha_i^k = \sum_{j_f=0}^{N_f} \sum_{j_v=0}^{N_v} \sum_{j_w=0}^{N_w} \begin{bmatrix} N_f \\ j_f \end{bmatrix} \begin{bmatrix} N_v \\ j_v \end{bmatrix} \begin{bmatrix} N_w \\ j_w \end{bmatrix} y_f^{j_f} y_v^{j_v} y_w^{j_w} (1-y_f)^{N_f-j_f} (1-y_v)^{N_v-j_v} (1-y_w)^{N_w-j_w}$$

$$j_f + j_v + j_w = i; \quad i = 0, 1, ..., R_F \tag{3.30}$$

where:

$$y_f = \frac{M_f \lambda_f S_f}{N_f} \tag{3.31}$$

$$y_v = \frac{M_v \lambda_v S_v}{N_v} \tag{3.32}$$

$$y_w = \frac{M_w \lambda_w S_w}{N_w} \tag{3.33}$$

are the various request traffic per request slot, and where the multinomial in (3.30) reflects independence among file data, voice, and video requests. S_f is a factor that maps C traffic to request traffic (of the order of 1/(file data holding time in units of frames)). The same definition applies to S_v and S_w.

The independence of the members of each service is assumed and justified and the condition $(j_f + j_v + j_w) = i$ reflects the convolution of the 3 binomial distributions. It is also assured that no priority exists in the uplink requests process, each call request leads to one call arrival to the satellite manager process, and the aggregate of call request arrivals from all N_f, N_v, and N_w slots lead to the probability α_i^k above. R_F is the maximum number of request calls that the satellite manager can grant in any frame, i.e.,

$$R_F = N_f + N_v + N_w \tag{3.34}$$

The number of slots per call and the bit rates differ with service but this has been taken care of already by the mappings of (eqns (3.20) - (3.22)). Moreover, following these mappings because of the absence of any priority structure on the uplink requests, the numbers of subslots N_f, N_v, and N_w have to be consistent with the average amount of traffic of each kind, i.e.

$$\frac{N_f}{(N_f+N_v+N_w)} = R_f \tag{3.35}$$

$$\frac{N_v}{(N_f+N_v+N_w)} = R_v \tag{3.36}$$

$$\frac{N_w}{(N_f+N_v+N_w)} = R_w \tag{3.37}$$

The stability criteria for the Markov chain of Fig.3.1 follows easily, i.e.

$$(\lambda_f M_f+\lambda_v M_v+\lambda_w M_w) \le NF \tag{3.38}$$

The average request success probabilities (averaged over k) on the uplink beam are now given by:

$$P_{sf} = \sum_{k=0}^{NF} P_k P_{sf,k} \tag{3.39}$$

$$P_{sv} = \sum_{k=0}^{NF} P_k P_{sv,k} \tag{3.40}$$

$$P_{sw} = \sum_{k=0}^{NF} P_k P_{sw,k} \tag{3.41}$$

where P_{sf}, P_{sv}, and P_{sw} are the average request success probabilities for file data, voice,

and video users respectively.

Given the transition probabilities, it is possible to solve the simultaneous equations resulting from the Markov chain of Fig. 3.1 and to obtain the steady state distribution P_k i.e. the probability of the total number of reserved data slots on the uplink beam.

After solving the state equations, the blocking probability is just the ratio of the blocked traffic given any state k to the total average traffic given the same state k; in other words:

$$P_B = \sum_{k=0}^{N_F} P_k \frac{\sum_{j=0}^{R_F} \left[\max(k+j-1-N_F,0)\right] t_j^k}{\sum_{j=0}^{R_F} j t_j^k} \tag{3.42}$$

where t_j^k is the transmission probability of the state diagram of Fig.3.1 given by:

$$t_{kj} = SS_k(\alpha_{j-k+1}^k)+(1-SS_k)(\alpha_{j-k}^k) \qquad \begin{array}{l} j>k\ ,\ j-k<R_F \\ k=0,1,\ldots,N_F-1 \\ j=0,1,\ldots,N_F \end{array}$$

$$t_{kj} = \alpha_0^k SS_k \qquad \begin{array}{l} k=j+1 \\ k=1,\ldots,N_F \end{array}$$

$$t_{kj} = (1-SS_k)(\alpha_{j-k}^k) \qquad \begin{array}{l} j-k=R_F \\ j=0,\ldots,N_F-1 \end{array}$$

$$t_{kj} = \alpha_o^o+\alpha_1^o SS_o \qquad k=j=0$$

$$t_{kj} = \alpha_o^k(1-SS_k)+\alpha_1^k SS_k \qquad k=j=1,2,\ldots,N_F-1$$

$$t_{kj} = 1-\alpha_o^{N_F} S_{N_F} \qquad k = j = N_F$$

$$t_{kj} = 1- \sum_{l=k-1}^{N_F-1} t_{kl} \qquad j = N_F$$

$$k \geq N_F - R_F + 1 \tag{3.43}$$

where SS_k is the service probability in each state k. The different call blocking probabilities for the different services are just the products of the blocking probability by the corresponding traffic ratios, i.e.

$$P_{bdf} = P_B \cdot R_f \tag{3.44}$$

$$P_{bv} = P_B \cdot R_v \tag{3.45}$$

$$P_{bw} = P_B \cdot R_w \tag{3.46}$$

The upper bound of the average total traffic throughput of all connection type users (video,voice, and file data) is given by:

$$\rho_{t1} = \frac{\sum_{k=0}^{N_F} P_k \cdot k}{(N_F + N_p)} \tag{3.47}$$

The number of data slots per frame allocated to the interactive data users has two components; The first represents the number of the data slots dedicated exclusively for the interactive data users and the second represents the number of unused slots by the connection type of users (N_F - k) where k is the number of occupied slots on the frame due to connection uesers. Now to compute the delay of the interactive user who will see a

combined pool of $(N_p + N_F - k)$ data slots as in equation (3.36). This user will follow a convertional slotted Aloha regime in which case the total interactive traffic throughput per slot becomes,

$$S_{di,k} = G_{di,k} e^{-G_{di,k}} \tag{3.48}$$

$$= (M_i \cdot \lambda_i)/(N_p + N_F - k) \tag{3.49}$$

and the interactive data users success probability per slot is given by :

$$S_{ui,k} = (S_{di,k}/G_{di,k}) \tag{3.50}$$

Morever the upper bound on the total average data throughput of all interactive users is given by:

$$\rho_{i2} = \sum_{k=0}^{NF} P_k \cdot (N_p + N_F - k) \cdot S_{di,k} \tag{3.51}$$

where P_k is the steady state solution of the state diagram of the Markov chain shown in Fig 3.1

The interactive user transmission delay to finally succeed to access the uplink channel is given by:

$$D_{di,k} = R_T \sum_{i=0}^{\infty} i \; S_{ui,k}(1 - S_{ui,k})^{i-1} \tag{3.52}$$

$$= \frac{R_T}{S_{ui,k}} \text{sec} \tag{3.53}$$

Needless to mention that each interactive data user tries only one slot per frame and each trial costs him a roundtrip satellite propagation delay R_T.

The actual interactive data throughput of this simple random access scheme given k is given by:

$$S_{di} = \sum_{k=0}^{N_F} P_k \cdot S_{di,k} \tag{3.54}$$

in units of useful data slots per frame.

The average delay for an interactive data user to access the satellite is obtained by averaging $D_{di,k}$ over the distribution of P_k:

$$D_{di} = \sum_{k=0}^{N_F} P_k \cdot D_{di,k} \tag{3.55}$$

Having found all the performance criteria of all kinds of services pertinent to the uplink, it is time now to group all the successful traffic contributions from all services into one traffic unit at the beam input to the satellite onboard switch and both are given by:

$$\rho_t = \frac{\sum_{k=0}^{N_F} P_k \cdot k}{(N_F + N_p)} + \frac{\sum_{k=0}^{N_F} P_k \cdot (N_F + N_p - k) \cdot S_{di,k}}{(N_F + N_p)}$$

$$= \rho_{t1} + \rho_{t2} \tag{3.56}$$

where P_K is the steady state solution of the number of occupied slots per frame. This is the total traffic per beam per frame per slot arriving at the switch input.

N_p plays a major role, its adjustment with ongoing amounts of interactive data traffic is

achieved by slowly generating flow control. For that objective another rule of thumb for adjusting N_p is

$$\frac{N_p}{(N_p+N_F)} = \frac{M_i\lambda_i}{(M_i\lambda_i+M_2\lambda_2+M_v\lambda_v+M_w\lambda_w)} \tag{3.57}$$

With the extra capacity left over from the connection (C) users (voice,video, file) used to advantage to compensate for the difference in (peak-average) traffic in bursty conditions. It is assumed that the satellite on-board processor relays the available left over (C) types of slots identities to the interactive data users (effectively, flow control). Also, it is assumed that high enough interactive traffic exists such that all available slots are occupied in which case the slotted Aloha throughput of eqns.(3.49)-(3.55) will be fixed at 0.32.

3.2 Results of the Uplink Multi-Access Technique

In this section, different results are obtained for different traffic scenarios. All results were obtained by assuming different traffic parameters that are subtituted in the different equations covered in the analysis. The uplink frame size is F_{UP} = 30. It is assumed that the different data bit rates for video, voice, file data, and interactive data users were 1.5 Mbps, 64 kbps, 64 kbps, and 9.6 kbps respectively. Since different traffic rates were used, it is assumed that the different users were multiplexed in frequency. For voice users, it is worth while to mention that we are dealing with packet type voice meaning that the packet could be lost without a lot of problems. It was also assumed that the different service rates are 1 service evey 10 minutes for video users, 1 every 3 minutes for voice users, 1 every 20 minutes for file data users, and 1 every 30 minutes for interactive data users. The different arrival rates were assumed to be 3 calls/day for video users, 3 calls/hour for voice users, 5 calls/day for file data users, and finally 3 calls/day for

interactive data users. For different traffic scenarios, the corresponding delays, reservation success probabilities as well as the throughputs of each traffic type are calculated and graphed. Figures 3.2 to 3.6 show the different delays of the four types of users for an increasing amount of file data traffic and a constant amount of video and voice traffic. It is also assumed that interactive data traffic satisfies constant delay and success probability operating at a near maximum slotted aloha throughput of 0.32. In other words, it is the maximum allowable interactive traffic that is shown in Fig.3.8; This maximum allowable traffic is obtained by assuming an operation at a constant per slot utilization (maximum throughput for slotted Aloha). For the remaining users, we notice that the file data connection delay increases sharply with the increase in the file data traffic (Fig.3.3). The delay for interactive data users is set to be constant. Video and voice users tend to have the same kind of connection delay (Fig.3.4, 3.5); At the beginning the delay tends to increase as the amount of file data traffic increases. The video and voice call establishment delays start to decrease when the amount of file data traffic reaches a certain threshold; an explanation of this phenomenon is that once it reaches a certain value, the file data traffic becomes too high to be handled by the reservation minislots N_f meaning that reservation does not successfully take place and that the potential number of file data users in the ground stations start on building up. From that moment and since the call establishment reservation process is a slotted aloha protocol, fewer file data users will be able to make a successful reservation request meaning that there won't be too much file data traffic on the uplink frame giving the chance for video and voice users who will see a less crowded frame and therfore they will suffer a lower call establishment dealy. The Blocking probability will normally increase (as shown in Fig.3.6). Fig.3.8 shows that the allowable interactive traffic decreases as the file data traffic increases. The sum of both kinds of traffic is shown in Fig.3.9 and represents the total traffic as the file data traffic increases.

In another scenario, the above case is compared with another case in which the voice traffic is increased by 50 percent and the video traffic is doubled. We increased the file data traffic and we found the corresponding delays, success probabilities and blocking probabilities as discussed in the previous case. For interactive data users, the reservation success probabilities and therefore the delays were constant as assumed and shown in Fig.3.10. The interesting part is the part related to the connection establishment delays for the different connection type of users. Starting with file data users, it is seen that the corresponding delay starts from exactly the same value (almost one round trip propagation delay) which is normal since there is space for low traffic. As the file data traffic increases, the delay increases but not exactly in the same way as it was increasing in the first case. The delay in the second case was almost the double of the delay in the first case and that is quite normal since there is more traffic in the second case than in the first case. For voice traffic, the delay increases at the beginning then decreases. In the second case, the voice call establishment delay starts at a higher value than that of the first case (which is normal since there is more voice traffic trying to get through the satellite) but it doesn't increase as was the case in Fig.3.4, instead it decreases continuously; the reason is that the threshold reached in the first case when file data traffic started to increase was already reached in the second case before having any file data traffic. The discussion of the results for the case of voice traffic also applies to the case of video traffic. In fact, in the first case (Fig.3.5), the call establishment delay for video users increases at the beginning than starts to decrease reflecting the fact that once it starts decreasing it means that more and more file data packets are being blocked and therefore being sent back to the Aloha reservation state giving the opportunity to video users to exploit more empty slots on the uplink frame and therefore less connection establishment delays. On the other hand when the amount of video traffic was doubled and that of voice traffic increased by 50 percent, it is noted that the call establishment delay for video users started at a much higher value then that of the first case (compare Fig.3.5 to Fig.3.13) which is very logical

since the amount of video traffic has doubled. Another observation is that this delay decreases as the amount of file traffic increases and therefore as more and more file data traffic gets blocked meaning less and less file data traffic on the uplink frame therefore more and more empty slots being exploited by video users. Once the blocking probability curves for both cases are compared, it is seen that both curves increase sharply as the amount of file data traffic increases with the difference that the second curve (Fig.3.14) shows a faster increase than the blocking curve shown in Fig.3.6 which is what should be expected since in the second case there is much more traffic than in the first case.

The combined traffic for the connection type of users (video, voice, and file data traffic together), the corresponding curves (ROT1 vs. file data traffic in Figs. 3.7 and 3.15) in both cases have the same shape and that the maximum channel throughput is almost the same (0.77 and 0.79). The only difference is that in the first case the starting value is 0.42 compared to 0.71 in the second case reflecting the additional amount of video and voice traffic.

It is also worth noticing that the allowable amount of interactive data traffic changes from the first case to the second case. The starting values are almost the same for both cases (0.065) representing the minimum service granted for interactive data users or the amount of interactive data traffic corresponding to the number of data slots NP dedicated exclusively for this kind of service. On the other hand we see that the upper limit or the starting point of the allowable interacti· data traffic changes dramatically; In fact there is no file data traffic we see that the allowable interactive data traffic starts at 0.19 or 19 percent of the capacity of the uplink channel for the first case compared to 0.09 for the second case. The difference in both values is due to the fact that from the first case to the second case the initial amount of connection type traffic was increased (more video and voice traffic) and we know that interactive data users do not compete for empty slots on the uplink frame and can only use those empty slots if there are no potential video, voice or file data users that would like to get through the channel.

Finally the last curves in both figures (Figs. 3.9 and 3.17) show the total throughput corresponding to all kinds of traffics. Again here both curves tend to converge to a certain maximum value around 0.85 (0.84 for the first case and 0.86 for the second case) after stabilization. There is however one difference since in the first case the starting value is 0.62 compared to 0.81 in the second case; Again this difference is due to the increase in the amount of video and voice traffic from the first case to the second case.

In another situation, the number of slots dedicated solely for interactive data users N_p is increased and the corresponding parameters are evaluated.

When N_p increases, it will be beneficial to interactive data users but not to the connection type of users since it is known that an increase in N_p means a decrease in the number of slots N_F available for other users on the uplink frame. From Fig.3.18 it is seen that the call establishment delay for interactive data users is always constant as assumed but the call establishment delay for the other types of users try to reach a certain minimum value which seems to be abnormal but it is not. In fact if a look is taken at Fig.3.22 which represents the blocking probability on the frame, it is seen that this probability reaches very high values for large N_p and therefore lesser available space on the uplink frame. In other words, the different types of services will be able to make their connection right away which explains the small call establishment delay but as soon as they get connected on the uplink frame they get blocked since there is no space left on the frame then they return for reservation which they will be granted right away, get through the frame, and again get blocked. As it should be expected it is seen that the increase in N_p is mainly beneficial to interactive data users whose total throughput increases constantly since those interactive data users will be able to exploit more and more available space at the expenses of the other connection types of users whose total throughput is shown in Fig.3.23 to be decreasing sharply as the number of slots N_p available for interactive data users increases.

Finally, we see that increasing N_p too much is not beneficial only for interactive data users (Fig.3.24) since the total throughput for all kinds of traffic decreases sharply as shown Fig.3.25.

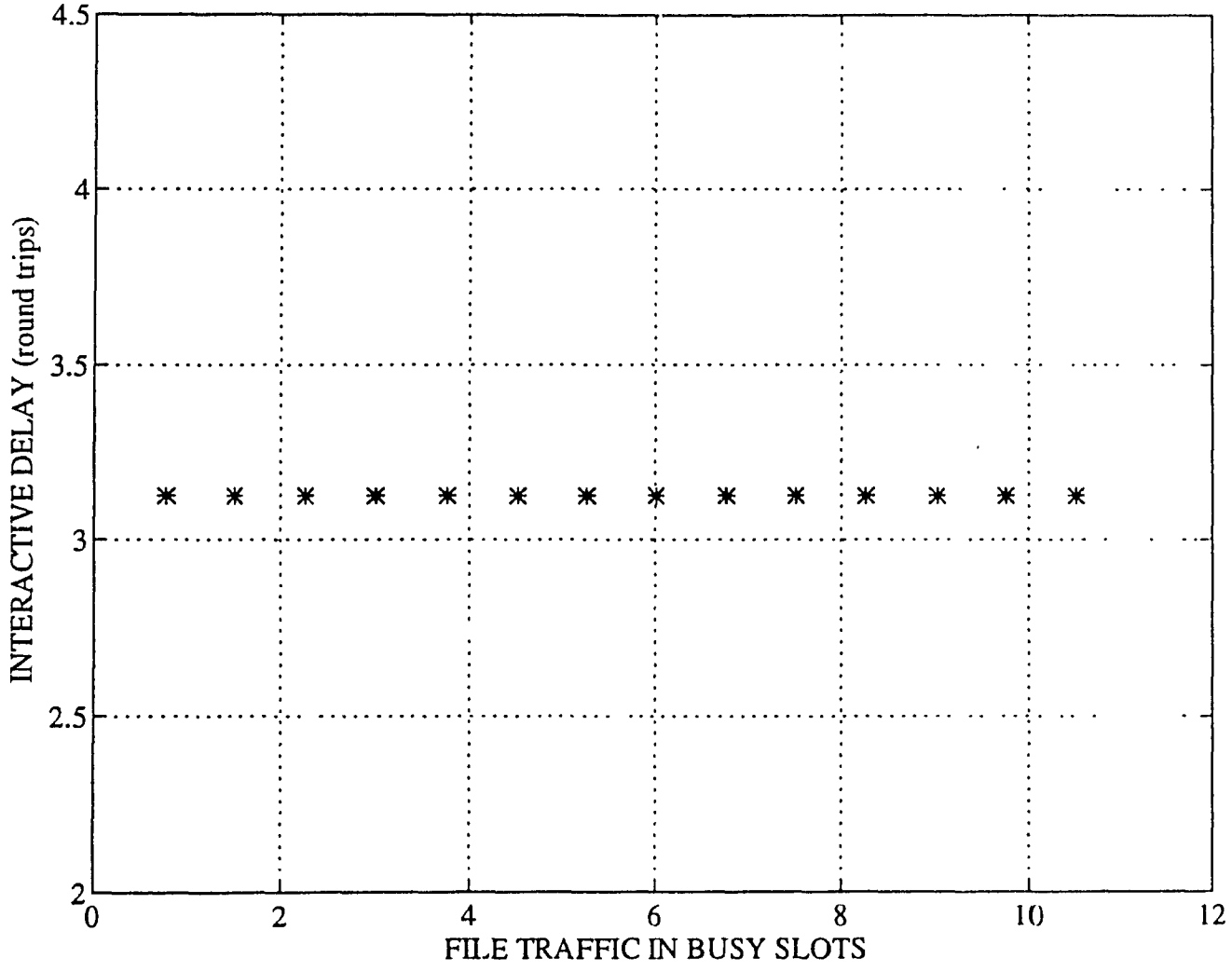


Figure 3.2 Interactive data Delay vs. File Traffic (case a)

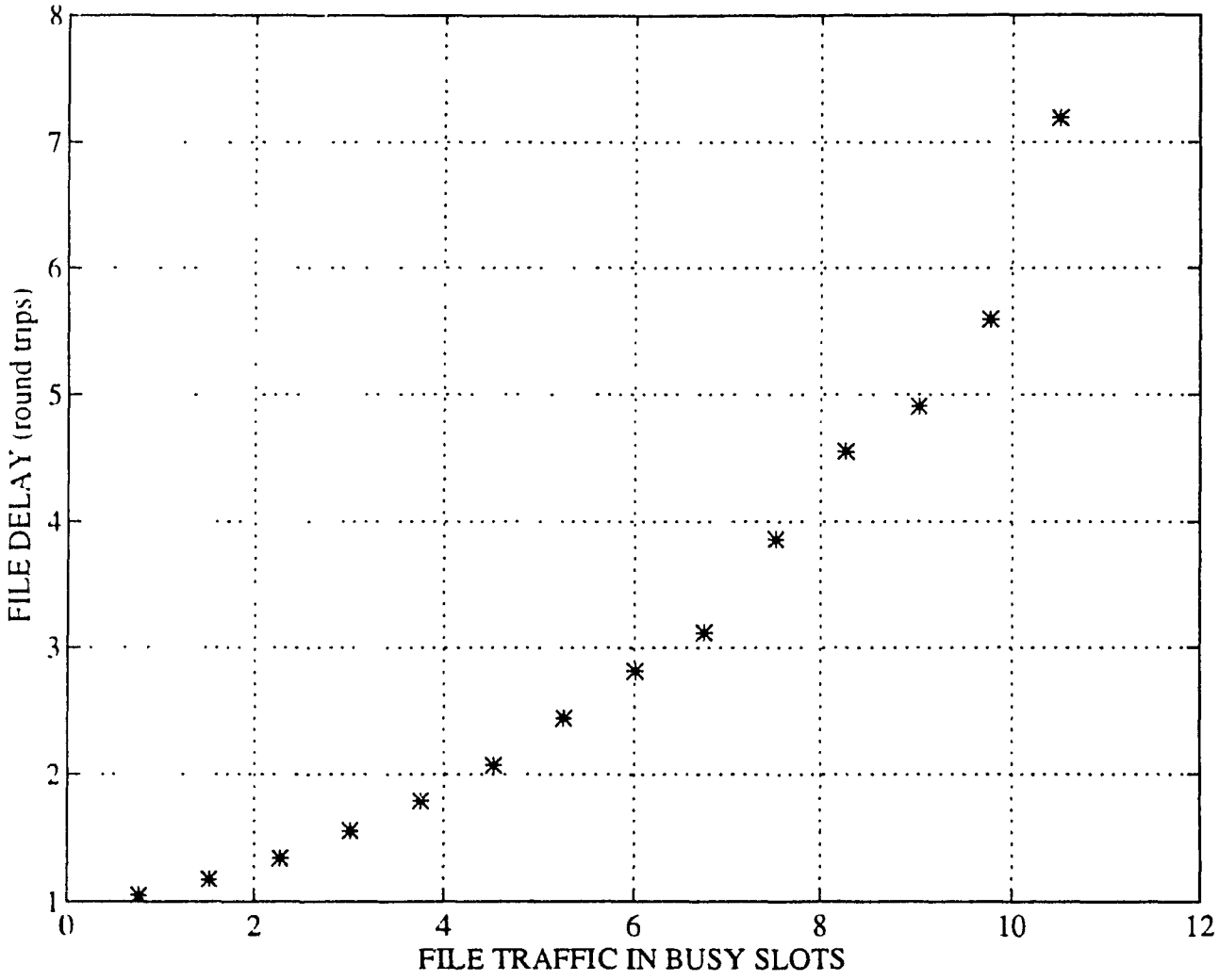


Figure 3.3 File Delay vs. File Traffic (case a)

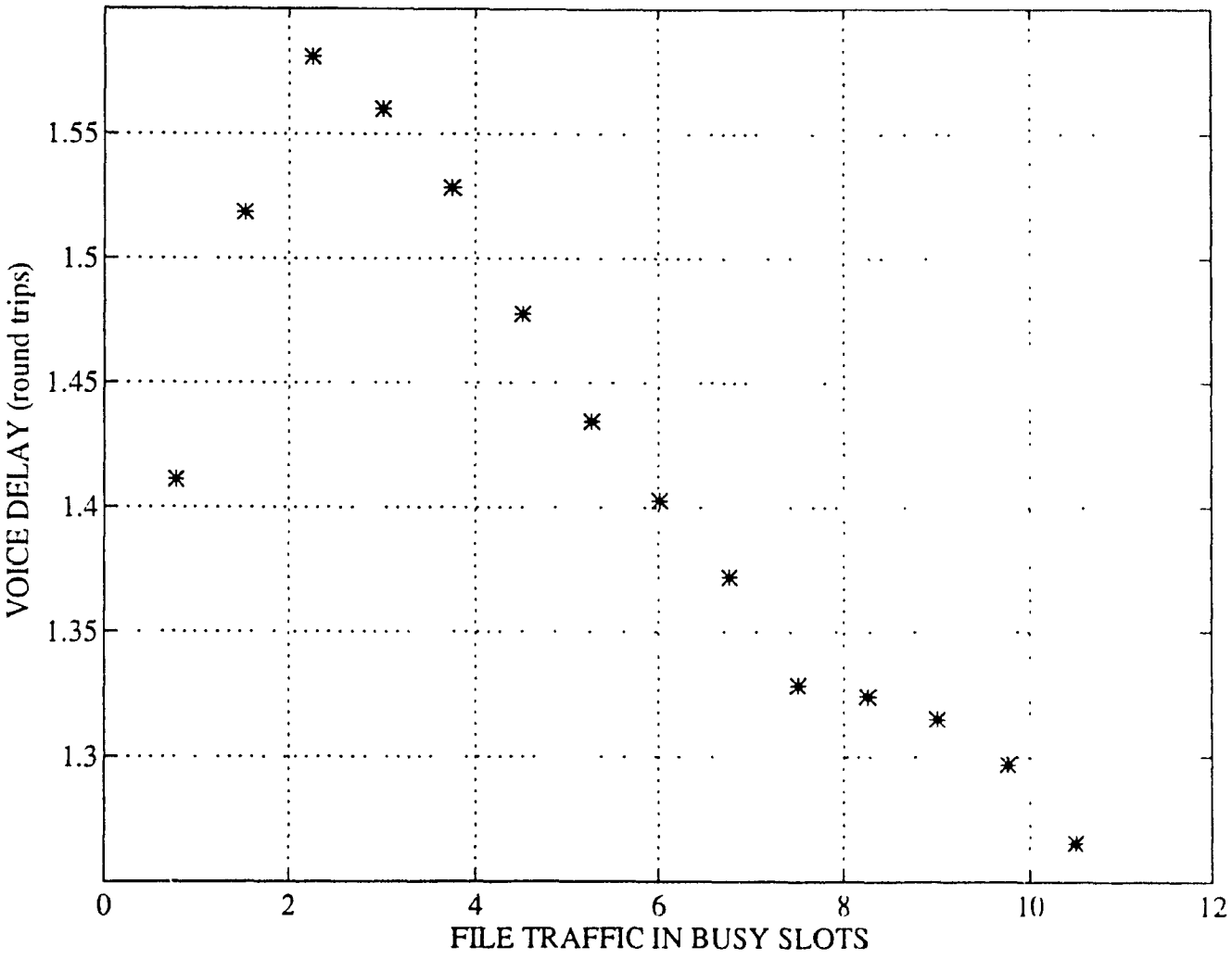


Figure 3.4 Voice Delay vs. File Traffic (case a)

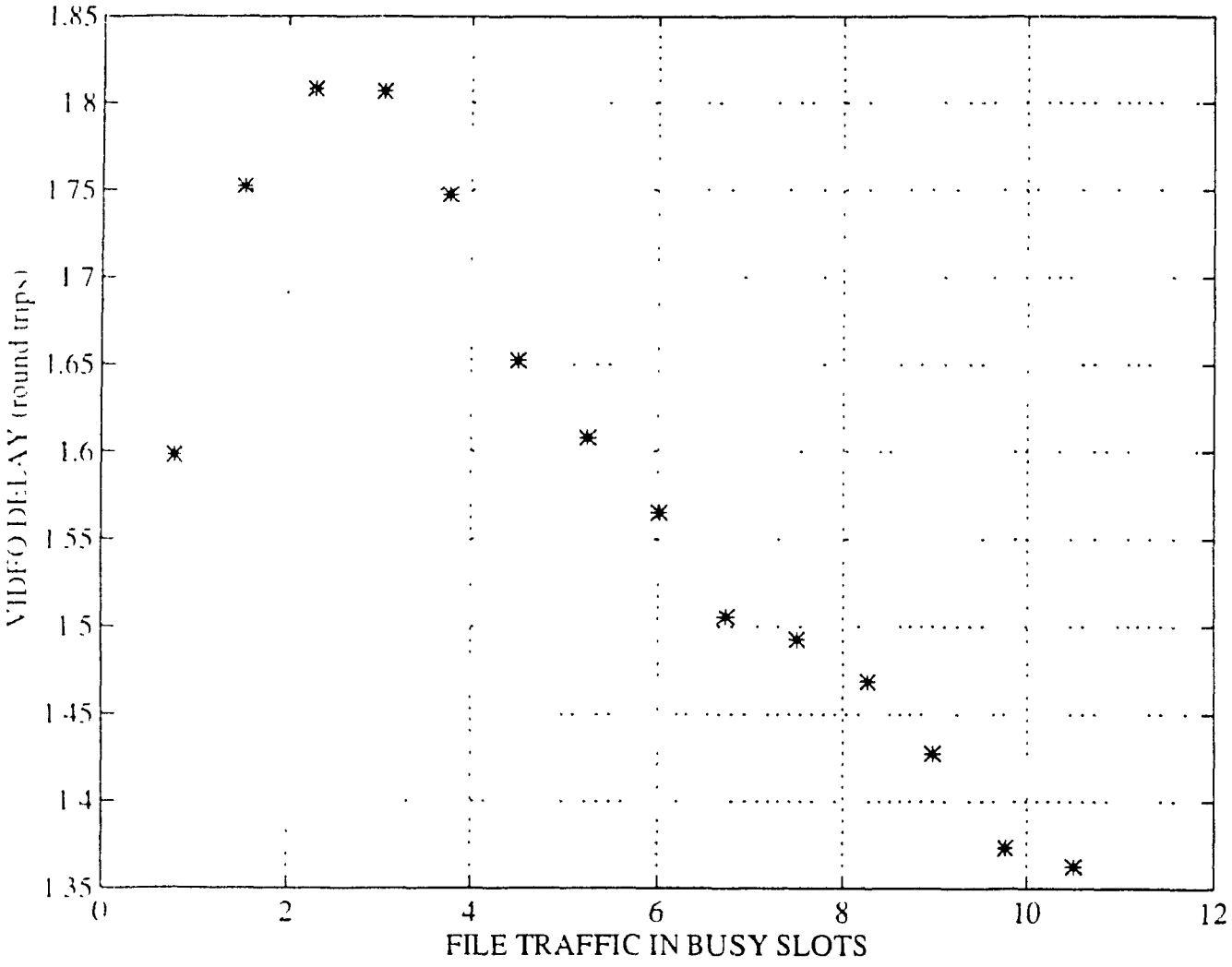


Figure 3.5 Video Delay vs. File Traffic (case a)

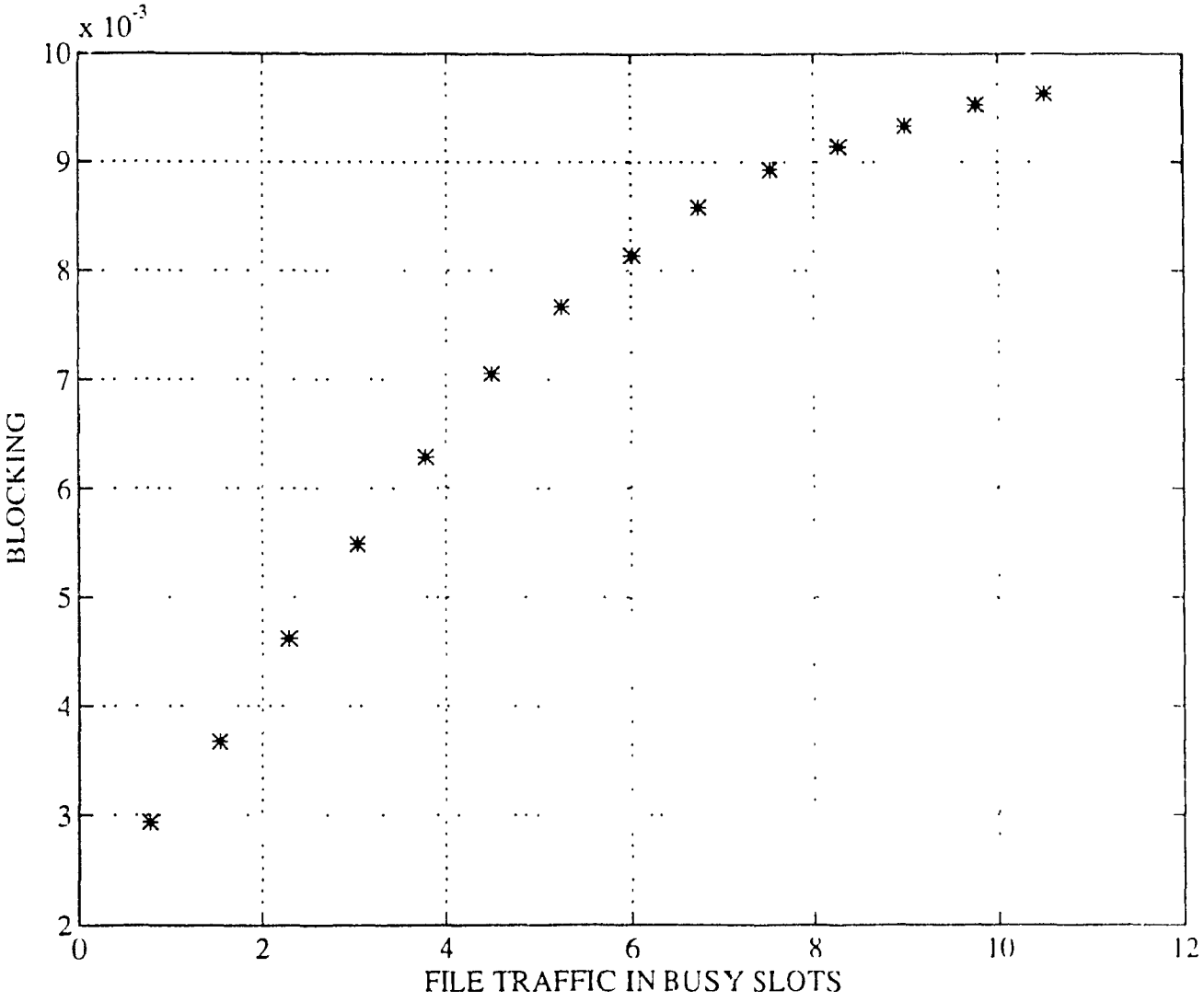


Figure 3.6 Blocking Probability vs. File Traffic (case a)

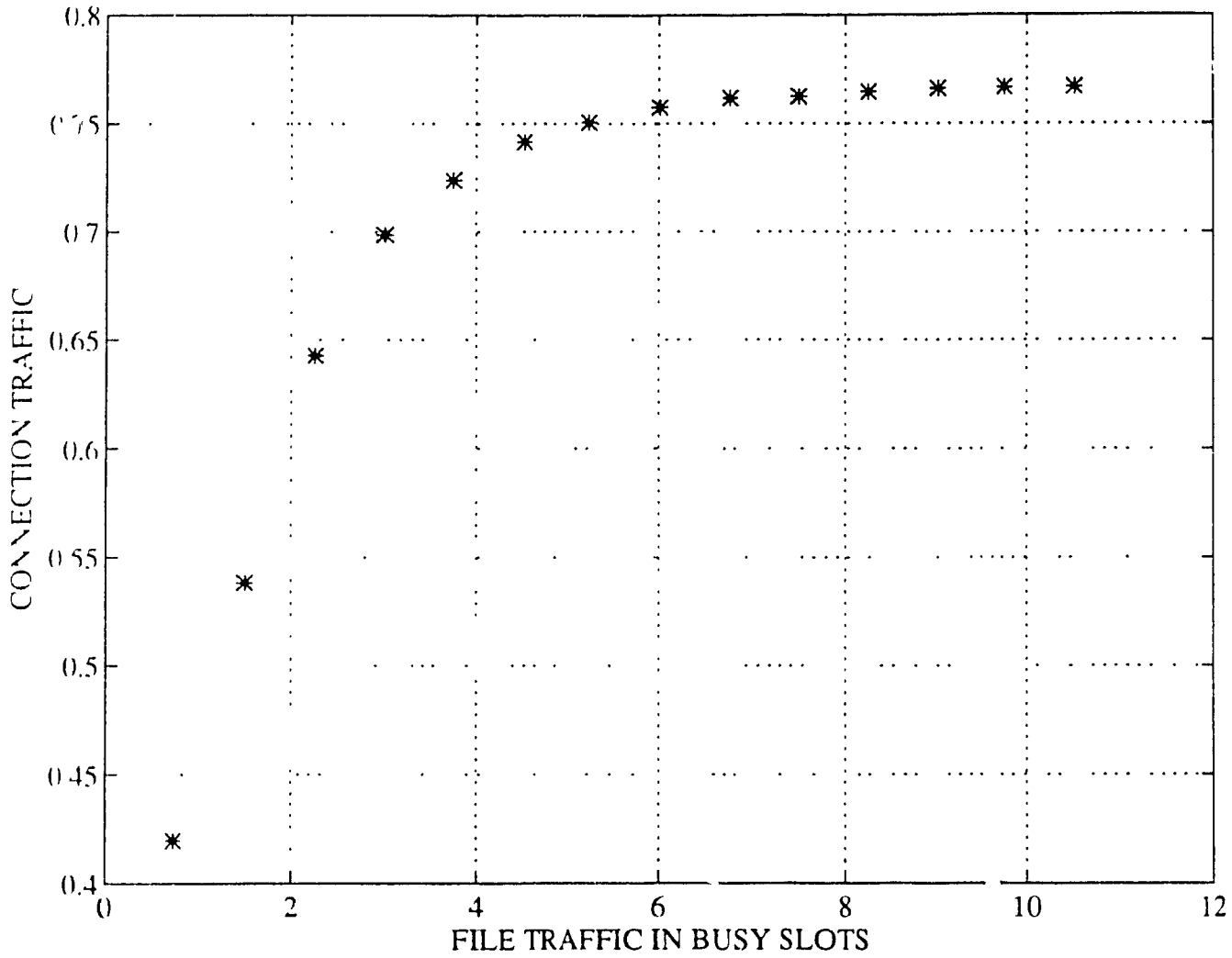


Figure 3.7 Throughput of Connection Traffic vs. File Traffic (case a)

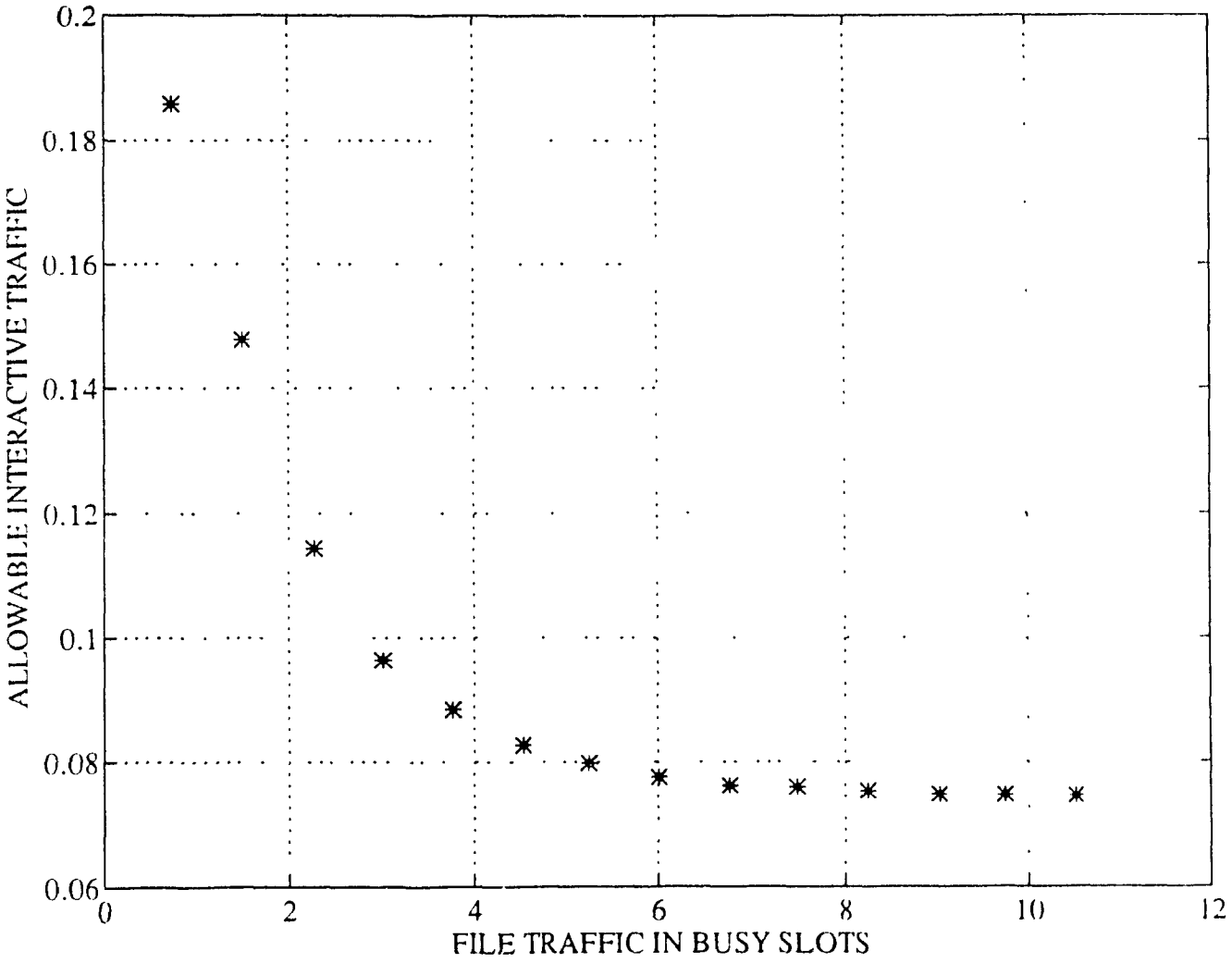


Figure 3.8 Allowable Interactive Traffic vs. File Traffic (case a)

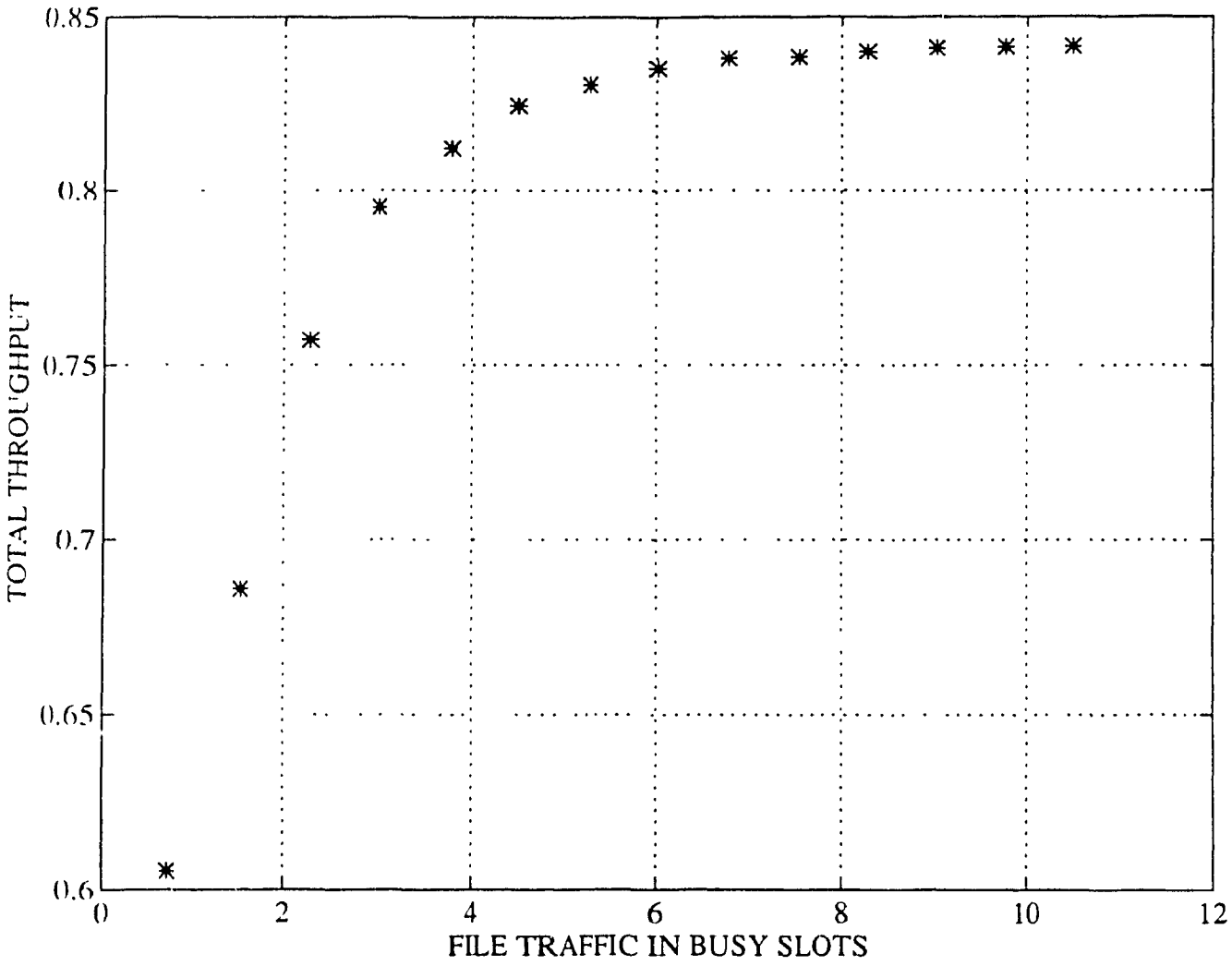


Figure 3.9 Total Throughput vs. File Traffic (case a)

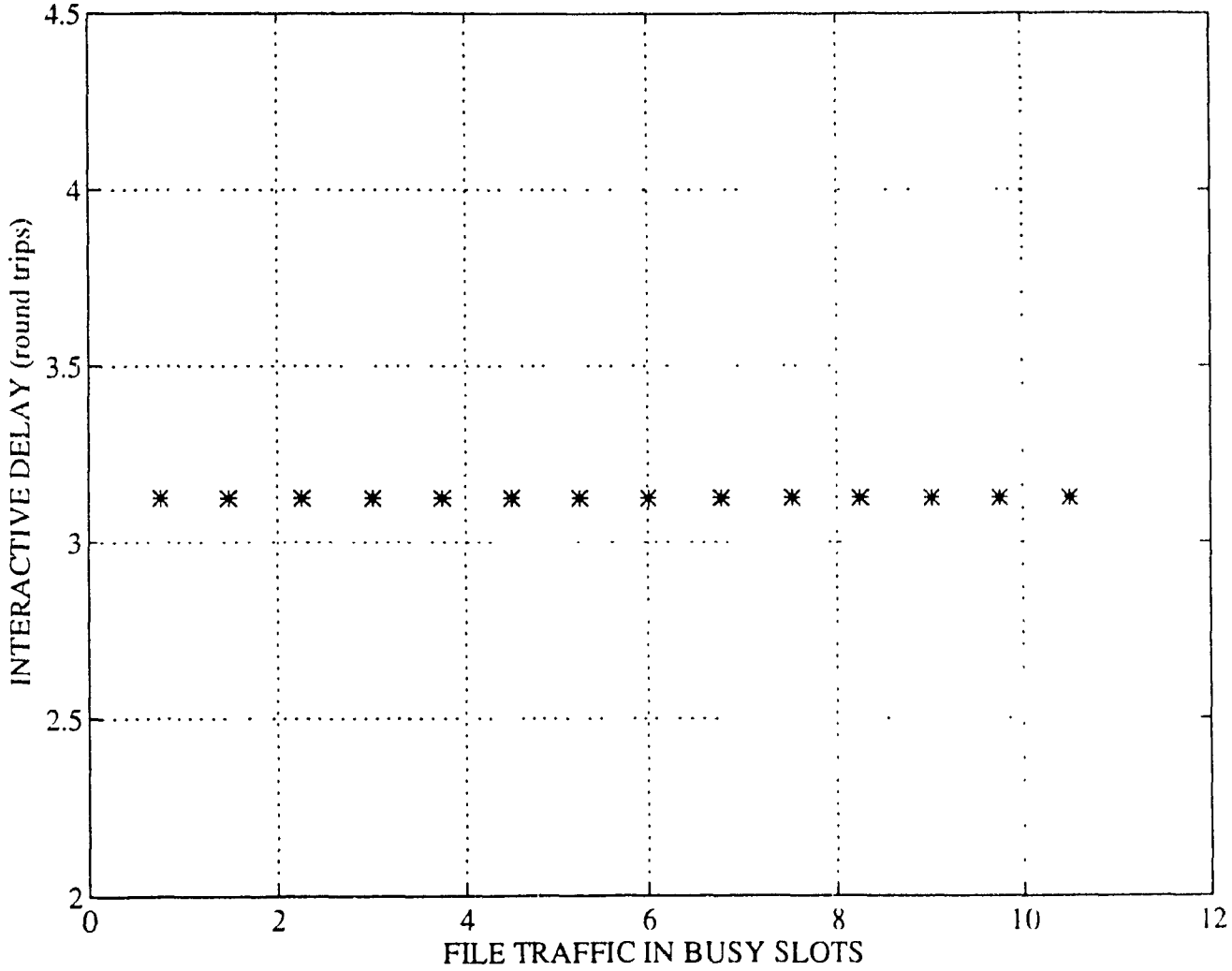


Figure 3.10 Interactive data Delay vs. File Traffic (case b)

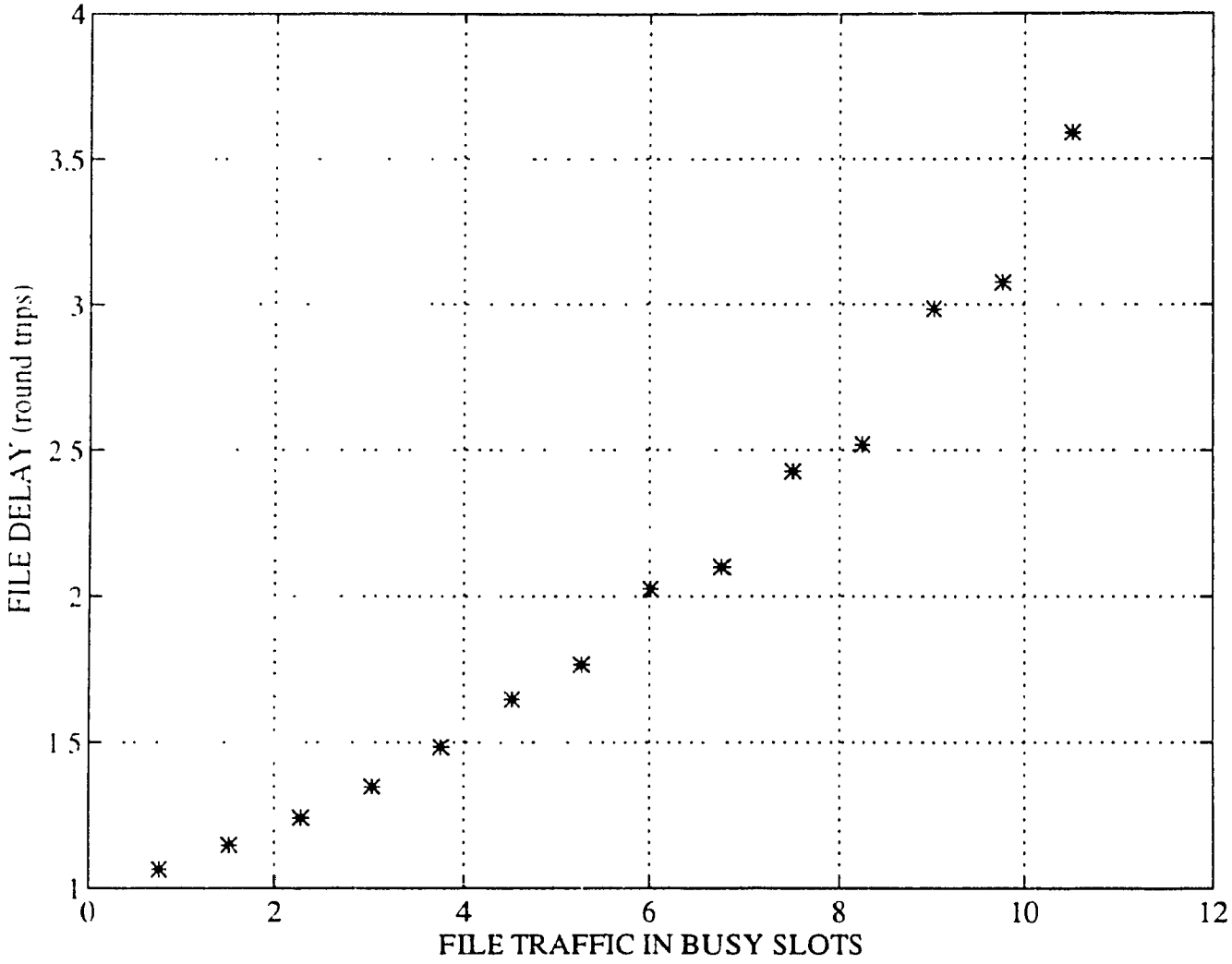


Figure 3.11 File Delay vs. File Traffic (case b)

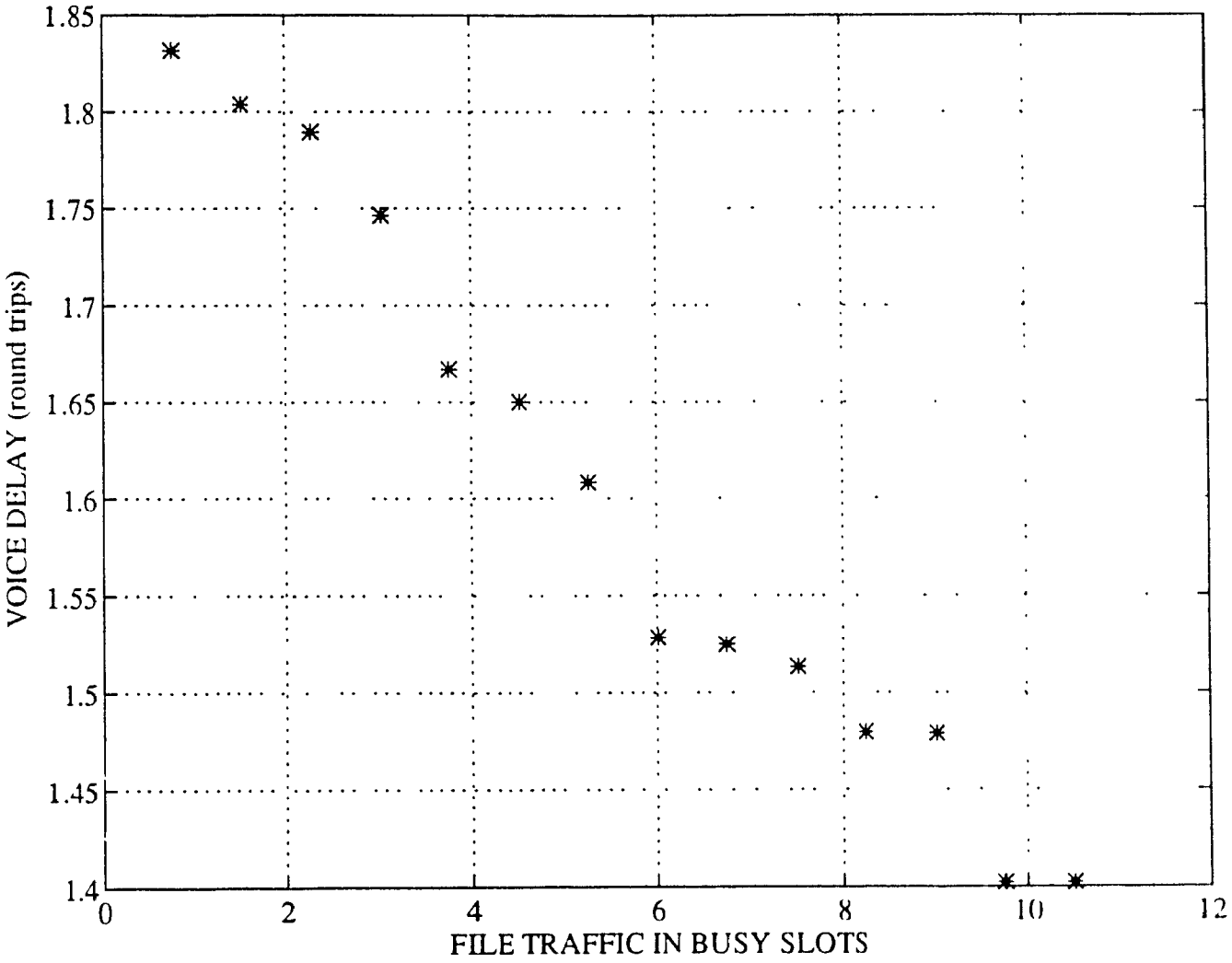


Figure 3.12 Voice Delay vs. File Traffic (case b)

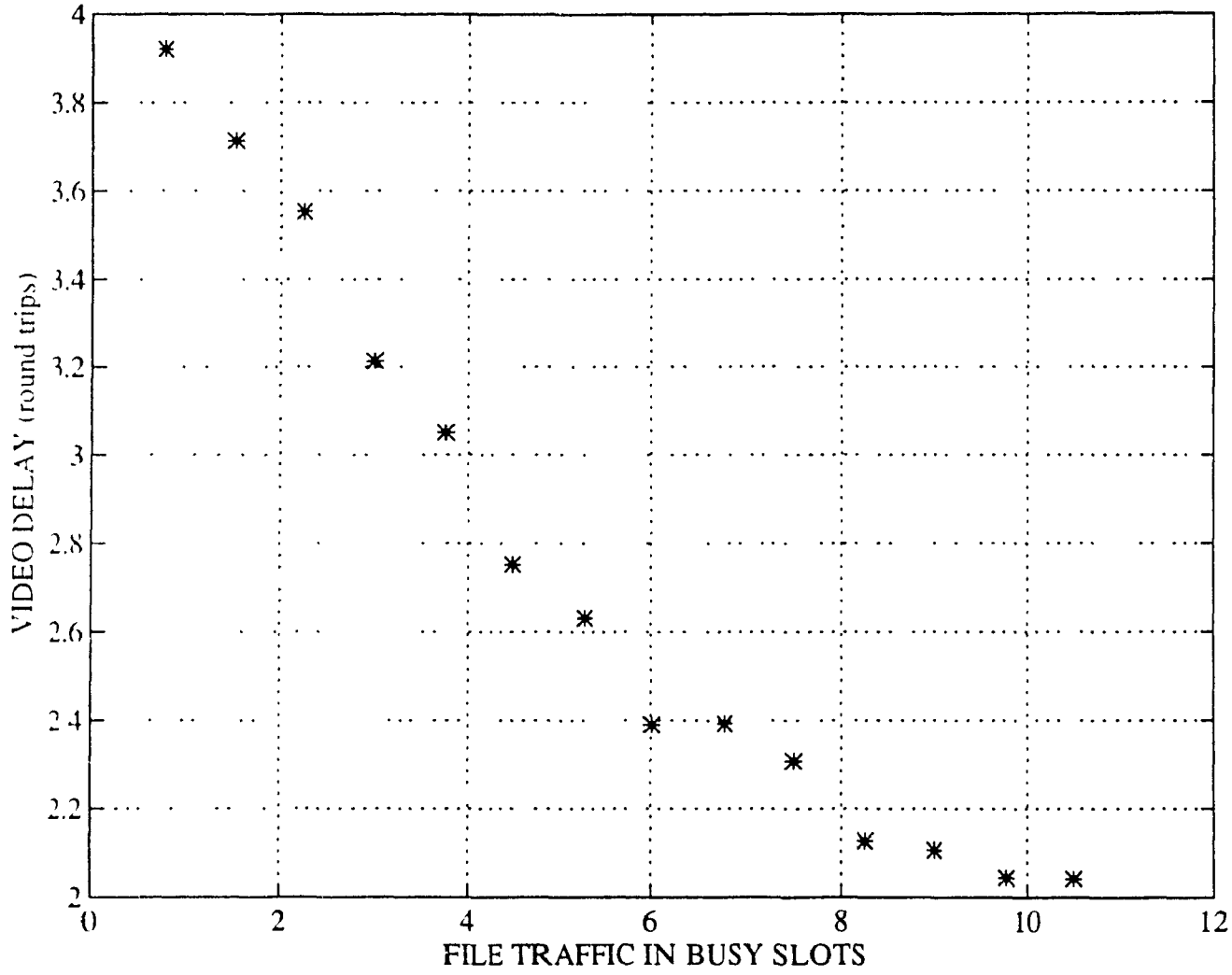


Figure 3.13 Video Delay vs. File Traffic (case b)

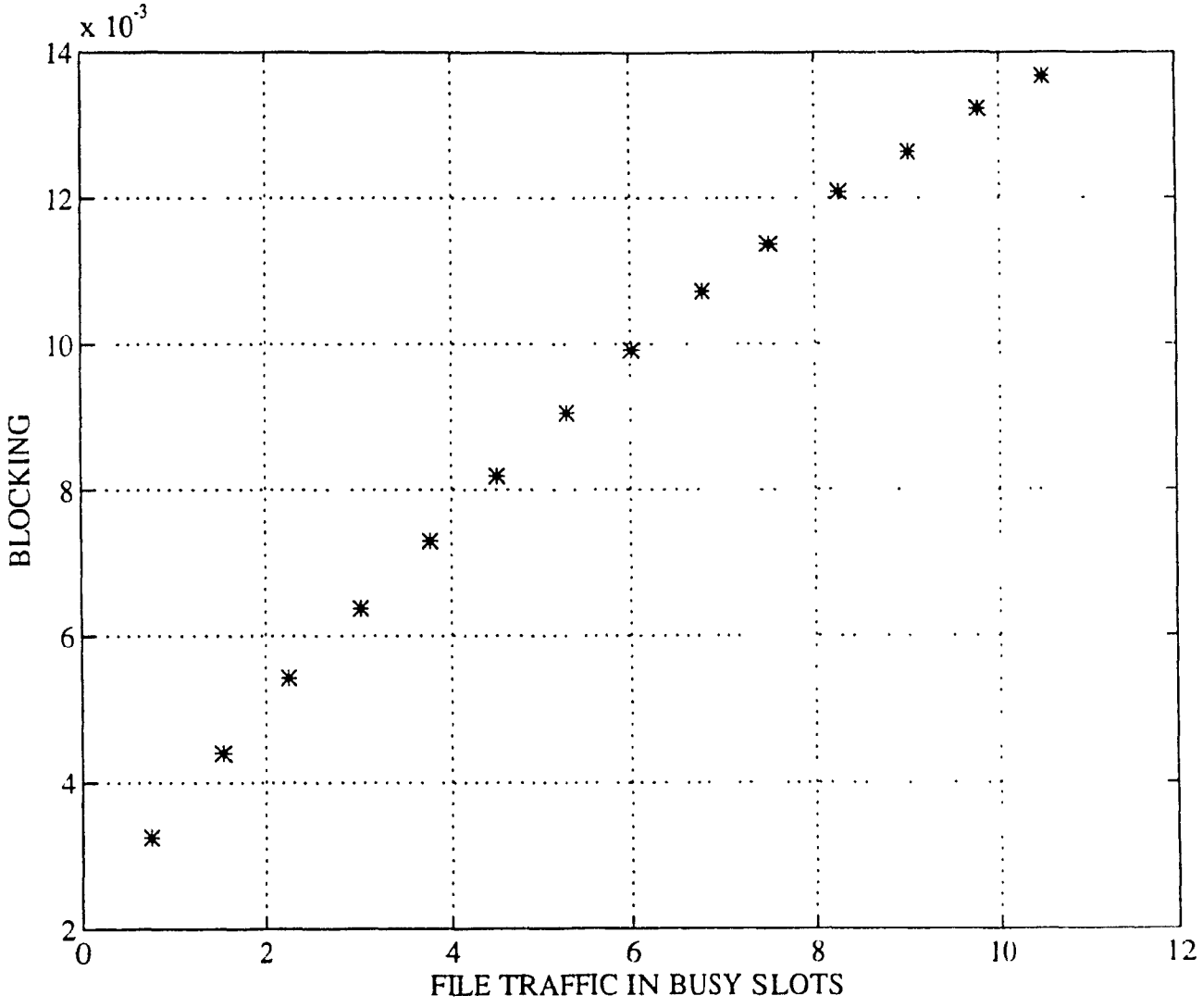


Figure 3.14 Blocking Probability vs. File Traffic (case b)

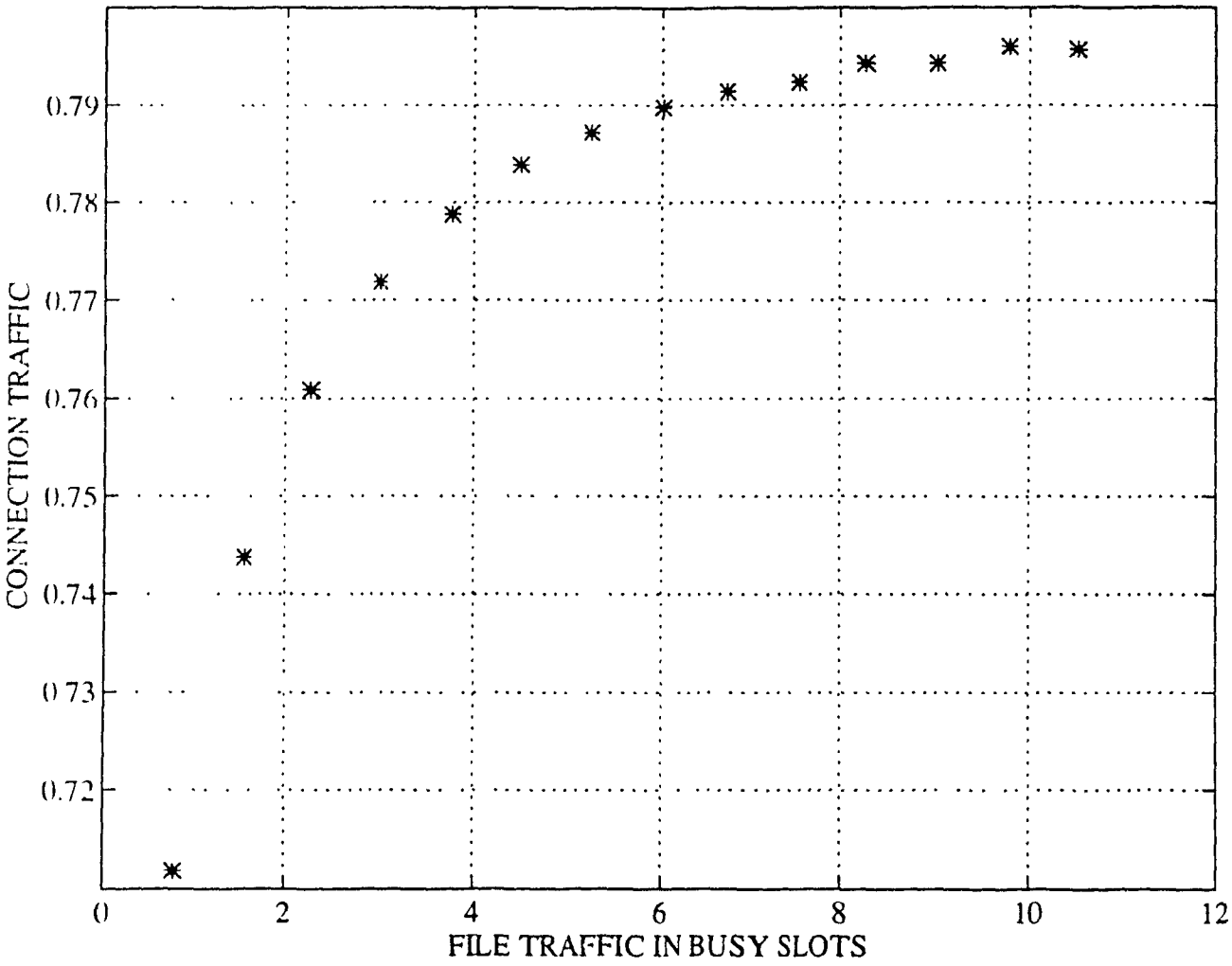


Figure 3.15 Throughput of Connection Traffic vs. File Traffic (case b)

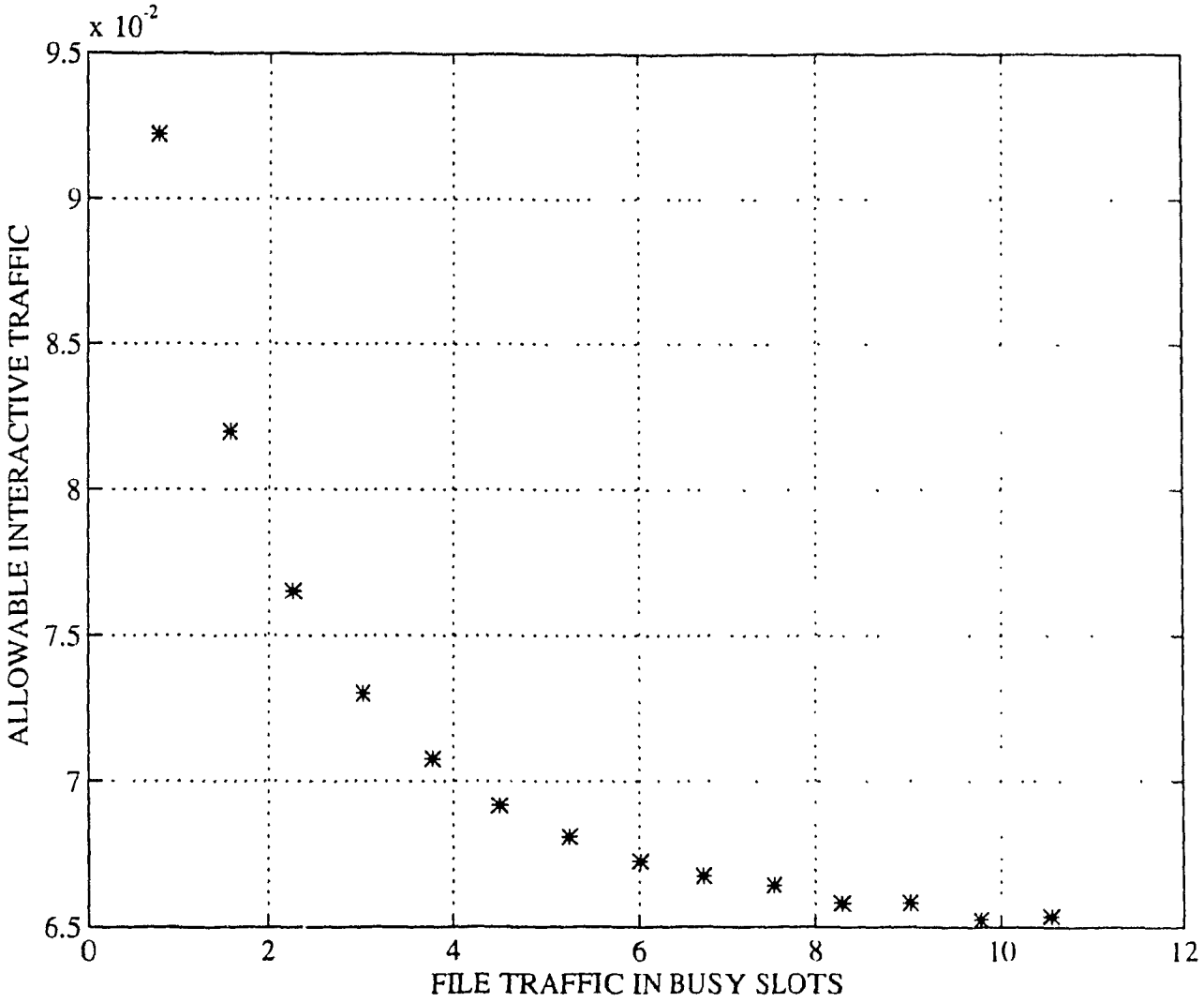


Figure 3.16 Allowable Interactive Traffic vs. File Traffic (case b)

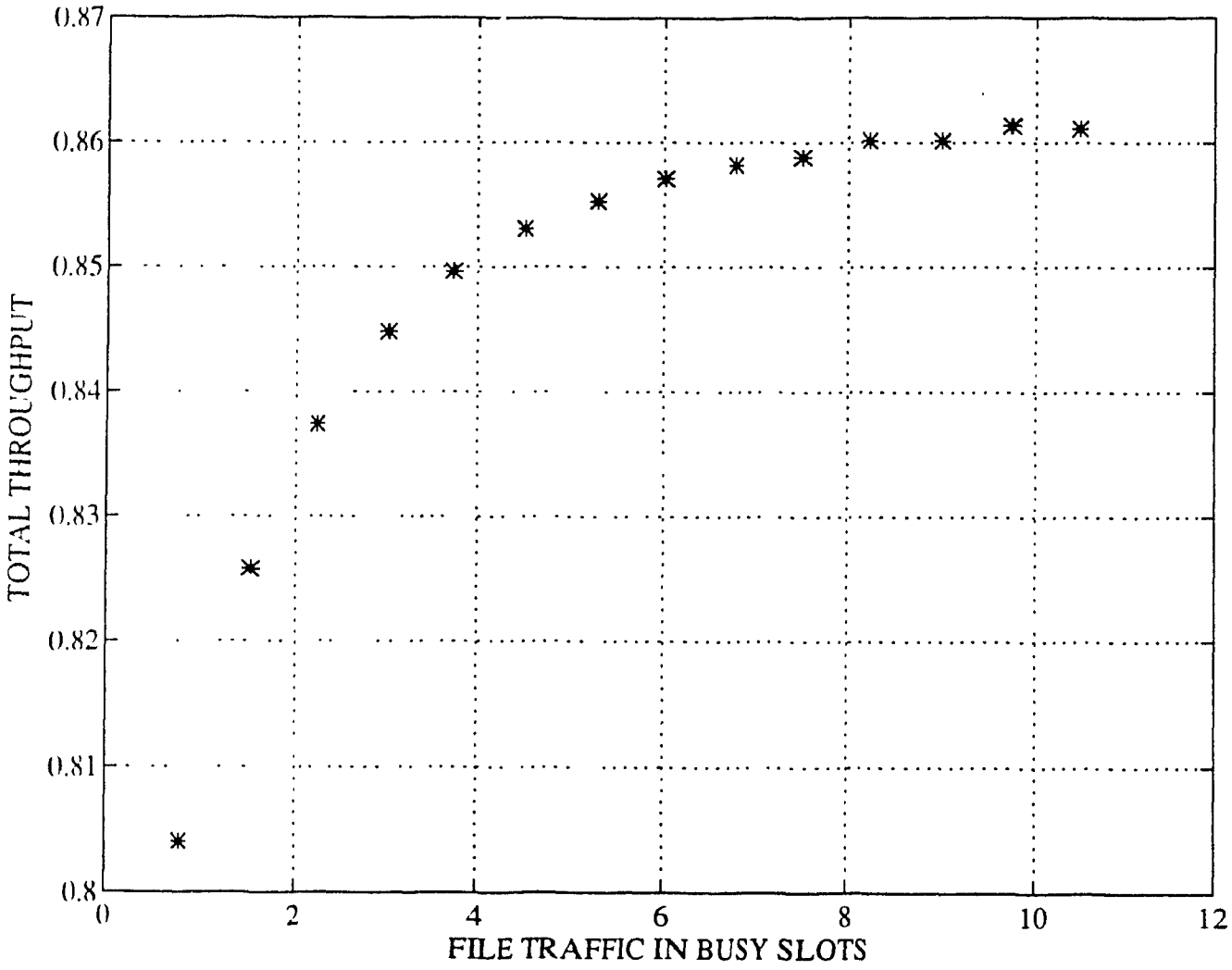


Figure 3.17 Total Throughput vs. File Traffic (case b)

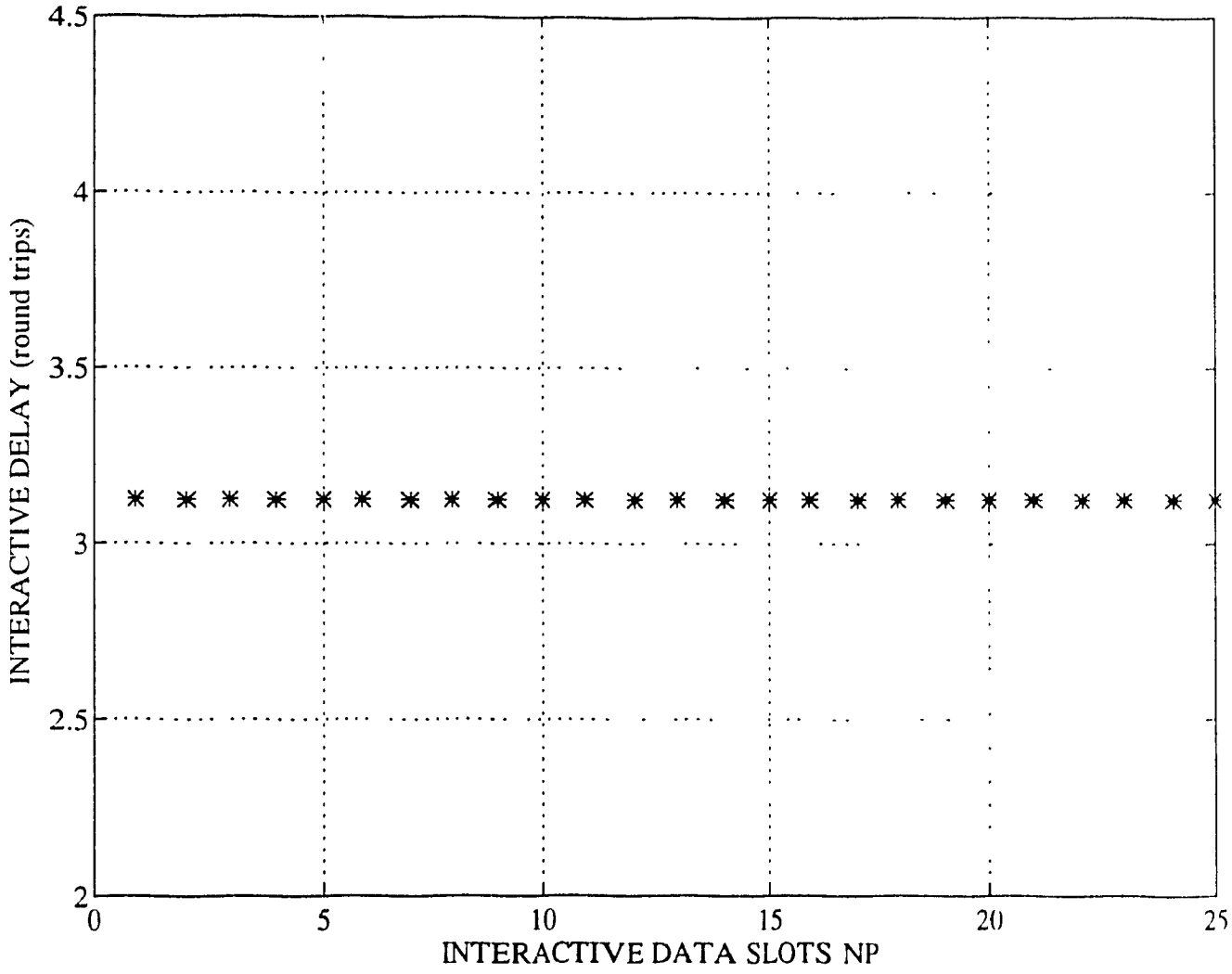


Figure 3.18 Interactive data Delay vs. Interactive Slots N_p

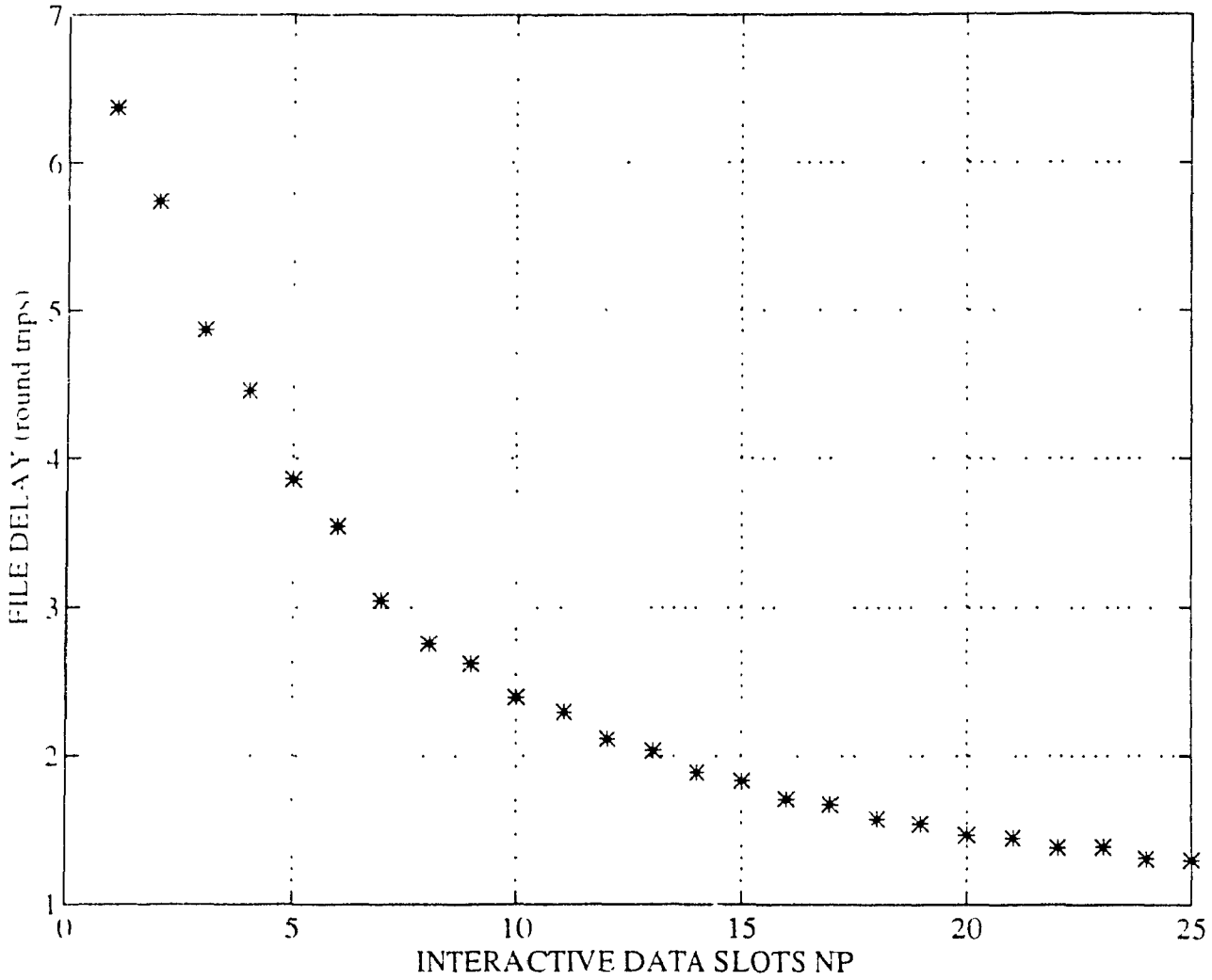


Figure 3.19 File Delay vs. Interactive Slots N_P

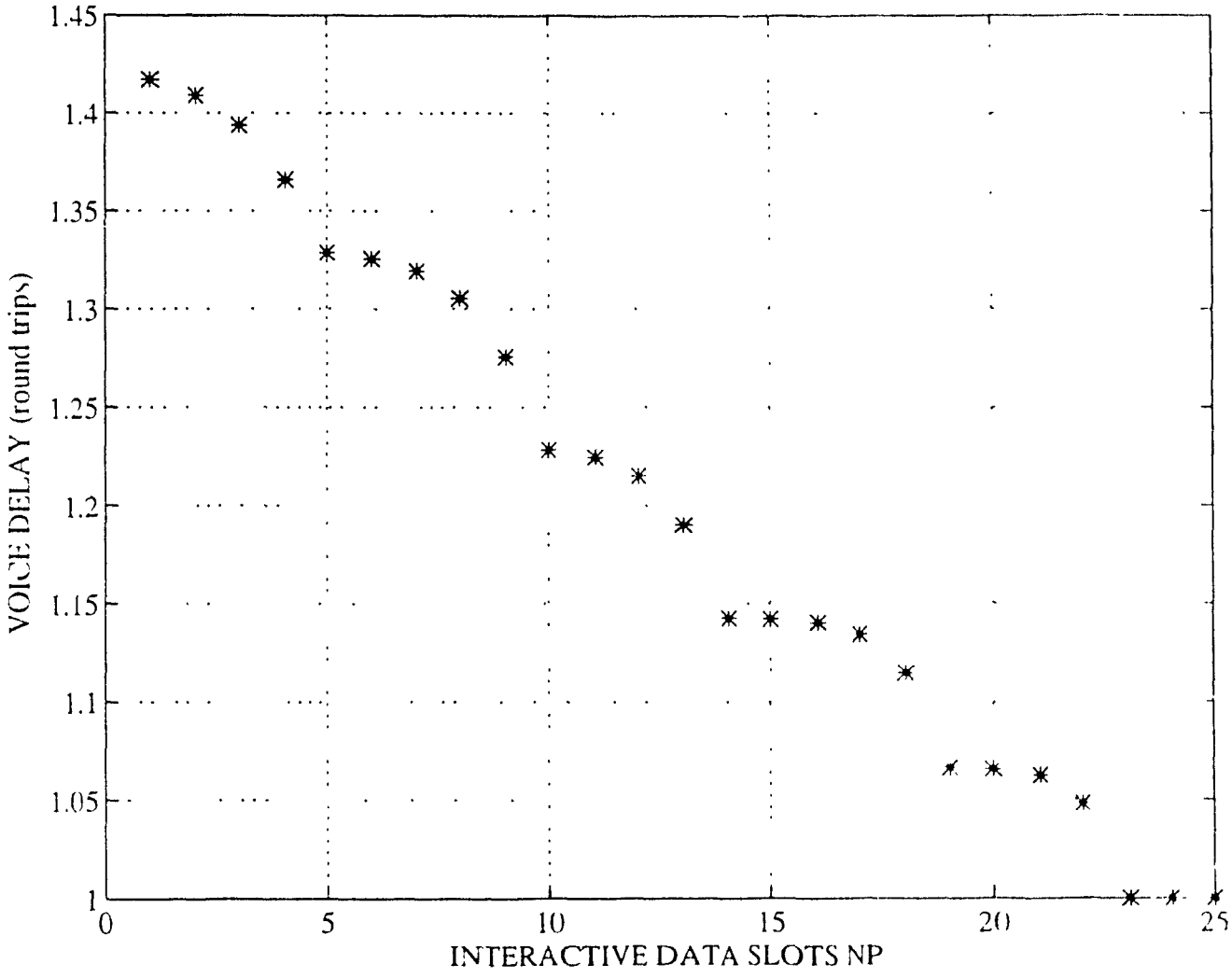


Figure 3.20 Voice Delay vs. Interactive Slots N_p

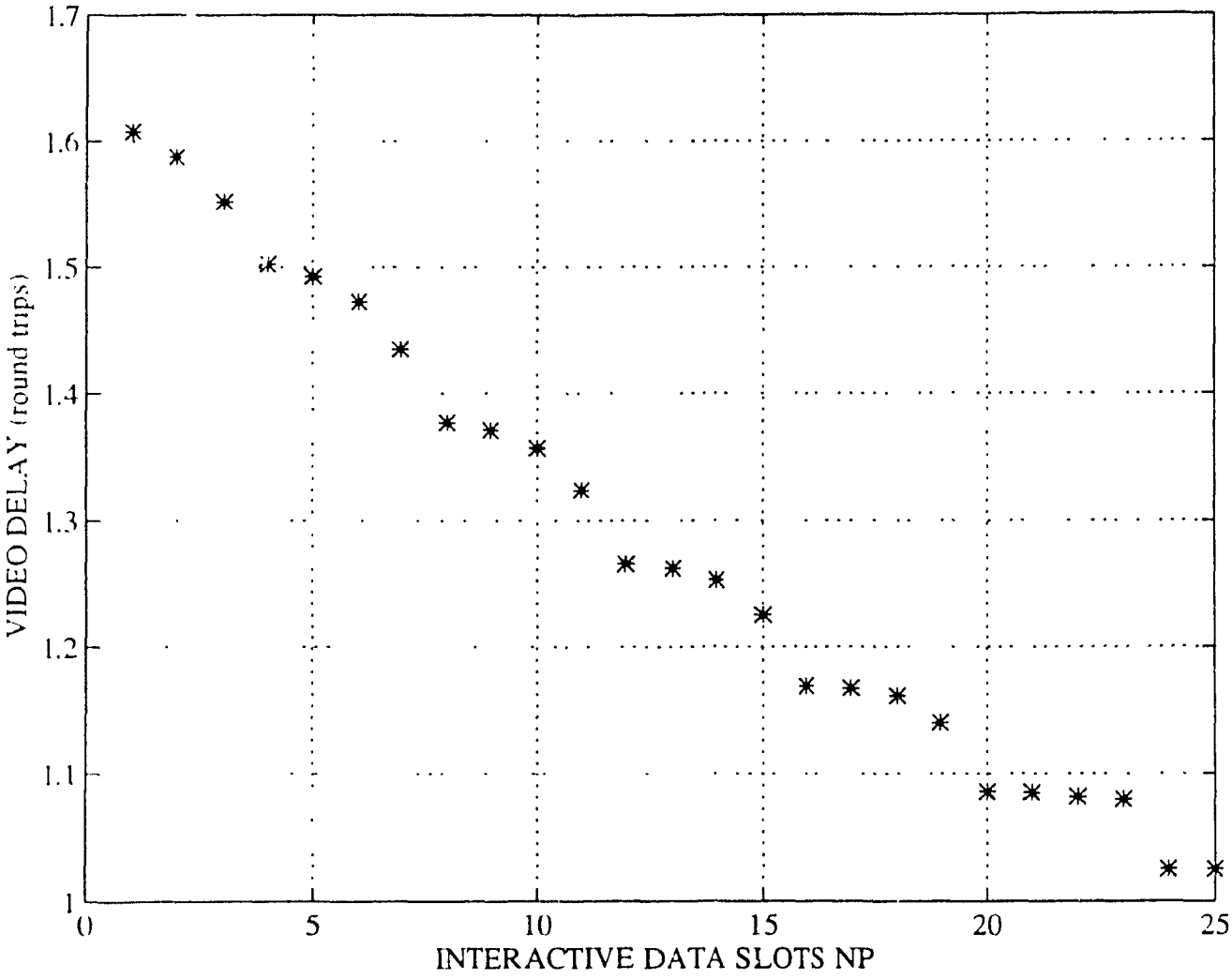


Figure 3.21 Video Delay vs. Interactive Slots N_p

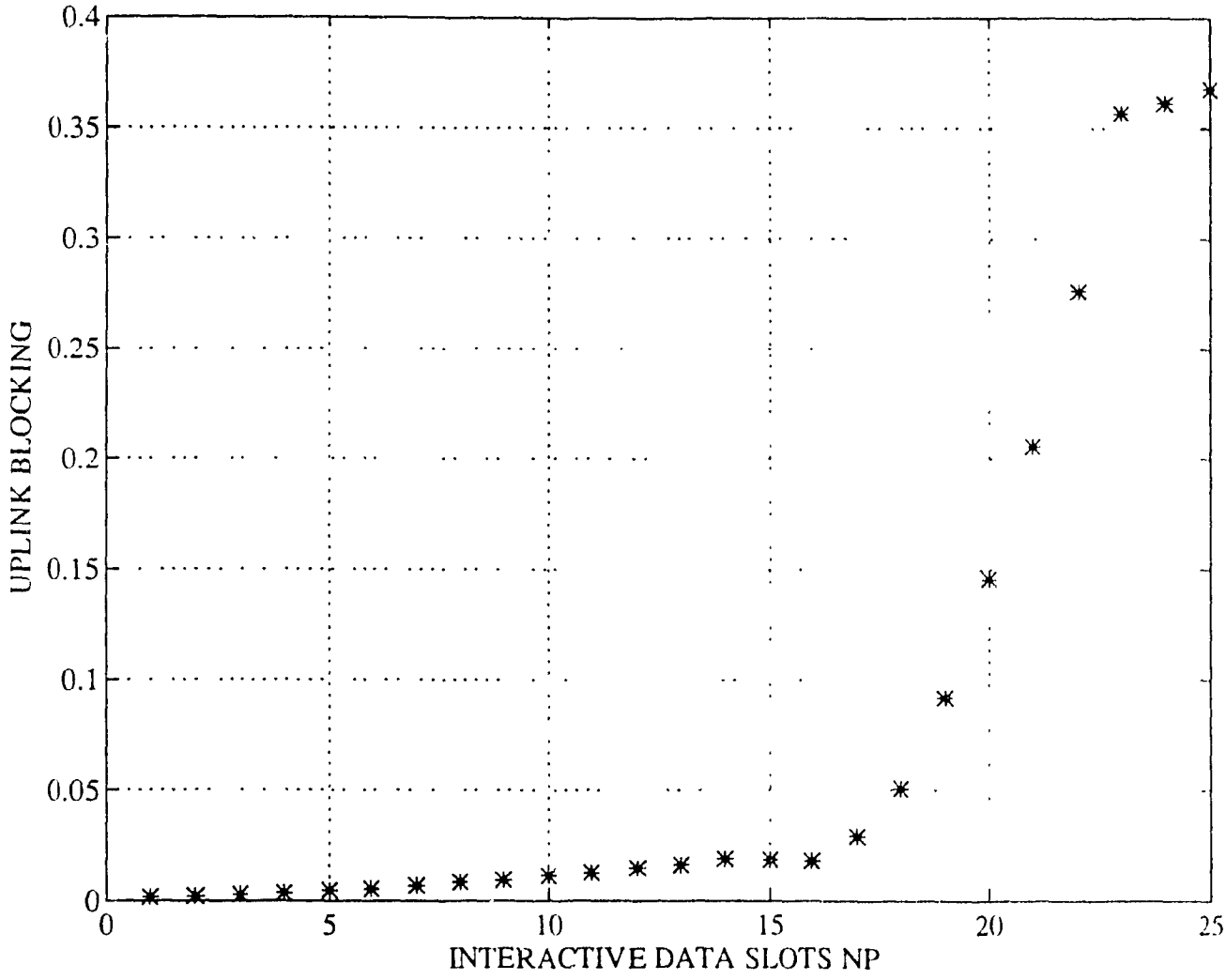


Figure 3.22 Blocking Probability vs. Interactive Slots N_I

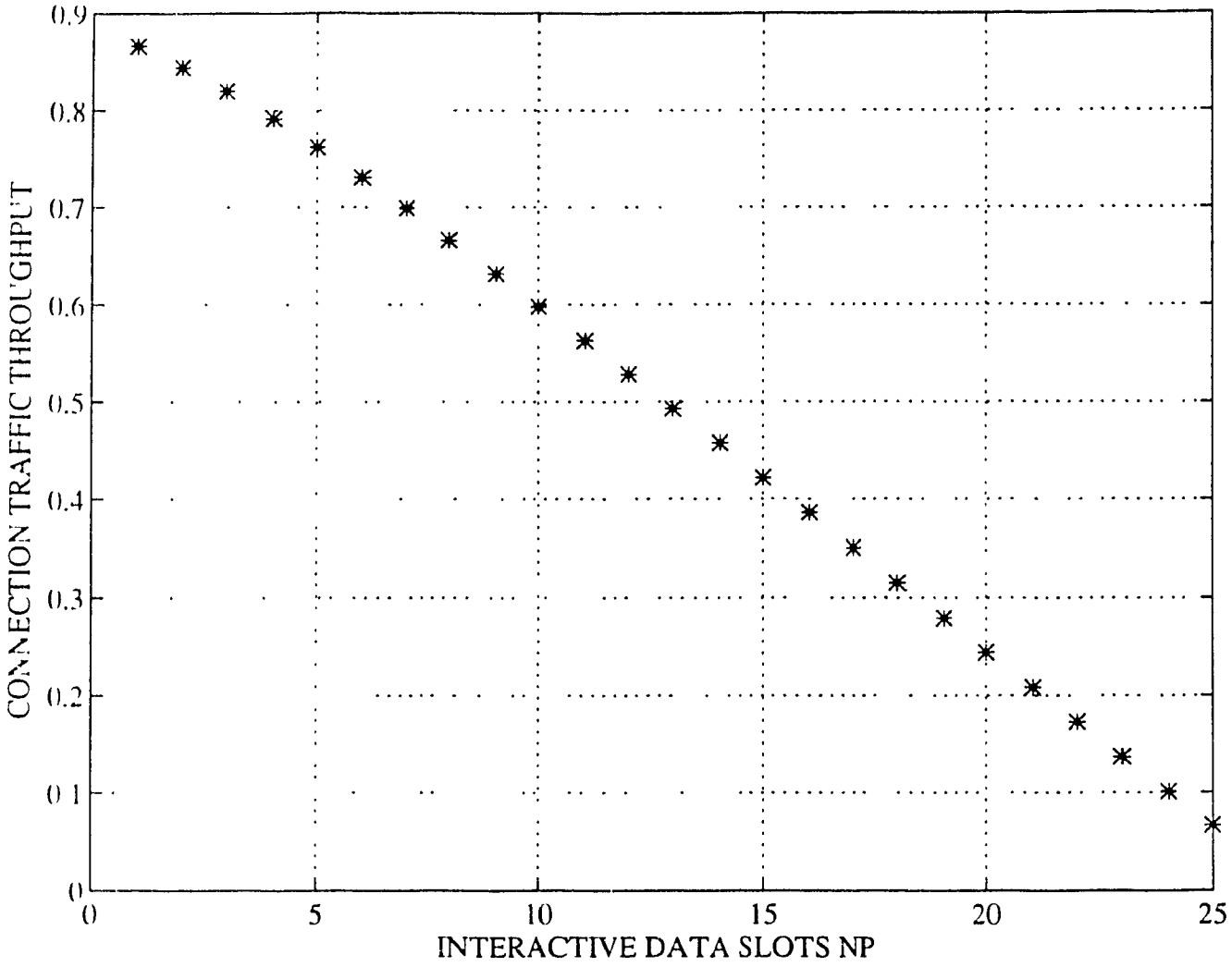


Figure 3.23 Throughput of Connection Traffic vs. Interactive Slots N_P

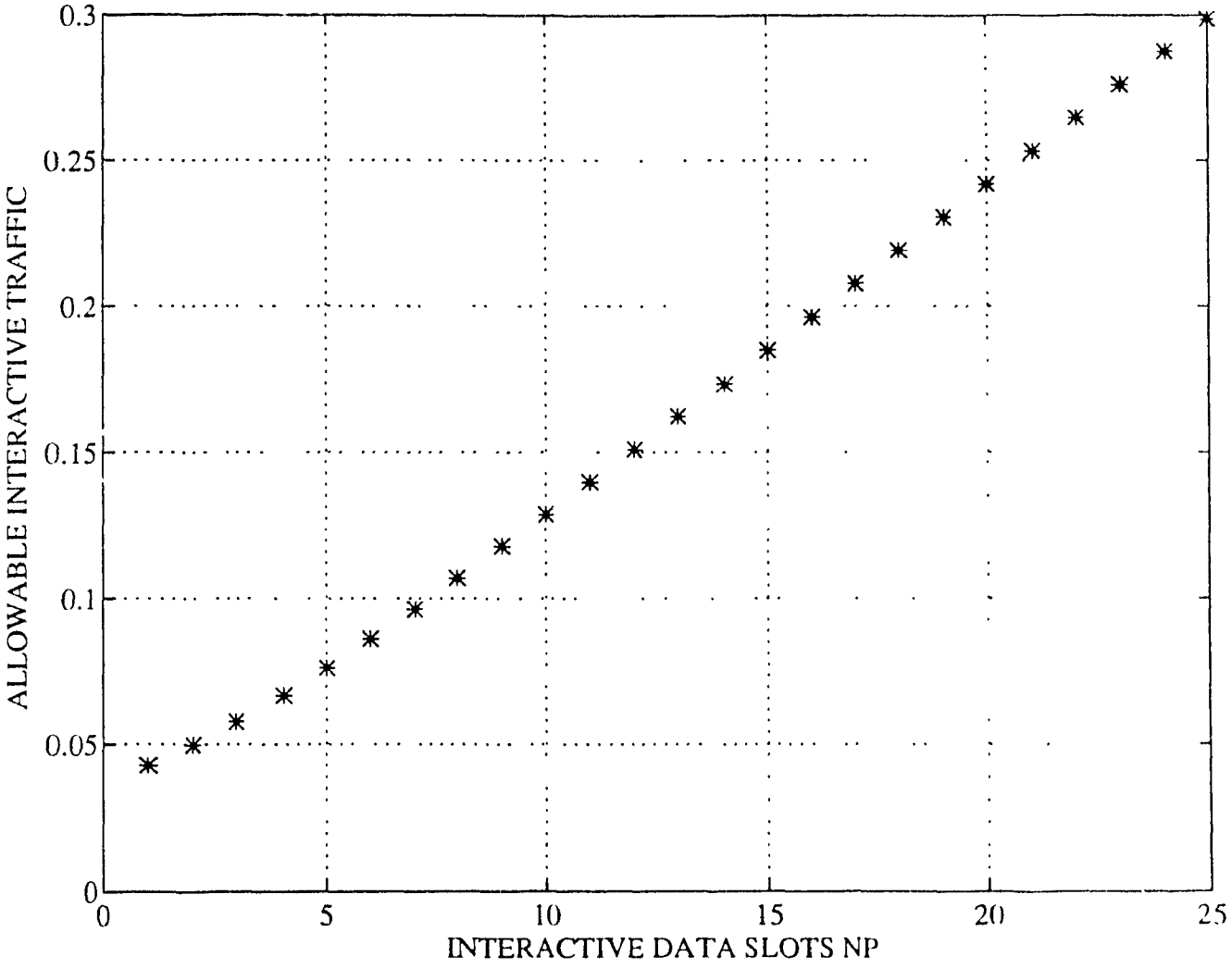


Figure 3.24 Allowable Interactive Traffic vs. Interactive Slots N_p

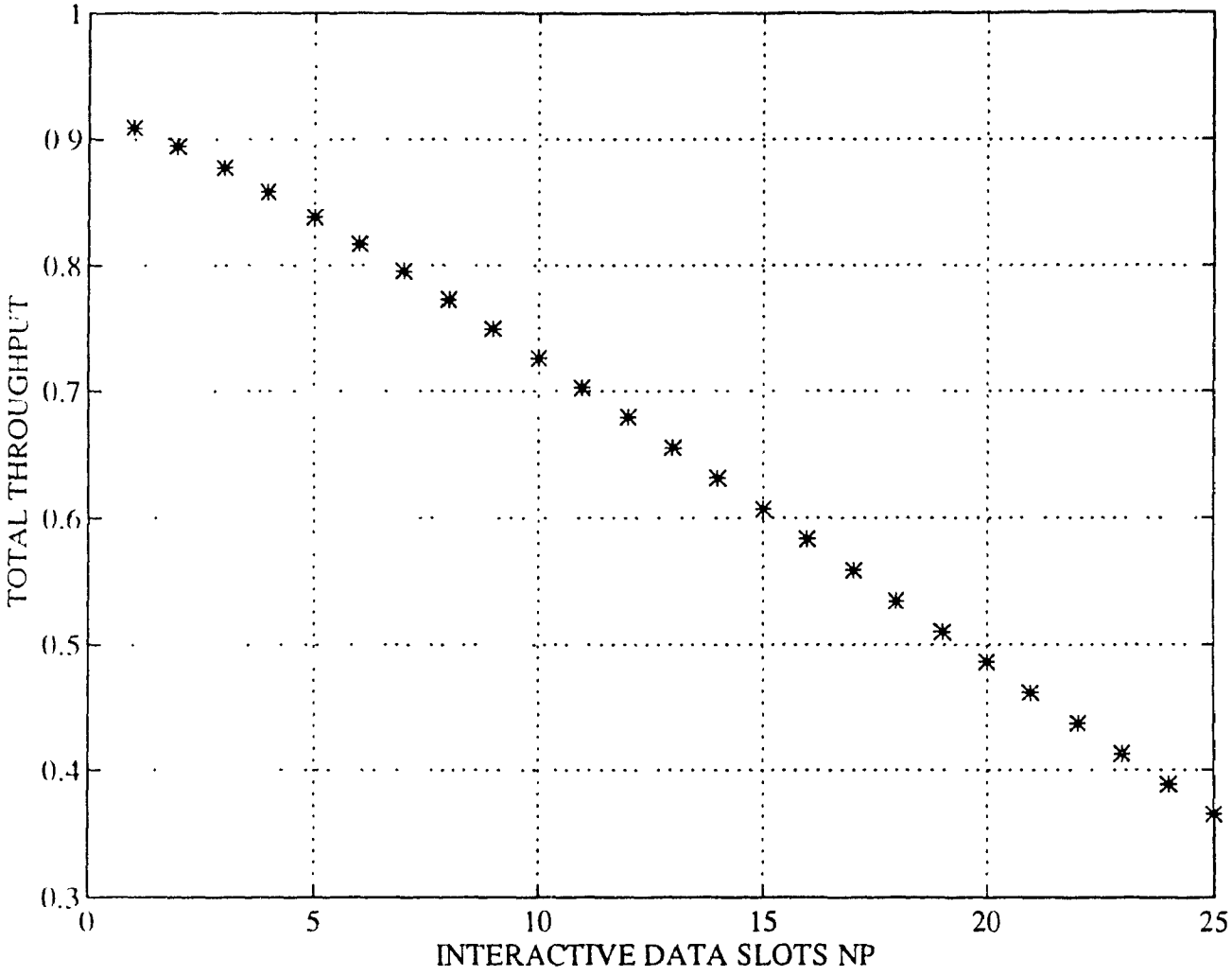


Figure 3.25 Total Throughput vs. Interactive Slots N_P

CHAPTER 4

ARCHITECTURE AND ANALYSIS OF THE PRIORITIZED KNOCKOUT SATELLITE SWITCH UNDER BALANCED LOAD AND HOL RESOLUTION

In this part we assume an internally nonblocking switch, however, it is possible that two or more input packets are destined to the same output. In this case we assume that the switch has an internal speed-up ratio similar to the knockout configuration coupled with a priority structure (Fig.4.1) and a Head Of Line (HOL) resolution technique to solve the problem of head of line blocking at the input buffers.

The switch employs both input and output buffering with the apparent advantage over the original knockout proposal in that priority structure is impeded here. Also to be introduced is a three phase resolution technique to accomodate priority and services integration as well as to improve the switch blocking characteristics

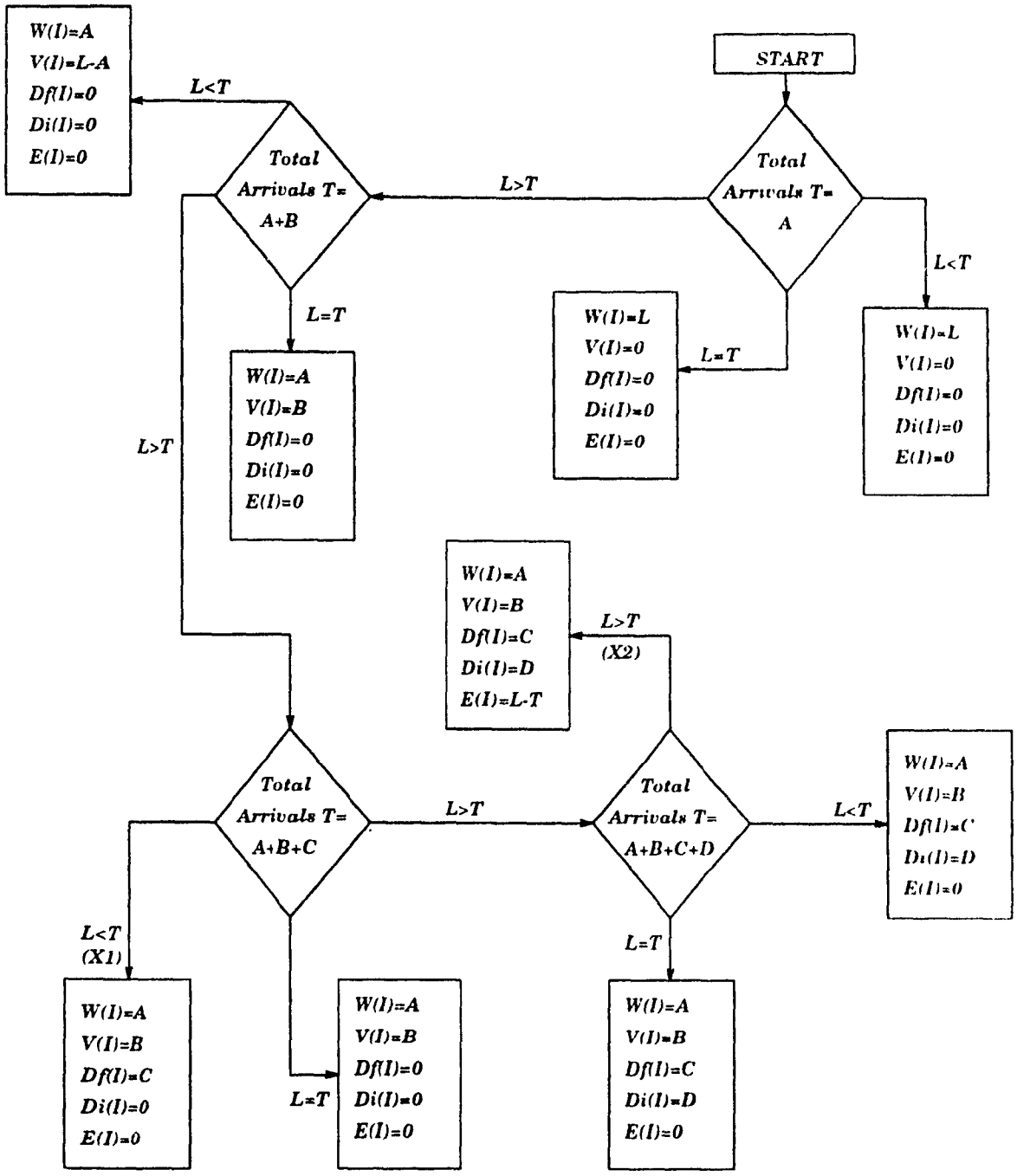
4.1. The Prioritized Knockout Switch

The best delay-thoughput performance is achieved when there is no blocking on the switch; in other words, in one time slot all input packets destined to one particular output port are delivered without any collision and then are queued in the output buffer and exited one by one. This configuration is extremely complicated and impractical in most cases [34]. To minimize the hardware and optimize the performance, the knockout switch was suggested [9]. The principle of the knockout switch is that within each time slot, up to L ($L << N$) packets can be transmitted to one particular output simultaneously. In an NxN switch there is the possibility of having n packets contending for one particular output($n \leq N$). The random selection of L packets out of n contending packets is the principle of the pure knockout switch; however in this work an algorithm is followed to

give a certain priority to some users over others. In other words the choice of the allowable L packets satisfy some rules.

The priority rules apply to different traffic types destined to the same output as in Fig.4 1. Priority is given to video signals over all other signals given that all signals contend for one output. Voice signals are given priority over file and interactive data signals if video signals are not among the contending signals for one output. Finally in case there are only file and interactive data signals contending for one output port, priority is given to file data users. In other words, if more than L (speed-up ratio) packets are destined to the same output port, blocking occurs for all types of packets (voice, file data, interactive data) plus the excess of the video packets exceeding L. Associated with each switch output I, 5 entities are stored in 5 registers W(I), V(I), $D_f(I)$,$D_i(I)$, and E(I) denoting respectively, accepted number of packets for video , voice, file data, and interactive data signals, as well as the number of empty spaces, in output buffer of output port (I) (I=1,....,N).

As stated above, the algorithm in Fig.4.1 implements a straight forward priority rule where the order of priority is video packets, voice packets, file data packets, and interactive data packets. The algorithm proceeds naturally, e.g., if there are A video packets, B voice packets, C file packets, and D interactive packets all at different input ports and all destined to the same output port I, and if L-A-B < C, then W(I) = A, V(I) = B, $D_f(I)$ = L-A-B , $D_i(I)$=0, E(I) = 0 and blocking occurs for C-(L-A-B) file packets and all arriving D interactive data packets (point X1 in Fig.4.1). In another scenario(point X2 in Fig.4 1), W(I)= A, V(I) = B, $D_f(I)$ = C, $D_i(I)$ = D, such that the sum of the accepted packets = A+B+C+D < L, in this case some empty spaces E(I) = L-A-B-C-D will be left in that output buffer and can be used to advantage as will follow shortly.



A=#of Video Packets; *B=# of Voice Packets;* *C=# of File Data Packets;* *D=# of Intve Data Packets,*

W(I)=Accepted # of Video Packets to one Output; *Df(I)=Accepted # of File Data Packets to one Output;*

V(I)=Accepted # of Voice Packets to one Output; *Di(I)=Accepted # of Intve Data Packets to one Output;*

E(I)= Leftover Switching Capacity out of L;

Figure 4.1: Knockout Switch Priority Flowchart

The analysis of this part corresponds to the different priority levels mentioned above. It is of course possible to change the hiearchy of the priority levels; For example we could give priority to voice users over video users etc. The analysis would be changed accordingly. In other words the analysis is still going to hold for different types of traffic and for different priority levels. Following through Fig.4.1, the packets blocking probabilities due to contention among packets of input beams are derived from the priority structure and given by:

$$\bar{P}_{bw} = \frac{\sum_{i=L+1}^{N} (i-L)\cdot P_A(i)}{\sum_{i=0}^{N} i\cdot P_A(i)} \tag{4.1}$$

$$\bar{P}_{bv} = \frac{\sum_{i=L+1}^{N} P_A(i)\left(\sum_{j=0}^{N-i} j\cdot P_B(j)\right) + \sum_{i=0}^{L-1} P_A(i) \sum_{j=L-i+1}^{N-i} (j-(L-i))\cdot P_B(j)}{\sum_{j=0}^{N} j\cdot P_B(j)} \tag{4.2}$$

$$\bar{P}_{bdf} = \frac{\sum_{i=L+1}^{N} P_A(i)\sum_{j=0}^{N-i} P_B(j) \sum_{k=0}^{N-i-j} k\cdot P_C(k) + \sum_{i=0}^{L} P_A(i)\sum_{j=0}^{L-i} P_B(j) \sum_{k=L-i-j+1}^{N-i-j} (k-(L-i-j))\cdot P_C(k)}{\sum_{k=0}^{N} k\cdot P_C(k)}$$

$$+ \frac{\sum_{i=0}^{L} P_A(i) \sum_{j=L-i}^{N-i} P_B(j) \sum_{k=0}^{N-i-j} k\cdot P_C(k)}{\sum_{k=0}^{N} k\cdot P_C(k)} \tag{4.3}$$

$$\bar{P}_{bdi} = \frac{\sum_{i=L+1}^{N} P_A(i) \sum_{j=0}^{N-i} P_B(j) \sum_{k=0}^{N-i-j} P_C(k) \sum_{l=0}^{N-i-j-k} l \cdot P_D(l)}{\sum_{l=0}^{N} l \cdot P_D(l)} +$$

$$\frac{\sum_{i=0}^{L} P_A(i) \sum_{j=0}^{L-i} P_B(j) \left\{ \sum_{k=0}^{L-i-j} P_C(k) \sum_{l=L-i-j-k+1}^{[N-i-j-k]} [l-(L-i-j-k)] \cdot P_D(k) + \sum_{k=L-i-j}^{N-i-j} P_C(k) \sum_{l=0}^{N-i-j-k} l \cdot P_D(l) \right\}}{\sum_{l=0}^{N} l \cdot P_D(l)}$$

$$+ \frac{\sum_{i=0}^{L} P_A(i) \sum_{j=L-i}^{N} P_B(j) \sum_{k=0}^{N-i-j} P_C(k) \sum_{l=0}^{N-i-j-k} l \cdot P_D(l)}{\sum_{l=0}^{N} l \cdot P_D(l)} \qquad (4.4)$$

where $\bar{P}_{bv}$, $\bar{P}_{bv}$, $\bar{P}_{bdf}$, and $\bar{P}_{bdi}$ are the packet blocking probabilities for video, voice, file data, and interactive data users respectively and P_A, P_B, P_C, P_D are the probabilities of arrival for the four kinds of traffic per slot per output. They are given by:

$$P_A(i) = \binom{N}{i} \left[\frac{\lambda_t \bar{R}_w}{N} \right]^i \left[1 - \frac{\lambda_t \cdot \bar{R}_w}{N} \right]^{N-i} \qquad (4.5)$$

$$P_B(i) = \binom{N}{i} \left[\frac{\lambda_t \bar{R}_v}{N} \right]^i \left[1 - \frac{\lambda_t \cdot \bar{R}_v}{N} \right]^{N-i} \qquad (4.6)$$

$$P_C(i) = \binom{N}{i} \left[\frac{\lambda_t \bar{R}_{df}}{N} \right]^i \left[1 - \frac{\lambda_t \cdot \bar{R}_{df}}{N} \right]^{N-i} \qquad (4.7)$$

$$P_D(i) = \binom{N}{i} \left[\frac{\lambda_t \bar{R}_{di}}{N} \right]^i \left[1 - \frac{\lambda_t \cdot \bar{R}_{di}}{N} \right]^{N-i} \qquad (4.8)$$

where

$$\bar{R}_w = \frac{M_w \lambda_w}{(M_t \lambda_t + M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{4.9}$$

$$\bar{R}_v = \frac{M_v \lambda_v}{(M_t \lambda_t + M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{4.10}$$

$$\bar{R}_{df} = \frac{M_f \lambda_f}{(M_t \lambda_t + M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{4.11}$$

$$\bar{R}_{dt} = \frac{M_t \lambda_t}{(M_t \lambda_t + M_f \lambda_f + M_v \lambda_v + M_w \lambda_w)} \tag{4.12}$$

are the overall traffic ratios. In other words $\bar{R}_w$ is the percentage of the traffic video users, $\bar{R}_t$ is the percentage of traffic due to voice users, etc.

The division by the number of outputs N, reflects the aggregate traffic to only one output. Also we note that the effects of the empty and collided packets in the different slots due to the uplink interactive data contention have been ignored. This is due to the fact that empty and collided packets are eliminated only at at the switch inputs. This may turn out to be a great advantage since it leaves some room to manoeuver for those contention packets and it keeps the order of video and voice packets within the frame.However our analysis provided a worst case in the sense of ignoring these advantages.

4.2. The Three Phase HOL Resolution Algorithm With Priority

4.2.1. Input Queue Analysis

The first phase of this algorithm (i.e. the prioritized knockout switch shown in Fig.4.1) has been already presented. The second phase assumes that blocked video and

voice packets are cleared, i.e., lost. This seems to contradict our earlier priority rules, however, we tried to guarantee data users some extra service (for sister packets in blocked inputs).

In the second phase the switch processor tries to resolve the HOL problem by locating blocked file users, selecting in sequential order a sister packet from that blocked input buffer and if its destination corresponds to one output with some empty space (E(I) ≠0), this sister packet is forwarded and the applicable destination E(I) is decremented by one. If unfortunately the sister packet of the HOL had a destination where E(I) =0, blocking is declared for the second time for this input buffer. Once all sister packets of all blocked HOL file input buffers are treated this way, we move to the third phase which is very similar to the second except that it treats blocked HOL packets of interactive type (the lowest priority). In other words, if the HOL is blocked and if there is no file data sister packet we look for interactive data sister packet whose destination is an output buffer that has some empty space; Once found, this interactive data sister packet is forwarded and the corresponding E(I) is decremented by 1.

For subsequent analysis, we need to define C_f to be the number of input buffers (out of N) whose head of line are file data packets and that are blocked from reaching a certain output following the execution of the algorithm of Fig. 4.1 The distribution of the number of input blocked to a certain output is given by:

$$P_{C_f}(i) = \sum_{a=0}^{N-i} \sum_{b=0}^{N-i-a} \sum_{c=0}^{N-i-a-b} P_A(a) \cdot P_B(b) \cdot P_C(c), \qquad i=0,\ldots,N-L \qquad (4.13)$$

where

$i = c - [L - a - b]^+$ and $[xx]^+ = 0$ if xx is negative.

Similarly, we define C_i to be the number of input buffers whose HOL are interactive data packets and that are blocked from reaching a certain output following the execution

of the algorithm of Fig.4.1 The distribution of the number of input interactive buffers blocked from reaching their destination is given by:

$$P_{C_i}(i) = \sum_{a=0}^{N-i} \sum_{b=0}^{N-i-a} \sum_{c=0}^{N-i-a-b} \sum_{d=0}^{N-i-a-b-c} P_A(a)\, P_B(b)\, P_C(c)\, P_D(d) \qquad (4.14)$$

$$i=0,\ldots,N-L$$

where

$i = d-[L-a-b-c]^+$, and $[xx]^+ = 0$ if xx is negative

All the above equations are derived under suitable independence assumptions and uniform load patterns at the input to the switch for all four kinds of services

The distribution of the empty places at one output port available for sister packets of the blocked input ports whose HOL is a file data slot is given by:

$$P_{e_f}(i) = \sum_{a=0}^{L-i} \sum_{b=0}^{L-i-a} \sum_{c=0}^{L-i-a-b} \sum_{d=0}^{L-i-a-b-c} P_A(a)\, P_B(b)\, P_C(c)\, P_D(d) \qquad (4.15)$$

$$i=0,1,\ldots,L$$

Similarly, the distribution of the empty places at one output port available for sister packets of the blocked input ports whose HOL is an interactive data slot is given by.

$$P_{e_i}(i) = \sum_{j=0}^{L} \sum_{k=0}^{N-L} \binom{k}{j-i} (1/N)^j (1-1/N)^{k-j} P_{C_f}(k)\, P_{e_f}(j) \qquad (4.16)$$

Now, if one looks at all file packets in input buffers, some would have transmitted successfully during the first knockout phase (Fig.3.1), some would have HOL blocking, some of those blocked would have sister packets successfully reaching the output destination during the three phase algorithm. So it is possible to pick at random one of the

input buffers and treat it as a queue with random service probability depending on the situation. Let $\bar{\mu}_f$ be the average service probability of file packet enumerated over all service scenarios. Now we have e_f empty slots in a certain output port picked at random, C_f file packets from all inputs blocked. If 2 or more users go for the same empty slot, we just pick one at random. There are i blocked file users trying to access e_f empty slots, i ranging from 2 to C_f. Enumerating over all possibilities, the overall average service for a typical file packet at one of the input ports conditioned over the number of contending packets C_f is given by:

$$\mu_f \Big|_{C_f} = 1 \text{ with probability } \beta_1 = (1-\bar{P}_{bdf})$$

$$= 1 \text{ with probability } \beta_2 = (\bar{P}_{bdf} \cdot \alpha_1 \Big|_{C_f}) \cdot \bar{R}_{df}$$

$$= 1 \text{ with probability } \beta_3 = (\bar{P}_{bdf} \cdot \alpha_2 \Big|_{C_f}) \cdot \bar{R}_{df}$$

where the conditional probabilities α_1 and α_2 are given by:

$$\alpha_1 \Big|_{C_f} = \begin{bmatrix} C_f \\ 1 \end{bmatrix} \left[1/N\right] \left[1-1/N\right]^{C_f-1} \cdot \left[1-(P_{e_f}(0))\right] \tag{4.17}$$

$$\alpha_2 \Big|_{C_f} = \sum_{i=2}^{C_f} \begin{bmatrix} C_f \\ i \end{bmatrix} \left[1/N\right]^i \left[1-1/N\right]^{C_f-1} \cdot \left[\sum_{j=i}^{L} P_{e_f}(j) + \sum_{j=1}^{i-1} (P_{e_f}(j)) \cdot j/i \right] \tag{4.18}$$

β_1 is the service probability for file HOL packet, β_2 and β_3 are the service probabilities for a file sister packet, and $\bar{R}_{df}$ is the probability that a sister packet of a blocked HOL packet happens to be a file data packet. Note that $\bar{R}_{df}$ multiplies only the services of the sister packet β_2 and β_3 but not β_1 which is for the HOL packet. The conditioning over C_f

is removed by averaging over the probability distribution of C_f thus obtaining

$$\bar{\mu}_f = \sum_{i=0}^{N-L} [(1-\bar{P}_{bdf}) + \bar{P}_{bdf} (\alpha_1 \Big|_{C_f} + \alpha_2 \Big|_{C_f}) \bar{R}_{df}] P_{C_f}(i) \qquad (4.19)$$

where $P_{C_f}(i)$ is given in eqn (68).

Now we follow the same procedure for interactive sister packets and arrive at at the follwoing conditional service probabilities.

$$\mu_i \Big|_C = 1 \text{ with probability } \overline{\beta_1} = (1-\bar{P}_{bdi})$$

$$= 1 \text{ with probability } \overline{\beta_2} = (\bar{P}_{bdi} \cdot \gamma_1 \Big|_{C_i}) \cdot \bar{R}_{di}$$

$$= 1 \text{ with probability } \overline{\beta_3} = (\bar{P}_{bdi} \cdot \gamma_2 \Big|_{C_i}) \bar{R}_{di}$$

where the conditional probabilities γ_1 and γ_2 are given by:

$$\gamma_1 \Big|_{C_i} = \binom{C_i}{1} \left[1/N\right] \left[1-1/N\right]^{C_i-1} \cdot \left[1-(P_{e_i}(0))\right] \qquad (4.20)$$

$$\gamma_2 \Big|_{C_i} = \sum_{i=2}^{C_i} \binom{C_i}{i} \left[1/N\right]^i \left[1-1/N\right]^{C_i-i} \cdot \left[\sum_{j=i}^{L} P_{e_i}(j) + \sum_{j=1}^{i-1} (P_{e_i}(j))\, j/i\right] \qquad (4.21)$$

and finally averaging over the distribution of C_i,we obtain,

$$\bar{\mu}_i = \sum_{i=0}^{N-L} [(1-\overline{P_{bdi}}) + \overline{P_{bdi}} (\gamma_1 \Big|_{C_i} + \gamma_2 \Big|_{C_i}) \overline{R_{di}}] \cdot P_{C_i}(i) \qquad (4.22)$$

where $P_{C_i}(i)$ is given by eqn (69).

Since video and voice slots are either serviced in case of no blocking or dropped in case of blocking the corresponding service probability for video and voice users is one for any given case i.e.

$$\bar{\mu}_v = 1 \tag{4.23}$$

for voice users and,

$$\bar{\mu}_w = 1 \tag{4.24}$$

for video users.

The average service to a packet in the input buffer (irrelative of its kind) is therfore given by:

$$\mu_{in} = \bar{R}_{di}\,\bar{\mu}_i + \bar{R}_{df}\,\bar{\mu}_f + \bar{R}_v\,\bar{\mu}_v + \bar{R}_w\,\bar{\mu}_w \tag{4.25}$$

μ_{in} represents the discrete service probability for an input buffer with discrete input probability per slot λ_t.

A packet, after reaching the HOL position, stays there until it is transferred to an output port. The duration of this stay is determined by the delay it experiences in the virtual queue it joins. If the HOL position is viewed as a service station and the time a packet stays at this position as its service time, input port can be analyzed by applying results for an M/G/1 queue [35]. For an input queue with infinite buffer size, the distribution of the input queue length denoted as N_i has a mean denoted as $\bar{N}_i$ given by the following equation:

$$\bar{N}_i = \lambda_i\,\bar{W}_i \tag{4.26}$$

where λ_i is the average arrival rate from the uplink beam given by eq.(3.54) and $\bar{W}_i$ is the mean input delay given by:

$$\bar{W}_i = \bar{W}_s + \frac{\lambda_i\,\bar{W}_s^2}{2(1-\lambda_i\,\bar{W}_s)} \tag{4.27}$$

As can be seen, $\overline{W}_t$ depends on the first and second moment of the delay function which is the queuing time and the service time. The first moment denoted by $\overline{W}_s$ is the mean delay in the system representing the delay experienced by a packet from the time it first arrives at the HOL of an input queue until it is successfully delivered to an oputput port For a " first come first serve policy (FCFS) ", $\overline{W}_s$ is given by :

$$\overline{W}_s = \frac{\mu_{in}(\lambda_t/\mu_{in})^L}{(L-1)!\,(L\mu_{in}-\lambda_t)^2} p_0 + 1/\mu_{in} \tag{4.28}$$

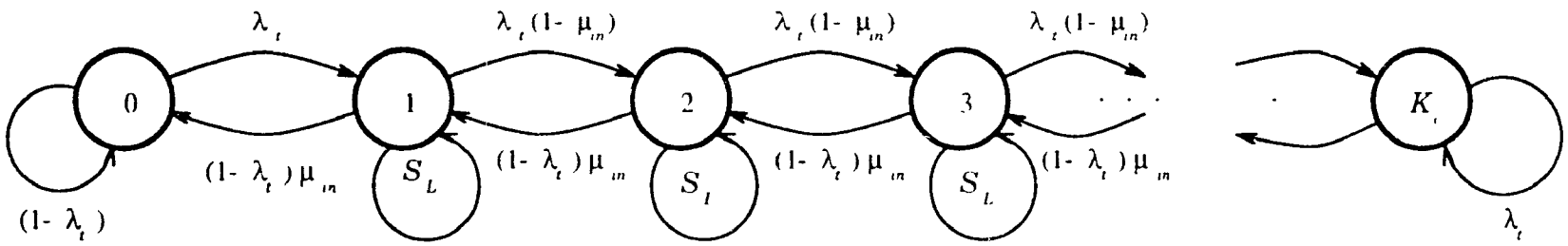
where p_0 is the steady state probability of zero packets in the system given by.

$$p_0 = \left[\sum_{n=0}^{L-1} \frac{(\lambda_t/\mu_{in})^n}{(n!)} + \frac{L\,(\lambda_t/\mu_{in})^L}{L!\,(L-\lambda_t/\mu_{in})} \right]^{-1} \tag{4.29}$$

and $\overline{W}_s^2$ is the second moment of the delay given by:

$$\overline{W}_s^2 = \frac{2(\lambda_t/\mu_{in})^L}{(L-1)!\,(L\mu_{in}-\lambda_t)^2} \cdot \left[\frac{\mu_{in}}{L\mu_{in}-\lambda_t} + 1 \right] p_0 + \frac{2}{\mu_{in}^2} \tag{4.30}$$

That was for an assumed infinite buffer. For a finite input buffer system of size K_t, there will be a blocking probability or a probability of buffer overflow. For such kinds of systems, the corresponding Markov chain is shown in Fig. 4.2 and is solved by solving the corresponding equilibrium equations [36]. Once the queue length distribution $P_t(t)$ is found, it will be easy to find the blocking probability. For the case of the input buffer system, a closed form expression of the state probability distribution is given by:



$$S_L = 1 - (1 - \lambda_t)\mu_{in} - \lambda_t(1 - \mu_{in})$$

Figure 4.2: Input Buffer System Chain

$$P_i(k) = \frac{1-\lambda_i/\mu_{in}}{1-(\lambda_i/\mu_{in})^{k+1}} \left(\frac{\lambda_i}{\mu_{in}}\right)^k \tag{4.31}$$

for $0 \le k \le K_i$

where K_i is the input port buffer size. The buffer overflow probability is given by:

$$P_{if} = P_i(K_i) = \frac{1-\lambda_i/\mu_{in}}{1-(\lambda_i/\mu_{in})^{K_i+1}} \cdot \left(\frac{\lambda_i}{\mu_{in}}\right)^{K_i} \tag{4.32}$$

The division of λ_i by μ_{in} in some of the previous equations gives the utilization per slot that takes all services and HOL algorithm into consideration. This means that all users will have on average the same input delay regardless of the user type.

4.2.2. Output Queue Analysis

Output queues are modeled as M/M/1 queues. Assuming uniform traffic load to all output ports of the switch,considering λ_i as the probability of having a packet to serve and μ_{out} as the probability of serving it (which all HOL services into account),the total traffic for file packets reaching a certain output is given by:

$$\bar{\lambda}_{df} = \lambda_i \cdot \bar{R}_{df} \cdot (1 - P_{if}) \tag{4.33}$$

similarly for interactive users

$$\bar{\lambda}_{di} = \lambda_i \cdot \bar{R}_{di} \cdot (1 - P_{if}) \tag{4.34}$$

for voice packets

$$\bar{\lambda}_{v} = \lambda_i \cdot \bar{R}_{v} \cdot (1 - P_{if}) \tag{4.35}$$

and for video packets, we have

$$\bar{\lambda}_{u} = \lambda_{t} \cdot \bar{R}_{u} \cdot (1 - P_{if}) \tag{4.36}$$

Therefore the total traffic utilization per output per slot is given by:

$$\bar{\lambda}_{o} = \frac{\bar{\lambda}_{di} + \bar{\lambda}_{df} + \bar{\lambda}_{v} + \bar{\lambda}_{w}}{\mu_{in}}$$

$$= \frac{\lambda_{t} (1 - P_{if})}{\mu_{in}} \tag{4.37}$$

Let W_{o} denote the queuing time at the output queue and $\bar{W}_{o}$ denote the mean output queuing delay given by:

$$\bar{W}_{o} = \frac{\lambda_{o}/\mu_{in}}{1 - \lambda_{o}} \tag{4.38}$$

With infinite buffer, the output queue length, denoted by N_{o}, has mean given by :

$$\bar{N}_{o} = \frac{\lambda_{o}/\mu_{in}}{1 - \lambda_{o}/\mu_{in}} \tag{4.39}$$

For a finite size output buffer of size K_{o} the output blocking probability can be approximated using an M/M/1/K_{o} model [37]. For such models, the output buffer overflow probability is given by:

$$P_{of} = \frac{(1-\lambda_{o})\lambda_{o}^{K_{o}}}{1-\lambda_{o}^{K_{o}+1}} \tag{4.40}$$

4.2.3. Blocking Probability

Let the P_{B} denotes the blocking probability of the whole system; in other words P_{B} is the probability that a packet is blocked somewhere in the system. It is given by [38]:

$$P_B = 1 - (1 - P_{if})(1 - P_{of}) \tag{4.41}$$

Where P_{if} is the input buffer overflow probability given by eq. (4.32) and P_{of} is the output buffer overflow probability given by eq. (4.40). It is possible to notice that P_B is a function of K_i and K_o as well as λ_i and μ_{in}.

4.2.4. System Delay

The mean delay of the system is the delay experienced by a packet between its arrival at an input port until its delivery to an output port. There are 2 components in this delay: the first being the mean input delay $\overline{W}_i$, and the second being the mean output queuing delay $\overline{W}_o$. Therefore the total system delay denoted by $\overline{D}_s$ is given by:

$$\overline{D}_s = \overline{W}_i + \overline{W}_o \tag{4.42}$$

where $\overline{W}_i$ is the mean input buffer delay given in (4.27) and $\overline{W}_o$ is the mean output buffer delay given in (4.39).

4.2.5. Maximum Switch Throughput

The maximum switch throughput denoted by, λ_{max}, is the minimum of the maximum input throughput λ^i_{max} and the maximum output throughput λ^o_{max}. Each output queue is served by a dedicated output link with unit service rate; Therefore, λ^o_{max} is determined from:

$$\lambda^o_{max} = \lambda^o_{max}/\mu = 1 \tag{4.43}$$

Similarly, λ^i_{max} can be obtained from:

$$\lambda^i_{max} = \lambda^i_{max}\,\overline{W}_s = 1 \tag{4.44}$$

where $\overline{W}_s$ is given in (4.28) and represents the mean packet service time in an input queue. Since $\overline{W}_s \geq 1/\mu$, as can be seen from (4.28), it follows that:

$$\lambda_{max} = \min(\lambda'_{max}, \lambda^{o}_{max}) = \lambda'_{max} \tag{4.45}$$

The maximum throughput is obtained by assuming that $\mu = 1$. For different speed-up ratio L, the maximum throughput as a function of L is shown in Table 4.1. The maximum througput of 0.58 for L=1 agrees with the known result for a crossbar switch without speed-up. It is also seen that, as expected, $\lambda_{max} \rightarrow 1$ as $L \rightarrow \infty$. In fact λ_{max} grows quite rapidly.

Table 4.1:

Maximum Throughput vs. L

L	Maximum Throughput
1	0.58579
2	0.88454
3	0.97550
4	0.99559
5	0.99931
6	0.99991
7	0.99999
8	≈ 1

4.3. Results And Discussions.

In this section, the total througput coming from the uplink and represented by Eq.(3.54) goes through the input buffer, the switch, and the output buffer from where it will be multiplexed in time and sent to the corresponding downlink stations. The total uplink traffic or λ_t will be increased uniformly and for every value, different parameters are computed and discussed. It is assumed that the switch size N = 16.

To show the benefits of the algorithm, two situations are compared: the first represents the results that would have been found without the algorithm and the second represents the case with the algorithm applied. The only difference between the two cases is the service probability. Throughout section 4.2, the analysis represents the second case. The analysis of the first case which would be the case if no algorithm was applied is the same as the analysis of the second case except that Eq.(4.19) would be $\overline{\mu_f} = 1 - \bar{P}_{bdf}$ and Eq.(4.22) becomes $\overline{\mu_t} = 1 - \bar{P}_{bdt}$. Two situations are studied corresponding to two different traffic scenarios. In the first case, it is assumed that the largest portion of traffic is composed by file data and interactive data traffic; The corresponding traffic ratio are $\bar{R}_w = 0.01$, $\bar{R}_v = 0.04$, $\bar{R}_{df} = 0.6$, and $\bar{R}_{dt} = 0.35$. In the second case, it is assumed that video and voice packets form the largest part of the traffic compared to file data and interactive data traffic; In this case $\bar{R}_w = 0.45$, $\bar{R}_v = 0.45$, $\bar{R}_{df} = \bar{R}_{dt} = 0.05$. To distinguish between the first and the second case, all the figures that will be discussed in the following sections will have two parts a and b. Figures a will represent the first case where the traffic is mostly file data and interactive data while figures b will represent the second case where the biggest part of the traffic is composed of video and voice traffic. Now for each case mentioned above, there are some important parameters such that the speed-up traffic ratio L and the input and output buffer sizes K_i and K_o. Also for those different combinations of those parameters, different results are to be found and discussed.

4.3.1 Switch Blocking Probability.

As the total input traffic increases, the probabilitiy of having an arriving packet for the different kinds of services increases. Figures 4.3.a, 4.3.b, 4.4.a, 4.4.b, 4.5.a, and 4.5.b show the different blocking probabilities corresponding to Eqs.(4.1) to (4.4) respectively. Different situations are examined corresponding to different traffic scenarios. For all the different traffic scenarios, it is possible to see that video traffic suffers the less switch blocking than any other kind of traffic and that is because of the priority which is the first phase of the algorithm. It is also possible to see that for any case and when L=1, the blocking probabilities are considerably high for large values of the coming uplink traffic λ_t. As the speed-up ratio increases, the different blocking probabilities decrease continuously still with the video blocking having the lowest value followed by the voice traffic, the file data traffic, and the interactive data traffic showing again the order of priority at the HOL ports of the switch input; In other words, showing the priority of video traffic over all kinds of traffic, the priority of voice traffic over file data and interactive data traffic, and finally the priority of file data traffic over interactive data traffic which has the lowest priority.

Figures 4.2.a and 4.2.b show the different blocking probabilities when the speed-up ratio is L=1. The same blocking probabilities are shown in figures 4.3.a and 4.3.b for L=2 as well as in figures 4.4.a and 4.4.b for L=3. When compared, all the figures show the importance of the speed-up ratio in lowering the different blocking probabilities.

4.3.2 Mean Input Buffer Size.

As was shown in the analysis, the input buffers are modeled as M/G/1 queues. The mean input queue length is given by Eq.(4.26). For the different situations dicussed above, the mean input queue length is graphed for varying values of of input traffic λ_t and speed-up ratio L. Comparing the corresponding figures, it is possible to see that the

input buffer size is highly dependant on the speed-up ratio of the switch. When the speed-up ratio is one (L=1), the mean input buffer size increases rapidly and tends to reach infinity for a relatively small value of the input traffic λ_t = 0.4 without the HOL algorithm. With the HOL algorithm, the critical value of λ_t is larger and is equal to 0.44 which represents a gain of 10 percent as shown in Fig (4.6.a). Now as the speed up ratio increases (L=2) and under the same conditions, it is possible to see that (Fig (4.7.a)) the mean input buffer size reaches infinity at a value of λ_t equal to 0.75 without the HOL algorithm and 0.78 with the algorithm. A further increase in L (L=3) leads to an operating λ_t at 0.9 without the HOL algorithm and λ_t = 0.92 with the HOL algorithm as shown in Fig.(4.8.a). All theses results were found for a large proportion of file and interactive data traffic: Specifically, $\bar{R}_{df}$ = 0.6, $\bar{R}_{di}$ = 0.35 , $\bar{R}_{w}$ = 0.01, ans $\bar{R}_{v}$ = 0.04

For the situation where most of the traffic is video and voice traffic, the corresponding results change considerably which is normal since the algorithm gives benefit mainly to the file data and interactive data users. As shown in Figures 4.6.b (L=1), 4.7.b (L=2), and 4.8.b (L=3) the algorithm has no effect and the two curves representing its presence and its absence overlap.

4.3.3 Mean Output Buffer Size.

It is assumed that output queues are simple M/M/1 queues. The mean output queue length is given by Eq.(4.39). For increasing values of λ_t, the output queue length was calculated and graphed and that's for the different cases of traffic and speed-up ratios discussed above. Figures (4.9.a), (4.10.a), and (4.11.a) show the mean output queue length corresponding to the case where the majority of the traffic is file and interactive data and with the speed-up ratio L equal to 1, 2, and 3 respectively. Figures (4.9.b), (4.10.b), and (4.11.b) represent the same situations as above except that video and voice traffic form the main part of the whole traffic. Comparing the a figures with the b figures, it is possible to see that the algorithm benefits mainly data traffic over video and voice traffic given

the same speed-up ratio Another fact is that as the speed-up ratio increases the algorithm looses almost all the benefits to yield to the benefits of the knockout switch with high speed-up ratio

4.3.4 Blocking Probability.

The system blocking probability represents the fact that one packet is blocked somewhere in the system at the input buffer, the switch, or the output buffer. The system blocking probability is given by equation (4.41) where P_{if} is the input buffer blocking probability and P_{of} is the output buffer overflow probability. To show the merits of the algorithm, the case where the file data and interactive data traffic formed most of the incoming traffic was used: In other words $\bar{R}_w = 0.01$, $\bar{R}_v = 0.04$, $\bar{R}_{df} = 0.6$, and $\bar{R}_{di} = 0.35$. It was also supposed that the input buffers and the output buffers have the same size ($K_i = K_o$). Two values of λ_t were used: in the first case $\lambda_t = 0.5$ and in the second case $\lambda_t = 0.95$. For each case, the values of K_i and K_o were increased up to a given maximum and the corresponding blocking probabilities were calculated. Also for each case the value of the speed-up ratio L varied between the values 1, 2, and 3. The results of this section are shown in figures (4.12.a&b), (4.13.a&b), and (4.14.a&b). When $\lambda_t = 0.5$, it is seen that the blocking probability decreases sharply as the input and output buffer sizes increase. In fact an input and an output buffer size of 6 yield a very low blocking probability for the different speed-up ratio no matter whether the algorithm was used or not as shown in figures (4.12.a), (4.13.a), and (4.14.a). As λ_t increases to higher values which is the case of the second set of figures, the benefits of the algorithm as well as the speed-up ratio and the buffer sizes are clearly demonstrated. For $\lambda_t = 0.95$ and L=1, the algorithm shows a decrease in the blocking probability of the order of one third which is very remarkable. As the speed-up ratio L increases, the blocking probability decreases further with or without the algorithm and again the algorithm yielding to a

lower blocking probability in every case as shown in figures (4.12.b), (4.13.b), and (4.14.b).

4.3.5 System Delay.

The system delay is the sum of the mean input delay and the mean output delay as given in eq. (4.42). Again in this section it is assumed that the switch size is N=16, and the input traffic λ_i increases linearly. Figure (4.15.a) shows the case where most of the traffic is file data and interactive data traffic and the speed-up ratio is L=1. Figure (4.15.b) represents the case where most of the traffic is either video or voice traffic and again the speed-up ratio is L=1. Figures (4.16.a and b), (4.17.a and b) represent exactly the same case as figures (4.14.a and b) except that the speed-up ratio is L=2 and L=3 respectively. The curves show the same pattern as the ones related to the mean input buffer size discussed in section 4.3.2. Again it is seen that the smaller the speed-up ratio and the larger the proportion of data traffic, the greater the benefits of the algorithm are.

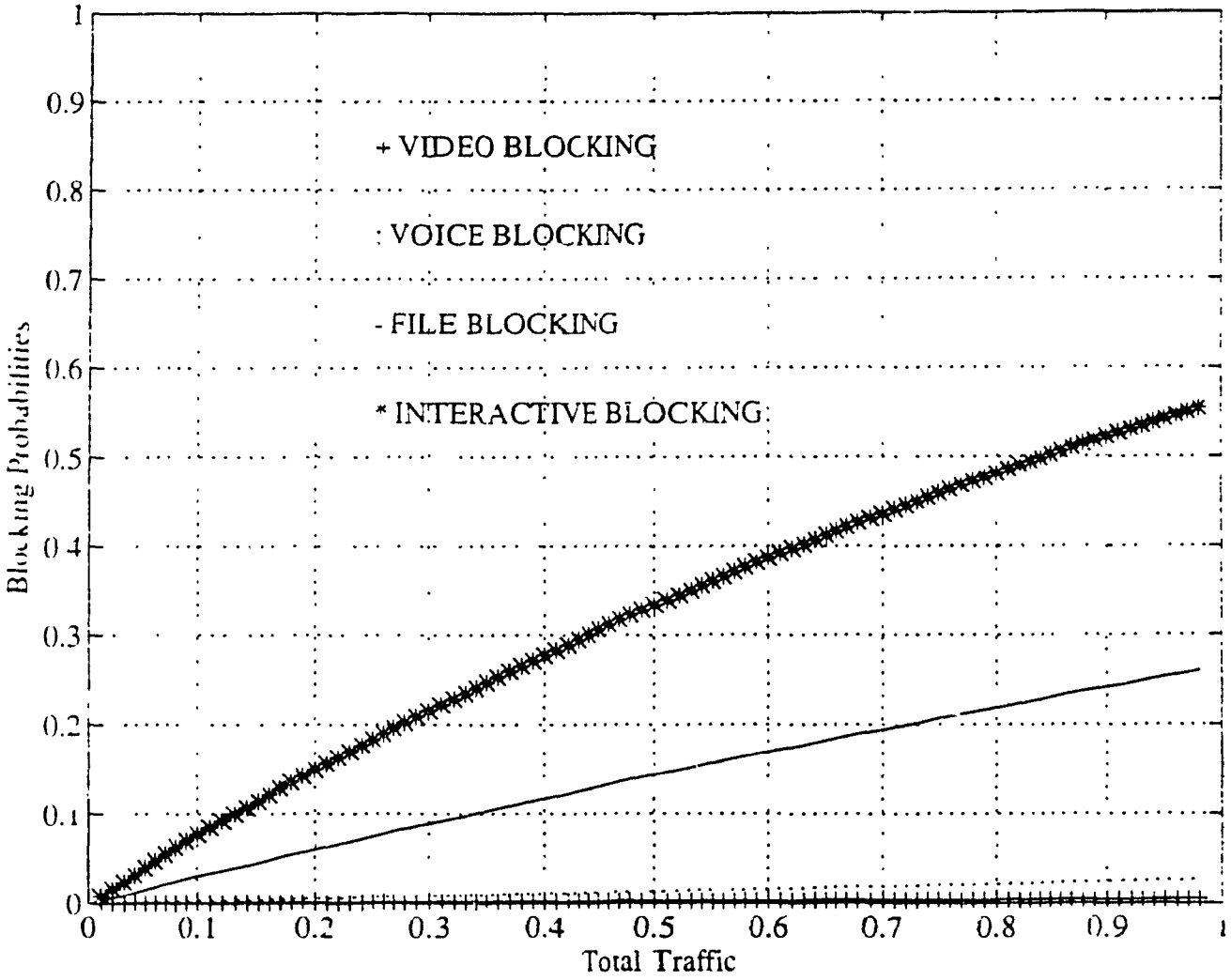


Fig. 4.3.a: Blocking Probabilities L=1, case a

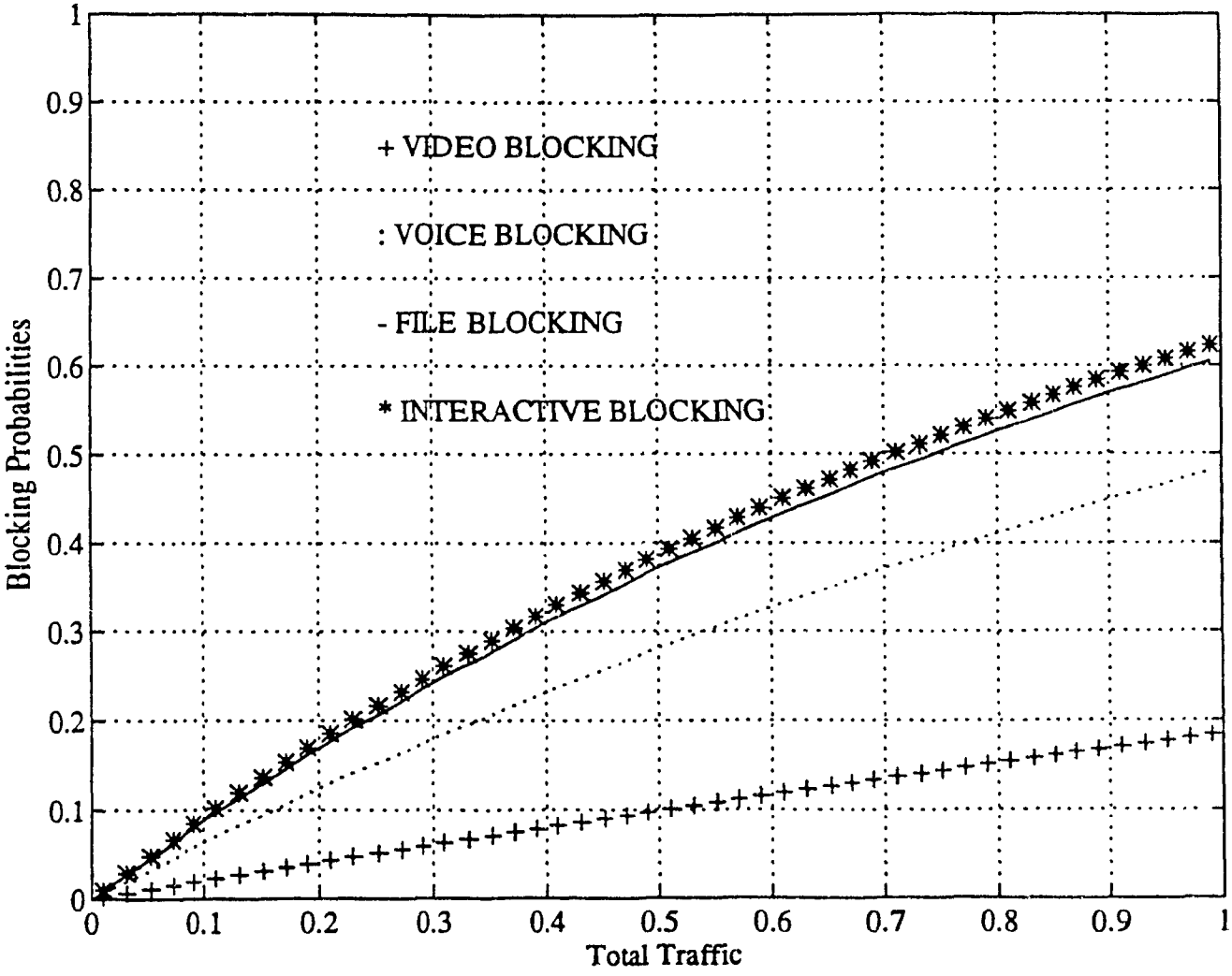


Fig. 4.3.b: Blocking Probabilities L=1, case b

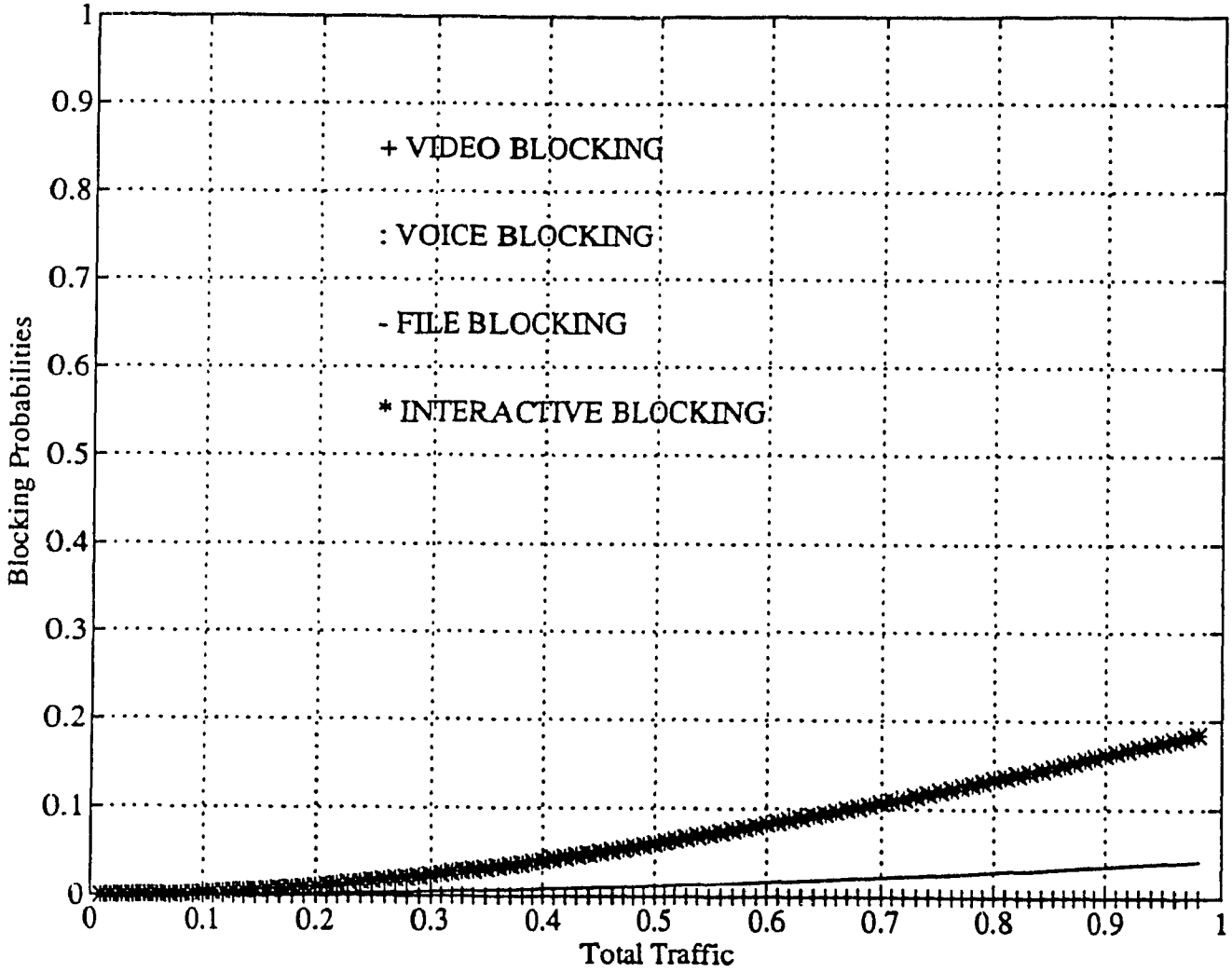


Fig. 4.4.a: Blocking Probabilities L=2, case a

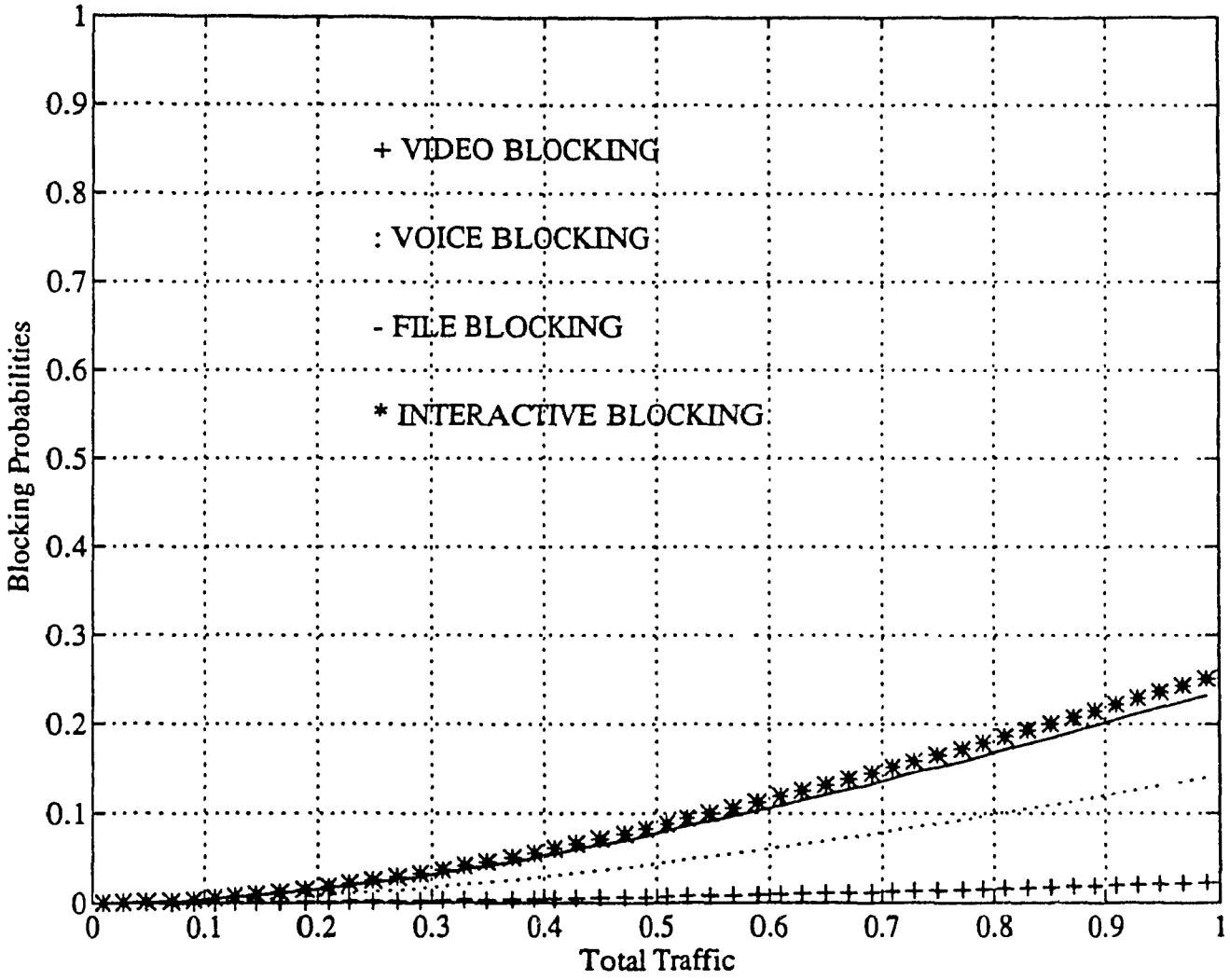


Fig. 4.4.b: Blocking Probabilities L=2, case b

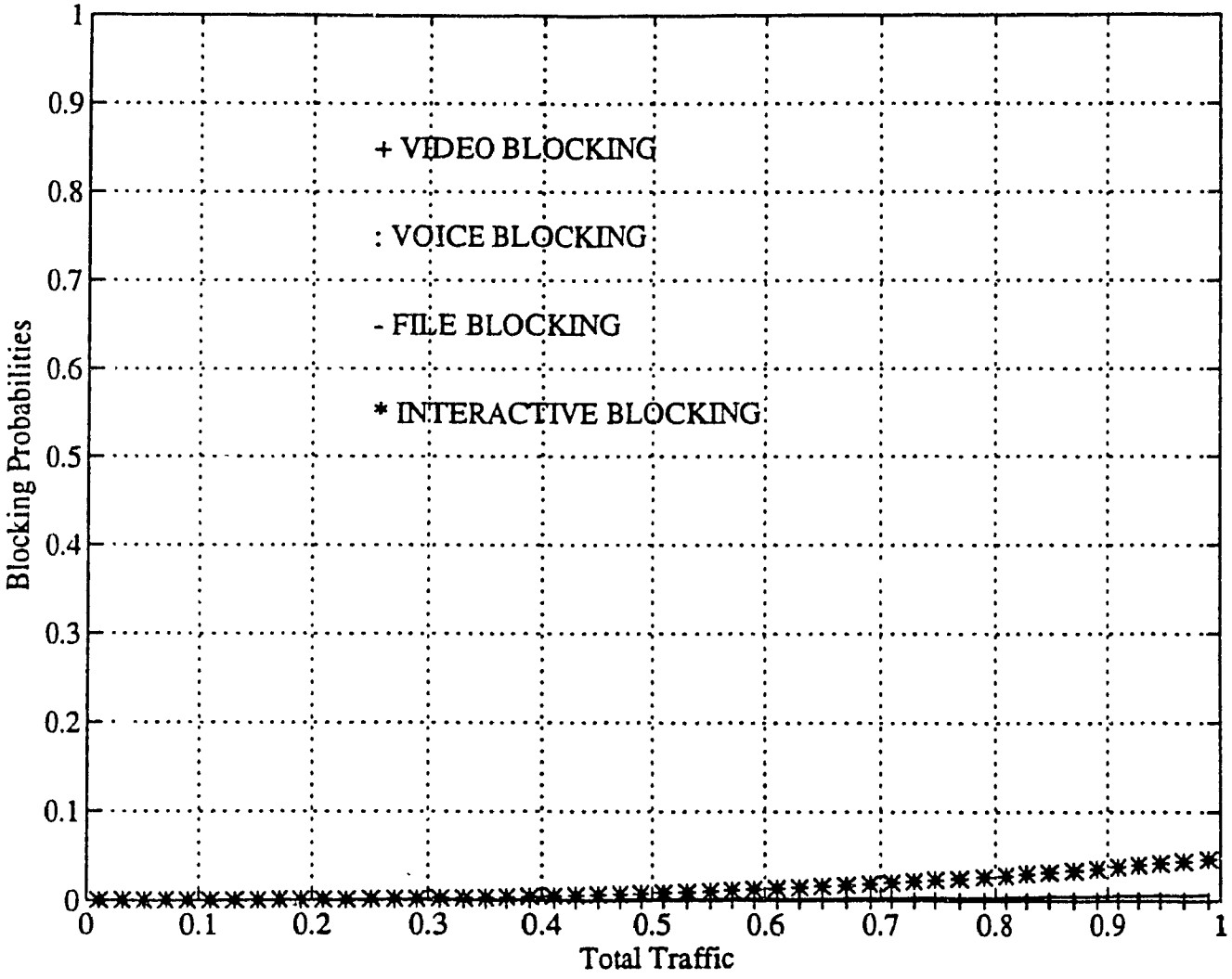


Fig. 4.5.a: Blocking Probabilities L=3, case a

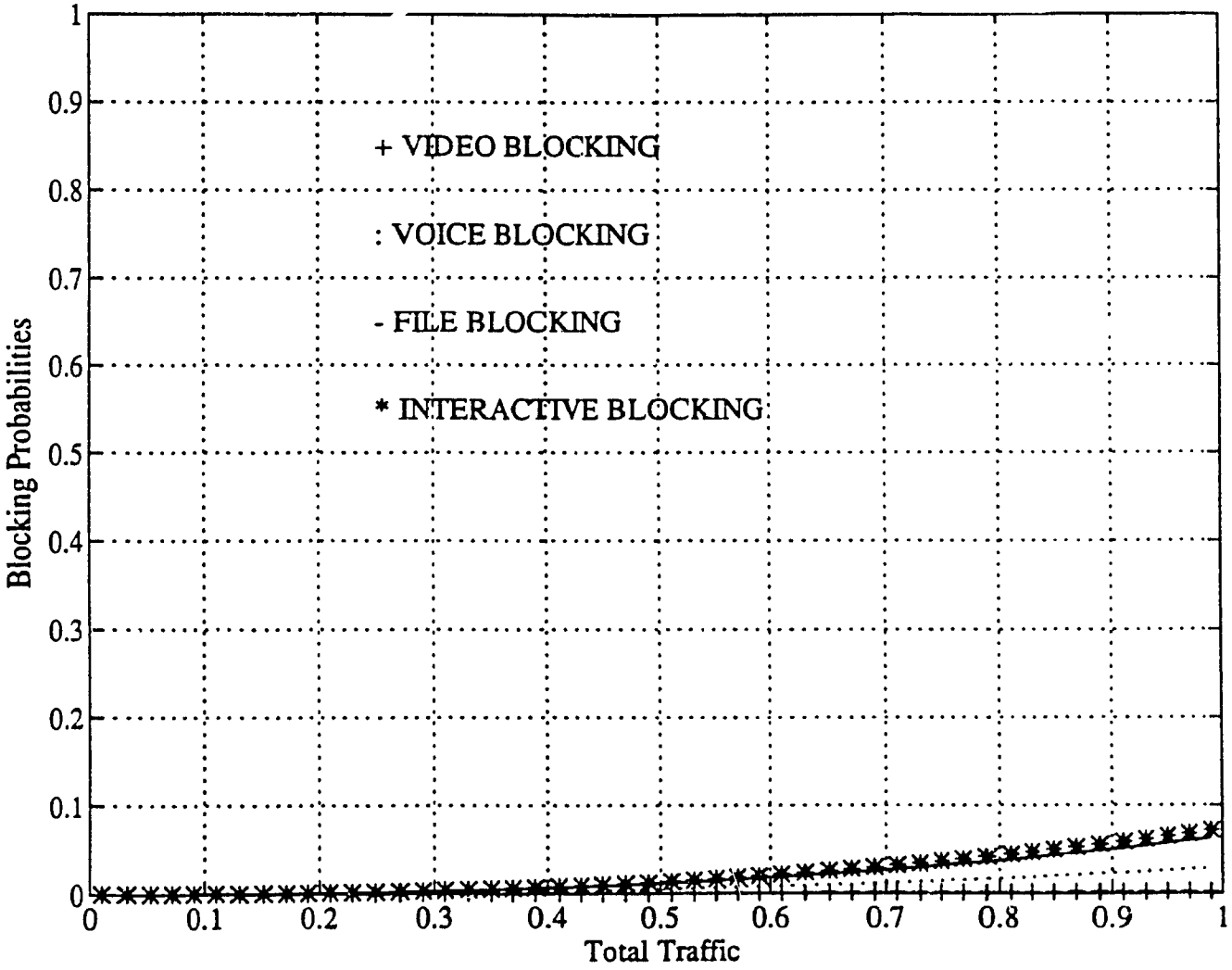


Fig. 4.5.b: Blocking Probabilities L=3, case b

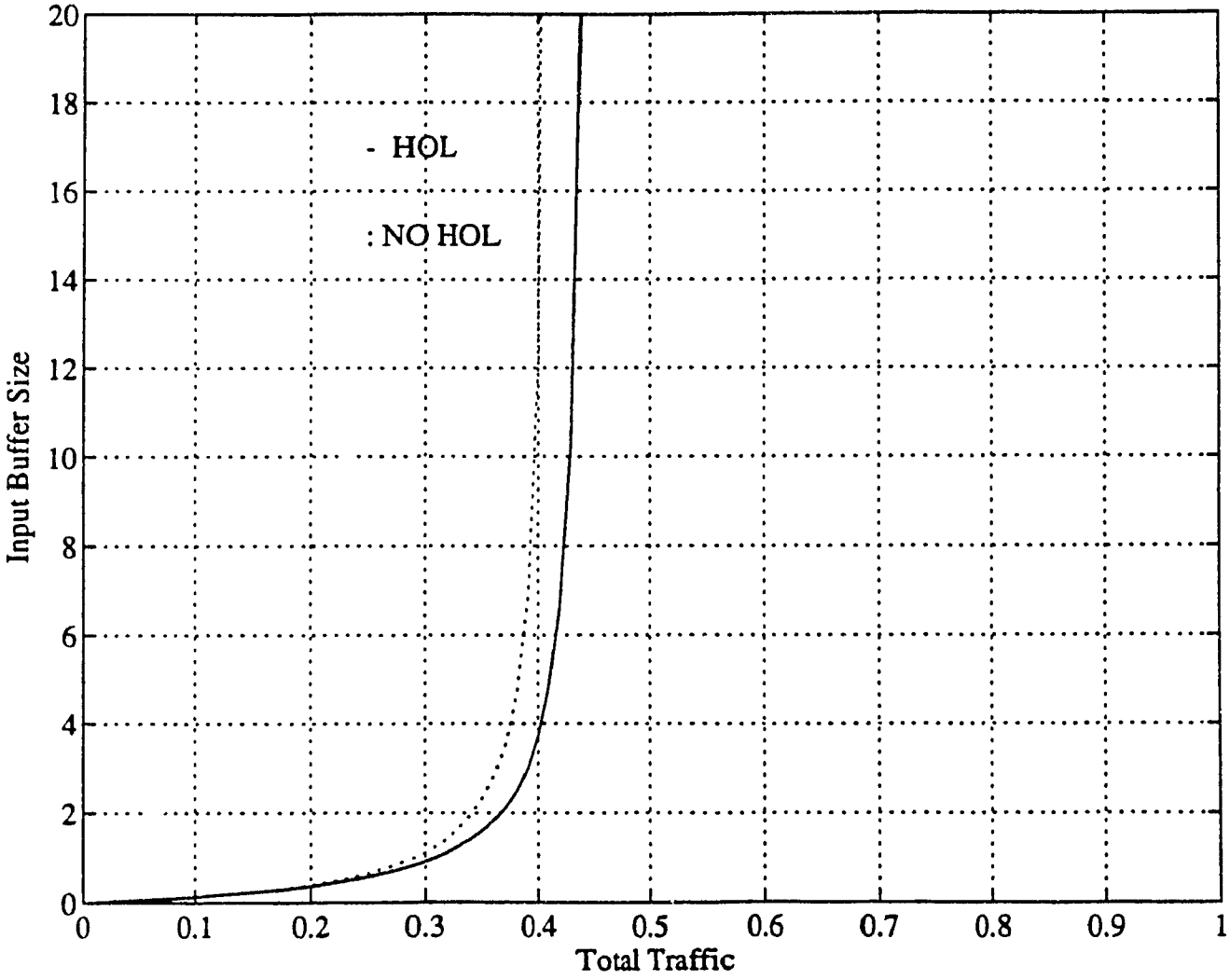


Fig. 4.6.a: Input Buffer Size. L=1, case a

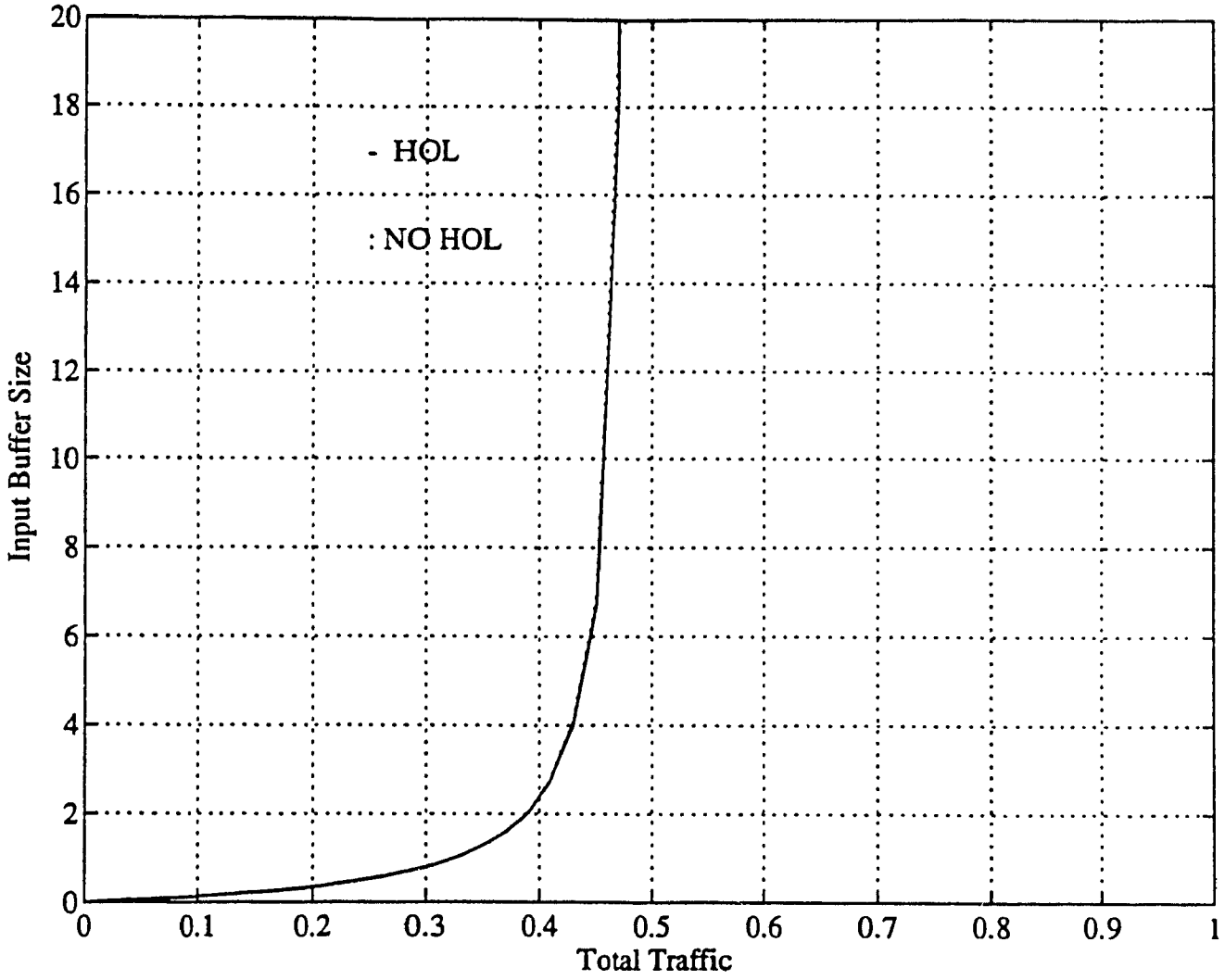


Fig. 4.6.b: Input Buffer Size. L=1, case b

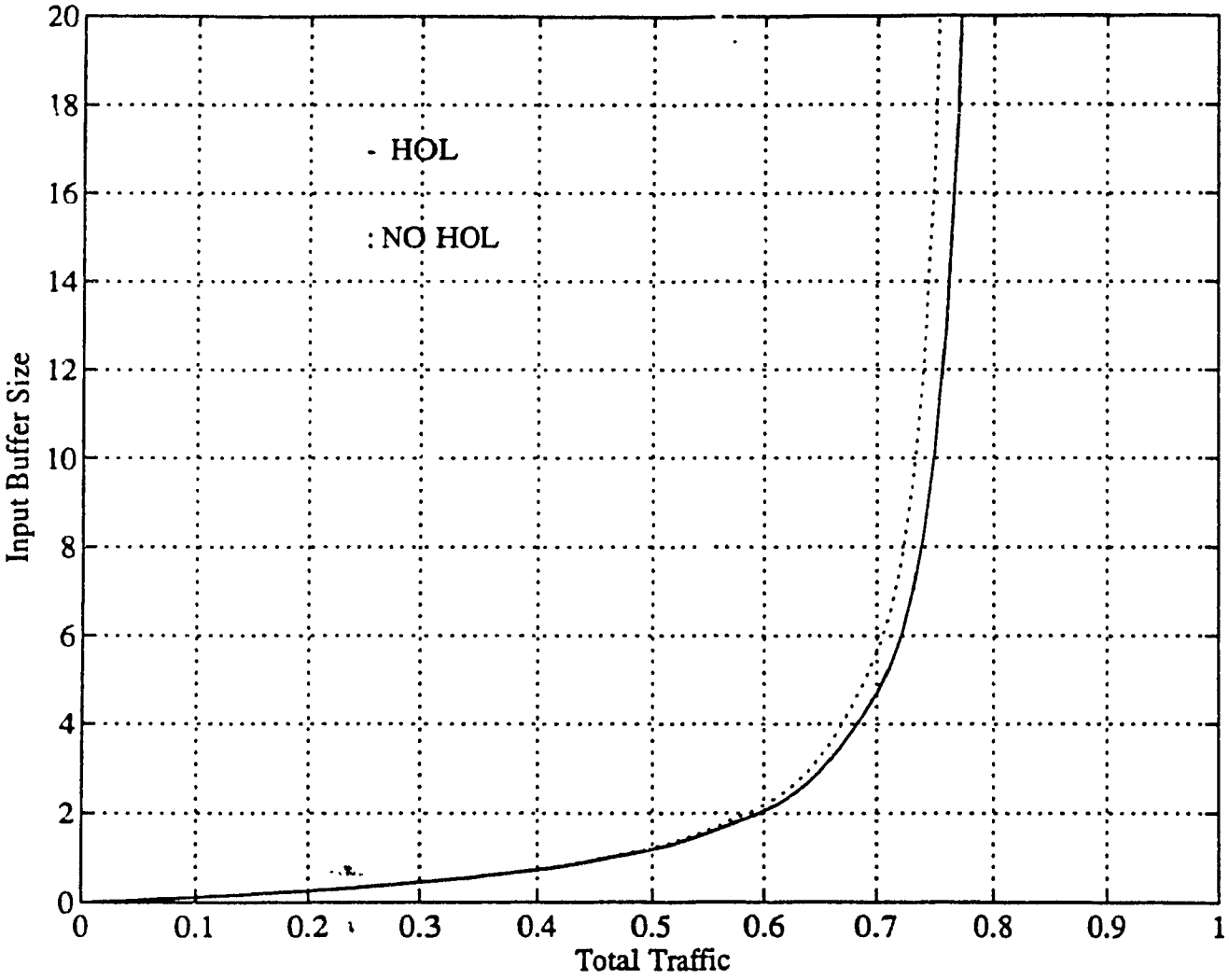


Fig. 4.7.a: Input Buffer Size. L=2, case a

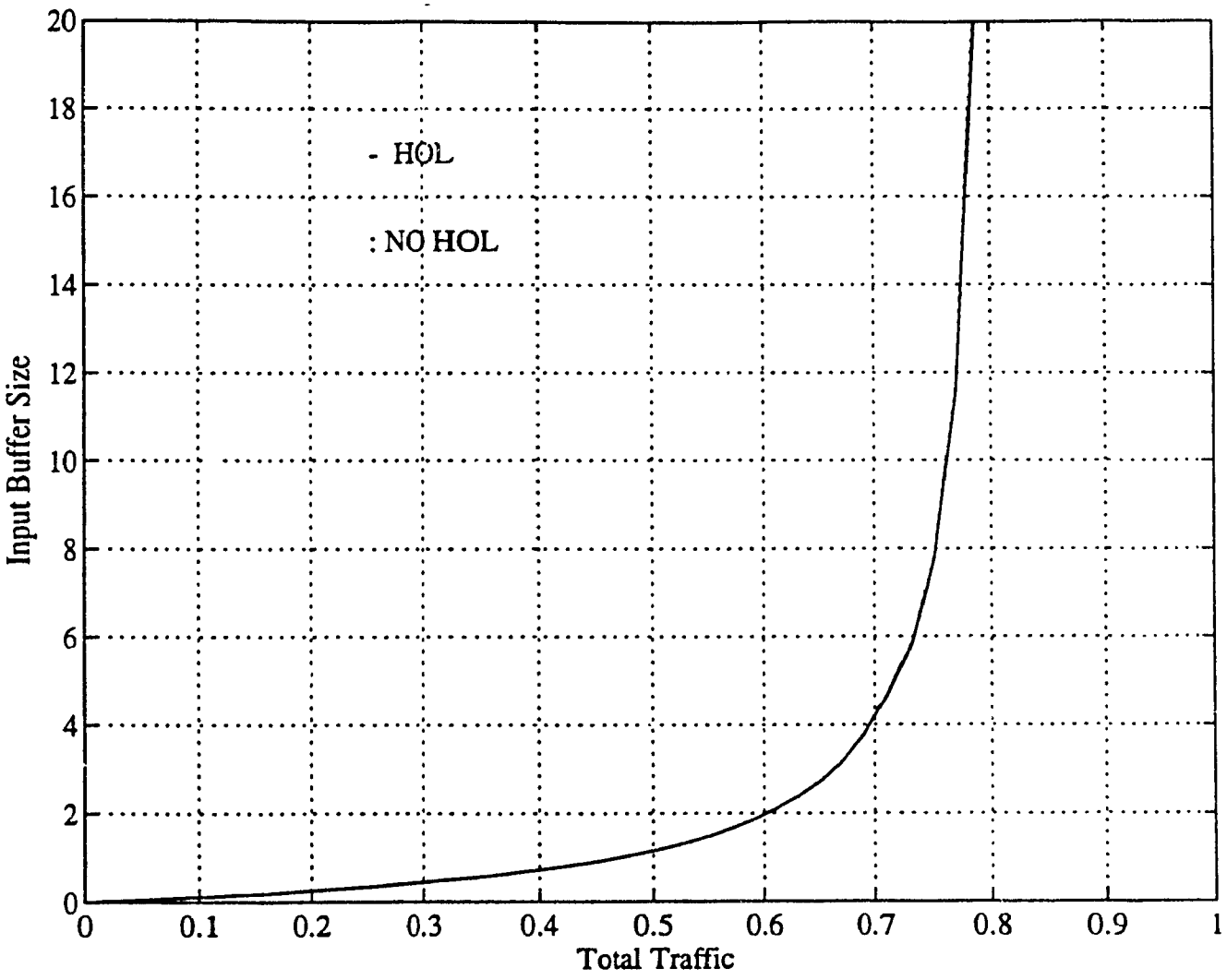


Fig. 4.7.b: Input Buffer Size. L=2, case b

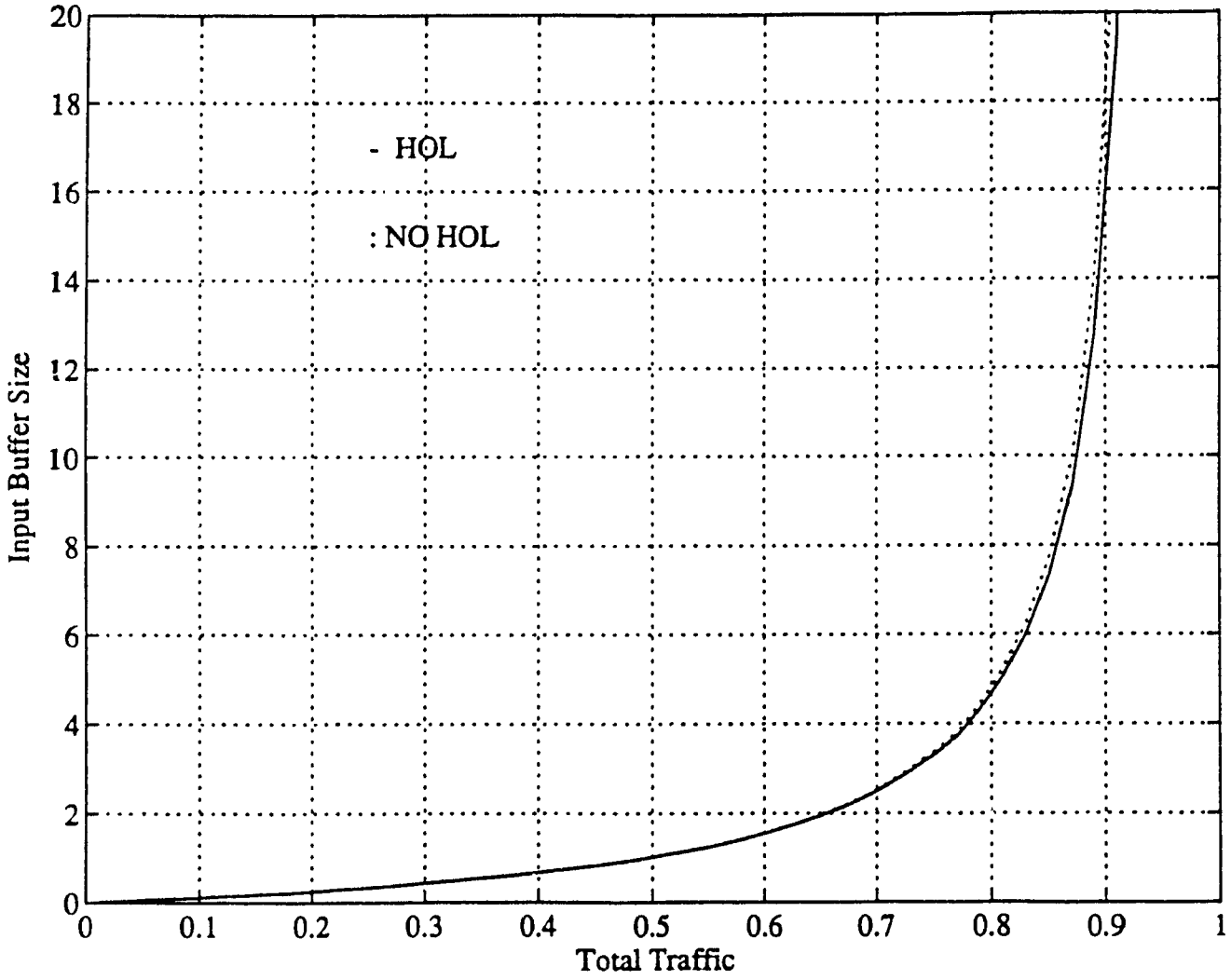


Fig. 4.8.a: Input Buffer Size. L=3, case a

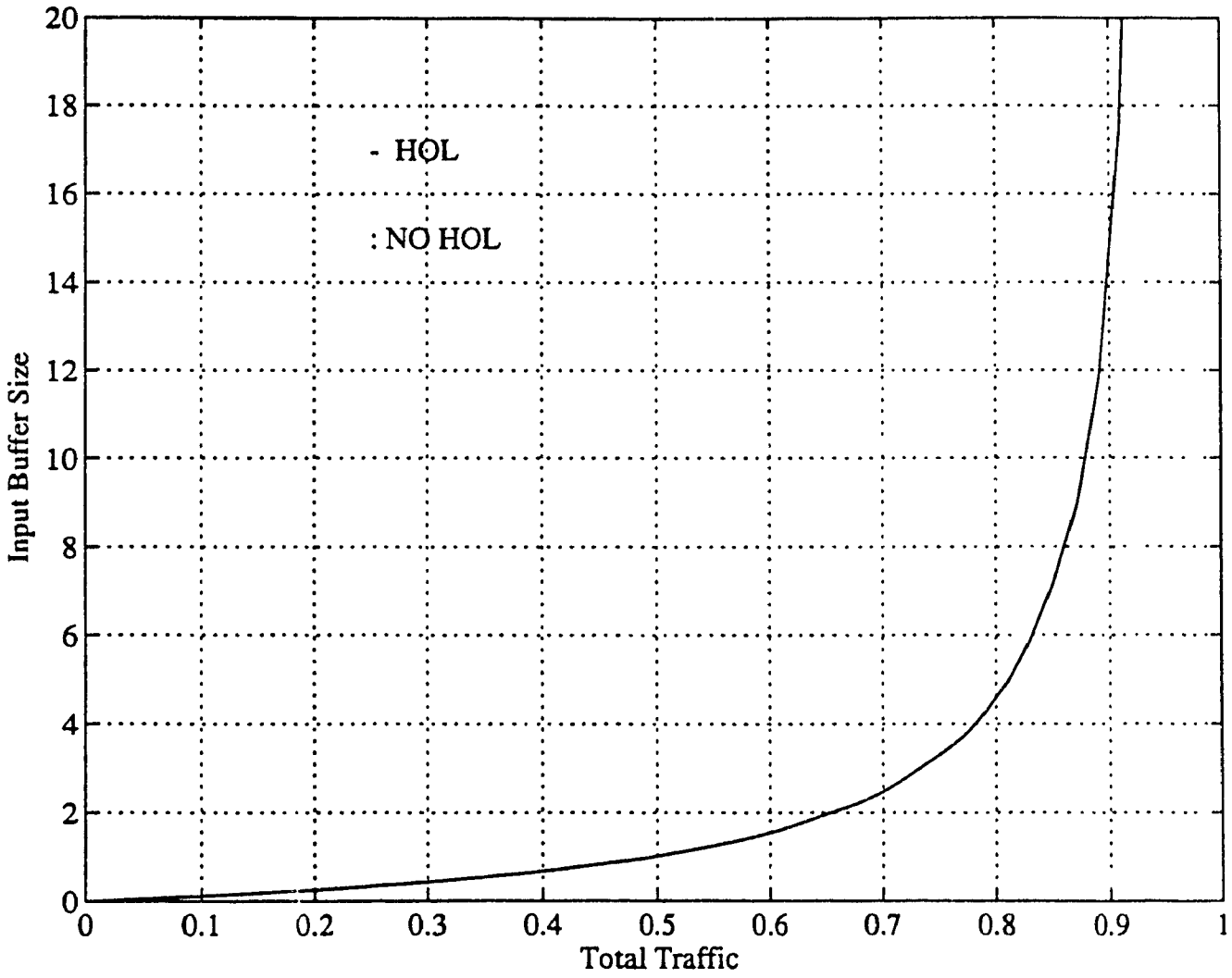


Fig. 4.8.b: Input Buffer Size. L=3, case b

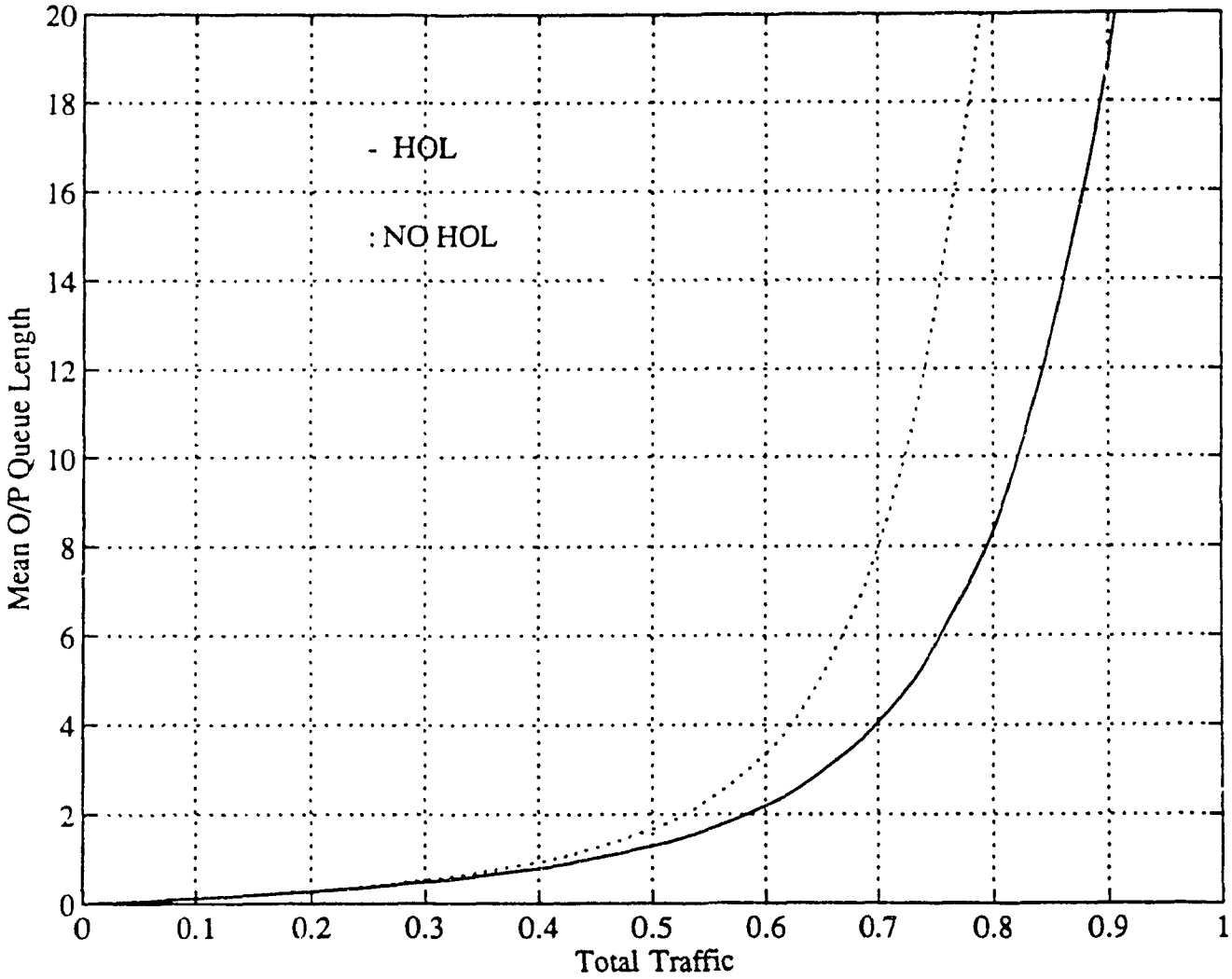


Fig. 4.9.a: Output Buffer Size. L=1, case a

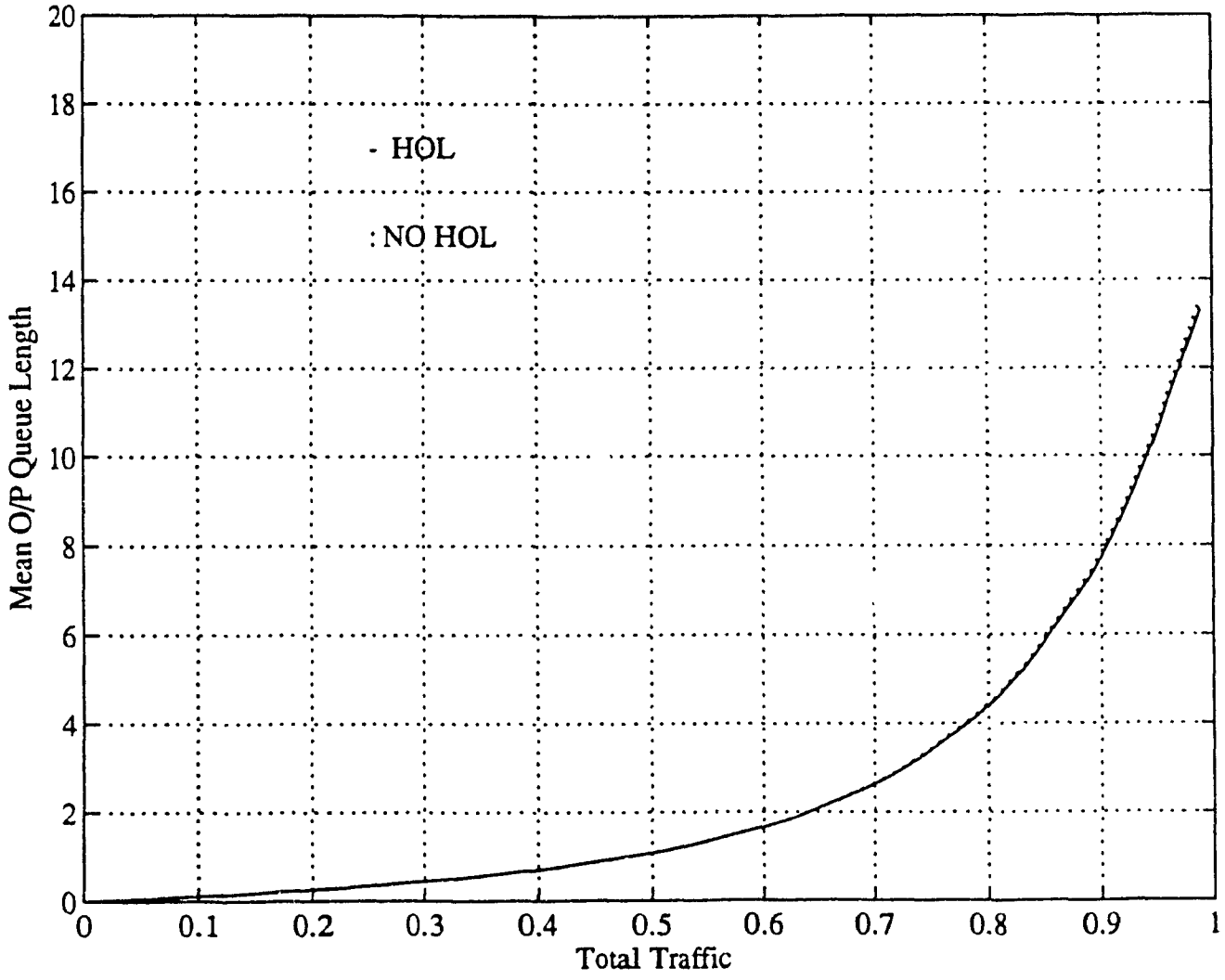


Fig. 4.9.b: Output Buffer Size. L=1, case b

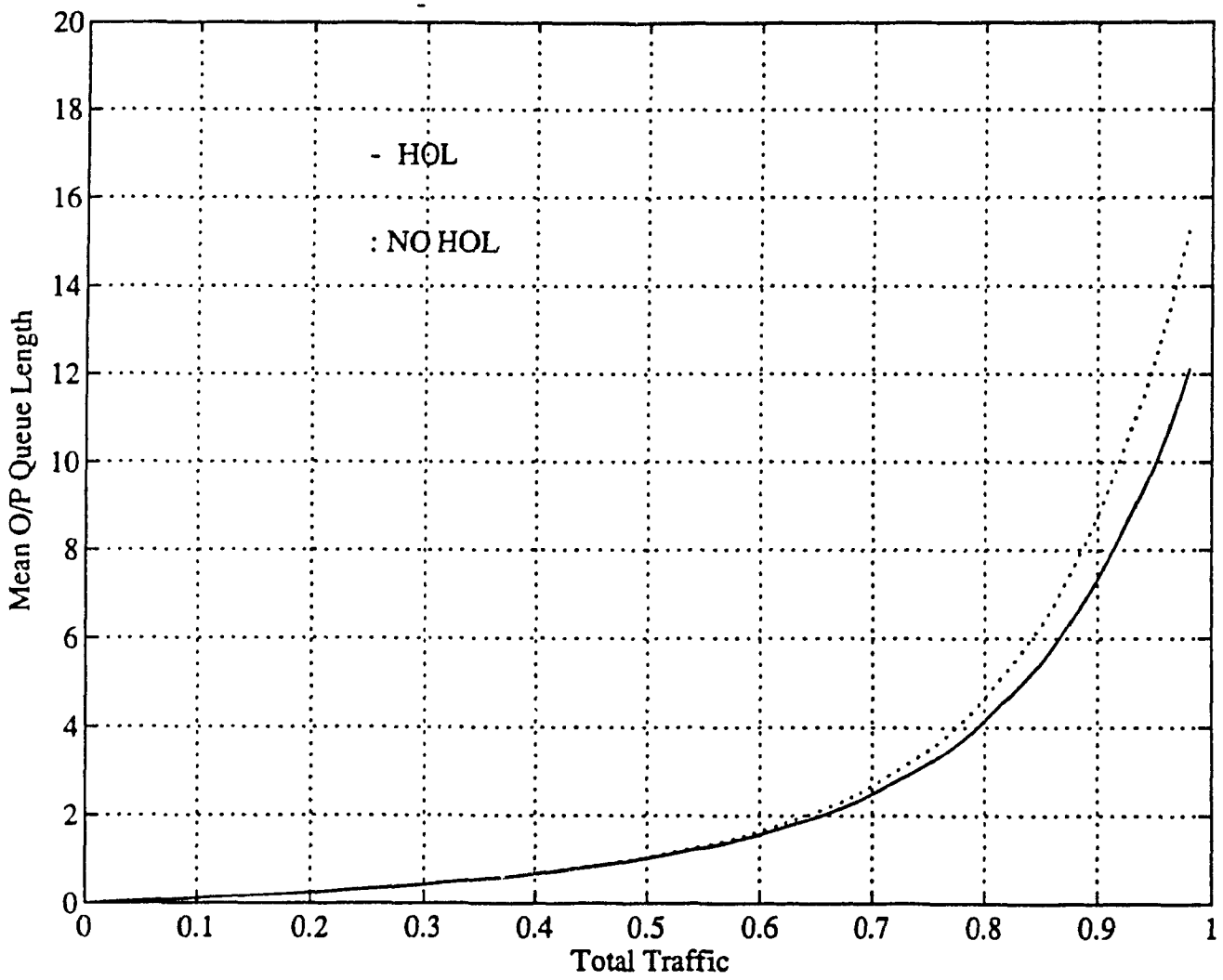


Fig. 4.10.a: Output Buffer Size. L=2, case a

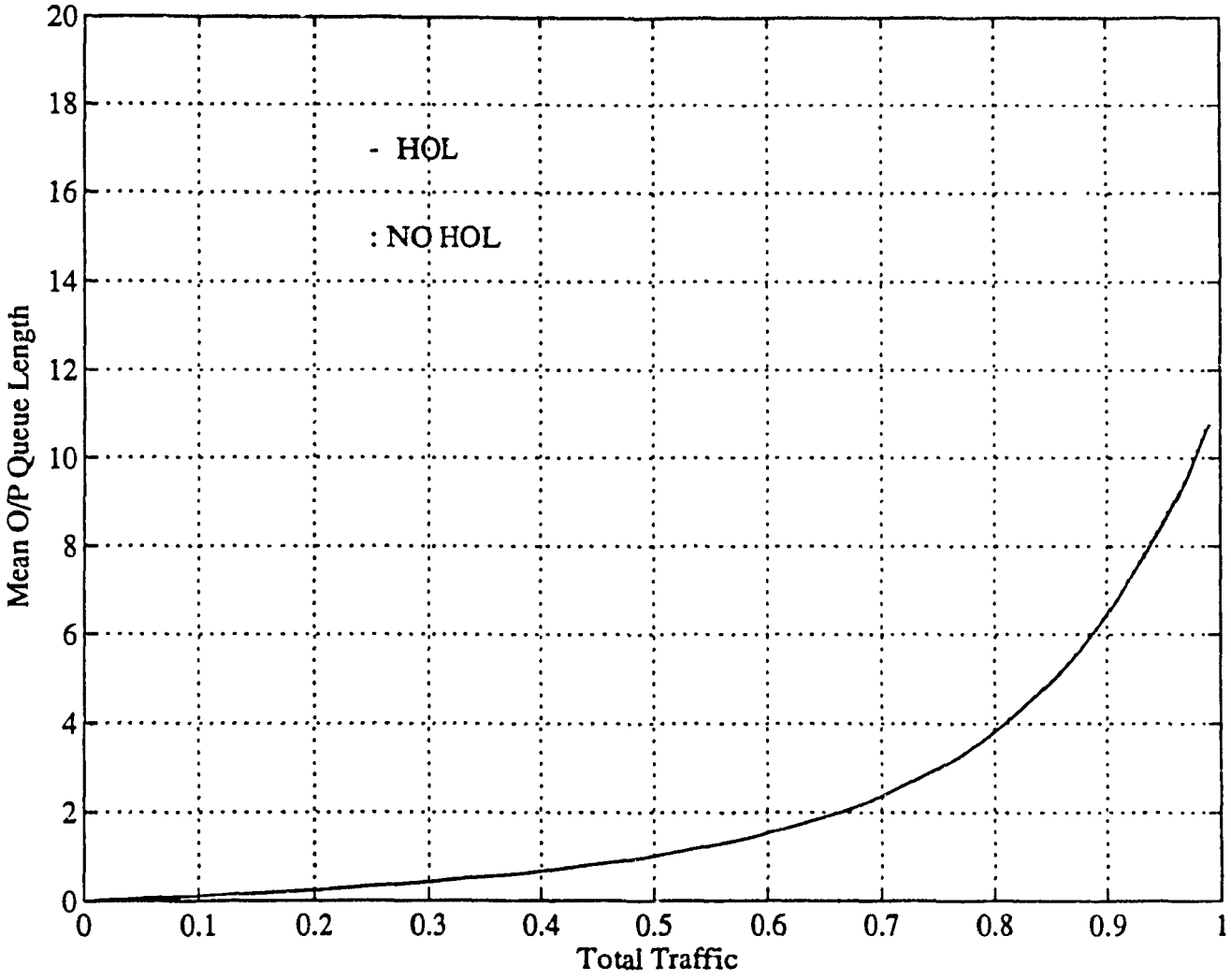


Fig. 4.10.b: Output Buffer Size. L=2, case b

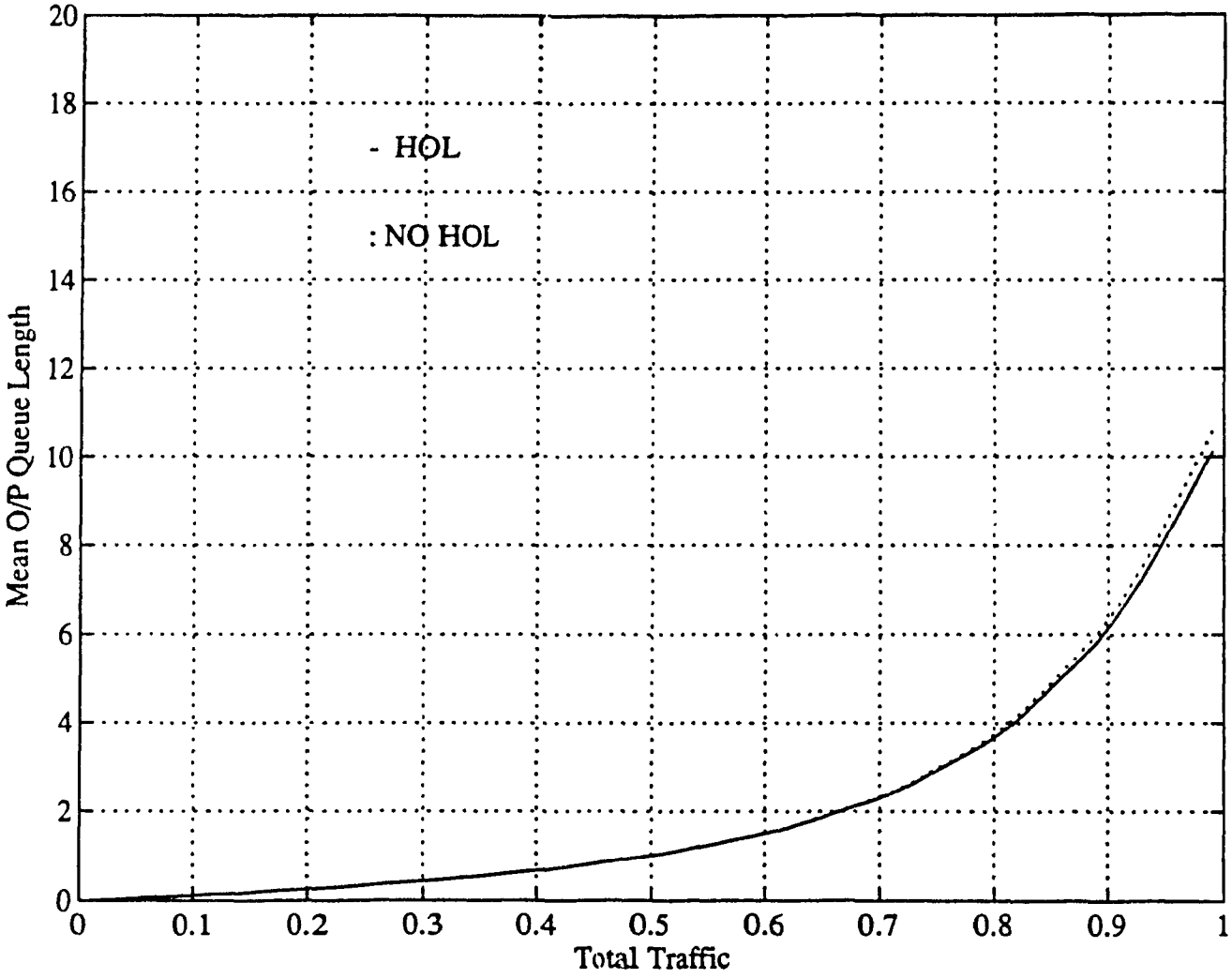


Fig. 4.11.a: Output Buffer Size. L=3, case a

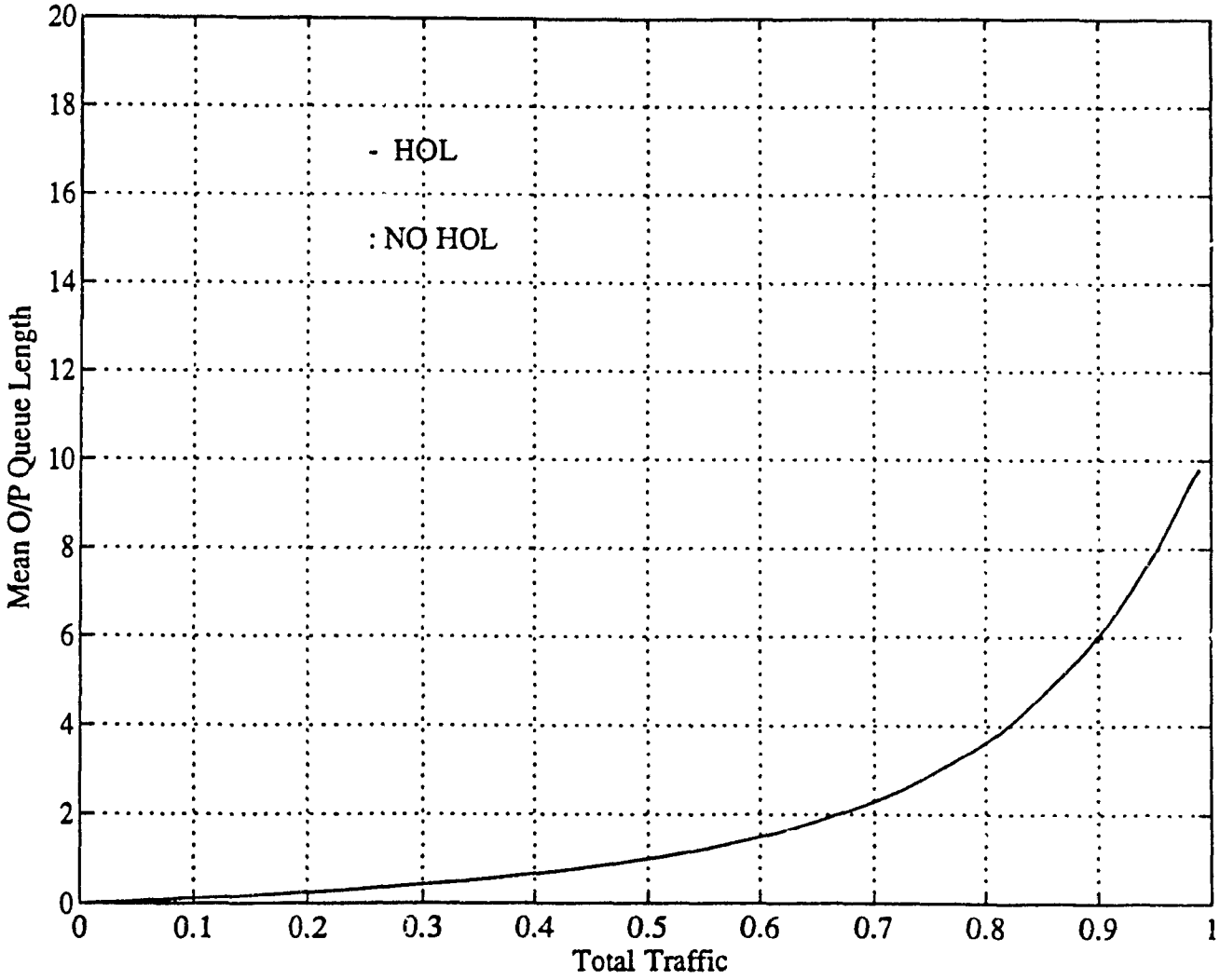


Fig. 4.11.b: Output Buffer Size. L=3, case b

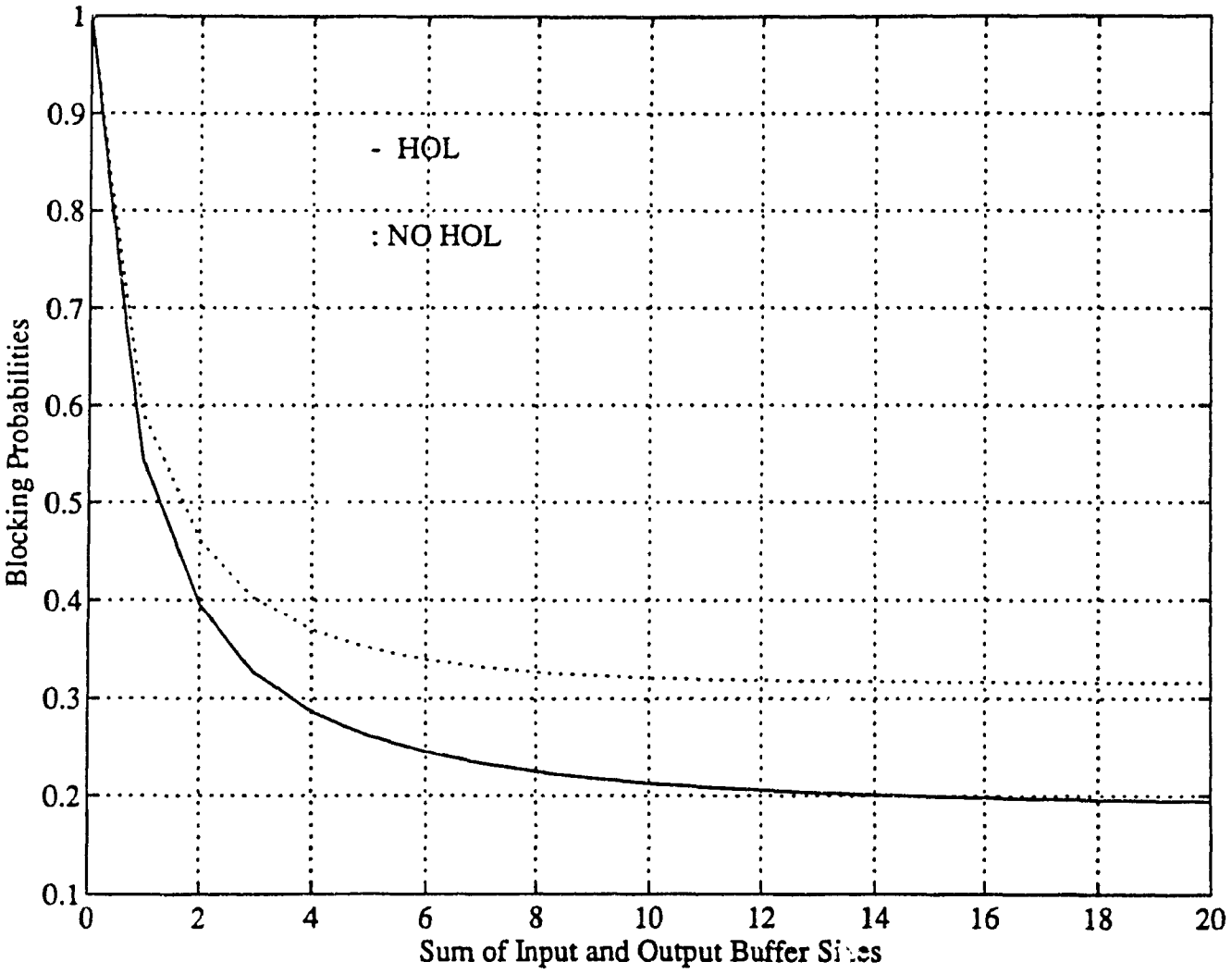


Fig.4.12.a: System Blocking Probability. L=1, lambda=0.95

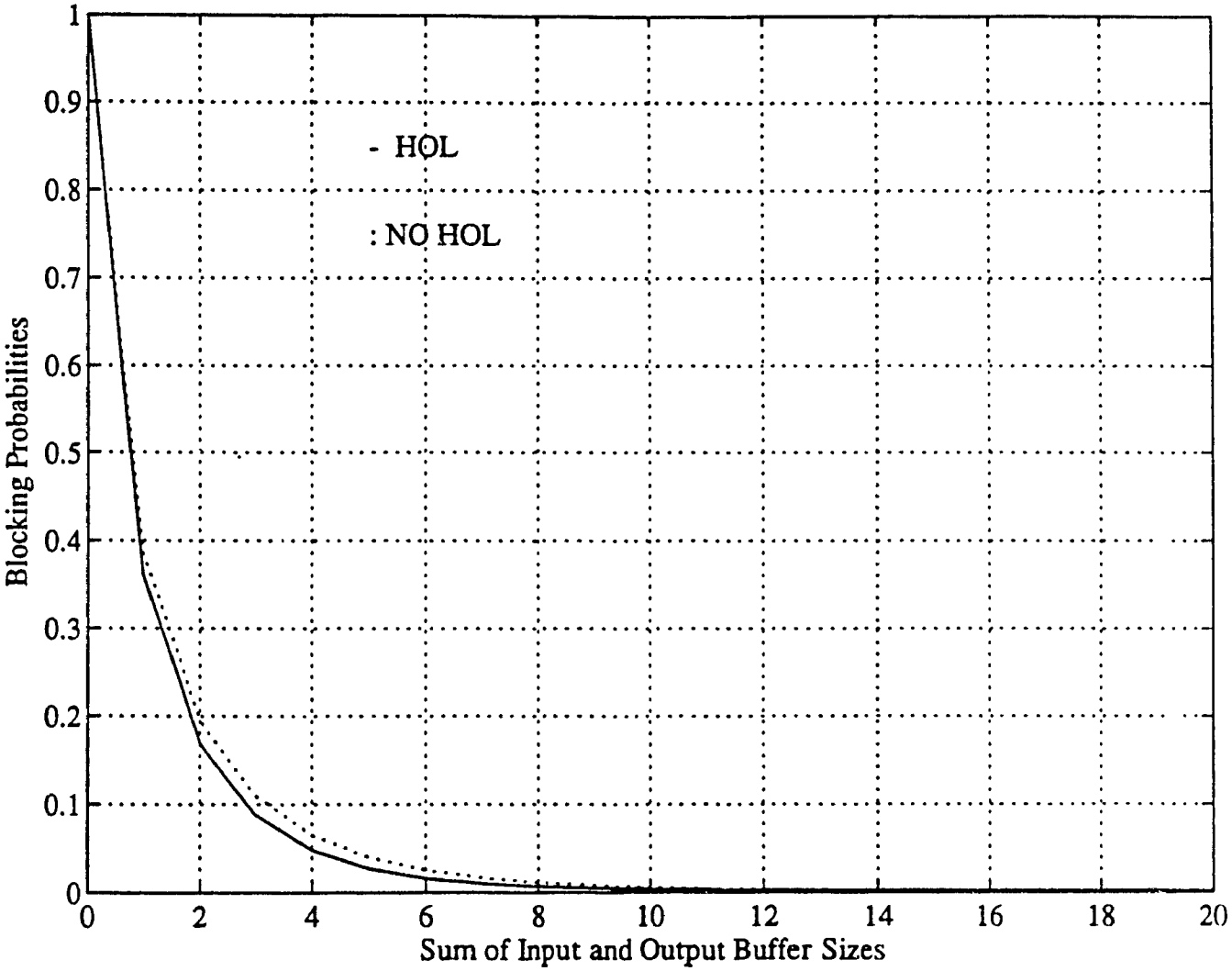


Fig.4.12.b: System Blocking Probability. L=1, lambda=0.5

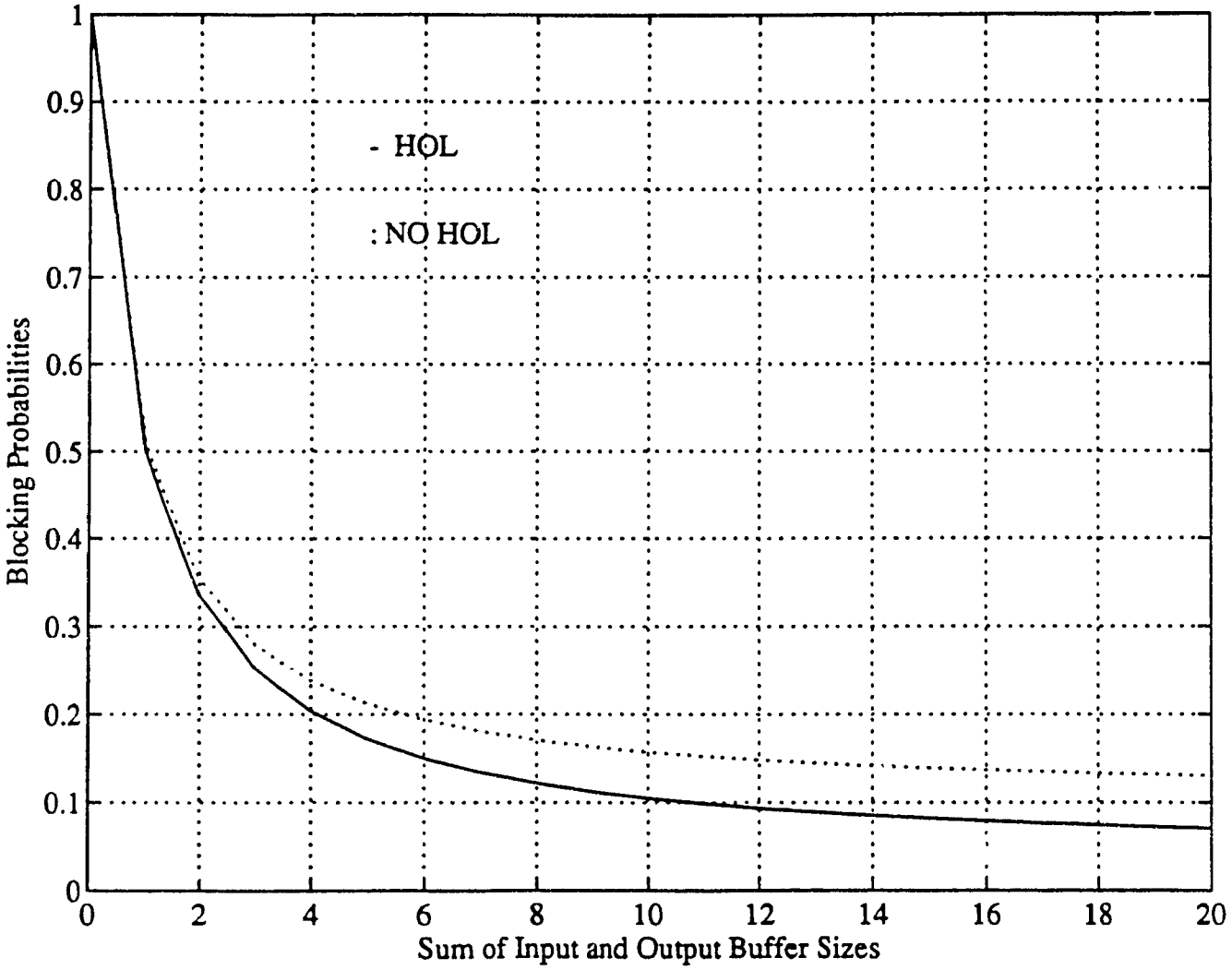


Fig.4.13.a: System Blocking Probability. L=2, lambda=0.95

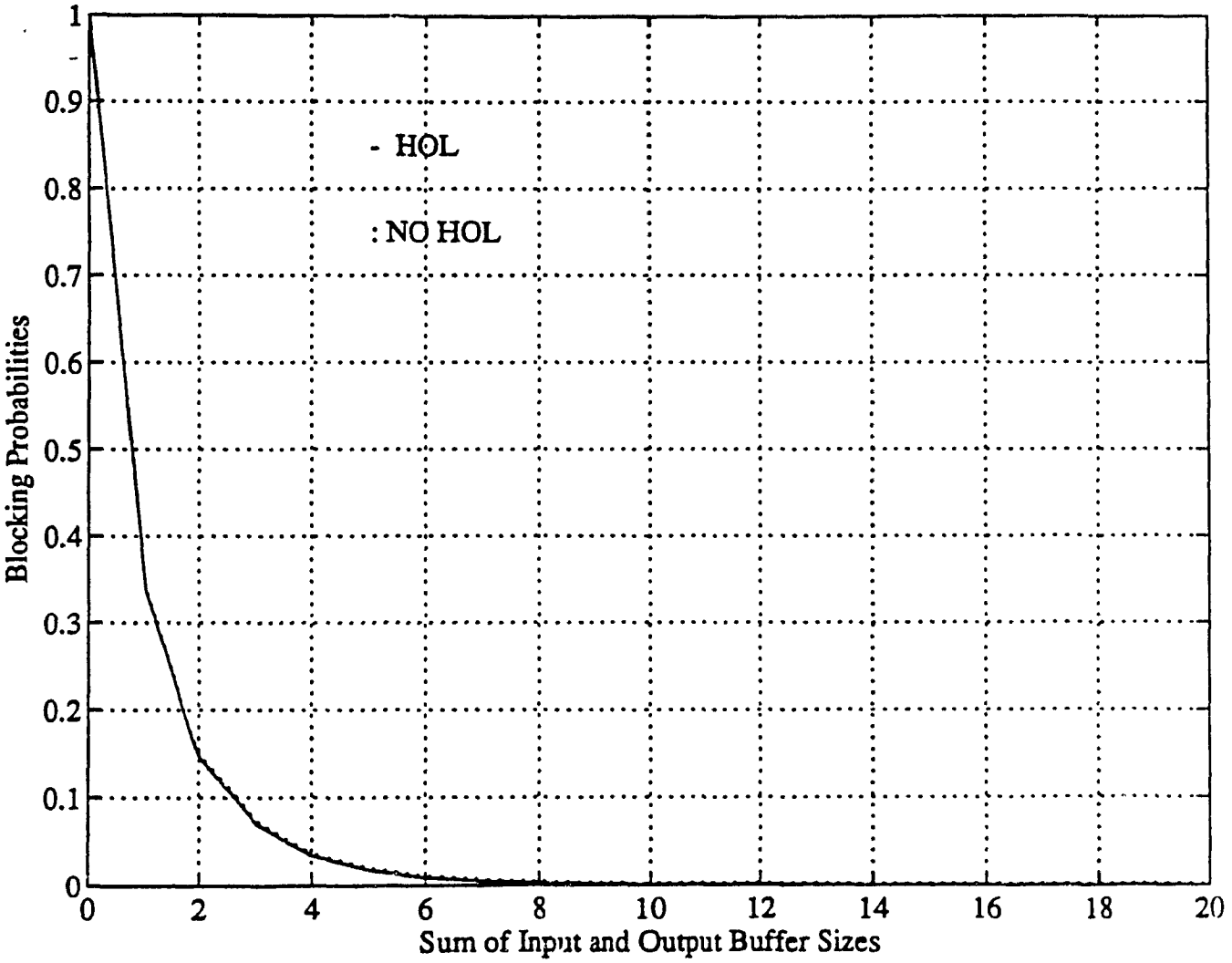


Fig.4.13.b: System Blocking Probability. L=2, lambda=0.5

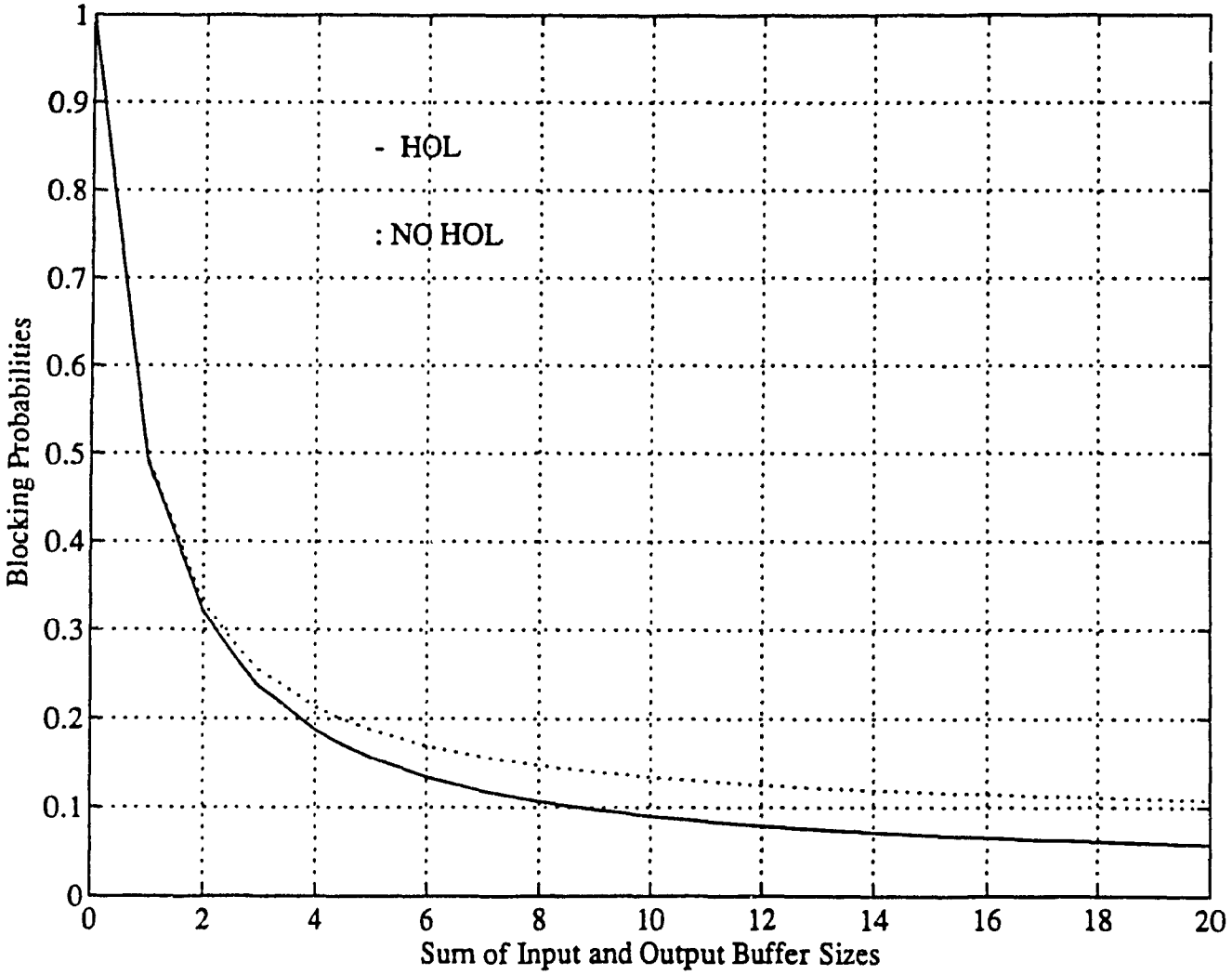


Fig.4.14.a: System Blocking Probability. L=3, lambda=0.95

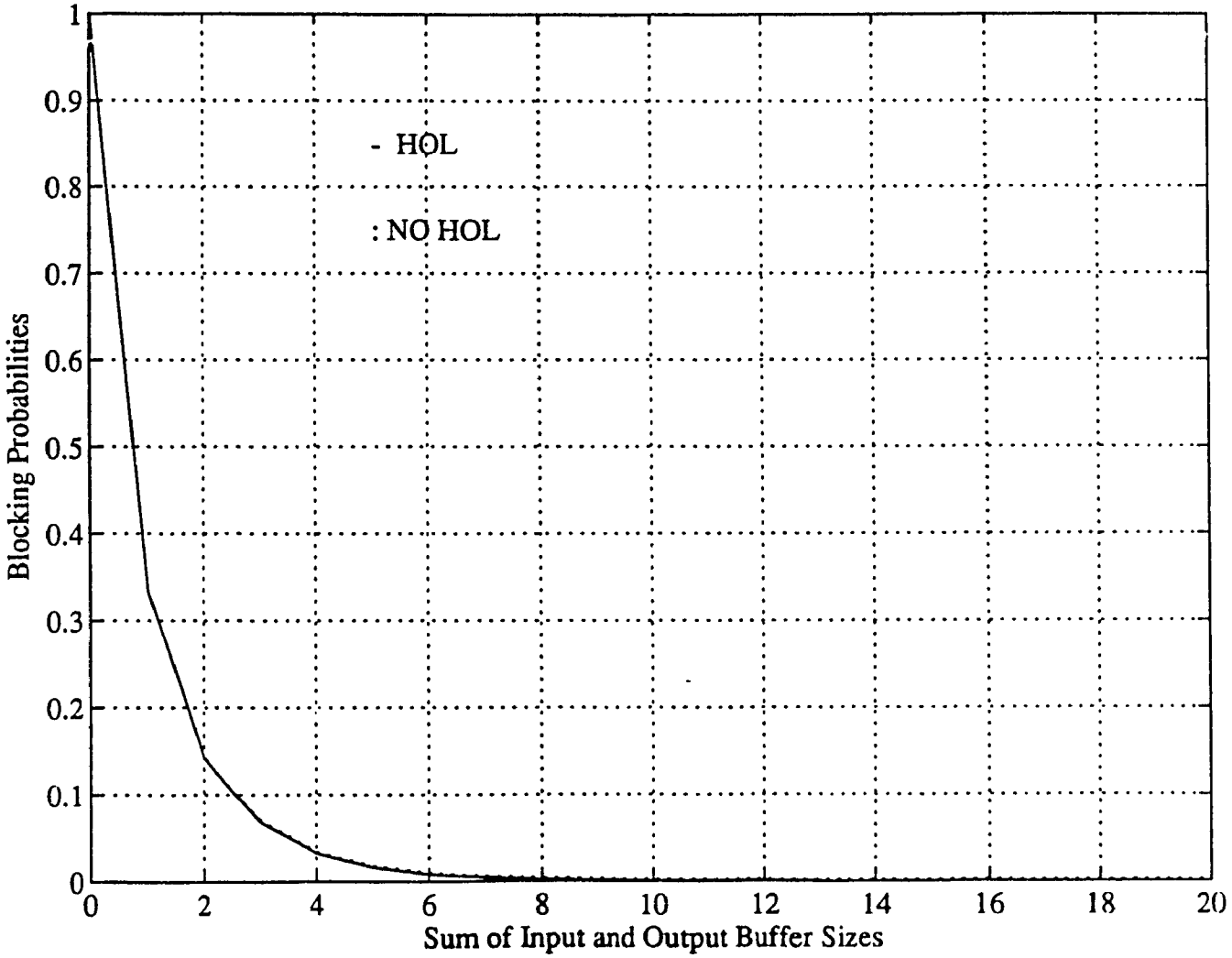


Fig.4.14.b: System Blocking Probability. L=3, lambda=0.5

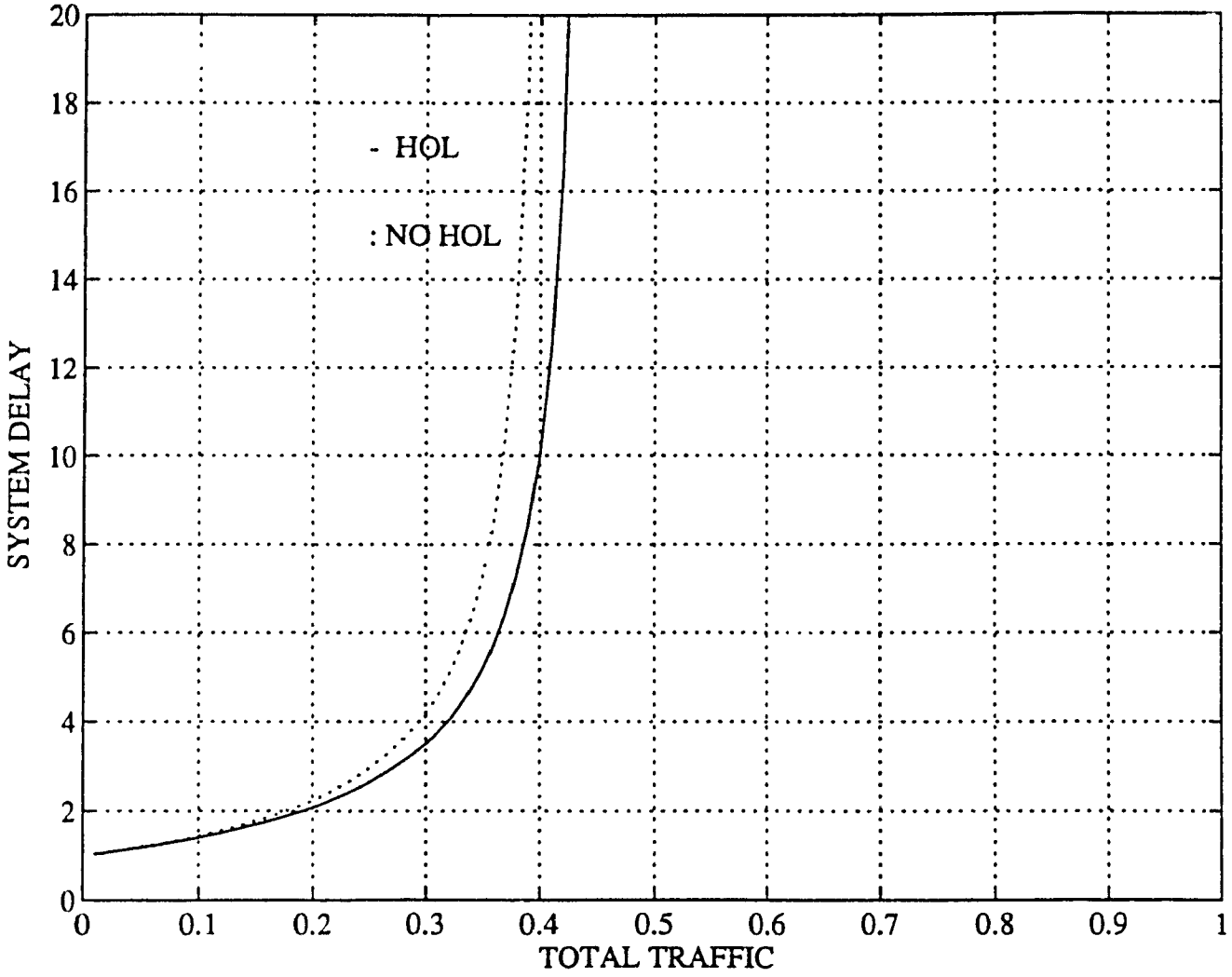


Fig.4.15.a: System Delay. L=1, case a

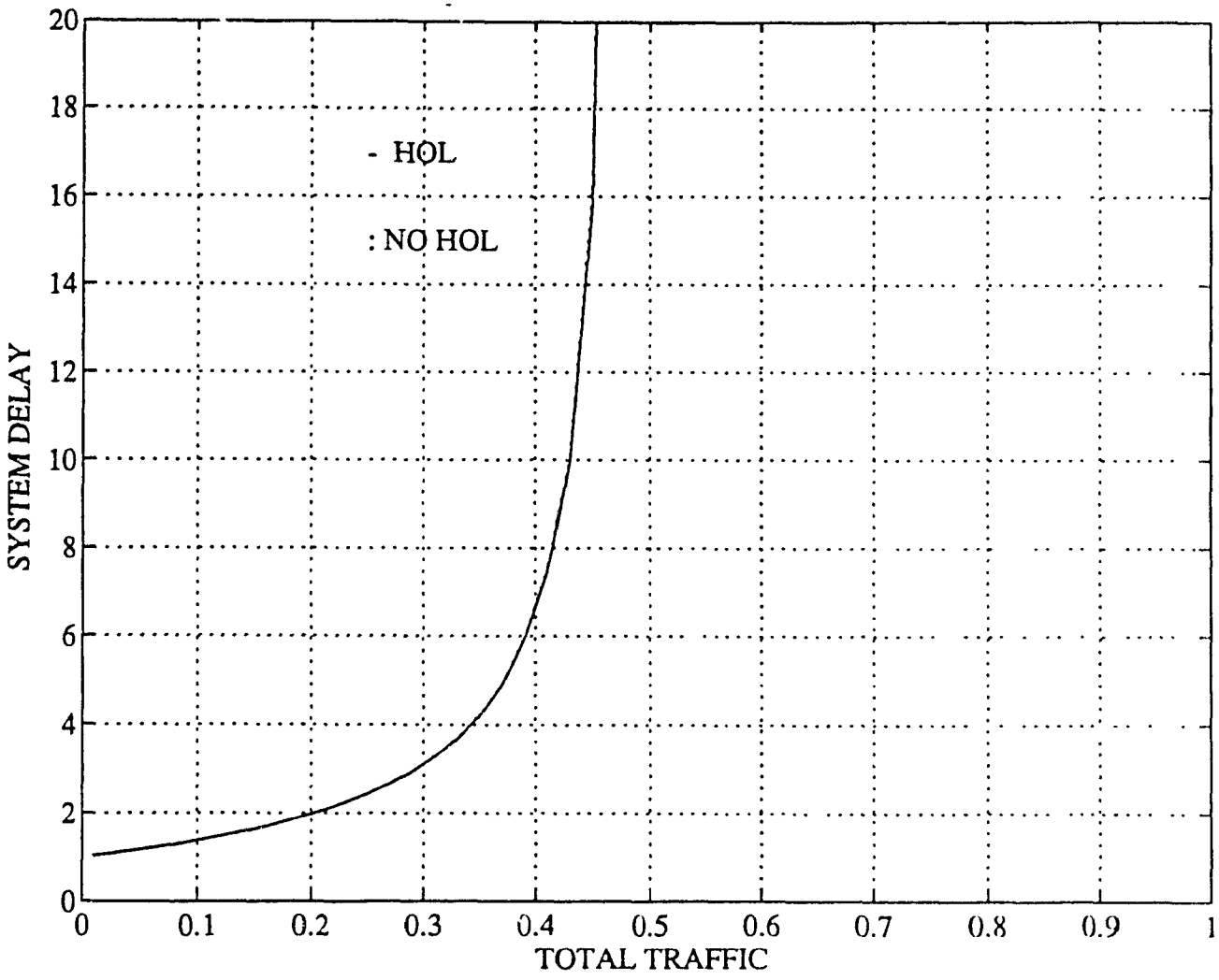


Fig.4.15.b: System Delay. L=1, case b

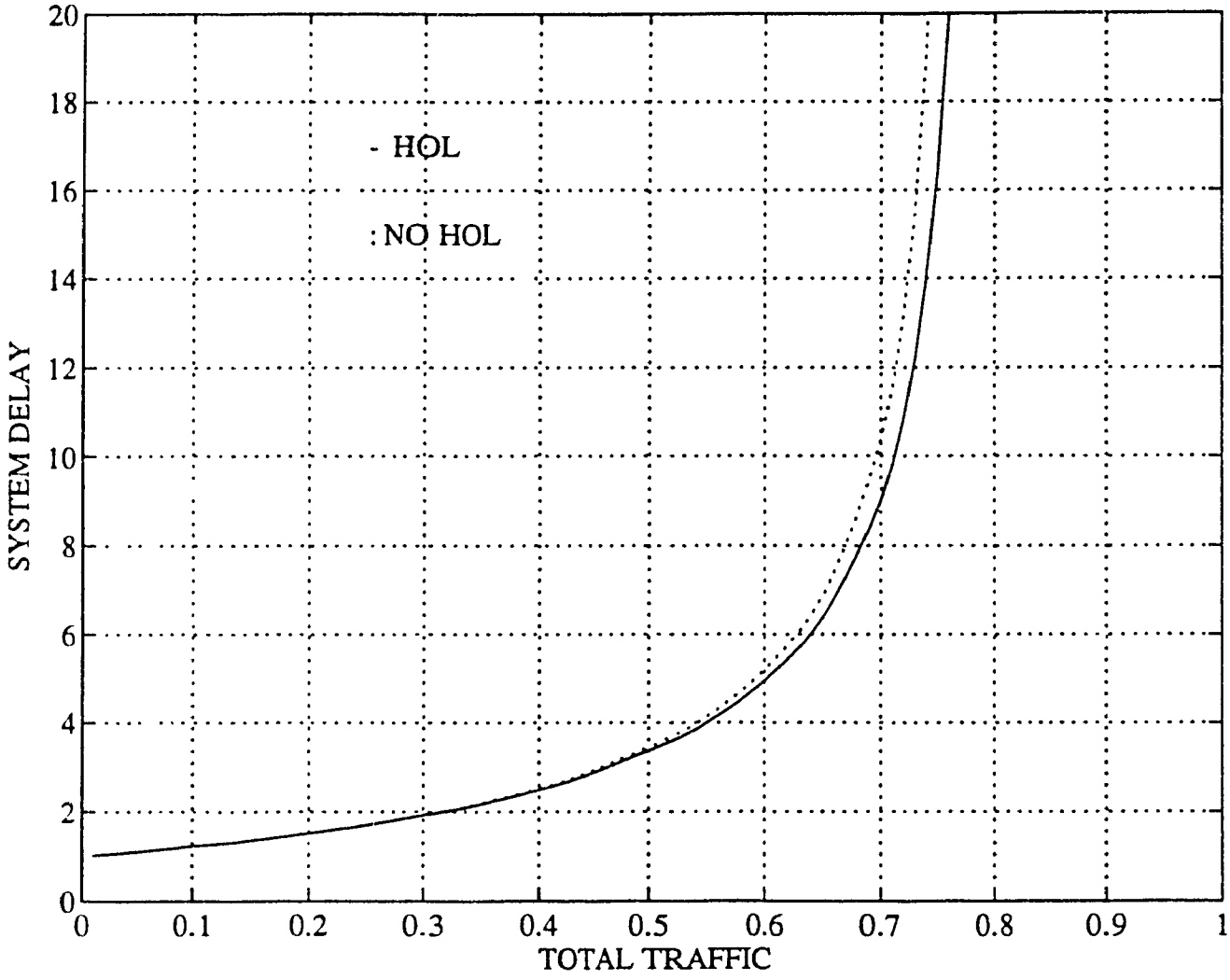


Fig.4.16.a: System Delay. L=2, case a

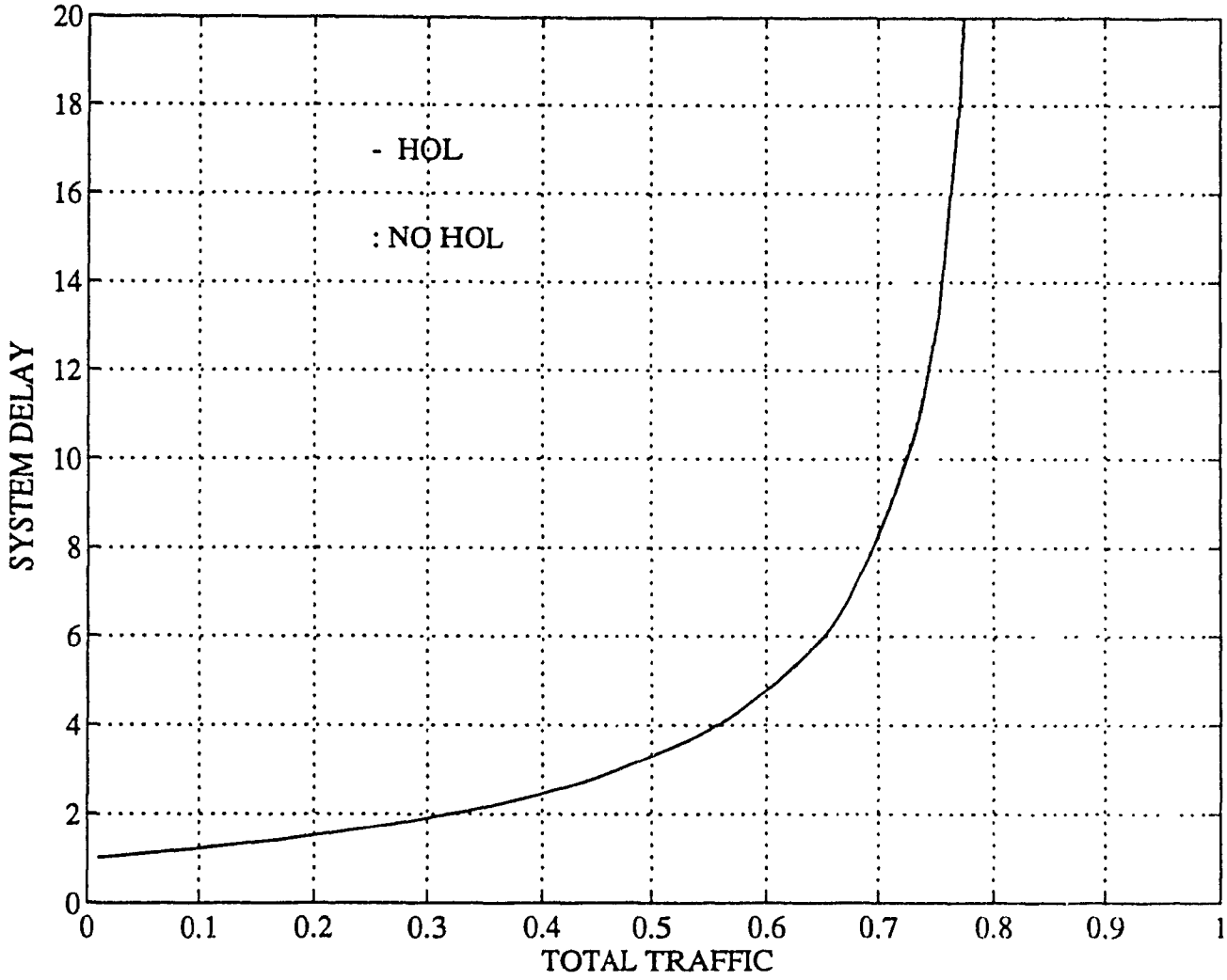


Fig.4.16.b: System Delay. L=2, case b

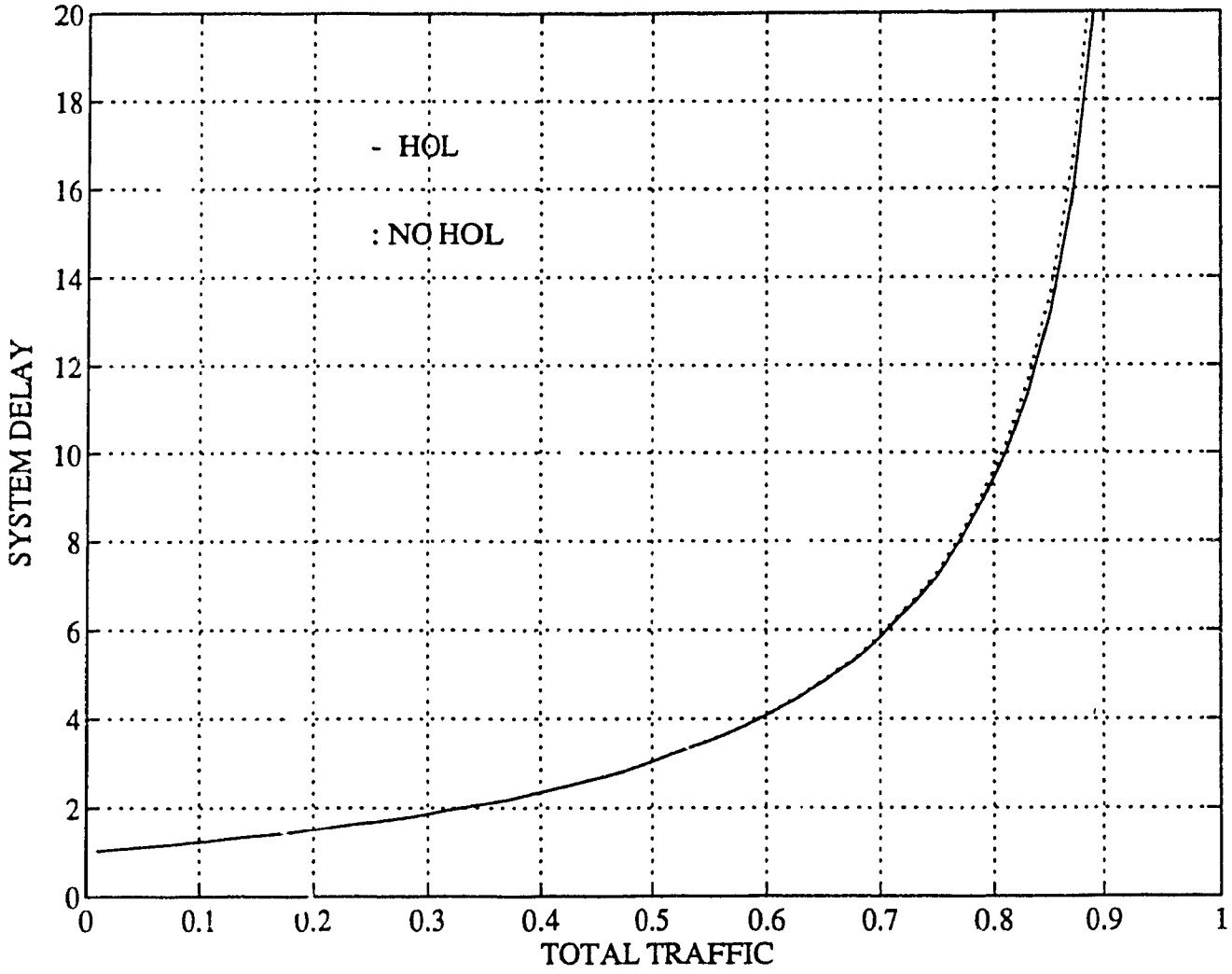


Fig.4.17.a: System Delay. L=3, case a

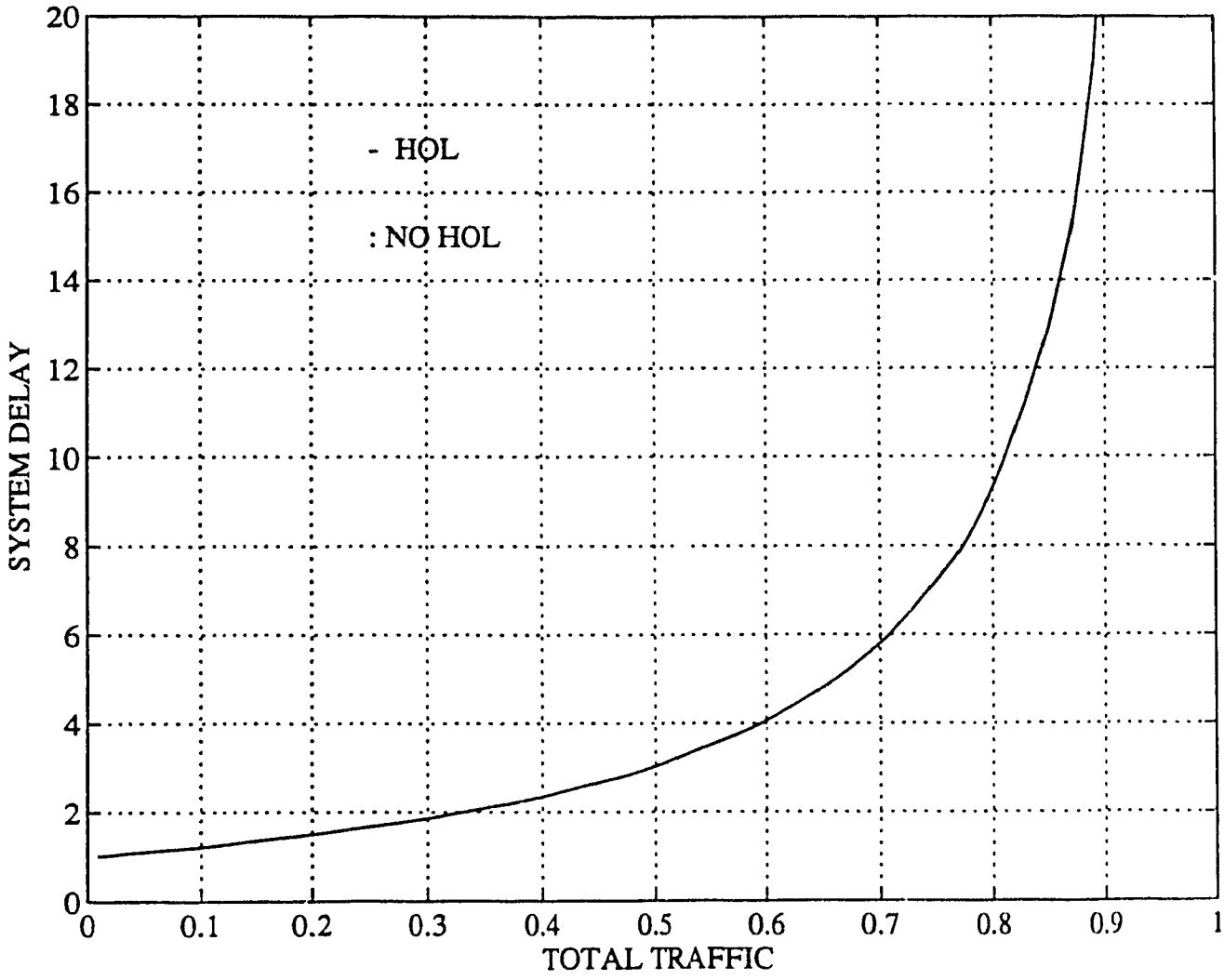


Fig.4.17.b: System Delay. L=3, case b

CHAPTER 5

CONCLUSION

A new Satellite Demand Assignment Multiaccess Technique of the moving frame boundary was presented. The technique satisfied the needs of different services including video, voice, and file data services which have relatively high utilization and long call durations. These services access the satellite by using one of the corresponding request slots available for each service based on a slotted Aloha protocol to get the reservation. Once reservation is granted, a user goes on and transmits its information in the reserved slots of the uplink frame. On the other hand bursty interactive data users are guaranteed some minimum space on the uplink frame which they exploit by using slotted Aloha protocol for transmission. Since it is always possible that some slots are unoccupied on the uplink frame, interactive data users get hold of those unused slots and transmit their data bits in those slots. By doing so, we tried to accomodate more interactive data users at no extra expense since those slots would have been empty anyway. This technique provided excellent call blocking, call establishment delays and greater amounts of interactive data throughputs at low traffic conditions of the circuit-switched classes. The results showed the trade-offs involved in selecting frame structures amenable to the different integrated services load variations. In Chapter four, a new HOL algorithm has been presented in conjunction with a nonblocking baseband satellite switch. The HOL was shown to provide a substantial improvement in the input buffer delay, packet blocking and buffer overflow probabilities. The HOL algorithm consisted mainly of accepting packets that are not Head of Line packets and that are sister packets of some HOL blocked packets. The idea is that if there is blocking in the head of the queue and if in that queue there is another packet whose output destination is not busy, we forward that sister packet even if the HOL packet is still blocked. It was also assumed that the sister packet had to be either

file data traffic or interactive data traffic only since video and voice packets were already given another level of priority in a preceding stage. Priority structures as well as accomodation of four different service classes were considered and it was shown that the combined priority and HOL alogorithm achieved the best balance between the different server classes. It was also noticed that increasing the speed-up ratio L leads to more improvements than using our new HOL resolution algorithm but we also recall that increasing the internal speed of the switch is a more costly venture. Finally we observed that having both advantages (L and HOL) leads to a dimishingly smaller blocking probabilities.

The analysis covered in Chapters III and IV was for one particular priority structure and as mentioned before it is possible, with some minor modifications, to cover different priorities and other types traffic. For example it is possible to analyse cases where we have more than four types of traffic as well as cases where the priority is not necessarily given to video over all other kinds of traffic, etc. Extension of this work to different unbalanced load conditions and application of various blocking switching techniques of the Banyan type to the integrated Satellite environemnt as well as the real time simulation of the different services and under the various conditions are some of the future research objectives.

REFERENCES

[1] K.M. Sundara Murthy and K.G. Gordon, " VSAT Networking Concepts and New Applications Development, " *IEEE Comm. Magazine*, pp. 48-49, May 1989.

[2] F. Ananasso and E. Soggese, " On Board Processing Concepts for Multibeam Communication Satellites, " *Alta Frequenza*, Vol. LVII-N.10, pp. 535-544, Dec. 1988.

[3] B.G. Evans, F.P. Coakley, M.H.M. El-Amin, S.C. Lu and C.W. Wong, " Baseband Switches and Transmultiplexors for use in an On-Board Processing Mobile/Business Satellite System, " *IEE Proceedings*, Vol. 133, Pt.F., No. 4, pp. 356-363, July 1986.

[4] R. M. Gagliardi, " *Introduction to Communications Engineering*, " Second Edition, Wiley Series in Telecommunications, pp. 524-526, 1988.

[5] M. B. Pursley, " Performance Evaluation for Phase Coded Spread Spectrum Multiple Access Communications - Part I: System Analysis, " *IEEE Trans. Comm.*, pp. 795-799, 1977.

[6] S. S. Lam, " Delay Analysis of Time Division Multiple Access (TDMA) Channel, " *IEEE Trans. Comm.*, pp. 1489-1494, 1977.

[7] T. Suda, H. Miyahara and T. Hasegama, " Performance Evaluation of an Integrated Access Scheme in a Satellite Communication Channel, " *IEEE J. Selec. Areas Comm* , Vol. SAC-1, No. 1, pp. 153-164, January 1983.

[8] N. Abramson, " The Aloha System - Another Alternative for Computer Communications, " *Proc. Fall Joint Comput. Conf. AFIPS*, Vol. 37, pp. 281-285, 1970.

[9] D. Raychaudhuri, " Stability, Throughput, and Delay of Asynchronous Selective Reject ALOHA, " *IEEE Trans. Comm.* Vol. 35, pp. 767-772, July 1987.

[10] K. Joseph and D. Raychaudhuri, "Stability Analysis of Asynchronous Random

Access CDMA Systems, " *Proc. IEEE Global Comm. Conf.*, Houston, pp. 48 1.1-7, Dec. 1986.

[11] B. Sklar, " *Digital Communications: Fundamentals and Applications*, " Prentice Hall, 1988.

[12] F. M. Nadori and W. W. Wu, " Advanced Satellite Concepts for Future Generation VSAT Networks, " *IEEE Comm. Magazine*, pp. 13-21, Sept. 1988.

[13] K. Koza, T. Muratani and A. Ogama, " On-Board Regenerative Receptors Applied to Digital Satellite Communications, " *Proc. IEEE*, Vol. 65, pp. 421-430, March 1977.

[14] M. Yabusabi and S.Suzubi, " Approximate Performance Analysis and Simulation Study for Variable-Channel-Per-Burst SS-TDMA, " *IEEE Trans. Comm.*, Vol. 38, No. 3, pp. 318-326, March 1990.

[15] L. C. Palmer and L. W. White, " Demand Assignment in the ACTS' LBR System, " *IEEE Trans. Comm.*, Vol. 38, No. 5, pp. 684-692, May 1990.

[16] M.J. Karol, M.G. Hluychi and S.P. Morgan, " Input Versus Output Queueing ona Space-Division Packet Switch, " *IEEE Trans. Comm.*, Vol. COM-35, No 12, pp. 1347-1356, Dec 1987.

[17] H. Vematsu and R. Watanabe, " Architecture of a Packet Switch Based on Banyan Switching Networks with Feedback Loops, " *IEEE J. Selec. Areas Comm*, Vol. 6, No. 9, pp. 1521-1527, Dec. 1988.

[18] S.Q. Li, " Nonuniform Traffic Analysis on a Nonblocking Space-Division Packet Switch, " *IEEE Trans. Comm.*, Vol. 38, No. 7, pp. 1085-1096, July 1990.

[19] L. R. Goke and G. J. Lipovski, " Banyan Networks for Partitioning Multiprocessor Networks, " *First Annual Intern. Symp. Comp. Arch.*, pp. 21-28, Dec 1973.

[20] J. H. Patel, " Performance of Processor-Memory Interconnections For Multiproces-

sors, " *IEEE J. Selec Areas Comm.*, Vol. 7, pp. 1091-1103, Sept. 1989.

[21] M. Kumar and J. R. Jump, " Performance Of Unbuffered Shuffle-Exchange Networks, " *IEEE Trans. Comput.*, Vol. C-35, No. 6, pp. 573-577, June 1986.

[22] Y. Jenq, " Performance Analysis of a Packet Switch Based on a Single-Buffered Banyan Network, " *IEEE J. Selec. Areas Comm.*, Vol. SAC-1, pp. 1014-1021, Dec. 1983.

[23] T. Theimer, E. Rathgeb and M. Huber, " Performance Analysis of Buffered Banyan Networks, " *IEEE Trans. Comm.*, Vol. 39, pp. 269-277, Feb. 1991.

[24] M. Huber, E. Rathgeb and T. Theimer, " Banyan Networks in an ATM Environment, " *Proc. ICCC'88*, Tel Aviv, Israel, pp. 167-174, 1988.

[25] D. M. Dias and J. R. Jump, " Analysis and Simulation of Buffered Delta Networks, " *IEEE Trans. Comput.*, Vol. C-30, No. 4, pp. 273-281, 1981.

[26] J. J. Degan & al., " Fast Packet Technology for Future Switches, " *AT&T Technical Journal*, pp. 36-50, March/April 1989.

[27] A. Huang and S. Knauer, " STARLITE: A Wideband Digital Switch, " *Proceedings Globecom, Atlanta, Georgia*, pp. 121-124, Nov. 1984.

[28] Y. S. Yeh, M. G. Hluychi ,and A. S. Acampora, " The Knockout Switch: A Simple, Modular Architecture for High-Performance Packet Switching, " *IEEE J. Selec. Areas Comm.*, Vol. SAC-5, No. 8, pp. 1274-1283, October 1987.

[29] Y. Oie and al, " Effect of Speedup in Nonblocking Packet Switch, " *Proc. ICC'89*, pp. 410-414, 1989.

[30] K.Y. Eng, " A Photonic Knockout Switch for High-Speed Packet Networks, " *IEEE J. Selec. Areas Comm.*, Vol. 6, pp. 1107-1115, Aug. 1988.

[31] T.T. Yee, " Nonblocking Copy Networks for Multicast Packet Switching, " *IEEE J Selec. Areas Comm.*, Vol. 6, No. 9, pp. 1455-67, Dec. 1988.

[32] M. Schwartz, " *Computer Communication Network - Design and Analysis*, " Prentice Hall, 1977.

[33] L. Kleinrock, " *Queuing Systems Volume 2: Computer Applications*, " John Wiley and Sons, p.9, 1976.

[34] H. Ahmadi and W. Denzel, " A Survey of Modern High-Performance Switching Techniques, " *IEEE J. Selec. Areas Comm.*, Vol. 7, pp. 1091-1103, Sept. 1989

[35] J. S. Chen and T. E. Stern, " Throughput Analysis, Optimal Buffer Allocation, and Traffic Imbalance Study of a Generic Nonblocking Packet Switch, " *IEEE J. Selec Areas Comm.*, Vol. 9, No. 3, pp. 439-449, Apr. 1991.

[36] A. K. Elhakeem, et al. " Analysis of a New Multiaccess/Switching Technique for Multibeam Satellites in a Prioritized ISDN Environment, " *IEEE J. Selec Areas Comm.*, Vol. 10, No. 2, pp. 378-390, Feb. 1992.

[37] J. S. Chen and T. E. Stern, "Performance of a Generic Nonblocking Switch, " *Proc. of 4th Int. Conf. Data Comm. Syst.*, Barcelona, Spain, June 1990.

[38] Tri T. Ha, " *Digital Satellite Communications*, " Mcgraw Hill Series on Communications, 2nd Edition, 1990.

****