PERFORMANCE EVALUATION
OF
THE INTEGRATION OF VOICE AND DATA IN A
HIGH-SPEED LOCAL AREA COMPUTER NETWORK
---- THE EXPRESSNET

by

JACKSON YING-KIN LEE

B.A.Sc., The University of New Brunswick, 1984

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES
(Department of Electrical Engineering)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
October 1987
@ Jackson Ying-Kin Lee, 1987
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

DEPARTMENT OF ELECTRICAL ENGINEERING

The University of British Columbia
2075 Wesbrook Place
Vancouver, Canada
V6T 1W5

Date : October, 1987
ABSTRACT

A high-speed local area communication network--the Expressnet--is investigated in this thesis with regard to voice and data transmissions. Performance criteria, such as channel utilizations, delay characteristics, and queue lengths are determined from computer simulation and numerical calculation approaches. The protocol is particularly suitable for the transmission of packetized voice as it is able to guarantee an upper bound on the transmission delay for each packet. The network under study thus will find major application in future office automation, where large amounts of voice will be integrated with data.
# TABLE OF CONTENTS

LIST OF ILLUSTRATIONS ................................................................. v

GLOSSARY OF SYMBOLS ........................................................................ x

ACKNOWLEDGMENTS ........................................................................ xii

1. INTRODUCTION ............................................................................. 1

1.1 ISDN ......................................................................................... 1

1.2 Organization of Thesis ......................................................... 4

1.3 Integration of Voice and Data in Packet-switched Networks .......... 5

1.3.1 Advantages ...................................................................... 5

1.3.2 Performance Criteria for Data and Voice Traffic ........ 6

1.3.3 Transmission Procedures of Packetized Voice and Data .......... 8

1.3.4 Channel Access Protocols ............................................ 9

1.4 Scope of Thesis ..................................................................... 11

1.4.1 Broadcast Systems, Round Robin Schemes, and Expressnet .... 11

1.4.2 Basic Assumptions and Performance Measures of Interest ...... 13

2. TRANSMISSION OF DATA PACKETS IN EXPRESSNET .............. 16

2.1 General Description of Expressnet ..................................... 16

2.2 Assumptions Used in Simulation Model for Data Protocol ........ 18

2.3 Summary of Results ......................................................... 20

2.3.1 Analysis of Results for Single Buffer Case .......... 20

2.3.2 Analysis of Results for Infinite Buffer Case .... 32
<table>
<thead>
<tr>
<th>Illustration</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Conceptual view of ISDN connection features</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>An integrated distributed communications architecture</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Block diagram of integrated voice/data packet transmission system</td>
<td>10</td>
</tr>
<tr>
<td>1.4</td>
<td>Block diagram of multiple access schemes</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>Block diagram of Expressnet</td>
<td>17</td>
</tr>
<tr>
<td>2.2</td>
<td>Transmission procedure of Expressnet</td>
<td>19</td>
</tr>
<tr>
<td>2.3</td>
<td>Utl vs IAT, N = 20</td>
<td>23</td>
</tr>
<tr>
<td>2.4</td>
<td>Mean queue length vs Utl, N = 20</td>
<td>23</td>
</tr>
<tr>
<td>2.5</td>
<td>Mean wait time vs Utl, N = 20</td>
<td>25</td>
</tr>
<tr>
<td>2.6</td>
<td>Variance of wait time vs Utl, N = 20</td>
<td>25</td>
</tr>
<tr>
<td>2.7</td>
<td>Utl vs N, IAT = 12.5</td>
<td>27</td>
</tr>
<tr>
<td>2.8</td>
<td>Mean queue length vs Utl, IAT = 12.5</td>
<td>27</td>
</tr>
<tr>
<td>2.9</td>
<td>Mean wait time vs Utl, IAT = 12.5</td>
<td>29</td>
</tr>
<tr>
<td>2.10</td>
<td>Variance of wait time vs Utl, IAT = 12.5</td>
<td>29</td>
</tr>
<tr>
<td>2.11</td>
<td>Utl vs N, IAT = 12, 12.5</td>
<td>30</td>
</tr>
<tr>
<td>2.12</td>
<td>Mean queue length vs Utl, IAT = 12, 12.5</td>
<td>30</td>
</tr>
<tr>
<td>2.13</td>
<td>Mean wait time vs Utl, IAT = 12, 12.5</td>
<td>31</td>
</tr>
<tr>
<td>2.14</td>
<td>Variance of wait time vs Utl, IAT = 12, 12.5</td>
<td>31</td>
</tr>
<tr>
<td>2.15</td>
<td>Utl vs N, IAT = 28</td>
<td>34</td>
</tr>
<tr>
<td>2.16</td>
<td>Mean queue length vs N, IAT = 28</td>
<td>34</td>
</tr>
<tr>
<td>2.17</td>
<td>Mean wait time vs N, IAT = 28</td>
<td>35</td>
</tr>
<tr>
<td>2.18</td>
<td>Variance of wait time vs N, IAT = 28</td>
<td>35</td>
</tr>
<tr>
<td>3.1</td>
<td>Packet speech model</td>
<td>40</td>
</tr>
</tbody>
</table>
3.2 Maximum ETE delay requirement ........................................ 40
3.3 Voice transmission in Expressnet ....................................... 42
3.4 Utl vs N, V1/V2 ............................................................ 46
3.5 Mean queue length vs Utl, V1/V2 ...................................... 46
3.6 Mean wait time vs Utl, V1/V2 .......................................... 47
3.7 Variance of wait time vs Utl, V1/V2 .................................. 47
3.8 Mean queue length vs Utl, V1 .......................................... 48
3.9 Mean queue length vs Utl, V2 .......................................... 48
3.10 Mean train size vs Utl, V1/V2 ......................................... 50
3.11 Variance of train size vs Utl, V1/V2 ................................ 50
3.12 PDLP vs Utl, V1/V2 ...................................................... 51
3.13 PDLP vs N, V1/V2 .......................................................... 51
3.14 Markov chain for V2 protocol .......................................... 52
3.15 Model for an active voice source ...................................... 53
3.16 VPDR vs Utl, V1/V2 ....................................................... 55
3.17 VPDR vs N, V1/V2 .......................................................... 55
3.18 Call arrivals and distribution on-off model .......................... 56
3.19 Actual vs estimated # active stations, \( \alpha = 0.1 \) ............... 59
3.20 Actual vs estimated # active stations, \( \alpha = 0.5 \) ............... 59
3.21 Actual vs estimated # active stations, \( \alpha = 0.9 \) ............... 60
3.22 Utl vs N, theoretical / simulation .................................... 60
3.23 Utl vs N, different \( \alpha \) .............................................. 63
3.24 Mean wait time, mean queue length vs N ............................ 63
3.25 Variance of wait time vs N .............................................. 64
3.26 % packet discarded, blocking probability vs N .................... 64
4.1 Configuration of integrated V/D Expressnet

4.2 Utl vs N, VDl, IAT = 12

4.3 Mean queue length vs Utl, VDl, IAT = 12

4.4 Mean wait time vs Utl, VDl, IAT = 12

4.5 Variance of wait time vs Utl, VDl, IAT = 12

4.6 Mean train size vs Utl, VDl, IAT = 12

4.7 Variance of train size vs Utl, VDl, IAT = 12

4.8 PDLP vs Utl, VDl, IAT = 12

4.9 PDLP vs N, VDl, IAT = 12

4.10 Utl vs IAT, VDl, N = 14

4.11 Mean queue length vs IAT, VDl, N = 14

4.12 Mean wait time vs IAT, VDl, N = 14

4.13 Variance of wait time vs IAT, VDl, N = 14

4.14 Mean train size vs IAT, VDl, N = 14

4.15 Variance of train size vs IAT, VDl, N = 14

4.16 Utl vs N, fixed IAT, K_{data}

4.17 Mean queue length vs Utl, fixed IAT, K_{data}

4.18 Mean wait time vs Utl, fixed IAT, K_{data}

4.19 Variance of wait time vs Utl, fixed IAT, K_{data}

4.20 Mean train size vs Utl, fixed IAT, K_{data}

4.21 Variance of train size vs Utl, fixed IAT, K_{data}

4.22 % D, V pks discarded vs Utl, fixed IAT, K_{data}
4.23 % delayed voice packets vs Utl, fixed IAT, K_{data}, \varphi ........... 84
4.24 Utl vs \varphi, varying K_{data}, fixed IAT, N ......................... 85
4.25 Mean queue length vs \varphi, varying K_{data}, fixed IAT, N ........ 85
4.26 Mean wait time vs \varphi, fixed K_{data}, IAT, N ..................... 86
4.27 Variance of wait time vs \varphi, fixed K_{data}, IAT, N ............. 86
4.28 Mean wait time vs \varphi, varying K_{data}, fixed IAT, N .......... 87
4.29 Variance of wait time vs \varphi, varying K_{data}, fixed IAT, N .... 87
4.30 Mean train size vs \varphi, varying K_{data}, fixed IAT, N .......... 89
4.31 Variance of train size vs \varphi, varying K_{data}, fixed IAT, N ... 89
4.32 % Data packets discarded vs \varphi, varying K_{data}, fixed IAT, N ... 90
4.33 % Voice packets discarded vs \varphi, varying K_{data}, fixed IAT, N ... 90
4.34 % D, V packets discarded vs \varphi, fixed K_{data}, fixed IAT, N ... 91
4.35 % delayed V pks(PDLP) vs \varphi, varying K_{data}, fixed IAT, N ... 91
4.36 % data transmitted vs \varphi, varying K_{data}, fixed IAT, N ........ 92
4.37 % voice transmitted vs \varphi, varying K_{data}, fixed IAT, N ........ 92
4.38 Utl vs N, fixed K_{data}, \varphi, IAT .................................. 94
4.39 Mean queue length vs N, fixed K_{data}, \varphi, IAT .................. 94
4.40 Mean wait time vs N, fixed K_{data}, \varphi, IAT ..................... 95
4.41 Variance of wait time vs N, fixed K_{data}, \varphi, IAT .............. 95
4.42 Mean wait time of voice pks vs N, fixed K_{data}, \varphi, IAT ...... 96
4.43 Variance of wait time of voice pks vs N, fixed K_{data}, \varphi, IAT.. 96
4.44 Mean train size vs N, fixed K_{data}, \varphi, IAT ....................... 97
4.45 Variance of train size vs N, fixed K_{data}, \varphi, IAT .............. 97
4.46 % Data, Voice Transmitted vs N, fixed K_{data}, \varphi, IAT ....... 98
4.47 % Voice packets discarded vs N, fixed K_{data}, \varphi, IAT ....... 98
**GLOSSARY OF SYMBOLS**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>total number of stations connected to the cable or network</td>
</tr>
<tr>
<td>( C )</td>
<td>total available channel capacity (Mbps)</td>
</tr>
<tr>
<td>( B )</td>
<td>size of data packets (bits)</td>
</tr>
<tr>
<td>( v )</td>
<td>propagation velocity of signals in a coaxial cable (km/sec)</td>
</tr>
<tr>
<td>( L_c )</td>
<td>length of cable (km)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>arrival rate of data packets (packets/slot)</td>
</tr>
<tr>
<td>( \text{IAT} )</td>
<td>mean interarrival time of data packets (seconds)</td>
</tr>
<tr>
<td>( \tau )</td>
<td>one-way propagation delay (seconds)</td>
</tr>
<tr>
<td>( \tau_c )</td>
<td>propagation delay in connection cable (seconds)</td>
</tr>
<tr>
<td>( T )</td>
<td>transmission time of a packet (seconds)</td>
</tr>
<tr>
<td>( t_d )</td>
<td>time taken to detect presence or absence of carrier (seconds)</td>
</tr>
<tr>
<td>( T_{\text{lcom}} )</td>
<td>duration of locomotive (seconds)</td>
</tr>
<tr>
<td>( a )</td>
<td>ratio between one-way propagation delay and the packet transmission time</td>
</tr>
<tr>
<td>( \text{SIM} )</td>
<td>simulation (results)</td>
</tr>
<tr>
<td>( \text{TH} )</td>
<td>theoretical (results)</td>
</tr>
<tr>
<td>( U_{\text{tl}}(\phi) )</td>
<td>channel utilization</td>
</tr>
<tr>
<td>( \overline{Q}_1 )</td>
<td>mean queue length of system, mean number of packets waiting in the system</td>
</tr>
<tr>
<td>( \overline{Q}_2 )</td>
<td>mean number of packets in the system, including the one being served</td>
</tr>
<tr>
<td>( X )</td>
<td>transmission period of a packet, including overhead before each transmission to determine which user gets access to the channel, preamble transmission time, and data transmission time (seconds)</td>
</tr>
<tr>
<td>( Y )</td>
<td>the time that the channel becomes idle before rounds (interround overhead) (seconds)</td>
</tr>
<tr>
<td>( D_n )</td>
<td>network transit time of a bit (seconds)</td>
</tr>
</tbody>
</table>
maximum tolerable delay of an individual bit, defined as the lapse of time from the moment the bit is produced to the time when it is successfully transmitted to the receiver (seconds)

e number of bits in a full voice packet

packet formation time in originating terminal or packetization time of Bv bits (seconds)

mean overall waiting time (seconds)

mean waiting time of data packets (seconds)

mean waiting time of voice packets (seconds)

variance of overall waiting time of packets

variance of waiting time of data packets

variance of waiting time of voice packets

mean train size

variance of train size

percentage of voice packets delayed more than 20 ms

percentage of packets that are to be discarded if they are to incur a delay more than the time taken to transmit a full voice packet

full size of voice packet, given by pgftfull = Dn/(Bv/C)

maximum number of voice users supportable (C/Cv) such that the longest delay experienced by each bit is guaranteed to be less than Dn

length of voice packet occupied by overhead

length of voice packet consisting of actual information bits

number of voice stations that can be accommodated by system such that the maximum tolerable network delay (20 ms) is not exceeded at any time

bit rate of the vocoder in the originating voice terminal (kbps)

transmission time of a packet in the channel (seconds), given by Bv/C

arrival rate of voice calls, in calls/minute
ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to his supervising professor, Dr. H.W. Lee, for his unending patience, invaluable guidance, and meticulous supervision throughout the course of the project. Also to be expressed are the author's sentiments of appreciation and affection towards his parents, Mr. and Mrs. C.C. Lee, whose love and continuous support helped make the completion of this thesis possible.

Grateful acknowledgement is given to the author's devoted friends, Mr. I.L. Soeharjono and Miss B. Robles, for their constant encouragement, particularly to Miss Robles for proofreading the paper.

The author also extends his thanks to Ms. Rachel Rousseau for typing the thesis.
1. INTRODUCTION

1.1 ISDN

It is now widely recognized that the "Integrated Services Digital Network (ISDN)" concept is providing a useful framework for the development of future telecommunications networks and services.

An ISDN can be regarded as a general-purpose, intelligent digital network capable of supporting (or integrating) a wide range of services (voice and non-voice) using a small set of standard multipurpose user-network interfaces and digital transmission links. [ISDN 1-3]

The major argument in favour of digital technology is that the transfer of digital information is less sensitive to noise, cross talk, and distortion than that of analog data. Faded signals can also be regenerated without introducing cumulative degradations.

As the common carrier network evolves to a predominantly digital backbone plant, the network processing functions are able to dynamically allocate portions of the capacity to support an array of functionally independent networks and facilities, closely tailored to individual user characteristics.
The full complement of communication facilities would comprise voice, digital data, image (facsimile), and video communications. These functions would be integrated into a network with an overall capability as depicted in Figure 1.1. [ISDN 2]

High-capacity, high-quality communication channels, largely over satellite facilities, or by radio means, provide the major long haul component of the distributed architecture (Figure 1.2). [ISDN 1]

The principal benefits to the user can be expressed in terms of economy and flexibility. The ultimate objectives of integration are to accommodate various existing and new devices, provide a wide range of flexibility and versatility to terminal connectivity, and realize the economies of equipment commonality, large-scale integration, higher resource utility, and combined network operations, maintenance, and administrative policies.

Activities currently under way are leading to the development of worldwide ISDN. A number of standard organizations are involved in various aspects of ISDN, with CCITT being the major controlling body.

Among the full ISDN context, this thesis deals primarily with the section on the integration of packetized voice and data in a local area network, with emphasis on voice transmission and protocols.
Figure 1.1 Conceptual view of ISDN connection features

Figure 1.2 An integrated distributed communications architecture
Moreover, it is assumed that the different user nodes in the communication environment are stationary. For information on the integration of voice and data on satellite and mobile radio transmission, the reader is urged to refer to references under the "MOB" heading.

1.2 Organization of Thesis

The organization of this report is as follows. Motivated by ISDN developments discussed in Section 1.1, Section 1.3 highlights general transmission aspects of integration of voice and data. This part covers potential benefits, performance criteria, transmission procedures, and channel access protocols. The scope of this thesis is then presented in Section 1.4, in which a brief introduction to the Expressnet and various assumptions are included.

Section 2 deals with the data protocols of the Expressnet and provides description on the general transmission mechanism. Simulation results in form of graphs are appended at different places within the section.

Section 3 is devoted to the operation of the voice transmission algorithms. The descriptions of the integrated voice and data protocols occupy Section 4. Section 5 draws a comparison among the various protocols. All these sections observe the format employed in Section 2. A discussion on the
simulation results, future trends, and possible research areas are provided in Section 6, which summarizes the whole thesis. In conjunction with interpretations of results, diagrams are presented wherever the author deems appropriate. A glossary of symbols, a list of illustrations, as well as a bibliography section are also included for the convenience of the reader.

1.3 Integration of Voice and Data in Packet-Switched Networks

1.3.1 Advantages

The recent rapid growth of digital computer communication networks have made it possible to use the existing computer data communication technology for speech communication. Both silent periods and low bit rate of speech motivate the consideration of packetized voice.

A packet-switched integrated digital voice and data communication network system is interesting because it offers the following potential benefits:

- economical because switching and transmission resources are shared
- enhanced services for users who require access to both voice and data communications
- flexible internetworking allows intercommunication among voice users on different types of networks
- efficient channel utilization by transmitting other voice and data packets during silence interval
- allows convenient accommodation of voice terminals with different rates and data formats
- only the required channel capacity is used to transmit rather than a fixed minimum bandwidth increment typically used in circuit-switched networks
- security techniques can be applied to digitized voice
- variable-bit-rate voice transmission techniques are possible, either to reduce average end-to-end bit rate or dynamically adapt voice bit rate to network conditions, and
- flexibility in coping with the changing traffic patterns

1.3.2 Performance Criteria for Data and Voice Traffic

To facilitate design and evaluation of different network configurations, it is essential to define a set of performance criteria which would provide, in a concrete manner, a better view of related issues and possible tradeoffs between resources utilization and performance expectation. Extensive work has been carried out in recent years in this area. [PFM 1]

Gruber [PFM 1] addressed top-down, end-to-end user-oriented performance requirements for voice and digital data services.
According to Gruber, performance criteria can be classified as being either transmission-oriented or service-oriented. Falling under each category are quality, delay, and availability considerations. A subset of the data/voice performance parameters are tabulated in Appendix A.

Basically, evaluation criteria for performance of switches and networks that carry both voice and data are primarily the same as that for separate networks. One difference is that the various performance criteria must be satisfied simultaneously for different voice and data traffic mixtures. Performance criteria include the evaluation of delays, assessment of system traffic handling capacities, throughput, transmission capacity spent on control overhead, and analysis of blocking and expected waiting time.

The performance criteria are applied to each type of traffic. Among the various parameters discussed in Gruber's paper delay and error performance are the most crucial ones to be considered in an integrated voice and data network.

For the voice transmission case, noise has less detrimental effect on voice intelligibility than it does on data. Therefore error criteria for voice need not be as stringent as that for data. However, delay in delivery of real time voice above a certain threshold might cause annoyance from the user's point of view.
On the other hand, data signals require considerably better error performance than voice but are not as concerned about transmission delay.

In view of the above, the problem of integrated network design requires the determination of minimum cost network resources (nodes, links, capacities) which satisfy average end-to-end (ETE) delay for packet-switched traffic and reliability constraints. Thus, the design problem includes the subproblems of capacity assignment, routing, and topological design.

1.3.3 Transmission Procedures of Packetized Voice and Data

Since much work has been done on the transmission of data on Packet Switched (PS) networks, the emphasis of the following discussion will be on voice communication, with occasional reference to data traffic.

The term packetized voice means digitized voice carried in the form of packets over a packet switched communication network. This term usually implies real-time conversations between two persons at geographically dispersed locations.

Major performance criteria as applied to voice are:

- voice quality, including delay performance,
- probability of loss of calls, probability of clipping
  \( [PV \ 1-5, \ VN \ 1-2] \)
- throughput, and
- reliability.
Other aspects in this branch of communication include:

- buffer sizing of a Packet Voice Receiver (PVR) [PVR 1-2]
- optimal packet length
- traffic models (queueing analysis)
- impact of errors, and
- voice storage.

The basic scheme for packet voice transmission includes voice encoding, packetization, transmission, packet reassembly, and buffering followed by decoding and playing back [PVR 2] (Figure 1.3). [MOB 2,3, SW 1, PVR 1, VDN 1]

1.3.4 Channel Access Protocols

With regard to the integration of voice and data in a local area network, the most important area to be considered is the establishment of appropriate channel access protocols or scheduling schemes. This thesis describes a multiplexed traffic from several users and different applications on the same bandwidth-limited channel. In order to achieve a higher utilization, it is necessary that a multi-access scheme be responsive to the particular characteristics of each type of traffic, voice, and data.

For data traffic, much research has been done on multiple access methods and many of the proposed schemes have been
Figure 1.3 Block Diagram of Integrated Voice/Data Packet Transmission System

- **Effects of excessive delay**
  - transit queue trimming

- **Network nodes**
  - Receiver discard packets if "too late"
  - Packet format:
    - at point A: \[ \text{H D} \]
    - at point B: \[ \text{H TS SP} \]

- **Network transient time**
  - 40 ms

- **Processing delay**
  - 135 ms

- **Packet loading time**
  - 150 ms

- **Processing delay**
  - 5 ms

- **Processing delay**
  - 40 ms

- **Receiver queue**
  - D/A CODEC (decoder)

- **Packetizer**
  - SAD

- **A/D CODEC**
  - (vocoder)

- **Filter**

- **Voice Encoding**

- **Packetization**

- **TRANSMITTER TERMINAL**

- **TRANSMISSION**

- **NETWORK with stochastic behaviour (links, queues, switching devices) may include store-and-forward Radio, Cable, SAT**

- **RECEIVER TERMINAL**

- **Processing delay**

- **Packet loading time**

- **Network transient time**

- **Packet buffer delay**

- **S/H: Sample & Hold**
  - SP: Speech
  - SRC: Source
  - Spkr: Speaker
  - H: Header
  - TS: Timestamp
  - D: Data
  - SAT: Satellite
  - SAD: Speech Activity Detector
implemented in actual networks; an example of this is the popular Ethernet, making use of the CSMA/CD protocol.

However, as the applied load increases, or in cases where the propagation delay is relatively large, the performance of conventional schemes proves to be unsatisfactory, especially when voice traffic is involved. As a result, many new protocols have been proposed in an attempt to improve network performance while satisfying traffic requirements of each class. (Figure 1.4)

It is the purpose of this thesis to look into the performance of a new network protocol, the Expressnet, in an integrated voice and data environment.

1.4 Scope of Thesis

1.4.1 Broadcast Systems, Round Robin Schemes, and Expressnet

The communication architecture of interest in this thesis is broadcasting. Broadcast systems can be classified as either bidirectional or unidirectional. In the former case, Bidirectional Broadcast Systems (BBS), such as Ethernet, all the communicating devices are connected to a common cable on which transmission signals propagate in both directions.

In the latter case, commonly referred to as Unidirectional Broadcast Systems (UBS), transmission signals are forced to
Figure 1.4 Block diagram of multiple access schemes
propagate in only one direction of the cable. Examples of protocols used for this configuration are Round Robin (RR) protocols [TH 9] and their variations, such as Expressnet and Fasnet.

The conventional RR schemes allow the users to send only one packet at a time. It has a fixed overhead for each packet. Each subscriber, in a prescribed order, is given a chance to transmit; it transmits if it has a message ready and declines if it has not. This subscriber will not be given a second chance before all other subscribers have had their chance.

The theme of this thesis is the Expressnet, as proposed by Tobagi. [TH 1-4, 7-8]

1.4.2 Basic Assumptions and Performance Measures of Interest

As mentioned earlier in Section 1.3, delay characteristics and error performances are the major issues involved in an integrated voice and data network (LAN or LHN).

This thesis concentrates on the measurement of the former criterion, namely delay characteristics, especially for the voice transmission case, where it plays an important role in network design. Specifically, the relationship between packet discarding percentages and number of stations in the network is investigated. Call blocking probabilities are also examined.
Owing to the fact that the Expressnet configuration is serving a local area network, errors due to imperfections of physical transmission facilities tend to be sufficiently small and can be considered negligible. In other schemes, a possible source of error would arise from collisions of packets. However, the Expressnet makes use of a conflict-free multiple access protocol and hence there is no collision of packets.

Based on these assumptions, the error aspect of V/D transmission is ignored and the main thrust of the thesis rests on the investigation of delay performance of the integrated network.

As for the general environment of the simulation study, it is assumed that the data part of the input load is already in digital form. With regard to voice, the speech input was assumed to have, prior to transmission, been properly filtered, subjected to voice coding processes, and the resulting bits packetized only during talkspurt durations. Receiver buffering operation, decoding functions, and playback schemes have already been implemented for delivery of speech to the listener. The task at hand is to specify a network protocol that can accommodate data traffic on top of voice transmission on the same network, without violating the basic voice delay requirement.

The goals of the thesis are to examine the channel utilization, delay performance, mean queue length, and packet discard rate of the Expressnet—a high-speed local area network
that is deemed appropriate for integrated voice and data transmission applications.
2. TRANSMISSION OF DATA PACKETS IN EXPRESSNET

2.1 General Description of Expressnet

Expressnet is a high-performance local area communication network comprising of an outbound channel and an inbound channel on which a plurality of stations are connected. (Figure 2.1)

Stations transmit on the outbound channel and receive on the inbound channel. To achieve broadcast communication, the inbound channel is linked to the outbound channel via a connection cable so that all signals transmitted on the outbound channel are duplicated on the inbound channel. In order to transmit on the bus, stations utilize a distributed access protocol. They make use of a conflict-free round robin scheduling scheme.

Transmitters in the outbound channel are of the unidirectional type, rendering the Expressnet a unidirectional broadcast bus architecture (UBS). The communication medium may be a twisted pair, a coaxial cable, an optical fibre, or a waveguide.

This channel access protocol yields higher transmission efficiency than conventional Round Robin (RR) schemes even when the end-to-end propagation delay is a significant fraction of, or even larger than, the transmission time of a packet.

The major strength of Expressnet lies in its ability to provide an upper bound on the network transit delay experienced
Figure 2.1 Block diagram of Expressnet

TX: TRANSMIT
RX: RECEIVE
by a potential packet. This outstanding feature makes the Expressnet an attractive system to be considered for integration of voice and data in a local area office environment.

Detailed descriptions of the transmission mechanism are found in Tobagi. [TH 1-4] A highlight of the major procedures used in basic packet transmission can be found in Appendix B.

2.2 Assumptions Used in Simulation Model for Data Protocol

The following are assumptions made in the actual simulations:

(a) Non-gated sequential service discipline (NGSS) is the version to be implemented. (Figure 2.2)

(b) The error aspect is not covered in this thesis.

(c) In the first part, each station possesses a single buffer for the purpose of packet storage. The results obtained are later compared to and verified by literature values. Results on the performance of the infinite-buffer case are also obtained.

(d) The duration of the locomotive is very small and is taken to be 1% of the full packet transmission time T, i.e., T_{locom} = 0.01 T.

(e) The packet or the transmission unit is of fixed length.

(f) The preamble, assumed to be much smaller than the information section of the packet to be transmitted, is
Packets arrivals are modelled by a Poisson process, with $\lambda$ denoting the arrival rate of packets and IAT denoting the interarrival time of packets.

![Diagram](image)

Arrivals intermingle together; Arrival of one packet in 1 station triggers next arrival in same stn

Figure 2.2 Transmission procedure of Expressnet
The carrier detector is placed very close to the channel. Therefore, the time $t_d$ required for detection of carrier is infinitesimal.

Poisson distribution of arrival of data packets, i.e., interarrival times of data packets, are exponentially distributed.

The medium of communication is the coaxial cable, with the velocity of propagation of signal taken to be $2/3$ of the velocity of light in free space.

2.3 Summary of Results

2.3.1 Analysis of Results for Single Buffer Case

First of all, several parameters are defined and assigned specific values so that the results can be compared with figures obtained from literature.

The quantity "$a$", which is the ratio between the propagation delay and the packet transmission time $T$, is set to be unity, i.e.,

$$a = \frac{\tau}{T} = \frac{L_c / v}{B / C} = \frac{L_c}{v} \cdot \frac{C}{B}$$
Length of cable (256 km) / Velocity of propagation of signal in coaxial cable (2 \times 10^8 \text{ m/s})

Number of bits in one packet / Channel capacity (1 Mbps)

= 1.

Utilization

Initially, it is assumed that there is a total of 20 stations along the outbound cable, yielding a maximum possible channel utilization factor of 0.9086, as shown by the following calculation:

$$\rho_{\text{max}} = U_{\text{max}} = \frac{(N) \times (T)}{(N) \times (T + t_d) + (\tau + \tau + t_d) + T_{\text{locom}}}$$

Total time required for transmission of all packets

$$= \frac{(20) \times 1}{(20) \times (1 + 0.0) + (1.0+1.0+0.0) + 0.01}$$

= 0.9086

The average network throughput (utilization) for a given value of data arrival rate $\lambda$ is given by

$$\bar{\rho} = \bar{U}_{\text{t}} = \frac{\text{time used for transmitting information}}{\text{length of train}}$$

$$= \frac{nT}{nX + Y}$$

where $X$ is the transmission period of a packet, including overhead, preamble and data transmission time; $Y$ is the interround overhead.
where $\bar{n} = \sum_{n=0}^{N} n p_n$

$p_n$ = probability that there are $n$ packet transmissions in a round

From Figure 2.3 it can be inferred that as the inter-arrival time increases, the utilization gradually drops. This can be accounted for by the fact that as the number of transmissions in a train decreases, the effect of the propagation delay becomes more prominent, thus decreasing the overall effectiveness of the system.

It can also be seen that as $\lambda \to \infty$, $\bar{n} \to N$ and the throughput reaches a maximum given by $NT/(NX + Y)$.

To ensure stable network operation, the normalized instantaneous system load $N\lambda$ must be maintained below unity. To that effect, the arrival rate $\lambda$ is restricted in order to avoid system overload due to excessive outstanding packets at any one time.

The constraint is determined as follows:

$$\rho = N\lambda < \frac{NT}{N(T + t_d) + (\tau + \tau_c + t_d) + T\text{loc}}$$

$$0 < \lambda_{\text{max}} < \frac{T}{N(T + t_d) + (\tau + \tau_c + t_d) + T\text{loc}} \text{ packets/sec}$$
Figure 2.3 Utl vs IAT, N = 20

Figure 2.4 Mean queue length vs Utl, N = 20

(description on p.26)
Therefore, the interarrival time (IAT) has to be greater
than, or at least equal to, the reciprocal of $\lambda_{\text{max}}$. 

Delay Characteristics

The expected waiting time and variance of the waiting time
are given by

$$\overline{W} = E[w] = \frac{1}{n} \sum_{n=1}^{N} n \, p_n \frac{e^{\lambda[(n-1)X+Y]} - 1}{\lambda}$$

$$\sigma_{w}^2 = \text{Var}_w = E[w^2] - (E[w])^2$$

where

$$E[w^2] = \frac{1}{n} \sum_{n=1}^{N} n \, p_n \frac{e^{\lambda[(n-1)X+Y]} \left[e^{\lambda(n-1)X+Y-2/\lambda} + \frac{2}{\lambda^2}\right]}{e^{\lambda[(n-1)X+Y]} - 1}$$

Upon examining Figure 2.5, one can observe that the average
packet delay increases as utilization increases and is bounded
from above by the finite value attained at saturation, i.e.,
when $\lambda \to \infty$. This maximum delay is $N(T + t_d) + \tau + t_c + t_d + T_{\text{locom}} \sim (N + 2a)T$ slots.

As for the variance of the delay (Figure 2.6), it can be
seen that as $\lambda \to \infty$, the variance approaches zero. In each
round, each station will have a packet to send, rendering each
train a full cycle, consisting of transmissions from each and
**Figure 2.5** Mean wait time vs Utl, N=20

**Figure 2.6** $\sigma_w^2$ vs Utl, N=20

SIM : simulation
TH : theoretical
every station. In other words, all rounds are of full length and the packet delay becomes deterministic and equals to \(N(T + t_d) + \text{propagation delay}\). Thus we have a case in which the variance incurred in the network transmission is highest for throughput close to network capacity, while at network capacity the variance assumes the value zero. The simulation results resemble those from literature through formula evaluation.

**Mean Queue Length**

Figure 2.4 shows the relationship between the mean queue length of the system and the channel utilization. \(\overline{Q}_1\) is found to be approximately one less than \(\overline{Q}_2\). The former records the number of users in the queue awaiting transmission, while the latter maintains an account of the total number of users in the system, including the one currently involved in the transmission process. It can be seen that as \(\lambda\) increases the mean queue length \(\overline{Q}_2\) approaches \((N + t_d) + \tau + \tau_c + t_d + T_{\text{locom}} = (N + 2a)T\) slots, which is the same as \(\overline{W}\).

**Variation of Number of Stations at Fixed Interarrival Time**

As the number of stations gradually increases, the utilization increases until the system reaches saturation. (Figure 2.7) The mean waiting time, variance of the waiting time, and the mean queue length are found to be increasing, as the number of stations \((N)\) increases.
**Figure 2.7** Utl vs N, IAT = 12.5

**Figure 2.8** Q vs Utl, IAT = 12.5
The slope of the curve becomes increasingly steeper as the channel utilization becomes higher. (Figures 2.8-2.10)

Variation of Interarrival Time

Figures 2.11-2.14 demonstrate that a lowering in interarrival times will bring about increases in channel utilization, mean, and variance of waiting times, as well as the mean queue length.

Flowcharts for data protocol can be found in Appendix C.
Figure 2.9 $\bar{W}$ vs Utl, IAT = 12.5

Figure 2.10 $\sigma^2$ vs Utl, IAT = 12.5
Figure 2.11 $\text{Utl vs N, IAT = 12, 12.5}$

Figure 2.12 $\overline{Q}$ vs Utl, IAT = 12, 12.5
Channel Utilization (Utl)

Figure 2.13 $\overline{W}$ vs Utl, IAT = 12, 12.5

Variance of Wait Time ($\sigma_w^2$)

Figure 2.14 $\sigma_w^2$ vs Utl, IAT=12, 12.5
2.3.2 **Analysis of Results for Infinite Buffer Case**

In this case, the buffers in each station have infinite capacity and are thus capable of storing all data packets which arrive at the station.

The mean interarrival time of packets (of size 1280 bits) is assumed to be 28 slots (i.e., $28 \times \frac{1280}{1000000}$) or 0.035 seconds. The idea of using this figure is to examine the performance of the network under moderate load.

For this value of the interarrival rate, it was found that with 22 stations, the channel utilization approached 0.8, whereas the utilization figure declined to about 0.65 when the number of stations was only about 18. (Figure 2.15)

The mean queue length ranged from about 4 to about 8 as the number of stations grew from about 17 to 22. This indicates that more packets were awaiting transmission when the buffers are of infinite capacity. (Figure 2.16)

The mean and variance of the waiting time rise accordingly with increase in number of stations in the system. Once again, the mean waiting time bears a close relationship to the mean queue length which indicates the number of outstanding packets awaiting transmission. (Figure 2.17)

The variance of the waiting time displays a steep slope, reflecting the fact that wait duration becomes more unpredictable for increasing number of stations. (Figure 2.18)
It can be expected that a buffering of the various packets leads to increased waiting time for individual packets since a packet in a station has to wait for transmission of all the preceding packets before it can have a chance to transmit. Thus, as the number of stations or the load increases, the number of outstanding packets will increase quickly and flood the network.
Figure 2.15 Utl vs N, IAT = 28

Figure 2.16 Q vs N, IAT = 28
Figure 2.17 $\bar{W}$ vs $N$, IAT = 28

Figure 2.18 $\sigma^2$ vs $N$, IAT = 28
3. TRANSMISSION OF VOICE PACKETS IN EXPRESSNET

3.1 Voice Protocol

3.1.1 General Requirement of Voice Traffic

In contrast to data traffic, where delay in transmission is not a crucial design factor, the expected message delay is the most important criterion to be considered in voice communication since it affects the quality of real-time interactive speech significantly.

The total end-to-end (ETE) delay is measured from the instant a speech parcel is about to be generated to the moment when it is played to the listener at the receiver end. The ETE delay is composed of the following components:

1. originating terminal delay, made up of processing and packetization delays;
2. network delay, comprising of nodal queueing delay, switching delay, as well as transition time on the internodal links which includes transmission time and propagation delays;
3. delay at the receiving end, depending on the sort of playback schemes employed; and
4. other processing delays. \([PV, VN, PVR]\)

The human listener in a conversation has limited tolerance to the average delay and the fluctuation of delay. Subjective
evaluations indicated that an end-to-end delay of up to 300 ms is quite acceptable in the sense that the quality degradations thus incurred are hardly noticeable, while a higher delay figure can cause annoyance on the part of the listener.

The delay variance is also of importance because the receiver node has to provide a continuous stream of bits to the vocoder. That is, the variations in the packet arrival times have to be smoothed over by the receiver. This is done by delaying the packets in the receiver's buffer. If the receiver completes playing a packet to the user while the succeeding packet has not arrived yet, the receiver is forced to emit another signal, such as noise or silence. This, of course, causes degradations of speech quality. The smaller the variation of delay, the better the receiver buffer can be managed to provide the vocoder with the desired continuity.

A packet with too much delay has missed the playback time and is effectively lost. Thus, it is common measure to discard packets which would incur excessive delays. Anomalies less than 50 ms in duration tend to go unnoticed by the human hearing process. In fact, it was recommended that packet length of 10-50 ms of speech intelligibility could only be minimally affected by lost packets. There is an important tradeoff between network delay performance and lost voice packet rate. To maintain good speech quality, it is necessary to keep the packet discard rate below 1 to 2%. [CALL 2]
3.1.2 *Arrival Pattern of Voice Packets and Transmission Mechanism*

A popular model for voice traffic analysis is the Markovian model of interactive telephone speech comprising of alternate talkspurt and silence periods, the mean of the talkspurt duration making up about 40% of the conversation length. [MUX 1, SPCH 2]

In packetized voice networks, intelligible speech is transmitted at high rates, ranging from 1000 bps to 64 kbps. The latter figure is obtained from digitizing ordinary telephone speech (with bandwidth of approximately 4 KHz) using a conventional 8-bit PCM scheme at a sampling frequency of 8 KHz. The high rate voice yields high speech quality in that it is less susceptible to environmental degradations because of the high level of redundancy inherent in the coded speech.

For the voice transmission section of the thesis, the performance of Expressnet is evaluated for cases with and without the discarding of packets. Instead of an exponential distribution of interarrival times for all packets, as in the Data transmission case, the situation here is to be considered from two points of view, the global view and the local standpoint (i.e., at each station). From a global viewpoint, when a sufficiently large number of voice sources (> 10) are multiplexed, the distribution of interarrival times of talkspurt or silence periods closely resemble that of an exponential distribution.
However, if one looks at a particular station, then the arrival of packets becomes a deterministic process. New voice packets are generated by a terminal at regular intervals. The interval equals the packet length as long as a talkspurt is in progress.

The transmission mechanism investigated is the non-gated mode of transmission of the Expressnet.

3.2 Assumptions Used in Simulation Model for Voice Protocol

Basically, the assumptions in the Data section are still valid as far as the transmission mechanism is concerned. However, a few items pertaining to the characteristics of voice traffic are to be added to the list:

(a) Real-time speech (as in telephone conversation) is the voice type of traffic of interest.

(b) It is assumed that vocoders digitize voice at some constant rate of \( C_v \), where \( C_v \) is set to be 64 Kbps.

(c) Speech detection process had already been carried out and that during talkspurt periods, bits are grouped to form packets of size \( B_v \), which are to be transmitted in the network to the destination vocoder, as depicted in Figure 3.1.

(d) Figure 3.2 describes the assumption on the approximate maximum delay requirement for the various components of the total end-to-end delay.
$t_s$: speech detection time
$t_e$: speech detection hangover time

Figure 3.1 Packet speech model [MOB 3]

Figure 3.2 Maximum ETE delay requirement
The objective of the voice section is to devise a protocol for voice communication so that the transmission time of individual packets does not exceed 20 ms.

(e) Statistical independence between talkspurts is guaranteed. Durations of talkspurt and silence periods are exponentially distributed while the means of their distributions are 1.4 and 2.0 seconds respectively. These figures are taken from results by Brady [SPCH 3] for male speakers carrying on casual conversation while the speech activity detector (SAD) is operating. Talkspurt durations are thus about 40% of the talker's activity.

(f) Packets are generated only during talkspurt intervals, as shown in Figure 3.3.

Two major constraints have to be satisfied:

(i) The network transit time has to be smaller than or equal to the packet formation time so that the system is not overloaded. Therefore, it is required that $T_f \geq D_n$.

(ii) The sum of the packetization time and the network transit time cannot exceed the maximum tolerable delay of a bit, i.e., $T_f + D_n \leq D_v$.

Thus $N_{\text{max}}$ occurs when $T_f = D_n = D_v/2$, while the optimum $B_v(2)$ is given by $(D_v/2) \times C_v$. According to the Expressnet

* length of voice packet carrying actual information bits
Packet Arrivals

station 1

packet generation time:
20ms = 1280 bits/packet at \( C_v = 64 \text{kbps} \)

mean of talkspurt duration = 1.4 second
mean of silence duration = 2.0 second

station 2

station 3

SILENCE

station 4

Transmission Channel

1280 bits
\( \frac{\text{bits}}{\text{bps}} \) = \( \frac{1280}{c} \) seconds

periodic arrivals of voice packets within a talkspurt
non-gated service discipline

\( t + t_c + t_s \)

REALSTART

Tlocom

Figure 3.3 Voice transmission in Expressnet
protocol, $D_n$ has a value of $N(T_v^1 + t_d) + (\tau_c + \tau_c + t_d) + Tlocom$. Ignoring $t_d$ and $Tlocom$, $N_{\text{max}}$ can be formulated as $N_{\text{max}} = (D_v \times C)/(2B_v (1)^2 + D_v C_v)$, where $N_{\text{max}}$ is the maximum number of voice users supportable such that the longest delay experienced by each bit is guaranteed to be less than $D_n$. If the overhead bits are also ignored, then the expression reduces to $N_{\text{max}} = C/C_v$.

For our purpose, it is assumed that the packetization time is 20 ms. The source codes the bits at a rate of 64 kbps, i.e., the number of bits per packet is $B_v = T_f \times C_v = 20 \text{ ms} \times 64 \text{ kbps} = 1280$ bits.

In summary, $D_n$ has to satisfy the following relationship:

$$D_n = N_{\text{max}}(T_v + t_d) + (\tau_c + \tau_c + t_d) + Tlocom$$

$$= N_{\text{max}}(B_v/C + t_d) + (\tau_c + \tau_c + t_d) + Tlocom$$

$$\leq T_f = 20 \text{ ms}$$

3.3 Analysis of Results

A major quantity involved in the analysis is the number of stations that can be accommodated such that the maximum tolerable delay is not exceeded at any time.

The determination of this value is exemplified by the following calculation: (p.44)

1 : $T_v$ is the transmission of a packet on the channel
2 : $B_v^{(1)}$ is the length of packet occupied by overhead
\[ N_t = \frac{D_n}{(B_V/C)} \]

\[ = \frac{20\text{ ms}/(1280\text{ bits}/1\text{ Mbps})}{15.625} \]

This is also the figure for the full packet generation time "pgtfull" (in slots), since the packet transmission time \( T \) is normalized to be one. We are interested in finding out how many of these \( T \)'s can be packed into one packet generation time interval, hence the determination of \( N_t \). The key point in this operation is the fact that the actual channel capacity is much higher than the vocoder rate. An example of the system would be a 1 Mbps channel supporting transmissions from numerous vocoders operating at 64 kbps. The idea is that as long as \( N \) is less than or equal to \( N_t \), it is guaranteed that the \( D_n \) constraint is not violated.

To move one step further, one can exploit the fact that in a casual conversation, a speaker is active for merely 40% of the time. By sacrificing a tiny fraction of the full talkspurt, various other stations can be attached to the system, with minor degradation of speech quality. This can be implemented as discarding of packets. The objective is to determine the maximum number of stations that can be supported by the system such that the packet discard rate is within tolerable limits. As mentioned earlier, it was reported that a
packet loss percentage of 1 to 2% is quite acceptable for voice communications [CALL 2].

The rest of this section is organized into two subsections describing performance of the two cases, namely:

(V1) Transmission Without Discarding of Packets, and
(V2) Transmission With Discarding of Packets if the potential delays that would be incurred are larger than the time required to generate a packet.

(V1) Transmission Without Discarding of Packets

The channel utilization can be observed increasing rapidly as the number of stations increases (Figure 3.4). As the system approaches saturation, this increase levels off and finally settles to its steady state value.

The means of the delays (Figure 3.6) also manifest similar behaviours as those observed in the Data case. One can see that an increase in channel utilization in turn increases the average delay. However, there exists an upper bound for the delay, the expression of which is \((N + t_d)T + (\tau_c + \tau_c + t_d) + T_{locom} or (N + 2a)T\) slots.

It can be inferred that an increase in channel utilization causes a corresponding increase in the variance of the delay, with the curve becoming sharper as the utilization approaches unity (Figure 3.7).

The mean queue length (Figure 3.8) is found to closely match that of the mean waiting time and the mean number in the system is one more than that of the mean queue length.
Figure 3.4 $U_{tl}$ vs $N$, V1/V2

Figure 3.5 $\bar{Q}$ vs $U_{tl}$, V1/V2
Figure 3.6 Mean wait time vs Utl, V1/V2

Figure 3.7 $\sigma^2_w$ vs Utl, V1/V2
Figure 3.8 Mean queue length vs Utl, V1

Figure 3.9 $Q$ vs Utl, V2
As for the mean train size, the results in Figure 3.10 show that as the channel utilization exceeds 0.87, the mean train size is approximately given by $T_z = 0.4 N$. This is reasonable because talkspurts occupy only about 40% of a person's speech in a normal telephone conversation.

The variance of the train size can be observed increasing as the system becomes more heavily loaded. However, at a certain point, the curve abruptly declines from its climax and falls off quite rapidly. This is because as the utilization approaches unity, every station is transmitting a packet in each round (train) and the train size becomes more predictable. (Figure 3.11)

If 15 or more stations with lower indices have packets to be placed on the channel for delivery to the destination vocoder, some stations may have to wait for more than 20 ms (the maximum tolerable delay) for a chance to transmit. This hypothesis is substantiated by the figures found in the graphs plotting the relationship between percentage of delayed packets (PDLP) and the number of stations ($N$) as well as the channel utilization (Figure 3.12 and 3.13 respectively). One can see that PDLP increases from zero after the number of stations exceeds 15.

(V2) Transmission With Discarding of Packets

In this case, packets are discarded if the delay is larger than $T_f = 20$ ms. The channel utilization is found to be
Figure 3.10 $\bar{T}_z$ vs Utl, V1/V2

Figure 3.11 $\sigma_{T_z}^2$ vs Utl, V1/V2
Figure 3.12 PDLP vs Utl, V1/V2

Channel Utilization (Utl)

Figure 3.13 PDLP vs N, V1/V2

Number of Stations (N)
slightly lower than that in the V1 case. (Figure 3.4)

The means of the waiting times are smaller than those in the V1 case. (Figure 3.6) As for the variance of the delay, it is observed to be comparable to that of the V1 case at low utilization. (Figure 3.7) When more stations are attached to the system, some packets may be discarded and hence the delays are more predictable, i.e., less variable. Similar arguments hold for the mean queue lengths. (Figure 3.5 and 3.9)

The mean train size is slightly less than that in the V1 case. (Figure 3.10) The variance of the train size increases as the channel utilization increases. (Figure 3.11)

In fact, the protocol can be modelled by a 1-D Markov chain. (Figure 3.14) Talkspurt and silence durations are assumed to be exponential random variables. The rate of arrival of talkspurts is proportional to N-k, which is the number of offhook users not currently in talkspurt; while the rate of departure of talkspurt is proportional to k, the current number of active talkspurts.

\[(N-k+1)\lambda_t \approx (N-k)\lambda_t \]

\[k\mu_t \approx (k+1)\mu_t\]

\[\lambda_t = \text{Mean silence duration}\]
\[\mu_t = \text{Mean talkspurt duration}\]

Figure 3.14 Markov chain for V2 protocol [MUX 1]
(Talkspurt/silence model)
The model for an active voice source is illustrated in Figure 3.15. \( P_{on} \) and \( P_{off} \) are transition probabilities in the Markov transition matrix.

\[
P_{off} \quad 1 - P_{off} \quad P_{on}
\]

\[
1 - P_{on}
\]

**Figure 3.15 Model for an active voice source [CALL 1]**

The packet discarding rate (VPDR) is approximately given by the equation below, whose form bears a striking resemblance to the fractional speech loss equation of a Time Assignment Speech Interpolation (TASI) system. [SPCH 1-2]

\[
VPDR = \frac{\sum_{k=N_{t}+1}^{N} \left[ \binom{N}{k} P_{on}^k P_{off}^{N-k} \right] \left[ k - N_{t} \right]}{\sum_{k=0}^{N} \left[ \binom{N}{k} P_{on}^k P_{off}^{N-k} \right]}
\]

\[
= \frac{1}{N P_{on}} \sum_{k=N_{t}+1}^{N} \left[ \binom{N}{k - N_{t}} P_{on}^k P_{off}^{N-k} \right]
\]

\[
P \left[ \text{total # of packet arrivals} = k \right] = f \left( k, N, P_{on} \right)
\]
Results for this part is shown in Figures 3.16 and 3.17.

The packet discarding rate is found to increase slowly in step with increase in the number of stations in the system. This can be observed until \( N = 28 \), beyond which VPDR increases more rapidly, hinting excessive load on the system. It can be observed that for a 1% packet discard rate, the number of stations that can be supported is about 28. The discard rate becomes 2% when 31 stations are attached to the system. Flow charts for voice protocols can be found in Appendix C.

3.4 Call Estimation Implementation

3.4.1 Introduction

For interactive telephone conversations, it can be assumed that each telephone offers 0.25 Erlangs of traffic [CALL 2] in peak office hours.

In the model used, it is assumed that there is a call every 12 minutes (i.e., the arrival rate of call \( \lambda_c = 1 \text{ call/12 minutes} \)). The call arrivals are governed by an exponential distribution [SW 2]. Furthermore, the duration of calls is exponentially distributed with mean 3 minutes.

\[
\lambda_c = \frac{1}{\text{IAT}_{\text{CALL}}} = \frac{1}{12} = \frac{1}{4} = 0.25 \text{ Erlangs}
\]

\[
\mu_c = \frac{1}{\text{DUR}_{\text{CALL}}} = \frac{1}{3}
\]

\[
\rho = 0.25 \text{ for each station}
\]
Figure 3.16 VPDR vs Utl, V1/V2

Channel Utilization (Utl)

Figure 3.17 VPDR vs N, V1/V2

Number Of Stations (N)
The situation is depicted in Figure 3.18.

The distributed call-blocking mechanism described in (i) works as follows. A call which arrives at a station is blocked with a probability of \((1-p)\), where \(p\) is given by
\[ p = \min \left( \gamma \cdot \frac{\text{pgtfull}}{\left[ \frac{\tilde{N}_j}{(\text{DURCALL} \cdot \lambda_c)} \right]} \right), 1 \]

where

- \( \gamma \) is a constant
- \( \text{pgtfull} \) is the maximum number of stations allowed to guarantee zero packet discard and call blocking rates
- \( \tilde{N}_j \) is given by \( \tilde{N}_j = (1-\alpha)\tilde{N}_{j-1} + \alpha L_j/0.4; L_j = \text{actual length of previous train} \)
- \( \tilde{N}_j/(\text{DURCALL} \cdot \lambda_c) \) is the estimated number of active voice calls

The purpose of this section is to implement a call estimation scheme into the voice transmission aspect of Expressnet. Channel utilization, mean and variance of waiting times, call blocking probability, and packet discard rate are performance measures of interest. The idea is to devise a dynamic control scheme to regulate the amount of traffic in the system by making use of several parameters (number of stations, size of previous train, etc.) to achieve a desired level of performance.
3.4.3 Analysis of Results

The simulation time for the program is 6 hours, with $\gamma$ set to be 4.02. Three different values of $\alpha$, namely $\alpha = 0.1, 0.5,$ and $0.9$ are used in the simulation. The effectiveness of the individual cases are demonstrated in Figures 3.19, 3.20, and 3.21 respectively. They provide snapshots of the entire simulation process.

It can be seen from Figure 3.19 that in the $\alpha = 0.1$ curve the estimated number of active stations matches closely with the actual value throughout the run. Deviations from the actual values are about $\pm 1$, as can be observed from the diagram. Moreover, the transition from 16 to 17 is a smooth phenomenon.

However, for $\alpha = 0.9$, the curve produces a faithful replica for actual values of 10, but suffers relatively severe fluctuations at the transition of number of active stations from 10 to 11.

The set of points on the curve $\alpha = 0.5$ exhibited characteristics which lie between the $\alpha = 0.1$ and $\alpha = 0.9$ cases.

Hence, $\alpha = 0.1$ is the more favourable estimator for the number of active stations. This indicates that smaller $\alpha$ is preferable in satisfying the criteria of estimator precision.

In all the three cases mentioned, the estimated values seemed to be higher than the actual ones on average. This in

* $\gamma = 4.02$ : this one was chosen among others because it yields satisfactory performance
Figure 3.19 Actual vs estimated # active stations, $\alpha = 0.1$

Figure 3.20 Actual vs estimated # active stations, $\alpha = 0.5$
Figure 3.21 Actual vs estimated # active stations, $\alpha = 0.9$

Figure 3.22 Utl vs N, theoretical /simulation
turn implies that the percentage of voice calls blocked is slightly higher than necessary. If $\Psi$ can be adjusted to a lower value, one can expect the estimated values to come closely to the actual figures, with the outcome of even better performance.

Figure 3.22 serves to verify that the program is performing its job. It shows that the simulated utilization values lie in close proximity to the predicted ones, according to the formula:

$$Utl = \frac{DUR_{\text{CALL}}}{IAT_{\text{CALL}}} \ast \frac{N}{\text{pgtfull}} \ast \frac{\text{talkspurt duration}}{\text{talkspurt duration} + \text{silence duration}}$$

with the assumption that the blocking probability is zero.

Figures 3.23, 3.24, and 3.25 illustrate that $f$, $w$, and $\sigma_w^2$ increase at steady rates as $N$ increases, with the $\alpha = 0.9$ curve yielding highest values for all $N$. This is not unexpected because greater values of $\alpha$ display less prediction accuracy. Figure 3.24 also confirms the validity of the program for it shows that the mean waiting time exceeds the mean queue length by approximately one.

The behaviours of Packet Discarding Rate (PkDis) and Blocking Probability (PB) can be observed in Figure 3.26. Figure 3.27 and Figure 3.28 in turn examine their tradeoffs. In general, both measures amounted to more than 1% for $N = 75$ or higher. The $\alpha = 0.9$ case again produces larger values for
PB. This phenomenon can be explained by the fluctuations in estimated values.

For satisfactory channel utilization performance, subjected to the stringent packet discarding and call blocking constraints, an optimum range of 65 to 70 stations is recommended.
Figure 3.23 Util vs N, different α

Figure 3.24 Mean wait time, mean queue length vs N
Figure 3.25 Variance of wait time vs N

Figure 3.26 % packet discarded, blocking probability vs N
Figure 3.27 % packet discarded vs blocking probability %

Figure 3.28 % packet discarded vs blocking probability %
4. INTEGRATION OF VOICE AND DATA IN EXPRESSNET

4.1 Introduction

In recent years much research has been carried out to investigate the performance of integrated voice/data transmission systems. [VDN 1, CALL 1]

The major constraints that have to be satisfied include: i) delay constraints for voice packets, and ii) minimum bandwidth requirement for data. It is desirable to accomplish a network that can provide the bandwidth reserved for data on continuous basis. It is also required that the protocols be dynamic in allocating bandwidth to voice and data applications.

The main purpose of this section is to devise a local area network protocol. Such a protocol must be able to support a certain number of voice and data terminals, and yield reasonable channel utilization factor, while observing the delay constraint of voice traffic.

4.2 Features of the Proposed Network Configuration

The configuration under investigation is the Expressnet broadcast bus topology. It is assumed that there is a number of voice and data stations, alternating in position, along the full length of the coaxial cable, as shown in Figure 4.1.
The program models burst of interactive data arrivals as a Poisson process, while providing exponentially distributed talkspurt and silence durations for real time telephone conversations on the part of the voice users. The voice input is assumed to be coming from a single call.

A typical train consists of a mixture of voice and data packets. It is further assumed that the size of a data packet is identical to that of a full-size voice packet, i.e. 1280 bits.
4.3 Proposed Packet Transmission Protocols

Three protocols are investigated in this section of the thesis: VD1, VD2, and VD3.

Transmit-when-Ready (TWR) Protocol (VD1)

Regardless of the type of terminals it supports, if a station has a packet to send, it transmits the packet wherever its turn comes. Such transmission occurs irrespective of the potential delay that a packet may incur during the transmission. The size of the buffer of each station is assumed to be one.

Transmit-Data on a Discard Basis (TDS) Protocol (VD2)

In this case, a voice station transmits with probability $p$ whenever its turn comes. For a data station the corresponding transmission probability is $q$, where $p$ and $q$ are determined by expressions given below. A voice or data station discards its packet with probability $1-p$ and $1-q$ respectively. Each station is assumed to have a single buffer (of size 1) to hold a packet which is awaiting transmission.

Transmit-Data on a Delay Basis (TDL) Protocol (VD3)

As in TDS, each station computes the appropriate transmission probability and transmits accordingly. A voice station discards with probability $1-p$. However, the buffer size of a
data station is assumed to be infinite. Thus, data transmission can be realized on a delay basis. This implies that in case a data packet is prohibited from transmitting, it is placed as the first element in the queue for the station as the next outgoing packet. Future arrivals to the same station have to join the queue and wait in line.

The determination of the values of $p$ and $q$ is based on a channel utilization estimation scheme which consists of measuring the length of the previous train and the aggregate number of voice and data transmissions in the previous round. With this estimation mechanism, the values of $p$ and $q$ are given by:

$$
p_j = \min \left( \varphi \cdot \frac{\text{LT}_{j-1} + \text{NS}_{\max}}{\text{LT}_{j-1} \cdot \text{max} \cdot \text{max}}, 1 \right)
$$

$$
q_j = \min \left( \varphi \cdot \frac{\text{LT}_{j-1} + \text{NS}_{\max}}{\text{LT}_{j-1} \cdot \text{max} \cdot \text{max}}, 1 \right)
$$

where

$p_j$ = probability of transmitting a voice packet in the $j^{th}$ train

$q_j$ = probability of transmitting a data packet in the $j^{th}$ train

$\varphi$ = Overall Scale Factor
\[ W_{VV} = \text{Voice Factor (set to 1.0)} \]

\[ W_{dd_j} = \text{Data Factor in } j^{\text{th}} \text{ train} \]

\[ = \min \left( \frac{\text{NS}_{\text{max}}}{m_{\text{max}}} + K_{\text{data}} \times \text{NS}_{\text{max}}}{(Nv_{j-1} + N_{d_{j-1}}) + K_{\text{data}} \times \text{NS}_{\text{max}}} , 1 \right) \]

\[ K_{\text{data}} = \text{Data Weighing Factor} \]

\[ BF = \text{Basefactor (prevent fraction from going to infinity in case the previous train consists only of locomotive)} \]

\[ L_{\text{max}} = \text{Maximum Length of Train (excluding propagation delay)} \]

\[ LT_{j-1} = \text{Length of the } (j-1)^{\text{th}} \text{ Train, including propagation delay} \]

\[ \text{NS}_{\text{max}} = \text{Maximum Number of Stations connected to the cable such that the delay incurred in transmission of any packet is guaranteed to be less than 20 ms} \]

\[ Nv_{j-1} = \text{Total Number of Voice Transmissions in the } (j-1)^{\text{th}} \text{ Train} \]

\[ N_{d_{j-1}} = \text{Total Number of Data Transmissions in the } (j-1)^{\text{th}} \text{ Train} \]

With this mechanism, if the previous train is short, voice and data stations transmit with high probability. If the length of the previous train is long, transmission from voice and data stations are discouraged. Data stations are granted lower priority than voice stations in that \( W_{dd} \) never exceeds \( W_{VV} \).

Advantages of this network protocol and the configuration are:

a. Simplicity in implementation:
   - sophisticated central controller is not required
   - identical procedures of transmission in each train for each station
b. High channel utilization:

- Train type indicator, used if different types of trains are employed by the system, is not required. Thus overhead due to different train types is eliminated.
- The larger the number of stations connected to the system, the smaller the effect of the propagation delay.

c. Voice delay requirement satisfied:

- By discarding small amount of packets, train size prevented from being excessively long;
- By utilizing the round robin style of transmission, a maximum value of the delay is guaranteed;
- By restricting total number of stations connected to the system.

d. Priority system possible:

- Priority in transmission can be established, data transmission penalized by appropriate choice of weighing factor.

e. Fairness of channel usage:

- Equal share of bandwidth among users of the same class (namely voice or data users).

4.4 Summary of Results

This section consists of three sub-sections, namely the VD1, VD2, and VD3 sections. Their purpose is to describe the
performances of the protocols proposed in the earlier parts of Section 4.

4.4.1 Analysis of Results for VD1

For a fixed interarrival time (Figure 4.2), the channel utilization is found to increase with increasing number of stations connected to the Expressnet broadcast bus. The mean queue length is again one less than the total number of packets in the system. (Figure 4.3) The mean waiting time of data and voice packets are almost identical. (Figure 4.4) However, the variance of the delay encountered in data transmission resembles that of voice traffic only up to a utilization factor of 0.87, beyond which point the variance of data delay is significantly larger. (Figure 4.5) The mean train size (Figure 4.6) increases as the utilization increases and the variance of the train size (Figure 4.7) drops rapidly from its climax at Utl = 0.76. The percentage of voice packets (PDLP) delayed more than 20 ms is very small for Utl less than 0.8, corresponding to a system size of approximately 26 stations. (Figure 4.8, Figure 4.9)

For a fixed number of stations (in this case set to be 14), the channel utilization decreases as the interarrival times increases. (Figure 4.10) For an interarrival time of about 12 slots, the mean queue length is approximately 3. (Figure 4.11) The mean waiting time is found to drop as the
Figure 4.2 Util vs N

Channel Utilization (Util) vs Number of Stations (N)

Figure 4.3 Mean queue length vs Util

Mean Queue Length ($\bar{Q}$) vs Channel Utilization (Util)
Figure 4.4 Mean wait time vs Utl

Figure 4.5 Var of wait time vs Utl
Figure 4.6 Mean train size vs Utl

Figure 4.7 Var train size vs Utl
Channel Utilization (Utl)

Figure 4.8 PDLP vs Utl

Number of Stations (N)

Per Delayed V Pks (PDLP)

VD1
IAT = 12

0.0 25.71 51.42 77.14

0.0 22.85 45.71 68.57

0.4 0.64 0.88

Figure 4.9 PDLP vs N
Figure 4.10 Utl vs IAT

Figure 4.11 Mean queue length vs IAT
interarrival time of the data packet increases. (Figure 4.12) A similar trend can be observed with the variance of the waiting time (Figure 4.13), the mean, and the variance of the train size. (Figure 4.14 and Figure 4.15 respectively)

4.4.2 Analysis of Results for VD2

For a fixed IAT, \( K_{\text{data}} \), and \( \Phi \) the channel utilization increases rapidly with an increase in the total number of stations in the system. However, the increase levels off at \( N = 24 \). (Figure 4.16) The mean queue length (Figure 4.17), mean waiting time (Figure 4.18), and the variance of the waiting time (Figure 4.19) are found to increase in step with the channel utilization until Util equals 0.8, beyond which point the increase becomes much more significant. The mean (figure 4.20) and variance (Figure 4.21) of the train size show similar characteristics as in the VD1 case. The data packet discard percentage is observed to be higher than the voice packet discard rate by a fair amount. (Figure 4.22) The percentage of voice packets delayed (Figure 4.23) is quite low for a utilization up to 0.78.

For fixed \( N \), IAT, and \( K_{\text{data}} \), the channel utilization, the mean queue length, the mean waiting time, the variance of the waiting time, and the mean train size increase as \( \Phi \) becomes bigger. (Figures 4.24-30) Higher \( K_{\text{data}} \) increases the above quantities as well. The mean and variance of waiting time of
Figure 4.12 Mean wait time vs IAT

Figure 4.13 Var wait time vs IAT
Figure 4.14 Mean train size vs IAT

Figure 4.15 Var train size vs IAT
Channel Utilization (Utl)

Number of Stations (N)

Figure 4.16 Utl vs N

Mean Queue Length (Q)

Figure 4.17 Mean queue length vs Utl

IAT = 12
VD₂
KDATA = 5
φ = 1.3

VD₂
KDATA = 5
φ = 1.3
IAT = 12
Figure 4.18 Mean wait time vs Utl

Figure 4.19 Var wait time vs Utl
Figure 4.20 Mean train size vs Utl

Figure 4.21 Var train size vs Utl
Figure 4.22 % D, V pks discarded vs Utl

Figure 4.23 % delayed V pks vs Utl
Figure 4.24 Util vs $\varphi$

Figure 4.25 Mean queue length vs $\varphi$
Overall Scale Factor

Figure 4.26 Mean wait time vs $\varphi$

Overall Scale Factor

Figure 4.27 Var wait time vs $\varphi$
Overall Scale Factor

Figure 4.28 Mean wait time vs \( \phi \)

Figure 4.29 Var wait time vs \( \phi \)
data packets are larger for higher $K_{\text{data}}$. (Figure 4.28 and 4.29) In Figure 4.31, one observes that an increase in $K_{\text{data}}$ helps bring down the variance of the train size.

The percentage of data packets discarded can be reduced by increasing $K_{\text{data}}$ or by increasing $\varphi$, with all other quantities held constant. (Figure 4.32) As illustrated in Figure 4.33, higher $K_{\text{data}}$ gives rise to higher percentage of voice packets discarded. This is because there are more data packets in the previous train in general than when $K_{\text{data}}$ is small. For low $K_{\text{data}}$ values, percentage of data packets discarded is higher than the voice counterpart. (Figure 4.34) The reverse is true for higher $K_{\text{data}}$. Figure 4.35 demonstrates that more voice packets are delayed at $K_{\text{data}} = 5$ than when $K_{\text{data}} = 0.01$. An increase in $\varphi$ will also bring about an increase in PDLP since the system will be busier than before.

More data packets compared to voice packets is transmitted for an increase in $K_{\text{data}}$ while the reverse is true for the voice domain. (Figure 4.36 and Figure 4.37)

### 4.4.3 Analysis of Results for VD3

Since stations are now assumed to have infinite buffer size, the arrival rate of the data traffic has to be reduced in order to avoid excessive number of outstanding packets in the system. Hence, IAT is set to be 28 instead of 12. The channel
Overall Scale Factor

Figure 4.30 Mean train size vs $\phi$

Overall Scale Factor

Figure 4.31 Var train size vs $\phi$
Overall Scale Factor

Figure 4.31 % D Pks Dis vs \( \varphi \)

Figure 4.33 % V Pks Dis vs \( \varphi \)
Figure 4.34 & D, V Pks Dis vs $\phi$

Overall Scale Factor

Figure 4.35 PDL vs $\phi$
Figure 4.36 % data txd vs $\phi$

Figure 4.37 % voice txd vs $\phi$
utilization, mean queue length, mean waiting time, and variance of the waiting time increase with increasing size of the system. (Figures 4.38-4.41) The effect is more remarkable on the latter two quantities, since data packets which are not transmitted are now "stored" in buffer for later delivery. This implies an increase in the number of packets in the system. While the mean and variance of the data delays are large for N greater than 30, the mean and variance of the voice delays are well within reasonable bounds. (Figures 4.42-4.43) The mean and variance of the train size are also quite acceptable. (Figures 4.44-4.45) In fact, the two variances show declines for N greater than 30. Data traffic occupies about two-thirds of the total bandwidth at N = 30. (Figure 4.46) The percentage of voice packets discarded is about 1 when N is 30. (Figure 4.47) The probability of data packets delayed increases rapidly when the total number of stations in the system exceeds 30. (Figures 4.48-4.49)
Figure 4.38 Utl vs N

Figure 4.39 Mean queue length vs N

VD₃
KDATA = 5
φ = 1.3
IAT = 28
Figure 4.40 Mean wait time vs N

Figure 4.41 Var wait time vs N
Figure 4.42 Mean wait time of voice pks vs N

Figure 4.43 Var wait time of voice pks vs N
Number of Stations (N)

Figure 4.44 Mean train size vs $N$

$\bar{T}_z$

$K_{DATA} = 5$

$\phi = 1.3$

$IAT = 28$

Number of Stations (N)

Figure 4.45 Var train size vs $N$

$\sigma(T_z)^2$

$K_{DATA} = 5$

$\phi = 1.3$

$IAT = 28$
**Figure 4.46** % D, V Tx vs N

**Figure 4.47** % v pk discarded vs N
Figure 4.48 Prob data tx/delayed vs N

Figure 4.49 Prob data delayed vs N
5. COMPARISON OF THE DIFFERENT PROTOCOLS

The results for the different protocols are now plotted on the same graphs for easy comparison.

Comparing these with the cases where pure voice or pure data packets are transmitted, the integrated protocols yield much higher channel utilization factor. (Figure 5.1) This is due to the fact that more packets are ready for transmission, which in turn leads to more "full-size" trains.

The argument also explains the corresponding increase in mean queue length (Figure 5.2), mean waiting time (Figures 5.3, 5.5), and variance of the waiting time (Figures 5.4, 5.6). However, as \( N \) increases to values larger than 33, the VD2 will discard many of the packets, rendering the system a lower utilization. Similar trend is observed with the mean and variance of voice and data delays.

The mean train sizes of a V/D Expressnet employing the VD1 and VD2 protocols are much greater than those in the V1, V2 cases. (Figure 5.7) The variance of the train size (Figure 5.8) will increase as the number of stations increase, but will drop abruptly because at high utilization, the Expressnet behaves like a TDM system.

For a system size of about 30 stations (15 data stations and 15 voice stations), the packet discard rate for VD2 is about 1\%, while the amount of "late" packets is about 20\%. (Figures 5.9-5.10) The idea behind these figures is that with
15 voice stations alone, no packet discarding is required to guarantee that packets arrive within 20 ms. However, additional data stations appended to the system might violate this constraint. Now with a suitable protocol VD2, the Expressnet can accommodate almost double the original number of stations without incurring excessive amount of loss packets, and yet manage to transmit most of the packets on time to the destination.
Figure 5.1 Utl vs N

Figure 5.2 Mean queue length vs N
Figure 5.5 Mean wait time of v pks vs N

Figure 5.6 Var of wait time of v pks vs N
Figure 5.7 Mean train size vs N

- Number of Stations (N)
- Mean Train Size ($\bar{T}_z$)
- $VD_1$, $VD_2$
- $IAT = 12$
- $k_{data} = 5$
- $\varphi = 1.3$

Figure 5.8 Var of train size vs N

- Number of Stations (N)
- Var Train Size ($\sigma_{\bar{T}_z}^2$)
- $VD_1$, $VD_2$
- $IAT = 12$
- $k_{data} = 5$
- $\varphi = 1.3$
Figure 5.9: % voice pks discarded vs N

Figure 5.10: % delayed v pks vs N
6. CONCLUSIONS

The aim of this thesis is to study voice and data transmissions in Expressnet—a high-speed local area communication network with broadcast capability.

The major difference in performance requirement between the two classes of traffic is that while interactive data terminals require high level of accuracy, voice transmission can tolerate a higher error rate in exchange for a reduction in the network transit delay. This is possible because of the large amount of redundancy in human speech.

The data protocol for a 1-buffer case has been verified by theoretical figures. The infinite buffer case was also examined.

For voice transmissions, the end-to-end average network delivery time would have to be small and relatively invariant in order to avoid gaps that would reduce speech intelligibility. To maintain good speech quality, it is necessary to keep the packet discard rate below 1%. In fact, the packet discard rate in the second protocol proposed (V2) is given by an expression similar to the Time Assignment Speech Interpolation (TASI) speech loss equation. This second voice protocol was subjected to a maximum tolerable delay criterion, while the first protocol (V1) has no such constraint.

The affiliated section on call estimation has investigated the relationship between packet discarding rate and call
blocking probability, tradeoff between the two was observed. The effectiveness of the estimator was also looked into, through a comparison of the estimated versus the actual number of active voice calls. The estimators were found to be good ones.

For integrated voice and data transmission, three protocols have been proposed, investigated, and compared to the pure data, pure voice transmission cases. These include transmission without discarding of packets, discard voice packets/discard data packets, and discard voice packets/delay data packets schemes. With the latter two, schemes were devised to assign temporal priority on the transmission of voice packets, at the expense of delayed or discarded data packets. It was found that acceptable level of performance can be achieved for a fairly large number of stations.

Future research with regard to transmission in the Expressnet could be carried out in the following areas:

1. Implementation of voice call estimation scheme in integrated V/D protocols. The call estimation scheme devised in Section 3.4 can be applied to the integration of voice and data section in Section 4 through the following procedure. First of all, the $\alpha$ can be set, say, to 0.1. This implies that the incoming calls may be blocked according to the various parameters calculated from previous estimates. In case a call is accepted, then the individual voice
packet is transmitted or discarded according to the probability \( p_j \). Relationships between packet discard rates and call blocking probability can be determined.

2. Investigation into effects of preambles on the overall transmission efficiency, mean, and variance of waiting time, train size, etc.

3. Study of behaviour of protocols in which more than one packet may be transmitted by heavy traffic users.

4. Performance study of the system under different channel capacities.

5. Investigation of protocols which realize transmission of voice and data via different train types.

6. Proposal of other probability determination schemes.

7. Proposal of other transmission algorithms.

The topic of integration of voice and data in communication networks (ISDN) is of current interest and possesses enormous potential towards the design of better communications systems in the future. In order to reap the full potential of the integrated transmission plants, work has to be done on various networking aspects. The following are highlights of the possible research directions:

- Design of small, flexible, and economical voice terminals with capability to accommodate a variety of voice digitization algorithms, transmission rates, conferencing techniques, and security measures.
- Investigation of performance criteria and reliability of integrated packet voice and data systems.

- Examination and comparison of local distribution strategies and network configurations for packet-switched networks.

- There is much room for research in the area of interaction between voice and data flow controls in integrated systems.

Problems that can be foreseen include:

- The voice sections of the thesis as well as a lot of the work done so far were based on a 2-state model of alternative talkspurt and silence intervals. Other situations (e.g. double talk) were not extensively explored. However, the degree of complexity in simulation and analysis increase drastically with multi-state models.

- Stringent packet discarding rates and call blocking performance criteria essential for voice quality control pose design problems and limitations.

- There is always a tradeoff between voice and data transmissions. Voice transmission can only be favoured at the expense of data transmission (or vice versa), implying that there is always going to be contention of channel capacity among the different kinds of traffics.
Packet length optimization in terms of information throughput and mean time delay is not an easy task.

In summary, its round-robin style of transmission renders the Expressnet an advantage over the conventional multiaccess schemes (e.g. CSMA/CD). Analysis and simulations indicate that there is an upper bound on the maximum amount of the delay encountered by a (voice) packet. It is this unique feature of the Expressnet that makes it attractive for integrated voice/data transmission applications in a local area office environment. The system can dynamically allocate channel capacity to data traffic during idle periods of voice terminals, resulting in graceful transmission of voice as well as data packets.
### APPENDIX A1. CLASSIFICATION OF DATA PERFORMANCE PARAMETERS

<table>
<thead>
<tr>
<th>PERFORMANCE CRITERION</th>
<th>SERVICE</th>
<th>TRANSMISSION</th>
</tr>
</thead>
</table>
| Quality               | - User information loss probability  
                        - User information mis-delivery probability | - User information probability |
| Delay                 | - Access delay (call setup time) | - User information transfer delay |
| Availability          | - Access denial probability (network blocking)  
                        - Service availability | - User information transfer rate (throughput efficiency) |

### APPENDIX A2. CLASSIFICATION OF VOICE PERFORMANCE PARAMETERS

<table>
<thead>
<tr>
<th>PERFORMANCE CRITERION</th>
<th>SERVICE</th>
<th>TRANSMISSION</th>
</tr>
</thead>
</table>
| Quality               | - Incorrect treatment | - Echo  
                        - Quantization Noise  
                        - Interruptions  
                        - Errors |
| Delay                 | - Dial tone delay | - Absolute fixed delay  
                        - Variable speechburst delay |
| Availability          | - Call blocking | - Long interruptions  
                        - Call cutoff |
APPENDIX B. TRANSMISSION MECHANISM OF EXPRESSNET

In the Expressnet access protocol, the mechanism used in determining the access rights to users in a given round is made independent of the propagation delay $t$, in contrast to other multiple access schemes.

A global timing diagram and the flowchart of the procedure as carried out by an individual station are illustrated in Figures B1 and B2 respectively.

In the Expressnet access protocol (Figure B2) the time reference used is the end-of-carrier on the outbound channel (EOC(OUT)). The event EOC(OUT) is said to occur when $c(t,\text{OUT})$ undertakes a transition from 1 to 0. $c(t,\text{OUT})$ is a Boolean function denoting the presence or absence of carrier with a delay of $t_d$ seconds. Here $t_d$ represents the time required for the detection operation. Thus the gap between consecutive transmissions are reduced to values on the same order of magnitude as the time required to detect carrier.

A transmission unit (TU) consists of a preamble (for synchronization purposes at the receiver) followed by the information packet itself.

An ACTIVE user, who has a message to transmit, is said to be backlogged (or READY). Otherwise it is said to be idle (NOT-READY). The state of an ACTIVE station can be determined by examining whether its transmit buffer is empty or not. To that effect, a function $TB(t)$ is defined for each station:
TB(t) = \begin{cases} 
1 \text{ if its transmit buffer is nonempty} \\
0 \text{ otherwise.}
\end{cases}

An idle user does not contend for the channel. A backlogged user (a user wishing to transmit) waits until it detects EOC(OUT) on the outbound channel and then immediately starts transmitting its own packet (TU). Simultaneously, the user senses the outbound channel for activity on the upstream side. If activity is detected, the user aborts its current transmission. Otherwise, it completes it.

The succession of transmission units transmitted in the same round is defined as a "train." A train is generated on the outbound channel and is entirely seen on the inbound channel by all users. However, it can only be seen by a station on the outbound channel as long as the TU's in it are being transmitted by stations with lower indexes, i.e., stations on the upstream side of its transmitter. The maximum number of packets in the train is the maximum number of stations connected to the network.

According to the configuration and the transmission protocol, there is a single user, the one with the lowest index among those ready stations. The user is able to detect EOC(OUT). It does not have to abort its transmission and hence transmits successfully. Moreover, a user who has completed the transmission of a packet in a given round does
not encounter the event EOC(OUT) again in that round. This guarantees that no user can transmit more than once in a given round.

The detection of the presence of a train on the inbound channel is achieved by defining \( \text{TRAIN}(t,\text{IN}) = c(t-t_d,\text{IN}) + c(t,\text{IN}) \) whose value is 1 as long as the train is in progress. The transition \( \text{TRAIN}(t,\text{IN}): (1 \rightarrow 0) \) defines the end of train EOT(IN) and the transition \( \text{TRAIN}(t,\text{IN}): (0 \rightarrow 1) \) defines the beginning of a train BOT(IN).

The main feature of the topology rests on the fact that when the inbound channel is made exactly parallel to the outbound channel, the event EOT(IN) visits each user in the same order as they are permitted to transmit. Therefore, to start a new round, EOT(IN) is used as the synchronizing event. The mechanism allows the ready stations with the lowest index to be the first to complete transmission of its TU. Then the new train takes its normal course.

When the ALIVE station becomes ready, it does not have to wait for an EOT(IN) to start transmission. If an outbound train is observed, the station synchronizes transmissions with EOC(OUT). If, however, at the time it becomes ready, no train is observed on the outbound channel, then EOT(IN) becomes the synchronizing event.

To avoid losing the synchronizing event EOT(IN) if the train were to be empty, in the Expressnet a new round is always started by having all users transmit a short burst of
unmodulated carrier, termed "locomotive," of duration $t_d$ before attempting to transmit a packet. If the train is empty, then the locomotive constitutes the end of train, and EOT(IN) surely takes place.

From the above description, one can see that the Expressnet achieves a "conventional" round robin discipline where users are serviced in a prescribed order determined by their physical location on the network. If a user has not message in buffer when its turn comes up, it declines to transmit so it must wait for the next round before getting another turn. This type of discipline is referred to as the non-gated sequential service discipline (NGSS).
Signals and events as observed by station $j$, assuming that stations with indexes $i_1 < i_2 < j < i_3$ are nonidle.

Figure B1 Global timing diagram of Expressnet
Figure B2 Flowchart for the Expressnet access protocol
Flowchart for Main Program
( Data, Voice )
Flowchart for BOTRAIN (Data)
Flowchart for CHECKSERVE (Data)
Flowchart for VOICEGENERATE (Voice)
APPENDIX C5

START

CLOCK = BOT = exittime
trainsize = 0

qindx ≤ N?

front ≠ nil?

check next station

front ≠ entrytime ≤ exittime?

check next station

delay = exittime - front ≠ entrytime

delay ≤ pgtfull?

Increment discardpk

pktxd = pktxd + 1

place delay in LFD if delay is larger; update pkdelay, sqsv;
Both exittime and CLOCK incremented by TRANS[qindx] + t_d;
Busytime incremented by TRANS[qindx]; trainsize = trainsize + 1

pkremain = pkremain - 1
SUMREMAIN = SUMREMAIN - 1
discard(front)

tz : trainsize
pg : pgtfull

front ≠ nil?

check next station

nexttime = front ≠ entrytime

Flowchart for BOTRAIN (Voice, with discarding of packets)
LIST OF REFERENCES

ISDN

MOB

PFM

MUX

TH (Thesis)


SW (Switching)


PV (Packet Voice)


VN (Voice Network)


PVR (Packet Voice Receiver)


SPCH


CALL


VDN (Integrated Voice and Data Networks)