Packet Speech Program Review Meeting

June 3, 1982

Sponsored By

Department of Defense

Defense Advanced Research Projects Agency

Hosted By

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

Foreword

The Defense Advanced Research Projects Agency sponsored a meeting on June 3, 1982, at Lincoln Laboratory to review the results of the DARPA Packet Speech Program and to demonstrate packet speech technology to a group of invited visitors from the Department of Defense. The purpose of this document is to record the proceedings of the meeting by collecting the visual aids used in the technical presentations and in the demonstration.

TABLE OF CONTENTS

TAB #	
1	Packetized Speech Overview Lt. Col. Duane A. Adams, Defense Advanced Research Projects Agency
2	Network Voice Protocols Danny Cohen, Univ. of Southern California Information Sciences Institute
3	Speech Conferencing James W. Forgie, MIT Lincoln Laboratory
4	Wideband Integrated Voice/Data Network Clifford J. Weinstein, MIT Lincoln Laboratory
5	Packet Radio Speech Earl J. Craighill, SRI International
6	Algorithmic Achievements John Makhoul and Vishu Viswanathan, Bolt Beranek and Newman, Inc.
7	Very Low Rate Vocoder Richard Schwartz, Bolt Beranek and Newman, Inc.
8	Packet Voice Terminals and the LEXNET Local Network Gerald C. O'Leary, MIT Lincoln Laboratory
9	Compact LPC Vocoder Joel A. Feldman, MIT Lincoln Laboratory
10	Flexible Array Processors Glen Culler, CHI Systems, Inc.
11	VLSI Array Processor Edward Greenwood, Motorola Government Electronics Group
12	Single Chip LPC Robert Brodersen, Univ. of California/Berkeley

AGENDA FOR DARPA PACKET SPEECH MEETING

Lincoln Laboratory

3 June 1982

TIME	TOPIC	SPEAKER
0930-1100	Background, Goals, Key Results, and Demo	Adams (DARPA)
1100-1115	Break	
1115-1135	Network Voice Protocols	Cohen (ISI)
1135-1150	Speech Conferencing	Forgie (Lincoln)
1150-1205	Wideband Network	Weinstein (Lincoln)
1205-1225	Packet Radio Speech	Craighill (SRI)
1230-1330	Lunch	
1330-1350	Summary of Key Algorithm Developments	Makhoul (BBN)
1350-1405	Very Low-Rate Speech	Schwartz (BBN)
1405-1425	Packet Voice Terminal	O'Leary (Lincoln)
1425-1445	Break	
1445-1500	Compact LPC Vocoder	Feldman (Lincoln)
1500-1515	Flexible Array Processor	Culler (CHI Systems)
1515-1530	VLSI Array Processor	Greenwood (Motorola)
1530-1545	Single Chip LPC	Brodersen (U.C. Berkeley)
1545-1630	Future Plans and Discussion	Kahn (DARPA)

Attendees

Name	Affiliation
Abene, Capt. Peter	DCA
Adams, Lt. Col. Duane	DARPA/IPTO
Barna, Joseph	DCEC
Beauchemin, Edward	MIT-LL
Berger, Robert	MIT-LL
Bially, Ted	MIT-LL
Blankenship, Peter	MIT-LL
Blumenthal, Steven	BBN
Brodersen, Robert	U. California/Berkeley
Burchfiel, Jerry	BBN
Casner, Stephen	ISI
Cohen, Dan	ISI
Cole, Randy	ISI
Cowles, Jack	Western Union
Craighill, Earl	SRI International
Crawford, Richard	DCA
Culler, Glen	CHI Systems, Inc
Deckelman, Francis	NAVELEX
Falk, Gilbert	BBN
Falk, Col. Harold	ESD
Feldman, Joel	MIT-LL
Forgie, James	MIT-LL
Forsdick, Harry	BBN
Friedhoffer, Carol	NSA
Fussell, Jesse	NSA
Gilbert, Warren	ESD
Glazim, Bernard	Western Union
Gold, Bernard	MIT-LL
Greenwood, Edward	Motorola
Heggestad, Harold	MIT-LL
Hoversten, Estil	Linkabit
Johnson, Robin	OSD
Kahn, Robert	DARPA
Kang, George	NRL
Kantrowitz, William	MIT-LL
Kline, Lee	NRL
Krasner, Michael	BBN

Lippmann, Richard	MIT-LL
Luksevicius, Albinas	Ft. Monmouth
Lynn, Verne L.	DARPA
Makhoul, John	BBN
McCammon, Michael	CHI Sytems, Inc.
McElwain, Constance	MIT-LL
McLaughlin, Alan	MIT-LL
Moran, George	DCEC
Morris, Paul	Ft. Monmouth
Morrow, Walter	MIT-LL
Northrup, Richard	RADC
O'Leary, Gerald	MIT-LL
Parker, James	MITRE
Paul, Douglas	MIT-LL
Peatfield, C. Ross	MIT-LL
Peeler, Charles	Ft. Bragg
Proctor, William	CHI Systems, Inc.
Rettberg, Randy	BBN
Roucos, Salim	BBN
Sandy, G. Ferrell	MITRE
Schecter, Harold	ESD
Schwartz, Perry	MITRE
Schwartz, Richard	BBN
Secrest, Bruce	Texas Instruments, Inc.
Segota, Anton	ESD
Sharpe, Lt. Grady	ESD
Shippee, Bernard	Western Union
Sonini, Lt. Col. Joseph	ESD
Sues, Larry	RADC
Taylor, Lee	MIT-LL
Thomas, Robert	BBN
Tierney, Joseph	MIT-LL
Tomlinson, Raymond	BBN
Tremain, Thomas	NSA
Viswanathan, Vishu	BBN
Weinstein, Clifford	MIT-LL
Zimmermann, Lt. Col. Frank	DCEC

Organization Addresses

BBN
Bolt Beranek and Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

CHI Systems, Inc.
CHI Systems, Inc.
150-A Aero Camino
Goleta, CA 93017

DARPA
DARPA/IPTO
Architect Building
1400 Wilson Boulevard
Arlington, VA 22209

DCA
Defense Communications Agency
4135 Courthouse Road
Arlington, VA 22209

DCEC
Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090

ESD
Electronic Systems Division
L. G. Hanscom Field
Bedford, MA 01731

FT. BRAGG
Dir., ADDS Testbed
Fort Bragg, NC 28307

FT. MONMOUTH
Hqs., U. S. Army
Communications Electronics Command
Ft. Monmouth, NJ 07703

ISI
Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina Del Rey, CA 90291

Linkabit
Linkabit Corporation
10453 Roselle Street
San Diego, CA 92121

MIT-LL
MIT Lincoln Laboratory
P. O. Box 73
Lexington, MA 02173

MITRE
MITRE Corporation
Bedford, MA 01731

Motorola
Motorola
Government Electronics Division
8201 E. McDowell Road
Scottsdale, AZ 85251

NAVELEX
Naval Electronics Command
Washington, D.C. 20366

NRL
Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, D.C. 20375

NSA
Director
National Security Agency
9800 Savage Road
Ft. George G. Meade, MD 20755

OSD
Office Secretary of Defense
Washington, D.C. 20301

RADC
Rome Air Development Center
Griffiss AFB, NY 13441

SRI
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Texas Instruments, Inc.
P. O. Box 225936
Dallas, TX 75265

U. of California
U. of California/Berkeley
Dept. of EE/CS
Cory Hall
Berkeley, CA 94720

WU
Western Union
One Lake Street
Upper Saddle River, NJ 07458

PACKETIZED SPEECH OVERVIEW

3 JUNE 1982

LT. COL. DUANE A. ADAMS
DARPA

PACKETIZED SPEECH

OBJECTIVES

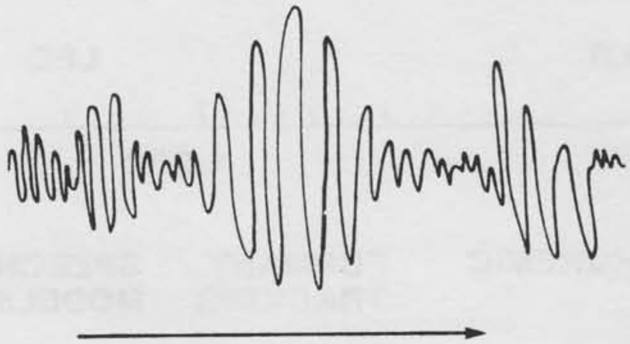
- ACHIEVE POINT-TO-POINT AND CONFERENCED NARROWBAND DIGITAL SPEECH THROUGH PACKET-SWITCHED NETWORKS
- DEVELOP ALGORITHMS FOR DEGRADED SPEECH ENVIRONMENTS
- DEMONSTRATE MULTI-USER INTEGRATED DATA/ VOICE TRANSMISSIONS
- MAINTAIN COMPATIBILITY WITH END-TO-END SECURITY AND INTERNET PROTOCOLS
- DEVELOP LOW-COST PACKET SPEECH HARDWARE

PACKETIZED SPEECH

MAJOR THRUSTS

- ALGORITHMS
- NETWORKS
- HARDWARE

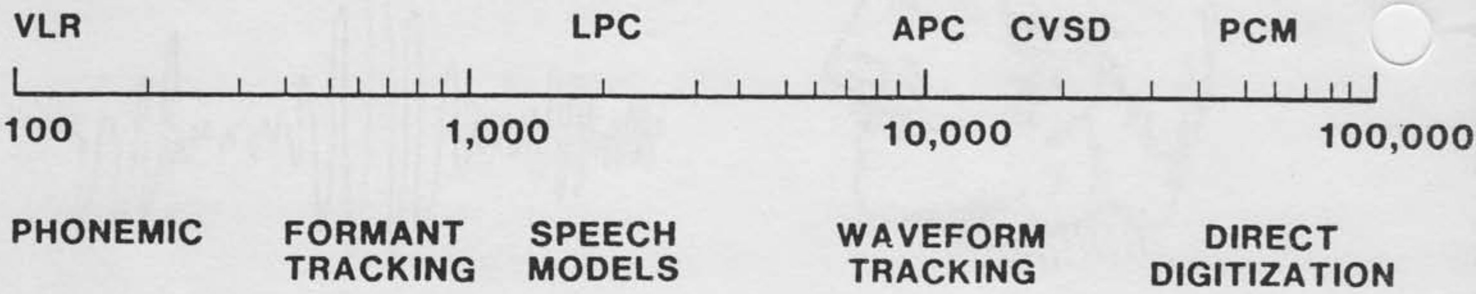
PACKETIZED SPEECH



SPEECH FREQUENCY RANGE: 0-4000 Hz

SAMPLING FREQUENCY: 8000 SAMPLES/SEC

PCM DATA RATE: 64,000 BITS/SEC



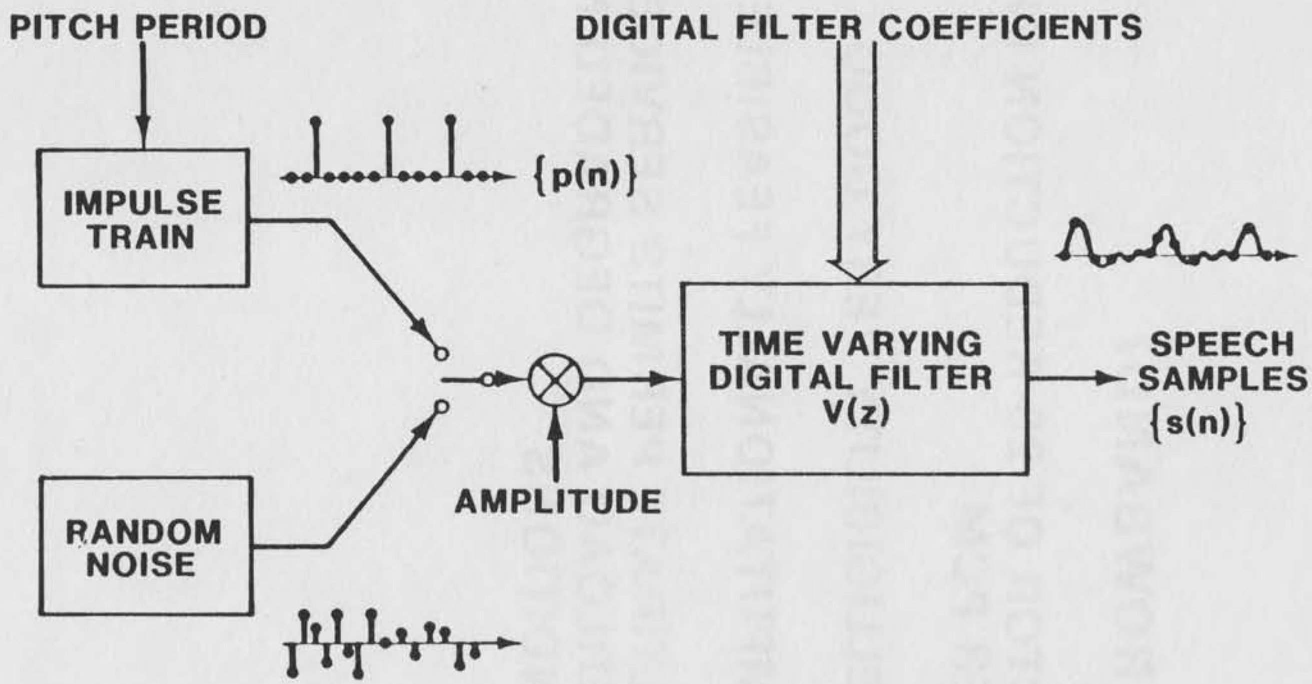
DIGITAL SPEECH REPRESENTATIONS

PACKETIZED SPEECH

WHY NARROWBAND?

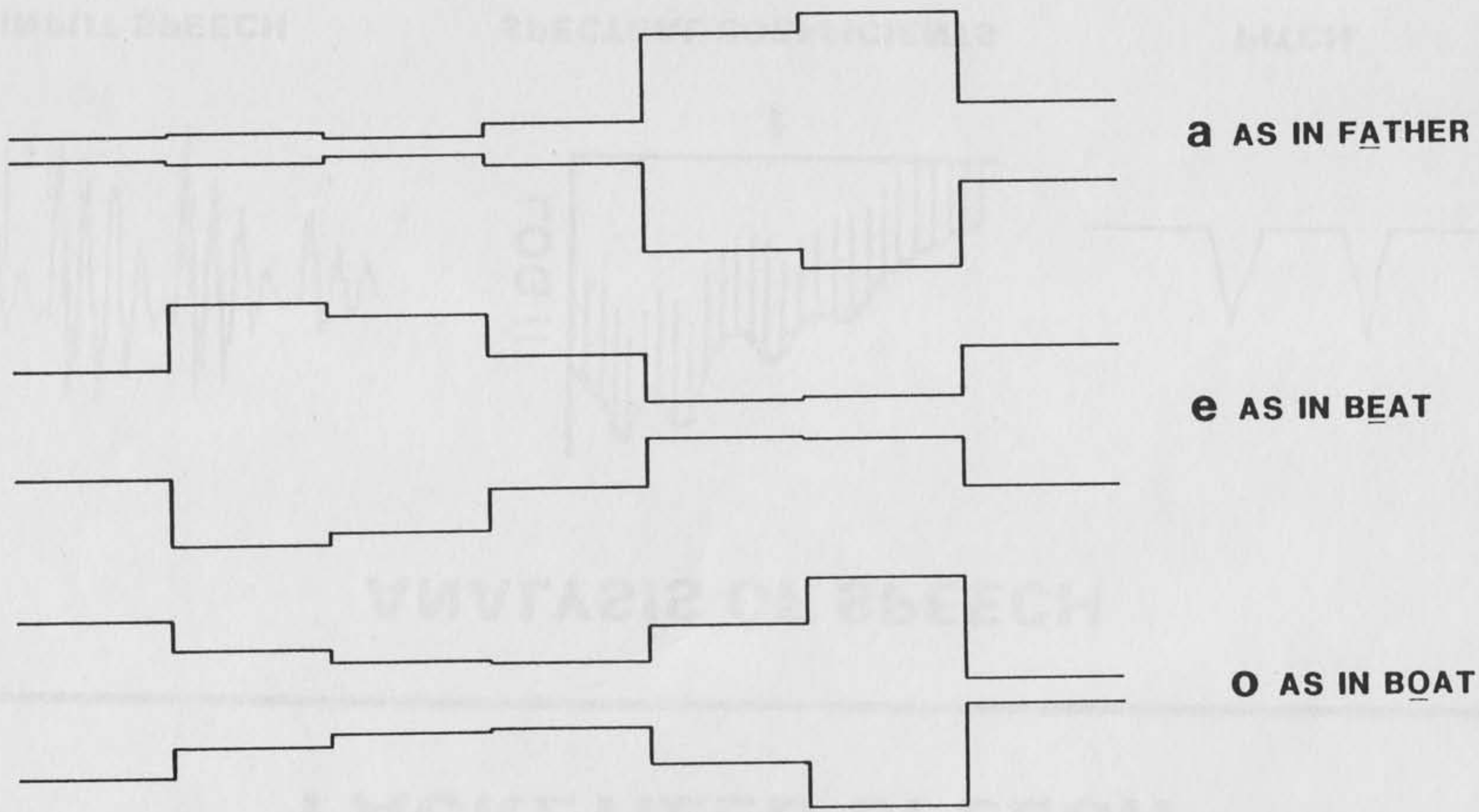
- FACTOR OF 25 REDUCTION IN BANDWIDTH OVER PCM
- INTELLIGIBILITY VERY GOOD
- COMPUTATIONALLY FEASIBLE
- MULTIRATE PERMITS SERVICE UNDER OVERLOAD AND DEGRADED NETWORK CONDITIONS

PACKETIZED SPEECH



SPEECH PRODUCTION MODEL

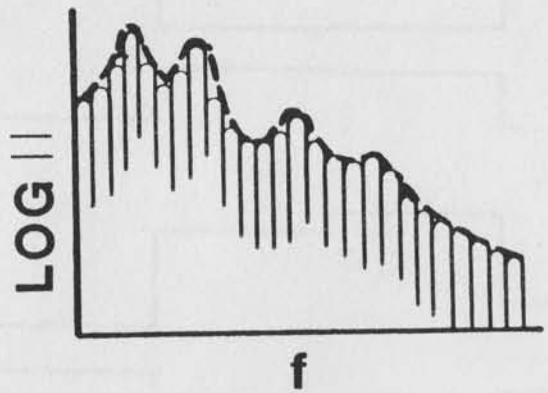
PACKETIZED SPEECH



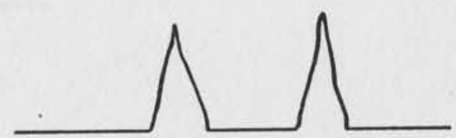
VOCAL TRACT MODEL FOR 3 VOWELS

ANALYSIS OF SPEECH

INPUT SPEECH



SPECTRAL COEFFICIENTS



PITCH

T = 10 TO 20 msec, 150 SAMPLES OF SPEECH
SPEECH ASSUMED TO BE STATIONARY OVER T

PACKETIZED SPEECH

VOCODER ALGORITHMS

- LINEAR PREDICTIVE CODING (LPC)
- CHANNEL VOCODER
- HOMOMORPHIC VOCODER
- SPECTRAL ENVELOPE ESTIMATION

PACKETIZED SPEECH

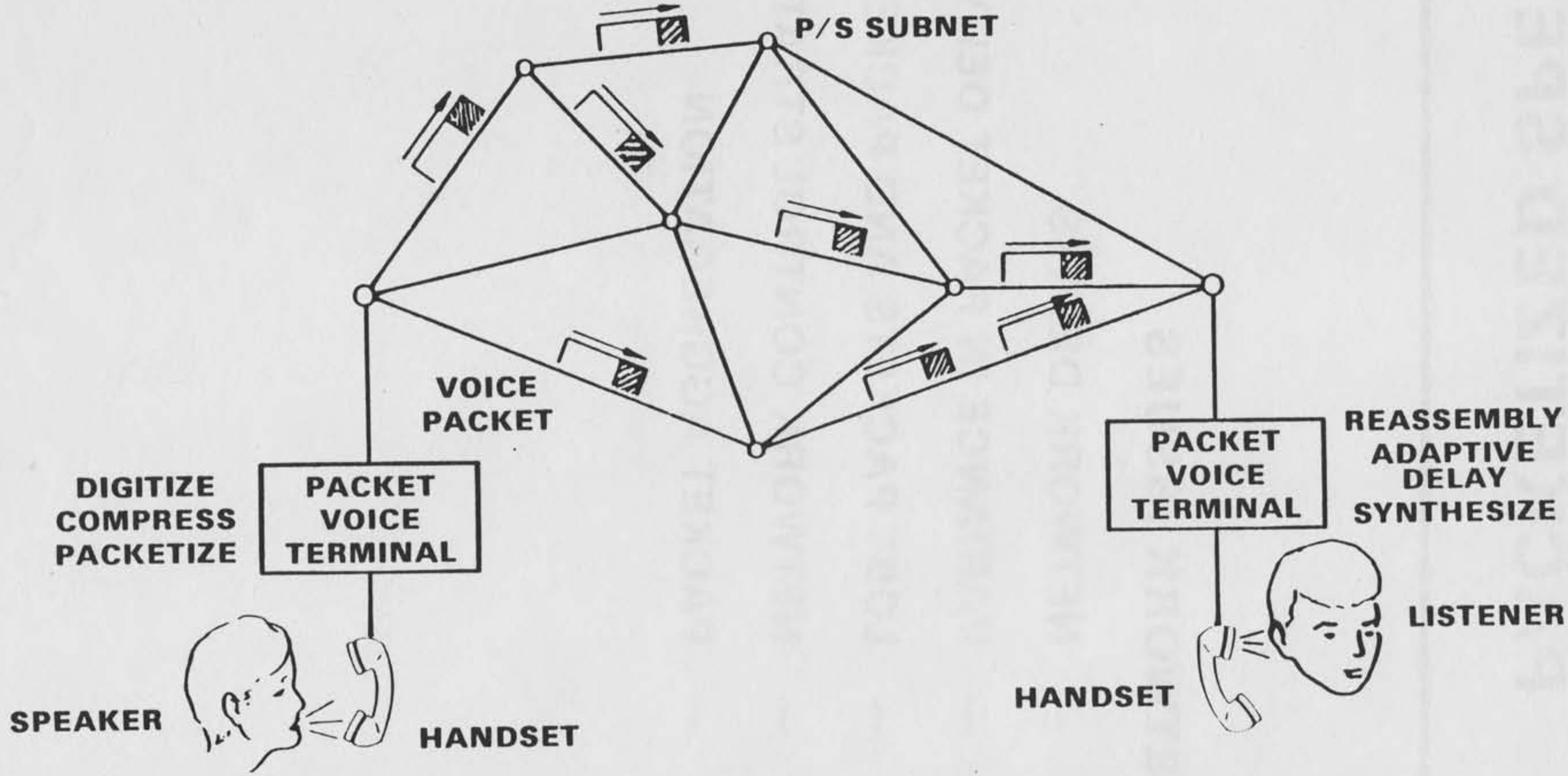
CONTROL PACKETS

HEADER	CALL SETUP

DATA PACKETS

HEADER	DIGITIZED SPEECH

PACKETIZED SPEECH



PACKET VOICE/DATA NETWORK

PACKETIZED SPEECH

NETWORK ISSUES

- NETWORK DELAYS
- VARIANCE IN PACKET DELIVERY
- LOST PACKETS AND PACKET REASSEMBLY
- NETWORK CONTROL STRATEGIES
- PACKET AGGREGATION

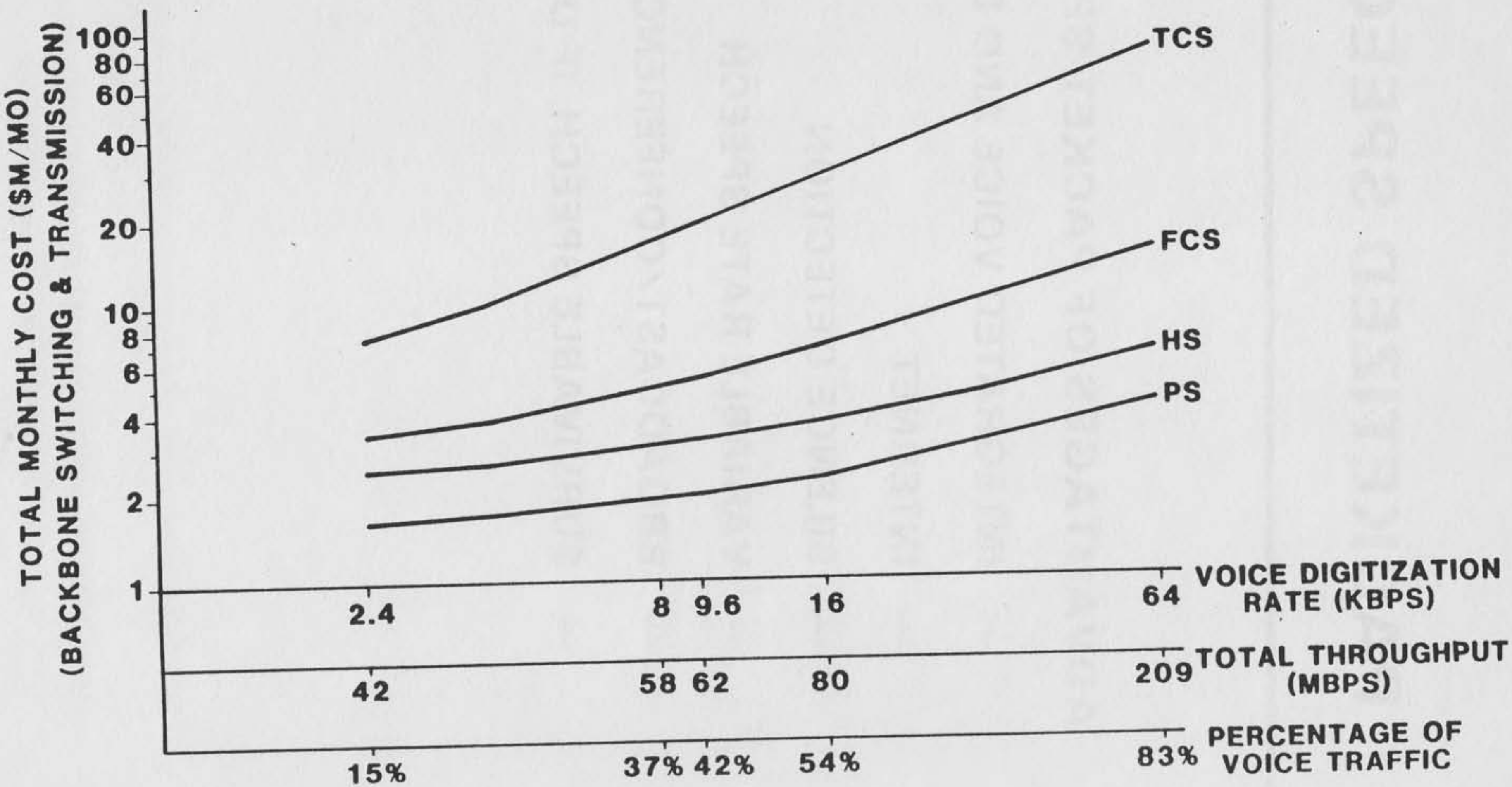
PACKETIZED SPEECH

ADVANTAGES OF PACKET SPEECH

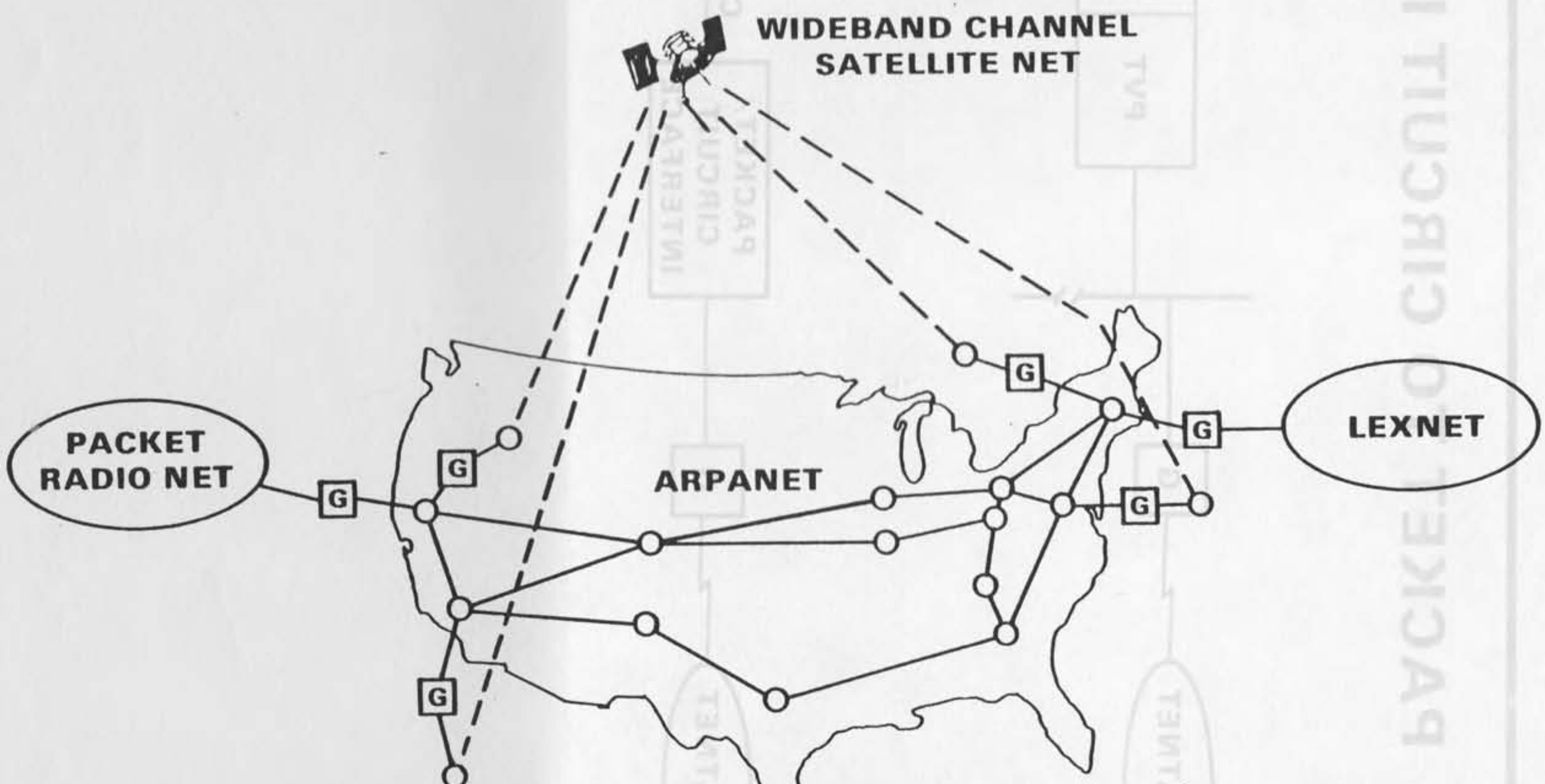
- INTEGRATED VOICE AND DATA
- INTERNET
- SILENCE DETECTION
- VARIABLE RATE SPEECH
- BROADCAST/CONFERENCING
- SURVIVABLE SPEECH, IF DESIRED

PACKETIZED SPEECH

TRAFFIC: 2,700 ERLANGS VOICE TRAFFIC
36.15 MBPS DATA TRAFFIC



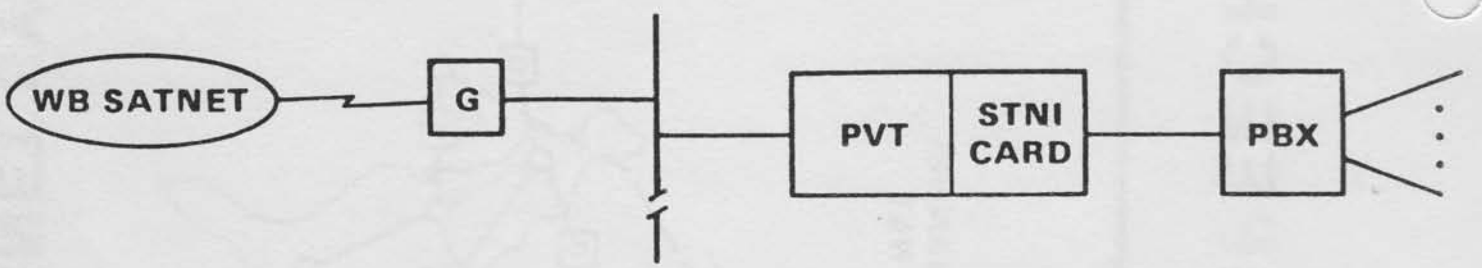
PACKETIZED SPEECH



EXPERIMENTAL INTERNET SYSTEM

PACKETIZED SPEECH

PACKET TO CIRCUIT INTERFACE



PACKETIZED SPEECH

MAJOR MILESTONES

1974	CVSD DEMONSTRATED ON ARPANET
	LPC DEMONSTRATED ON ARPANET
	NVP PROTOCOLS DEVELOPED
1976	LPC CONFERENCE ON ARPANET
1977	VARIABLE FRAME RATE LPC

PACKETIZED SPEECH

MAJOR MILESTONES (CONT.)

1978 CVSD DEMONSTRATED ON PACKET RADIO NET

SATNET CONFERENCING WITH LPC

SPECTRAL ENVELOPE ESTIMATION ALGORITHM

1979 REAL-TIME VARIABLE RATE CHANNEL VOCODER

INTERNET SPEECH DEMONSTRATED

1980 VOICE TERMINAL DEMONSTRATED ON LOCAL NET

PACKETIZED SPEECH

MAJOR MILESTONES (CONT.)

1981 NVP-II/ST PROTOCOLS

CHI-5 ARRAY PROCESSOR DEVELOPED

LPC SPEECH ON PACKET RADIO NET

PCM ON WIDEBAND CHANNEL

PACKETIZED SPEECH

MAJOR MILESTONES (CONT.)

1982 COMPACT LPC VOCODER (NEC)

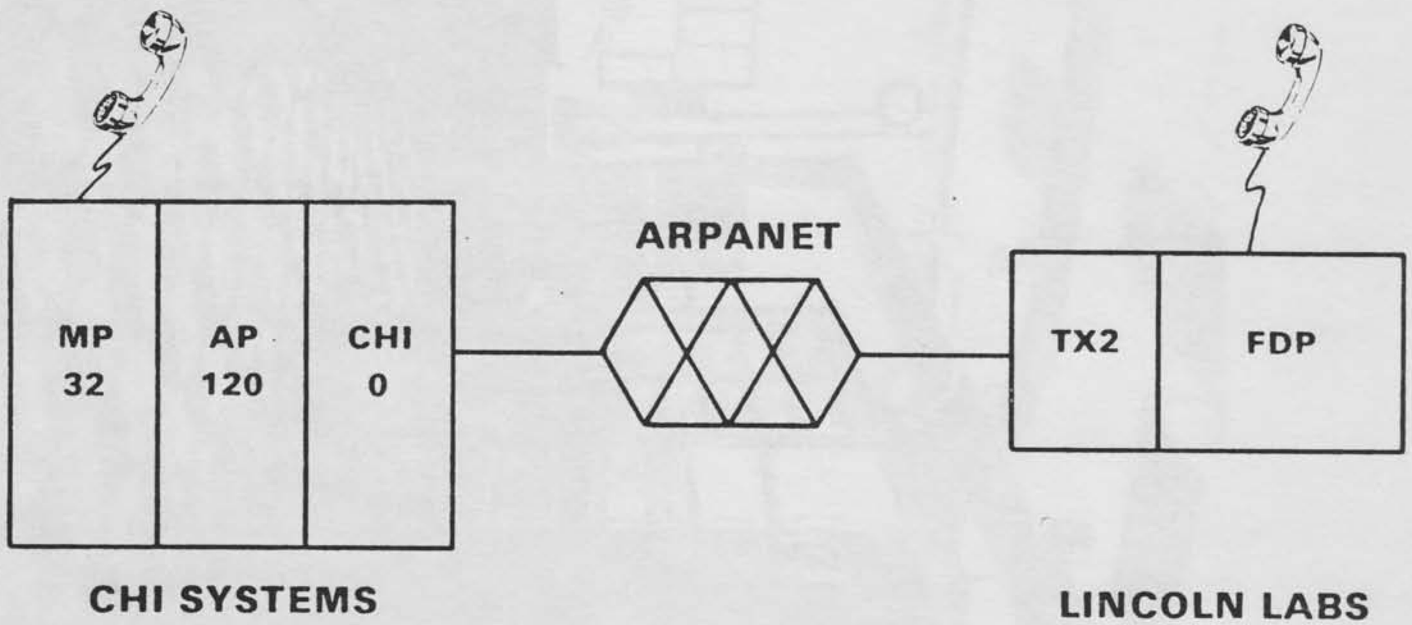
LPC ON WIDEBAND NETWORK

MULTI-RATE SPEECH ON WIDEBAND (PCM AND CVSD)

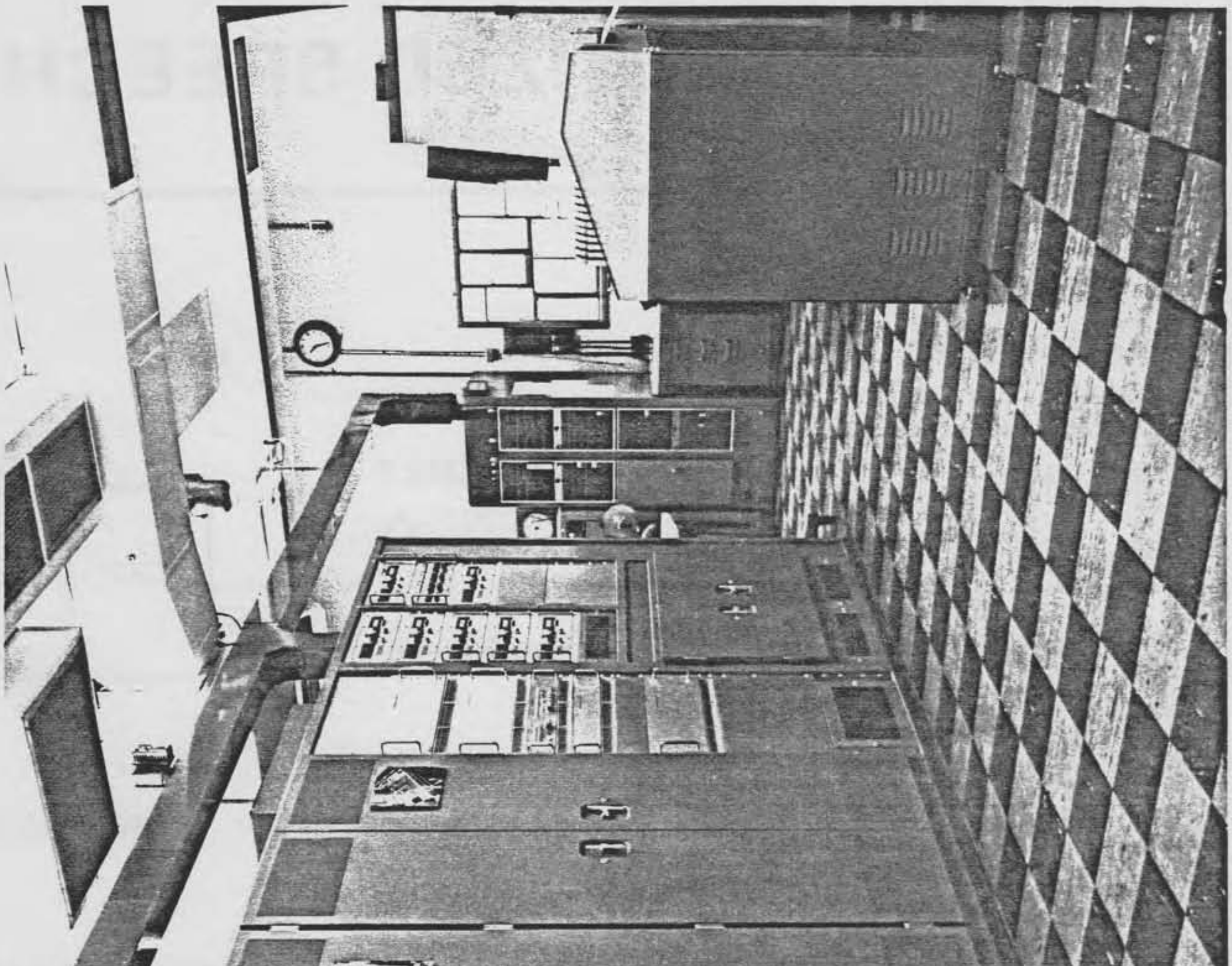
INTERNET CONFERENCING

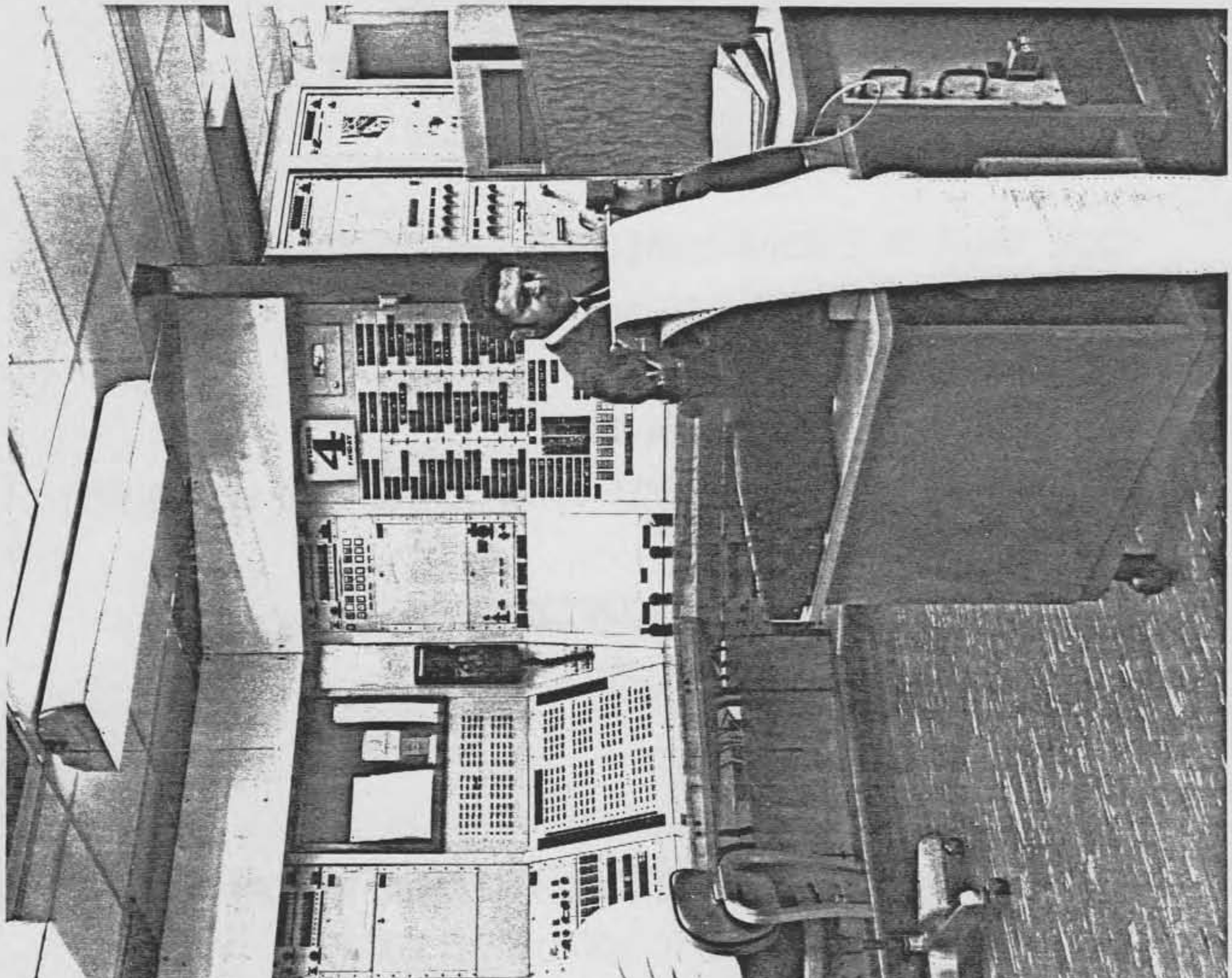
MOBILE LPC ON PACKET RADIO NET

PACKETIZED SPEECH



INITIAL NETWORK LPC DEMONSTRATION (December 1974)





SUMMARY OF

PACKET SPEECH DEMONSTRATION SEQUENCE

JUNE 3, 1982

LOCAL CALLS ON LINCOLN LEXNET

POINT-TO-POINT

1. PCM (64 KBPS)

2. ECVSD (16-64 KBPS)

3. LPC (2.4 KBPS) USING SINGLE-CARD LPC IN PVT)

CONFERENCE

4. PCM 4-PARTY CONFERENCE

CALLS OVER WIDEBAND SATNET

POINT-TO-POINT

5. PCM LL - ISI

6. PCM LL - SWITCHED TELEPHONE NETWORK INTERFACE AT ISI
(CALL TO LOCAL L.A. WEATHER)

CONFERENCE

7. 3-SITE, 4-PARTY LPC CONFERENCE

LL (2 PARTIES ON LEXNET)

ISI (ON LEXNET)

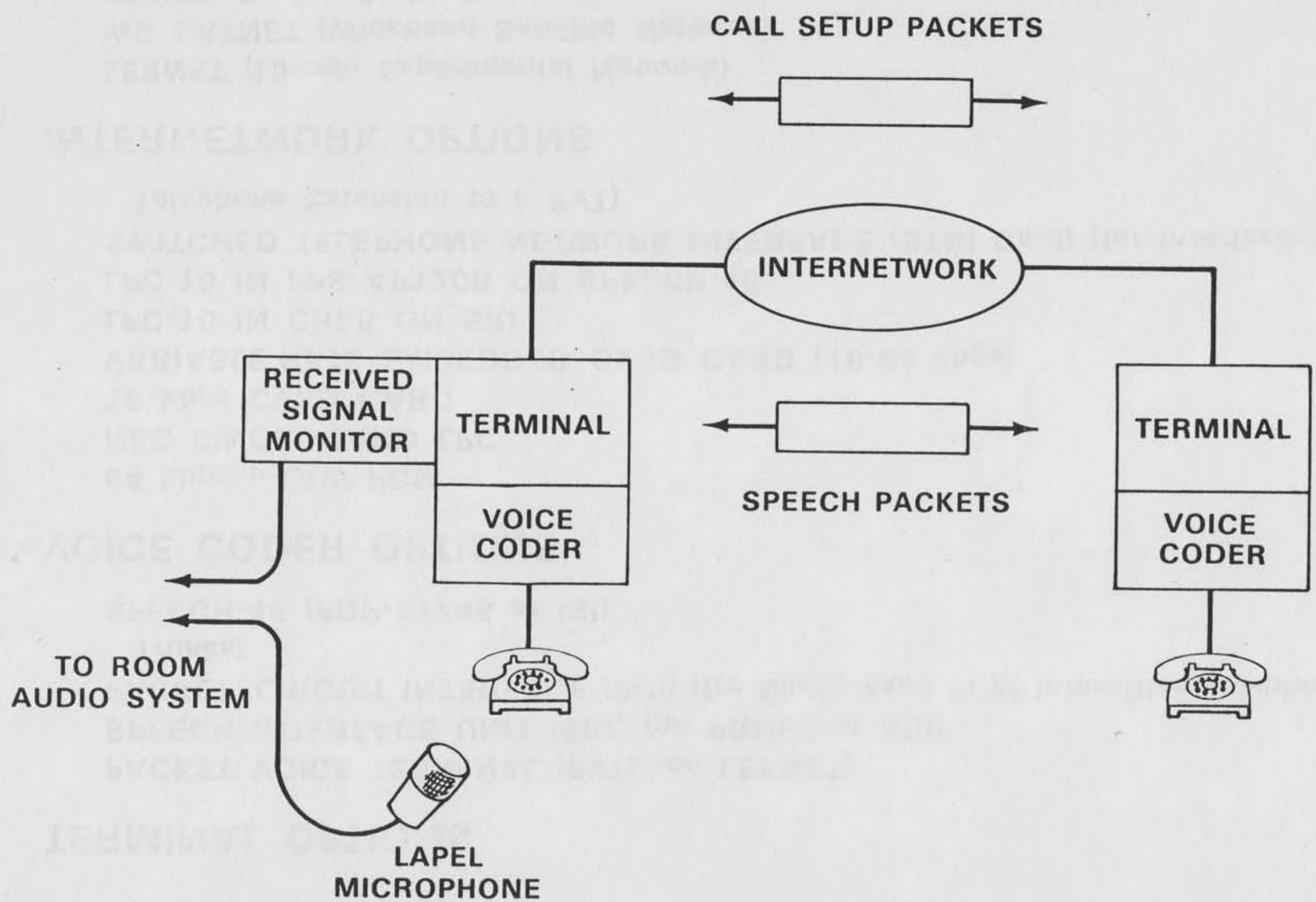
SRI (ON PRNET UNIT IN SPEECH LAB, USING CHI-V LPC
AND SIU)

POINT-TO-POINT (MOBILE)

8. LPC SRI (MOBILE PRNET UNIT) - LL (LEXNET PVT)

MOBILE VAN RUN DEMONSTRATING PRNET SPEECH AND
ALTERNATE ROUTING

GENERIC PACKET VOICE CALL DEMO SETUP



TERMINAL OPTIONS

PACKET VOICE TERMINAL (PVT) (on LEXNET)
SPEECH INTERFACE UNIT (SIU) (on PRNET at SRI)
PACKET/CIRCUIT INTERFACE (PCI) (for Multiplexed PCM Interoffice Telephone Trunks)
SPEECH-45 (PDP-11/45 at ISI)

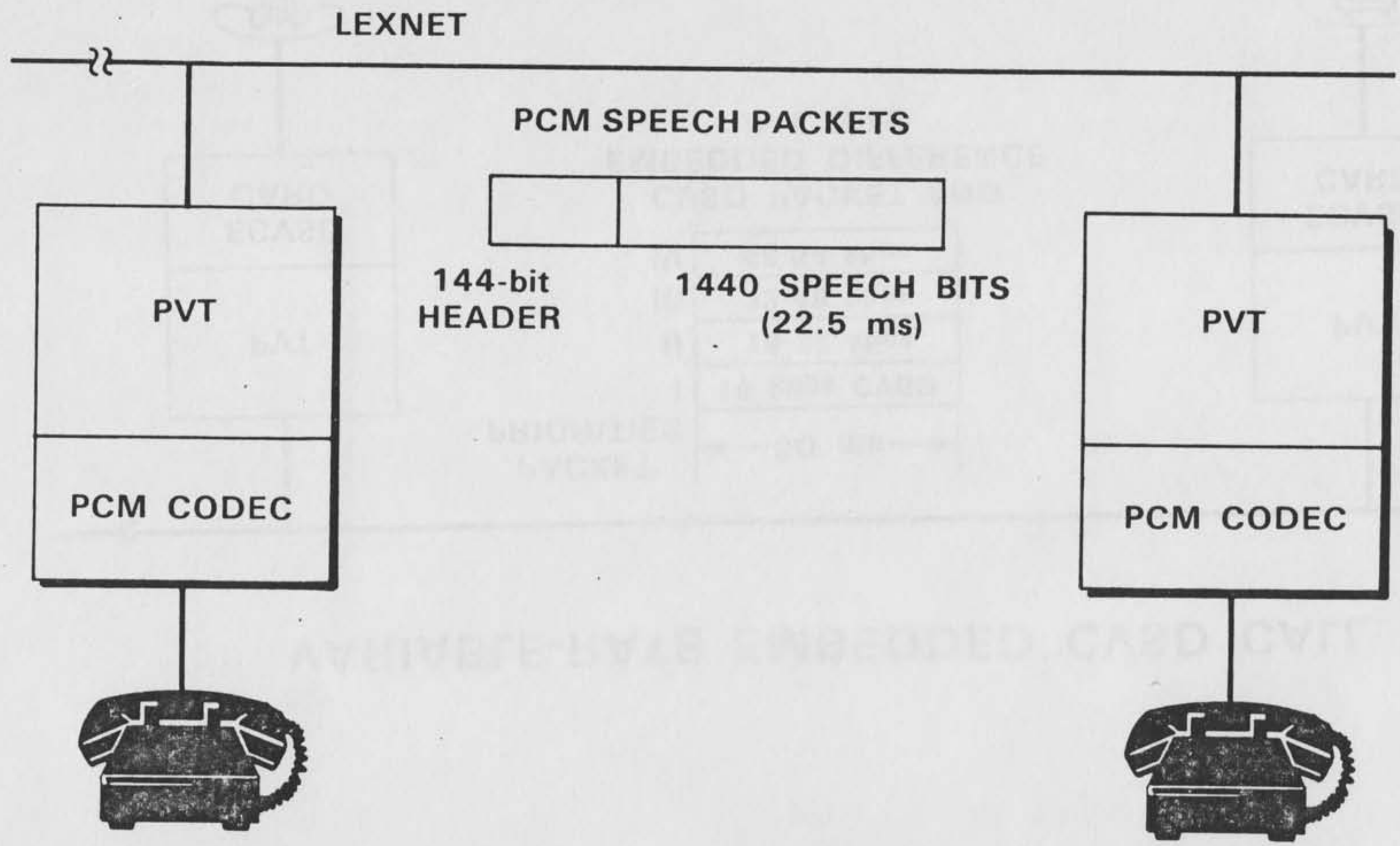
VOICE CODER OPTIONS

64 kbps μ-LAW PCM
NEC SINGLE-CARD LPC
16 kbps CVSD CARD
VARIABLE-RATE EMBEDDED CVSD CARD (16-64 kbps)
LPC-10 IN CHI-5 ON SIU
LPC-10 IN FPS AP120B ON SPEECH-45
SWITCHED TELEPHONE NETWORK INTERFACE (STNI Card) (for Interfacing a Telephone Extension to a PVT)

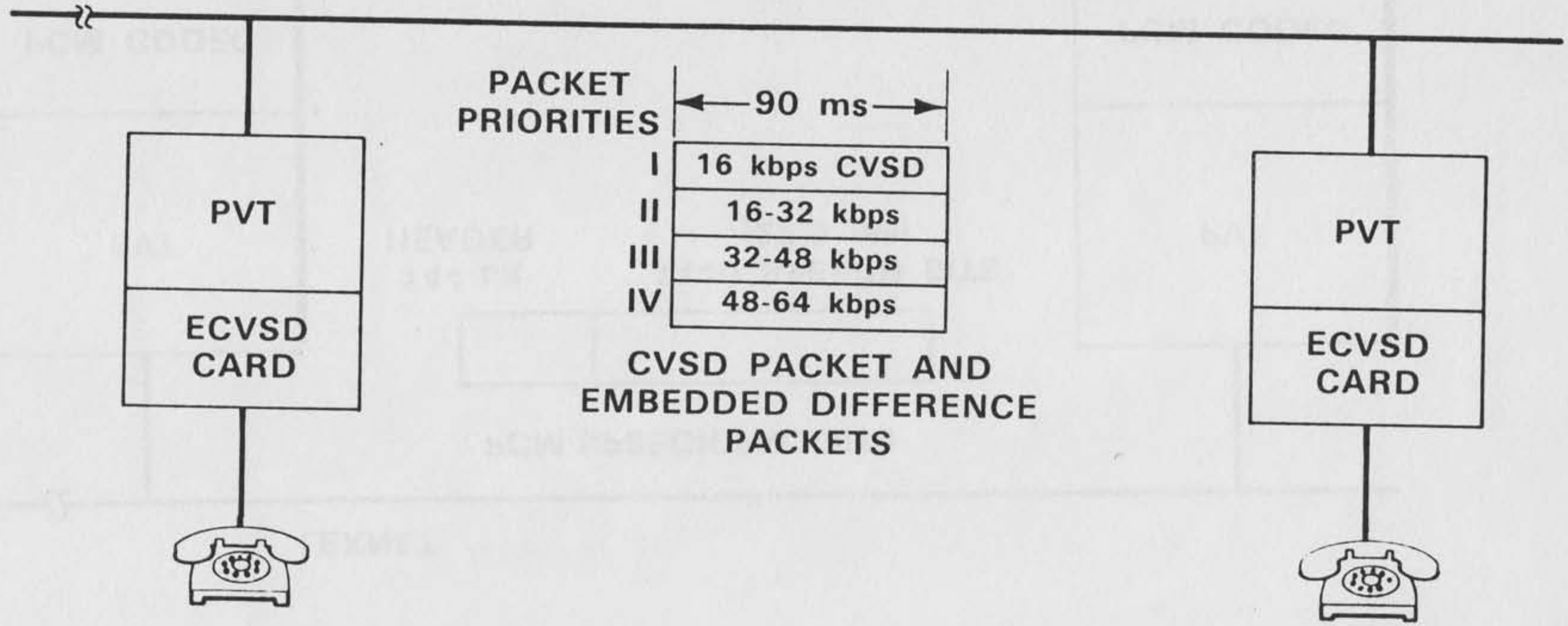
INTERNETWORK OPTIONS

LEXNET (Lincoln Experimental Network)
WB SATNET (Wideband Satellite Network)
PRNET (Packet Radio Network)
Switched Telephone Net

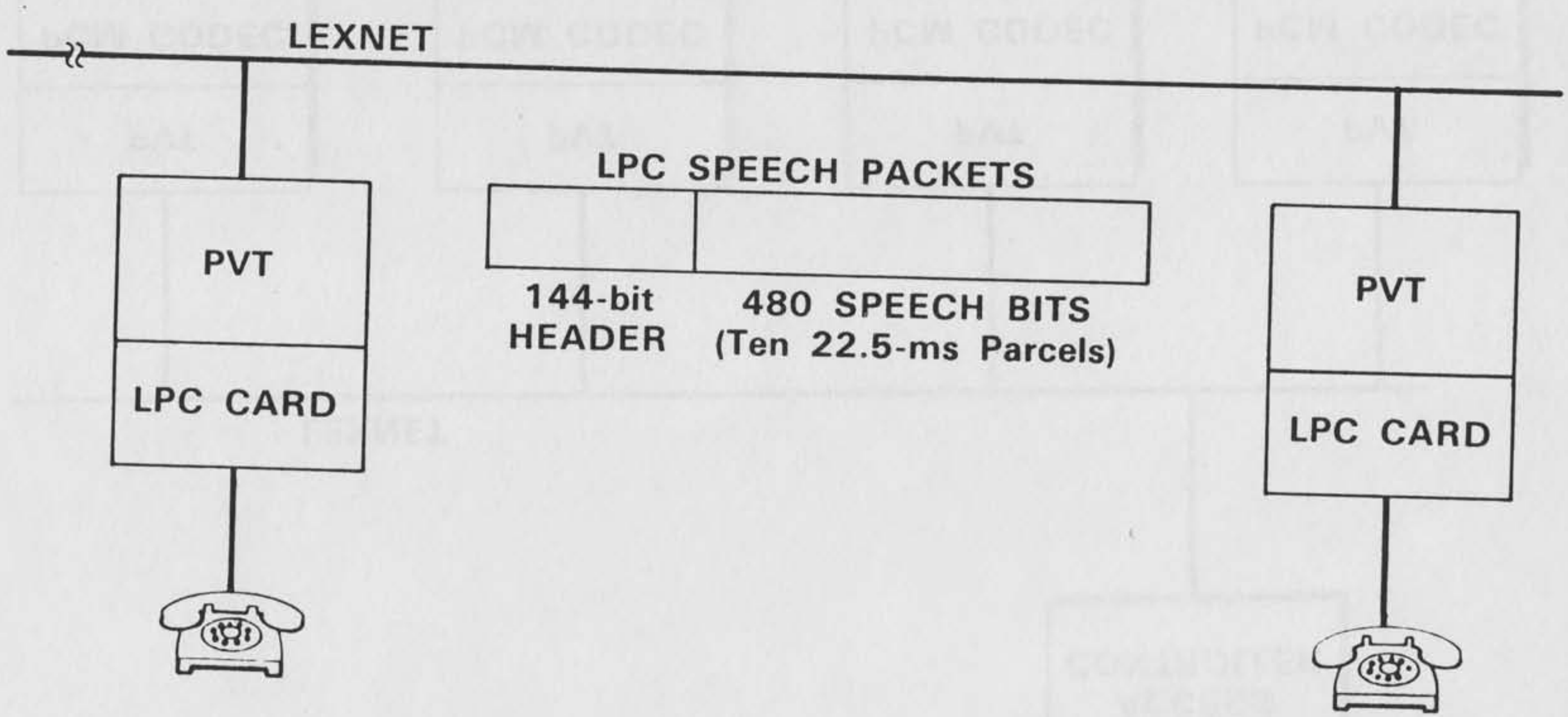
LOCAL POINT-TO-POINT 64 kbps PCM CALL



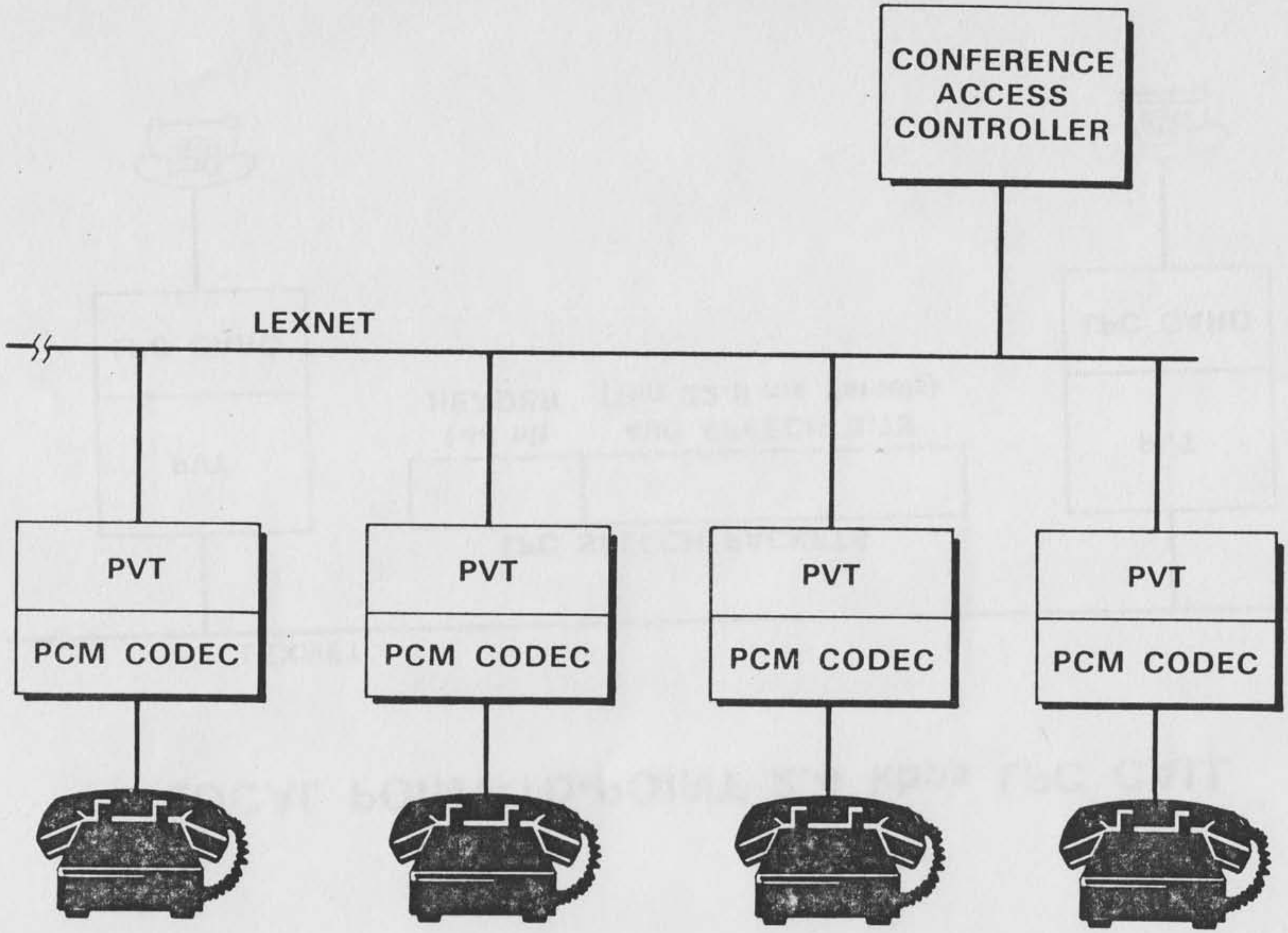
VARIABLE-RATE EMBEDDED CVSD CALL



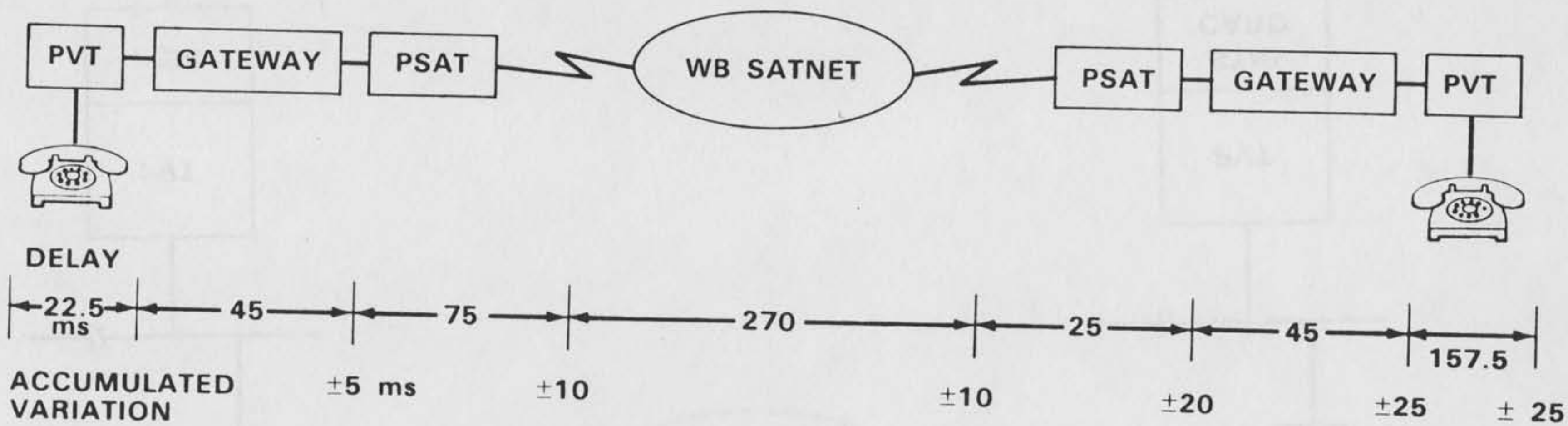
LOCAL POINT-TO-POINT 2.4 kbps LPC CALL



LOCAL PCM CONFERENCE

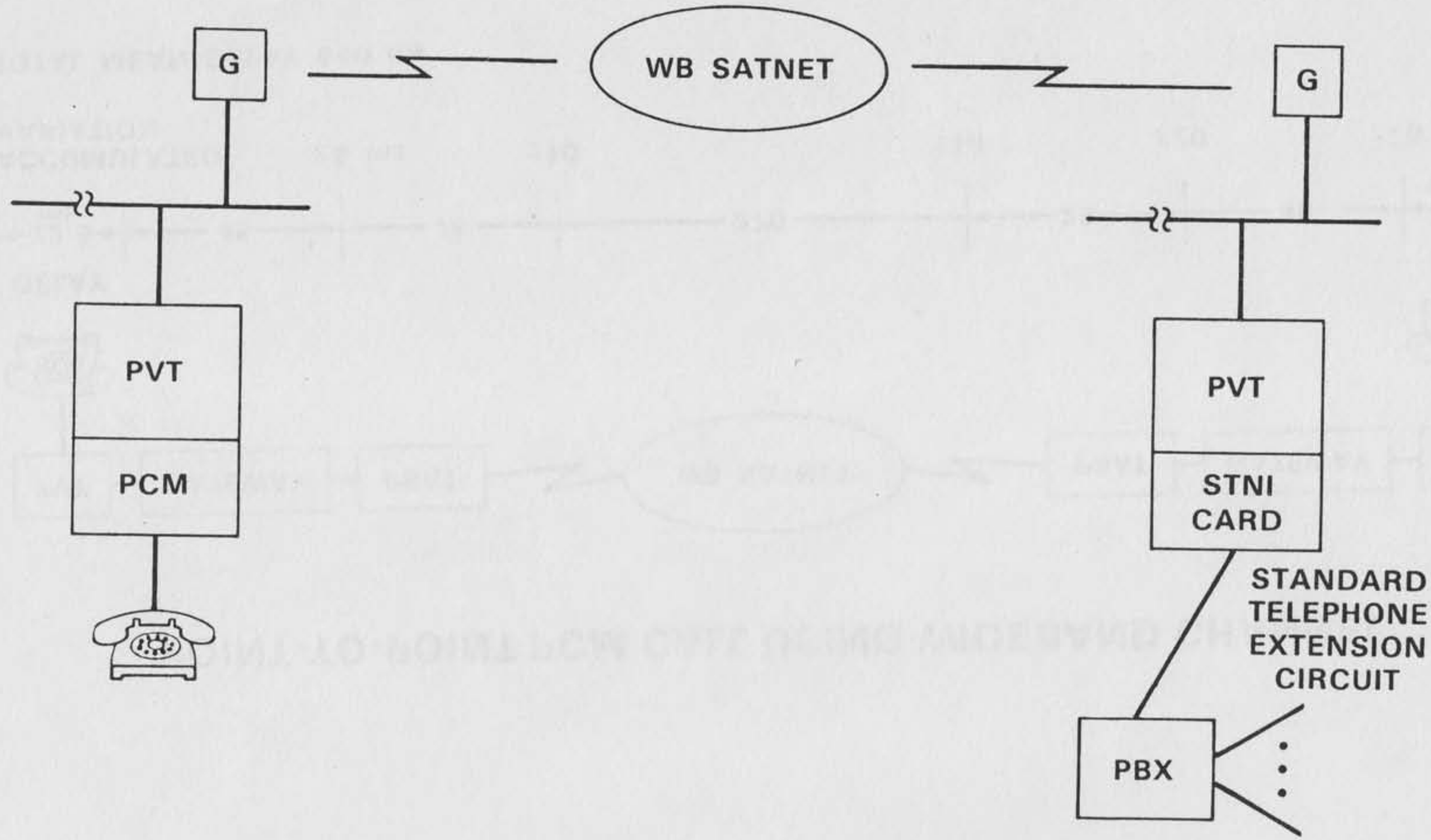


POINT-TO-POINT PCM CALL USING WIDEBAND CHANNEL

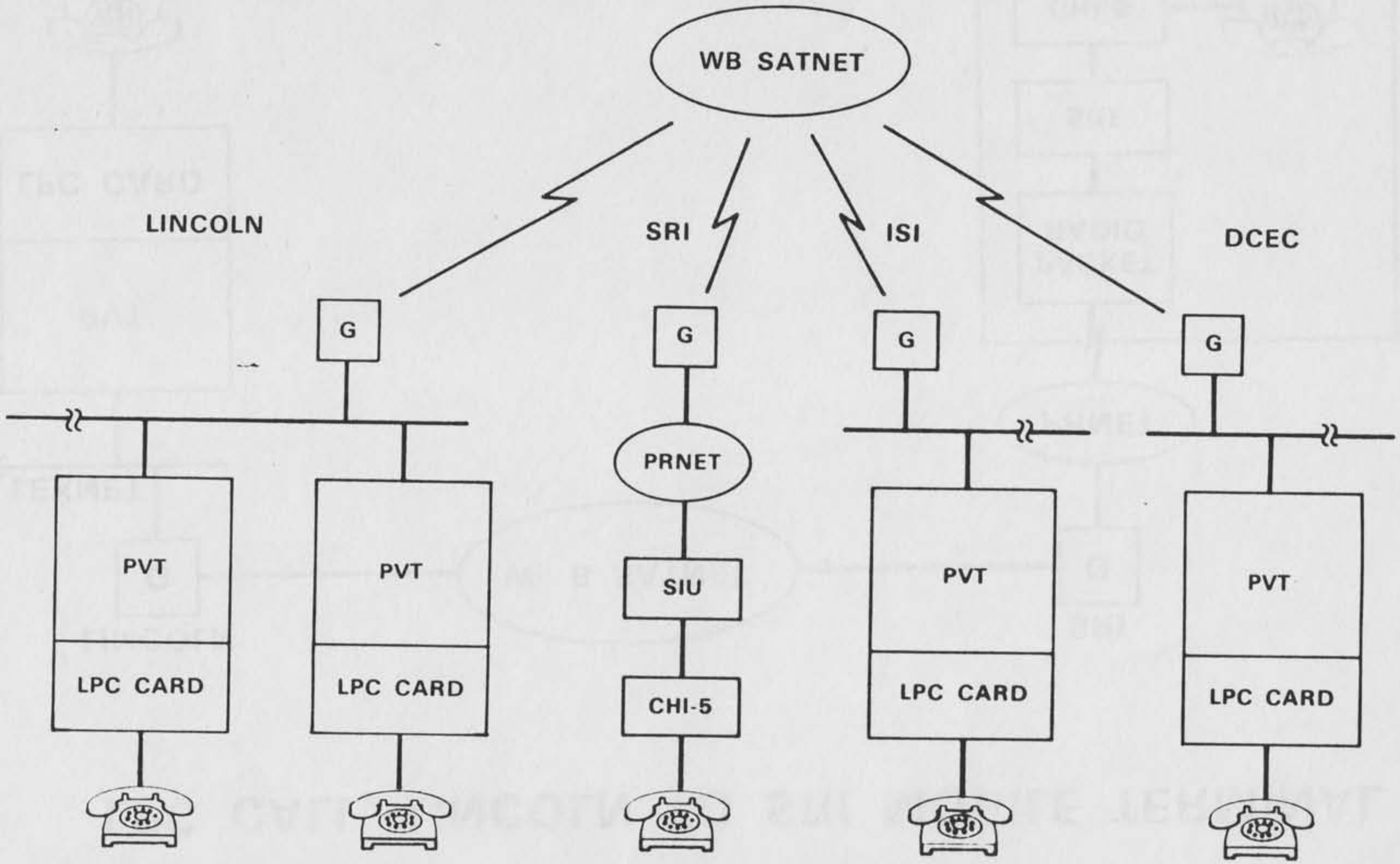


TOTAL MEAN DELAY 640 ms

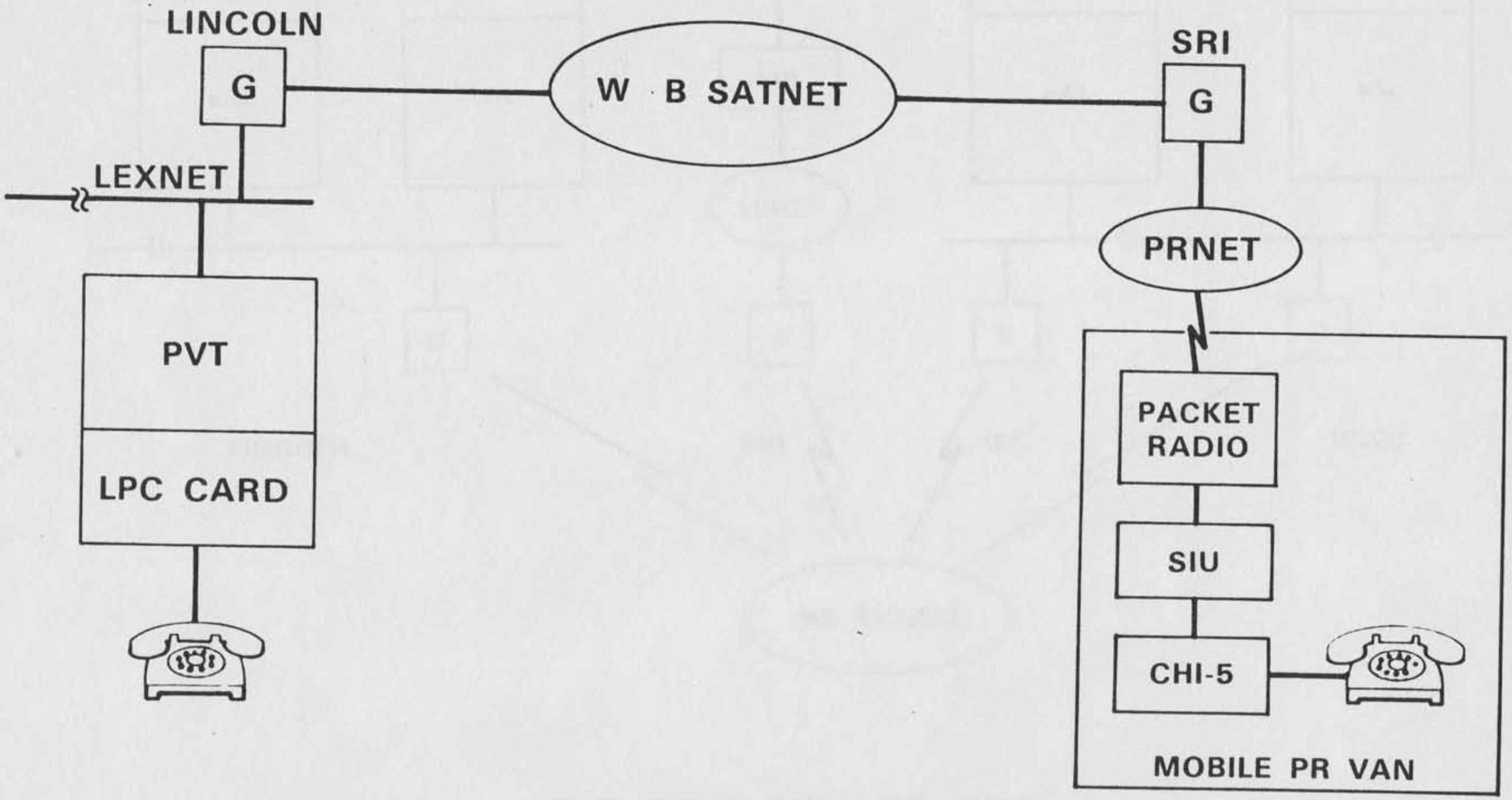
PCM CALL TO REMOTE TELEPHONE EXTENSION

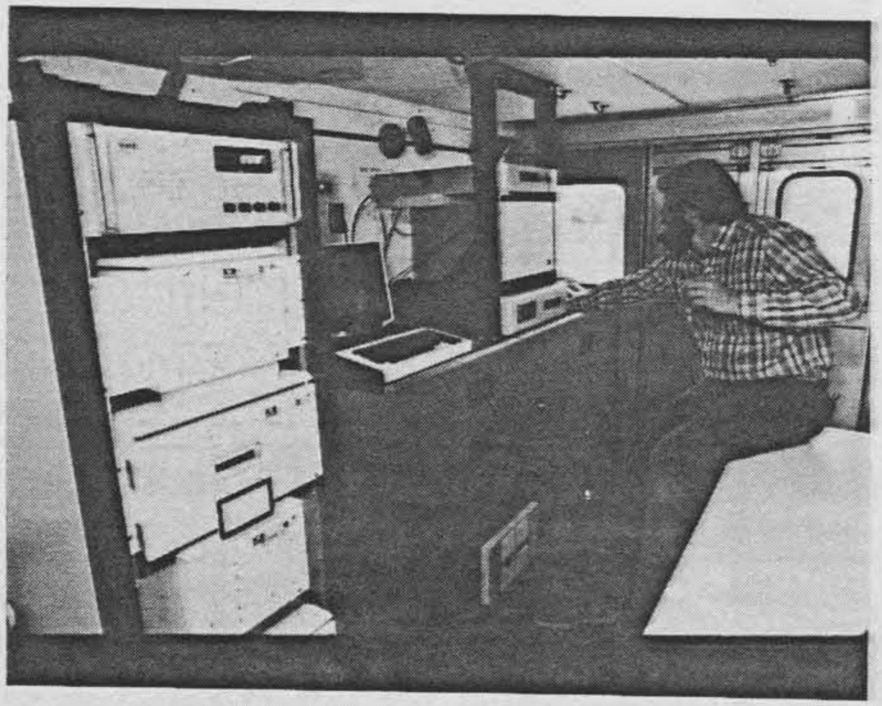


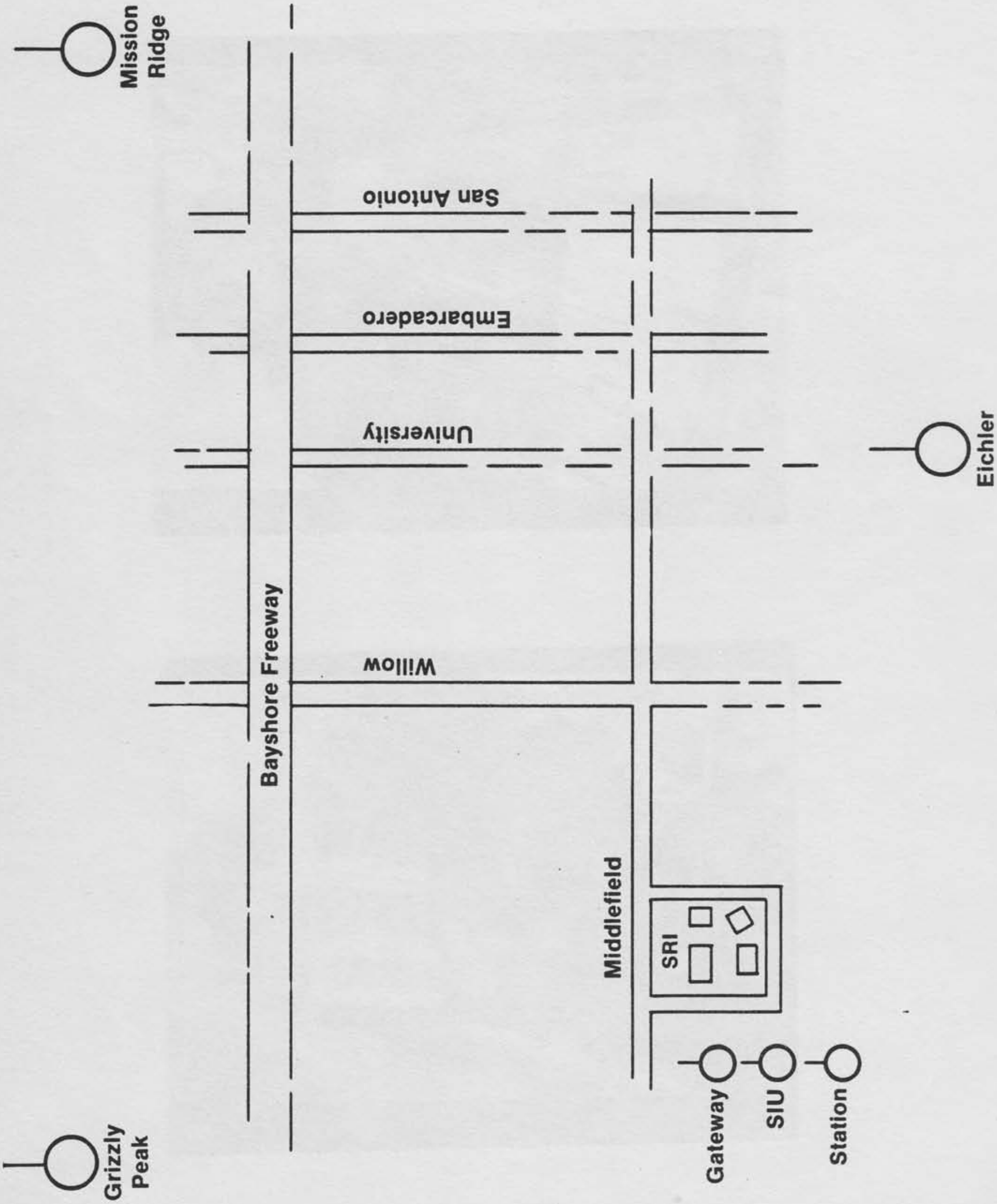
LPC CONFERENCE CALL



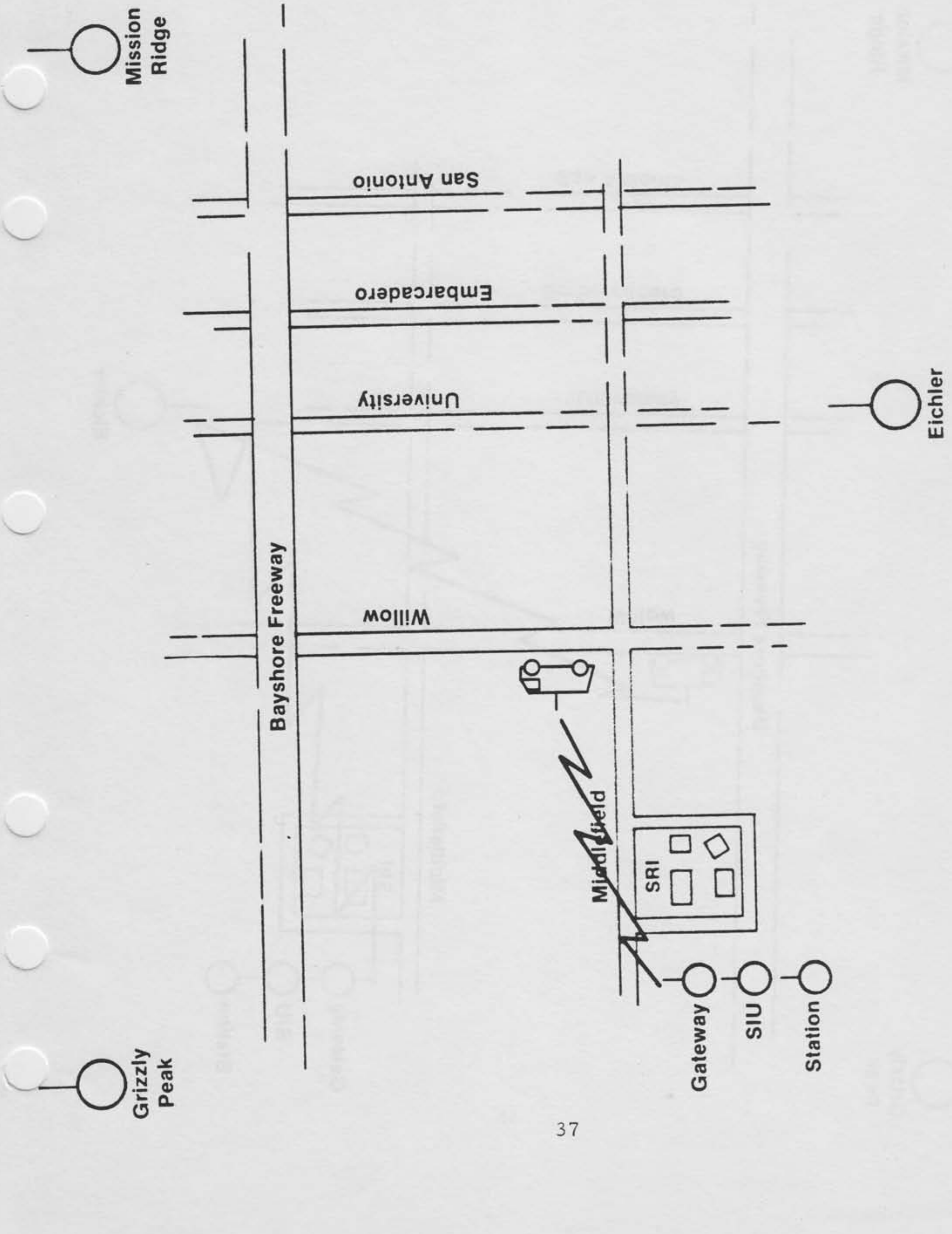
LPC CALL, LINCOLN TO SRI MOBILE TERMINAL

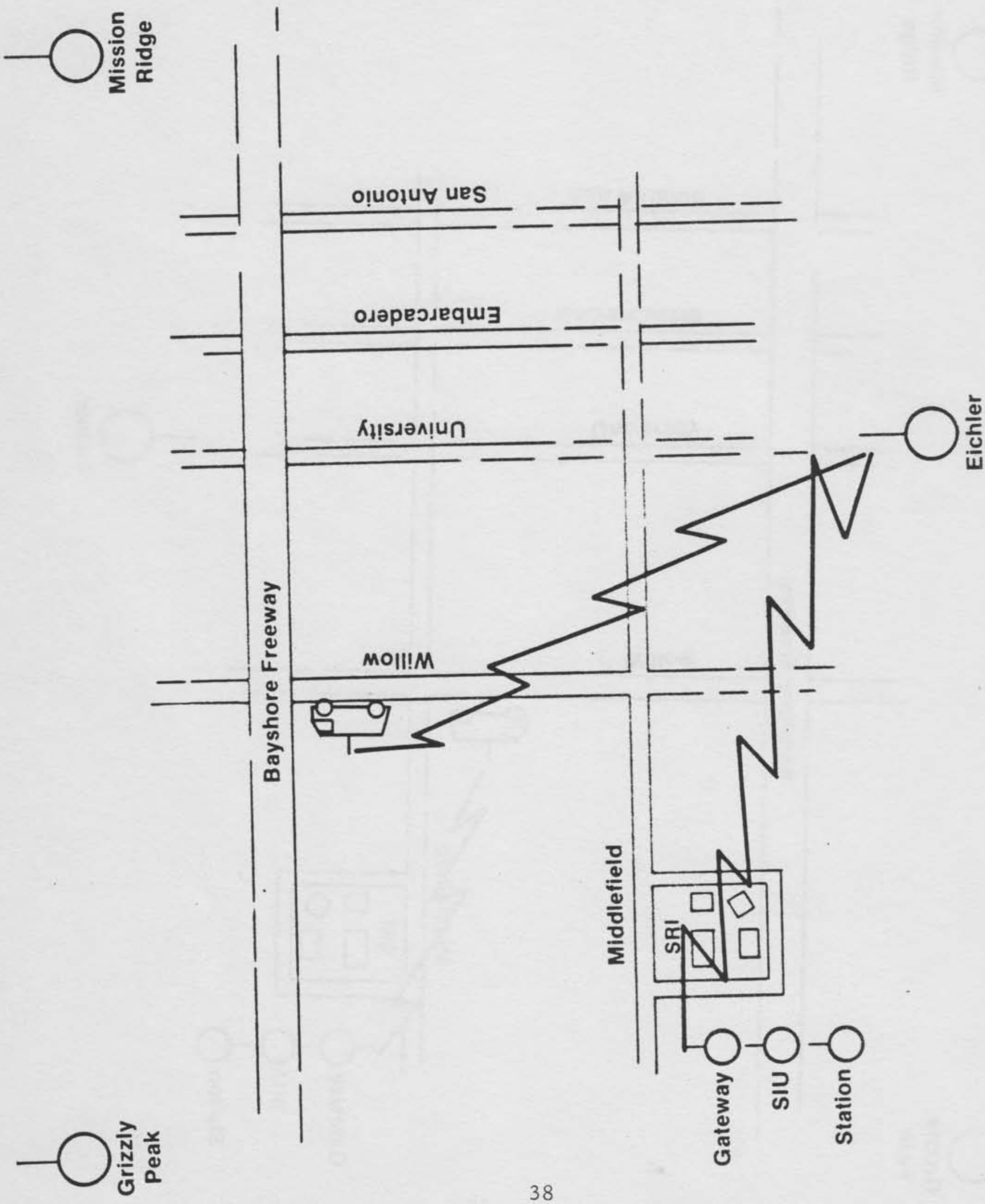


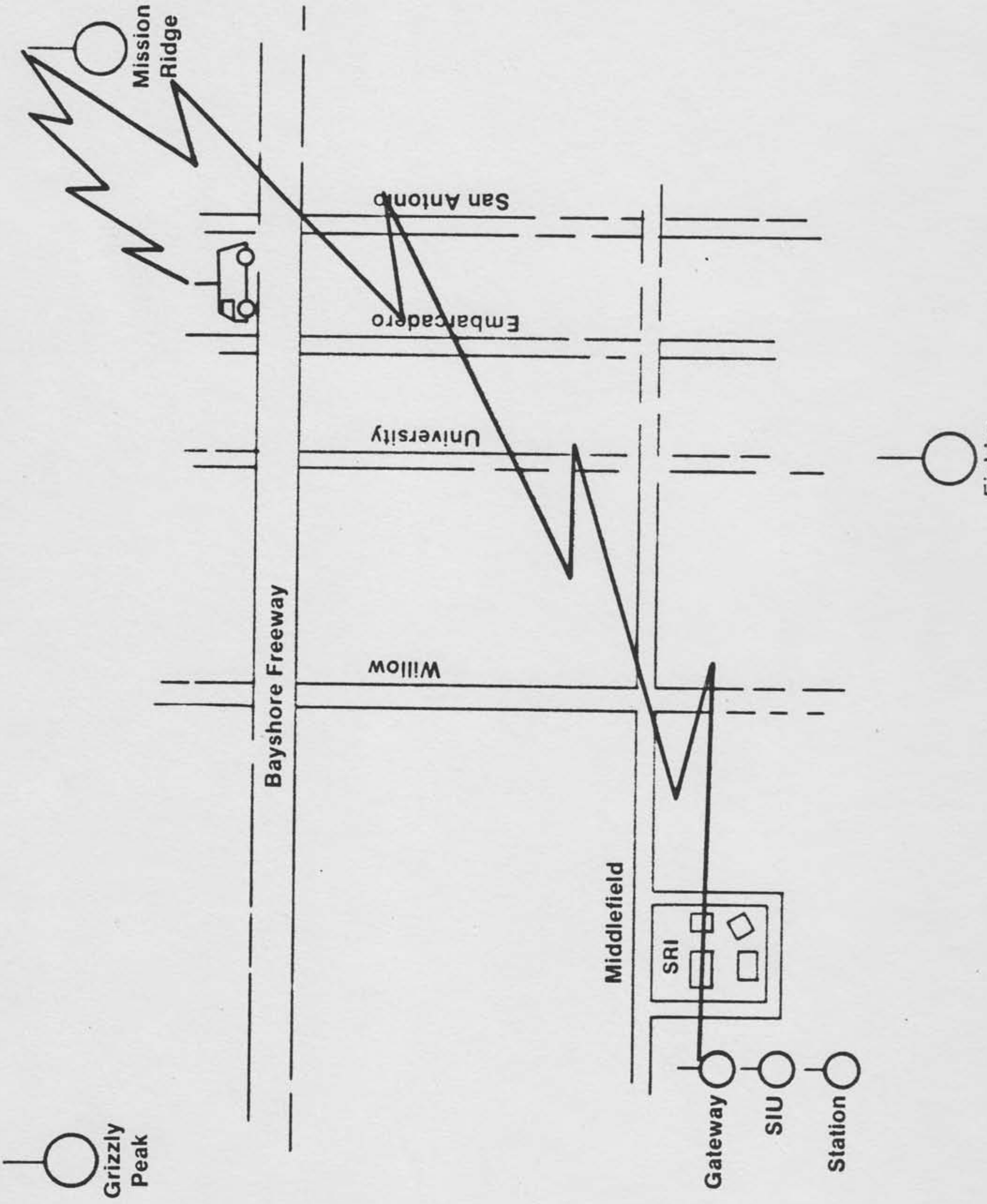




Mission Ridge
Grizzly Peak
Bayshore Freeway
San Antonio
Embarcadero
University
Willow
Middlefield
SRI
Eichler
Gateway
SIU
Station


Mission Ridge
Bayshore Freeway
San Antonio
Embarcadero
University
Eichler
Willow
Middlefield
SRI
Gateway
SIU
Station
Grizzly Peak


Mission Ridge
Grizzly Peak
Bayshore Freeway
San Antonio
Embarcadero
University
Willow
Middlefield
SRI
Eichler
Gateway
SIU
Station



Mission Ridge
San Antonio
Embarcadero
University
Willow
Bayshore Freeway
Middlefield
SRI
Eichler
Grizzly Peak
Gateway
SIU
Station

ISI

NETWORK VOICE PROTOCOLS

Danny Cohen

UNIVERSITY OF SOUTHERN CALIFORNIA
INFORMATION SCIENCES INSTITUTE
4676 Admiralty Way Marina del Rey California 90291

The History of Network Voice Protocols

In the beginning ARPA created the ARPAnet. Now the Network was formless and empty, darkness was over the surface of the deep....

And LGR said "Let there be protocols!" and there was NCP. ARPA saw that the NCP was good and said "Use it for Packet Voice!"...

Well,..... NCP was not exactly the right Host/Host protocol for online speech.

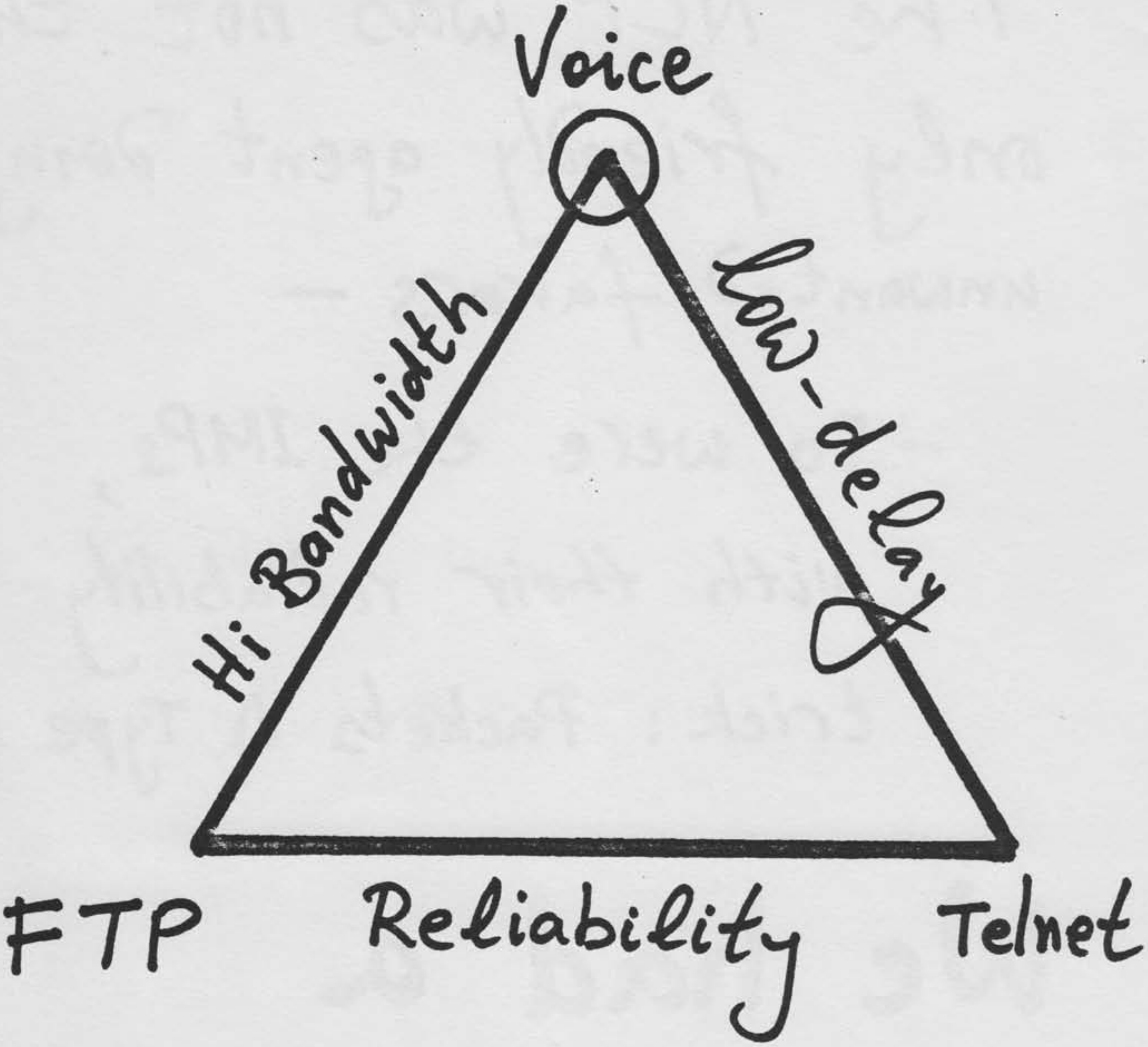
The NCP provided more favors than online-voice needed, at a non-trivial expense.

Therefore the need for NVP was recognized.

The NCP was not the only friendly agent doing unwanted favors —

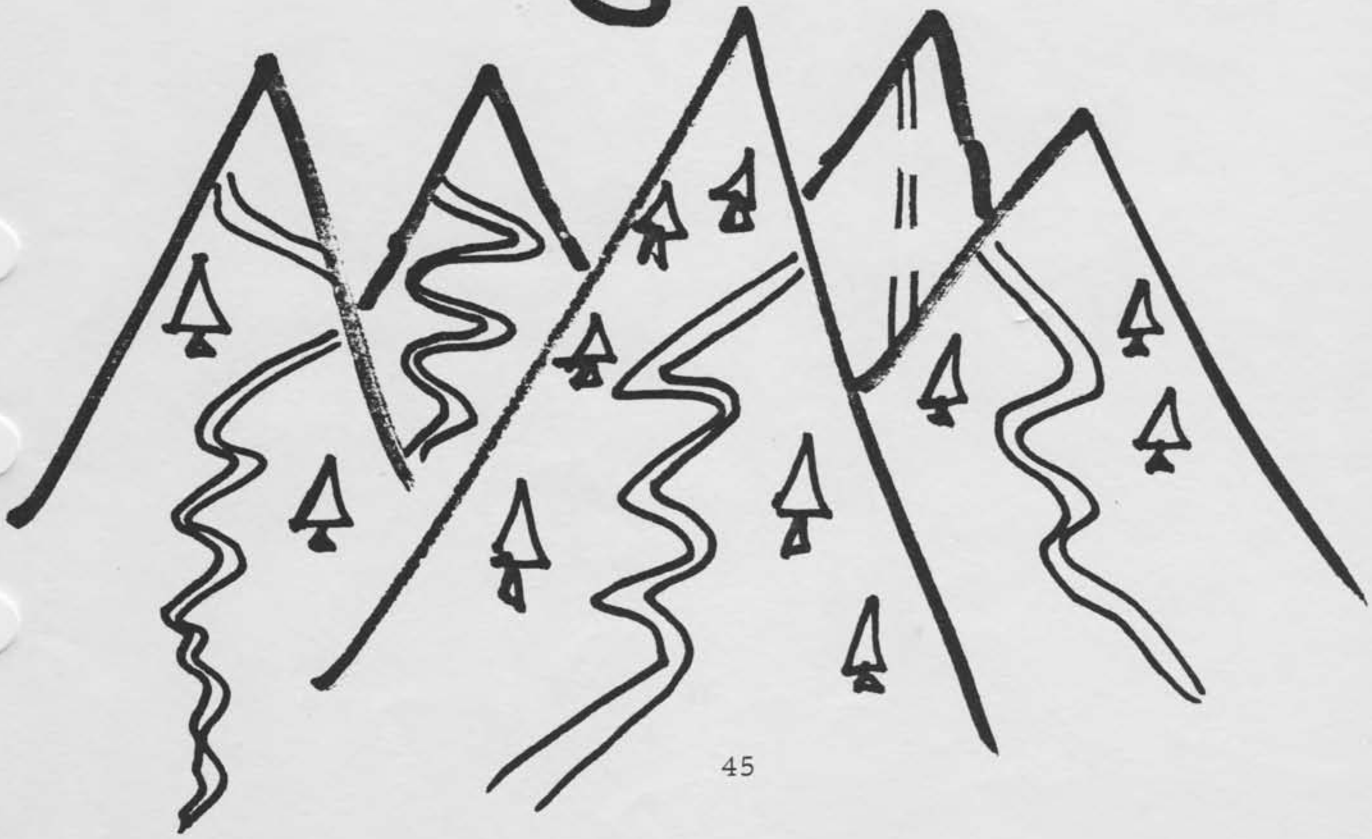
- So were the IMPs, with their reliability trick: Packets of Type 0

We had a problem...



The Arpanet protocols were not designed for real-time applications

So we had to think about it until the next meeting....



Objectives for NVP

* Real time data
* Extensible (vocoder indep)
* Network independent
* Separation of Data/Control

NVP-0 was the the first voice protocol, a precursor to NVP-1.

It was used for some CVSD experiments.

NVP-I was used for most of the experiments:

CVSD, LPC-I, LPC-II

NVP-I was extended to support Conferencing

NVCP

and voice-mail

When the PRNet was born we were expected to run online realtime voice to it by using

TCP

The thought of using TCP for this purpose was so underwhelming that the original TCP became

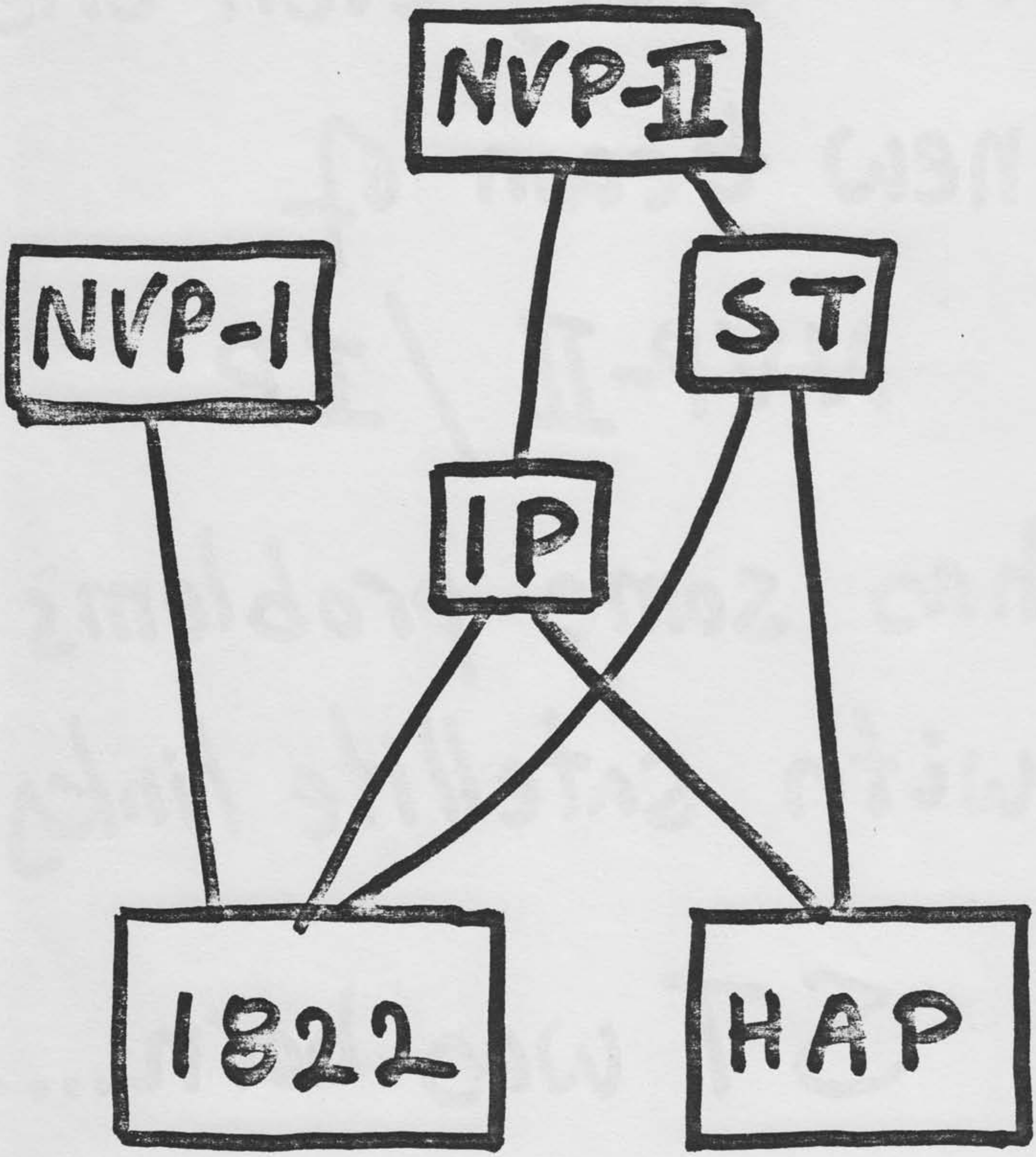
IP+TCP

NVP-I was modified to run on top of IP, and was also improved in some areas.

It became NVP-II

However, even the new team of NVP-II / IP has some problems with satellite links.

ST was born....


NVP-II
NVP-I
ST
IP
1822
HAP

LESSONS

packet communication is practical for online speech

Real-time data needs a different kind of protocol than the standard computer communication

(reliability, flow control...)

Motivated people and small committees can accomplish a lot.

Milestones in Packet Speech Communication

8/74: CVSD/ARPANet: ISI+LL

12/74: LPC/ARPANet: LL+CHI

1/76: LPC-CONFERENCE/ARPANet: ISI+LL+CHI+SRI

4/77: Flanagan (BTL) applies for a patent on packet transmission of speech

7/78: USA patent 4,100,377 granted

SUMMARY

The interest of the various phone companies in packet voice is a significant testimonial to our

SUCCESS

SPEECH CONFERENCING

JAMES W. FORGIE
MIT LINCOLN LABORATORY

OUTLINE

1. VOICE CONFERENCING SYSTEM DESIGN ISSUES

2. HISTORY OF PACKET CONFERENCING

3. CURRENT WB NET CONFERENCING SYSTEM

VOICE CONFERENCING SYSTEM DESIGN CHOICES

1. SUMMATION OR SIGNAL SELECTION

2. VOICE OR PUSH-BUTTON CONTROL

3. CENTRALIZED OR DISTRIBUTED CONTROL

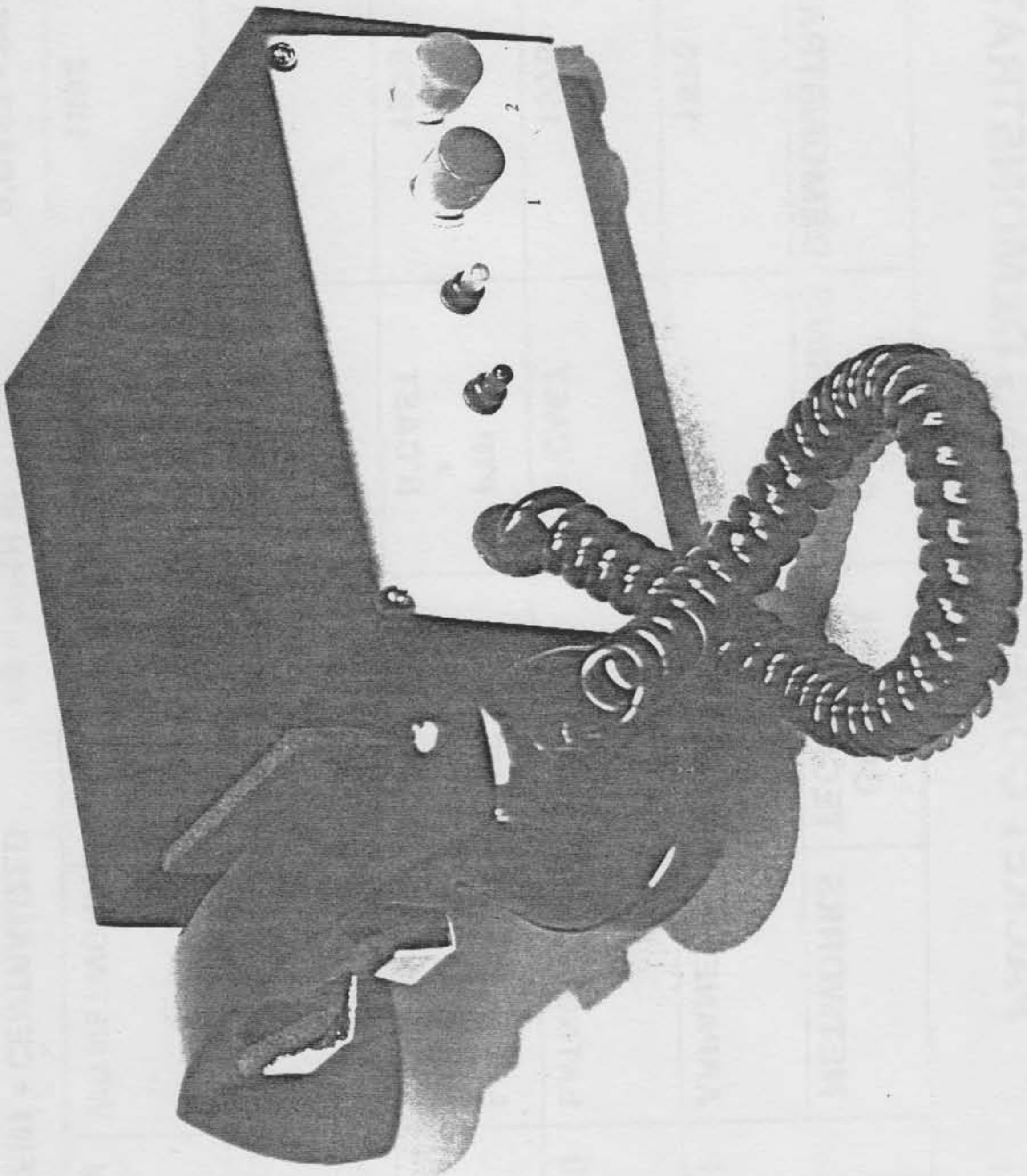
PACKET CONFERENCING DEMONSTRATIONS

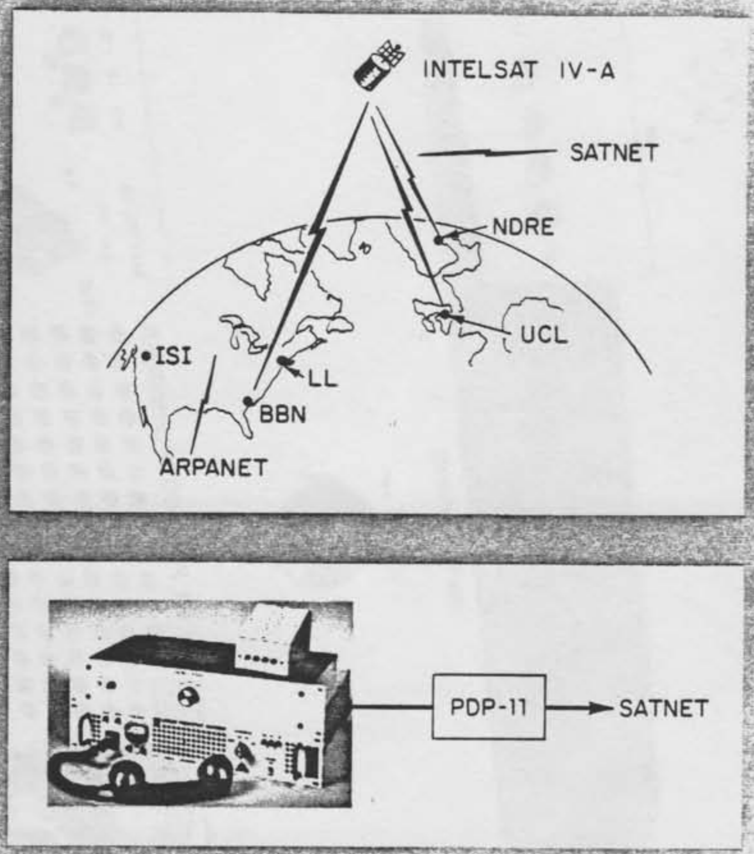
	NETWORKS	CONTROL TECHNIQUES		PACKET ADDRESSING	DEMONSTRATED	SITES
I	ARPANET	CENT	PB	PTP	1976	CHI, ISI, SRI, LL
II	SATNET	DIST	PB	B'CAST	1978	NDRE, UCL, BBN
III	ARPANET + SATNET	CENT + DIST	PB	PTP + B'CAST	1979	ISI, LL + NDRE, UCL
IV	SATNET	DIST	VOICE	B'CAST	1979	NDRE, UCL, BBN
V	WB NETWORK	DIST	VOICE	B'CAST	1982	ISI, SRI, LL

CENT = CENTRALIZED
DIST = DISTRIBUTED

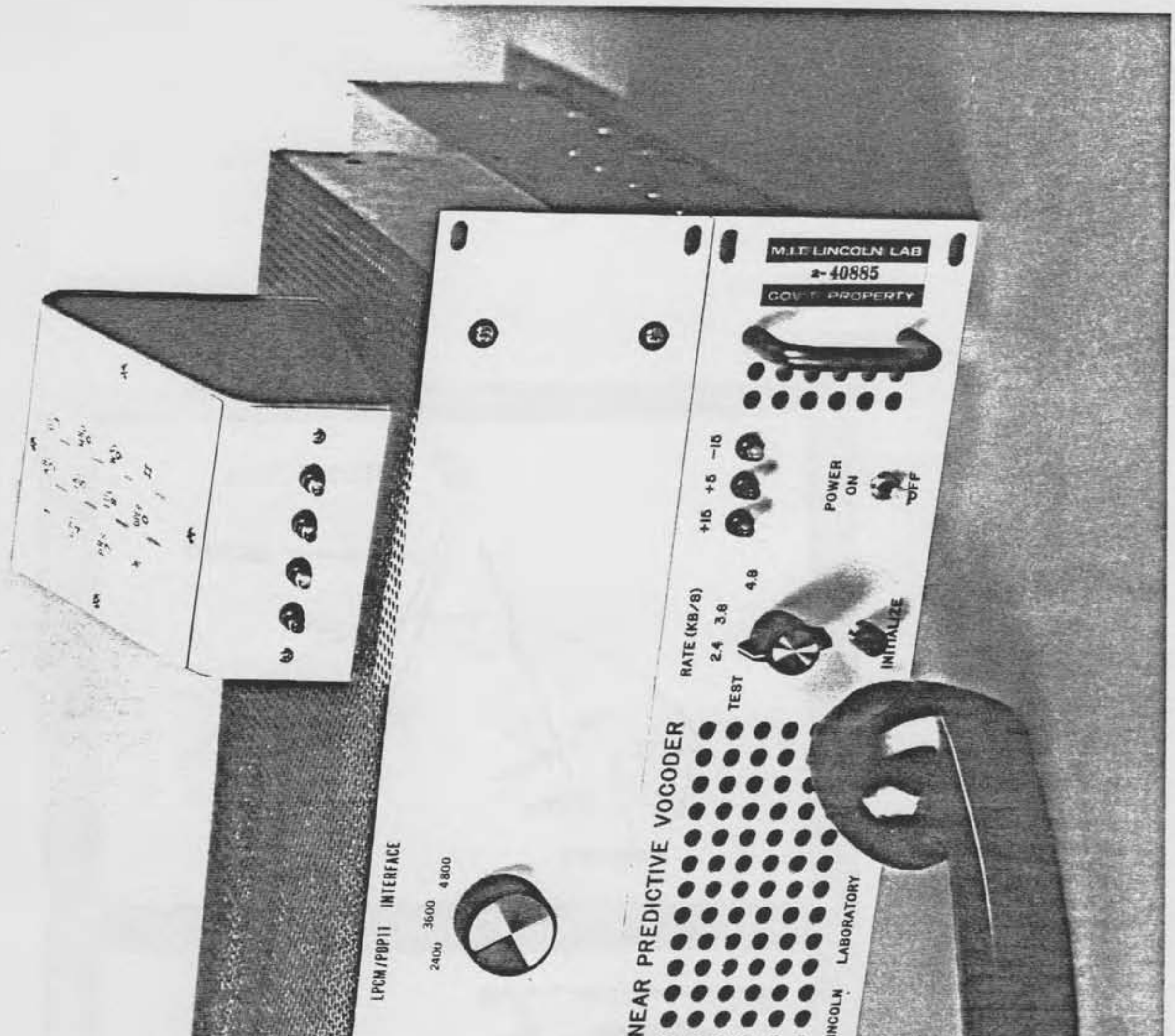
PB = PUSH BUTTON
PTP = POINT-TO-POINT

B'CAST = BROADCAST

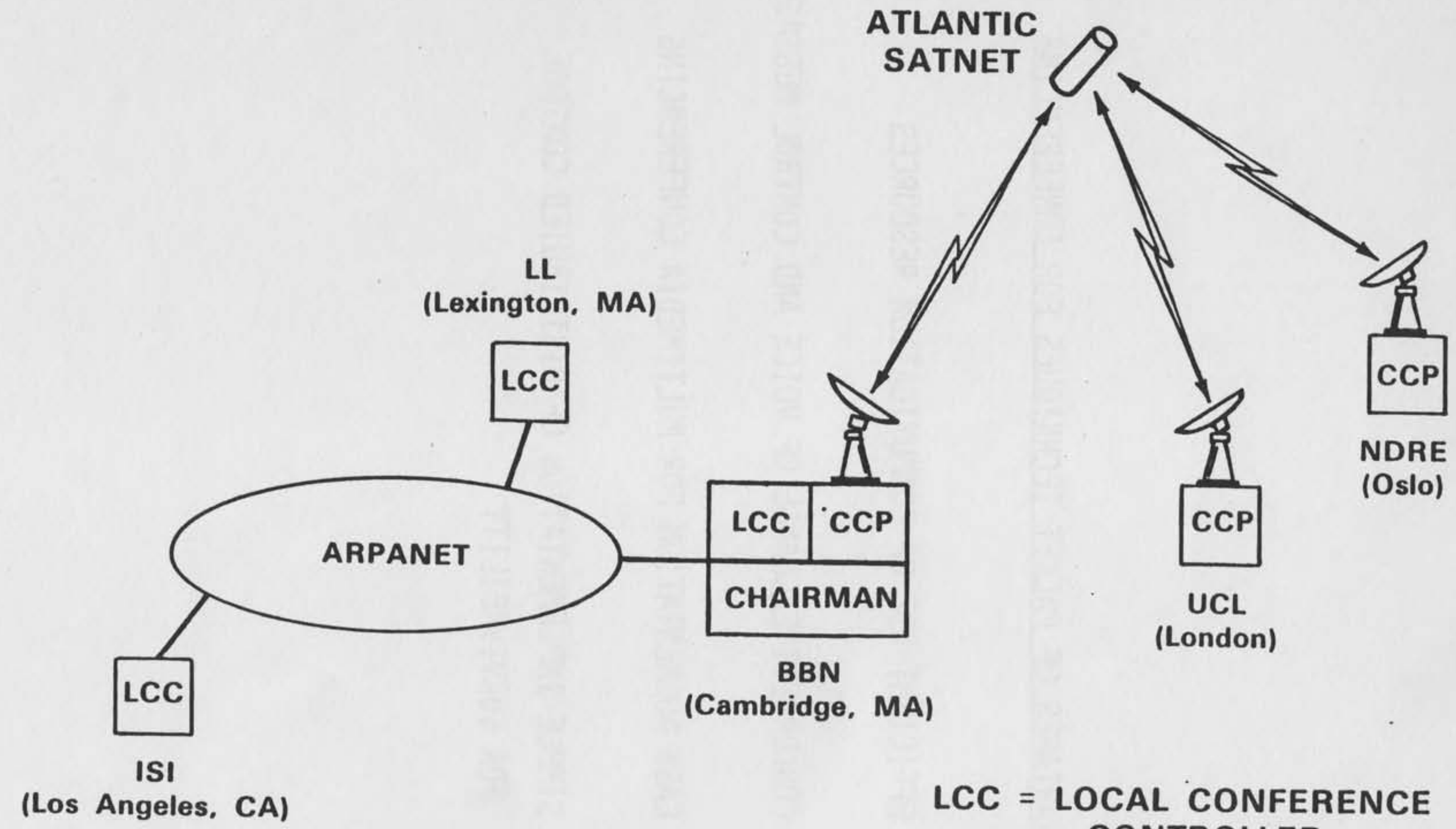




PACKET CONFERENCING EXPERIMENTS


LPCM/PDP11 INTERFACE
2400 3600 4800
ON
AC
RCVE
SEND
SEND LEVEL
VU
HANDSET
TEST
INT. CLK EXT
LINEAR PREDICTIVE VOCODER
M.I.T. LINCOLN LABORATORY
RATE (KB/S)
2.4 3.6 4.8
TEST
INITIALIZE
+15 +5 −15
POWER ON
OFF
M.I.T. LINCOLN LAB
2-40885
GOV'T PROPERTY

SATNET/INTERNET VOICE CONFERENCING



LCC = LOCAL CONFERENCE CONTROLLER

CCP = CONFERENCE CONTROL PROGRAM

ADVANTAGES OF PACKET TECHNIQUES FOR CONFERENCING

1. EFFICIENT USE OF COMMUNICATION RESOURCES

2. CONCURRENT EXCHANGE OF VOICE AND CONTROL MESSAGES

3. EASY AUGMENTATION FOR MULTIMEDIA CONFERENCING

4. SIMPLE IMPLEMENTATION OF DISTRIBUTED CONTROL FOR SURVIVABILITY

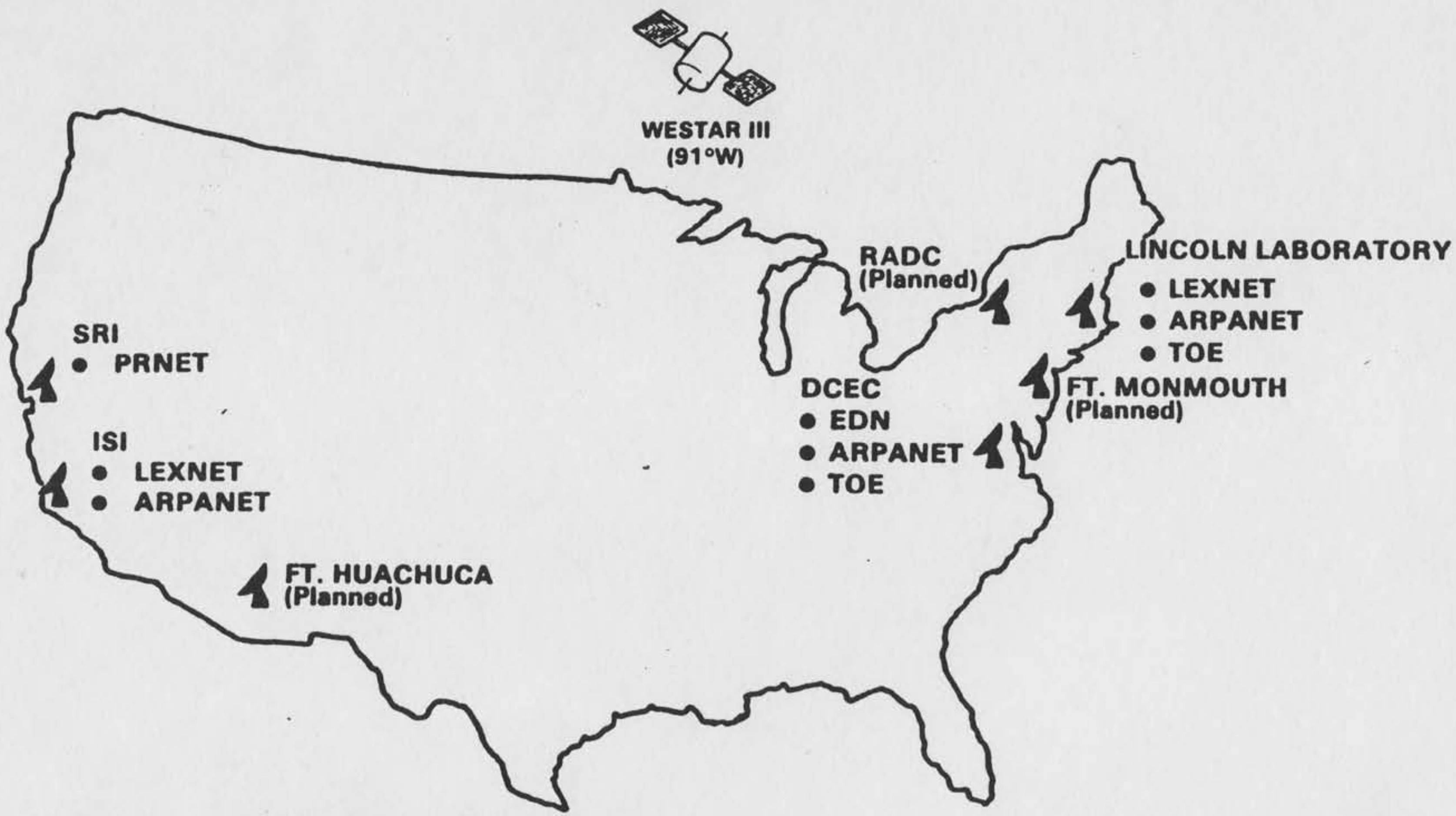
WIDEBAND INTEGRATED VOICE/DATA NETWORK

CLIFFORD J. WEINSTEIN
MIT LINCOLN LABORATORY

WIDEBAND INTEGRATED VOICE/DATA NETWORK

- MOTIVATION FOR WIDEBAND EXPERIMENT
- EXPERIMENTAL SYSTEM DESCRIPTION
- CHANNEL ASSIGNMENT AND PACKET MULTIPLEXING
- EXPERIMENT STATUS

WB SATNET TOPOLOGY

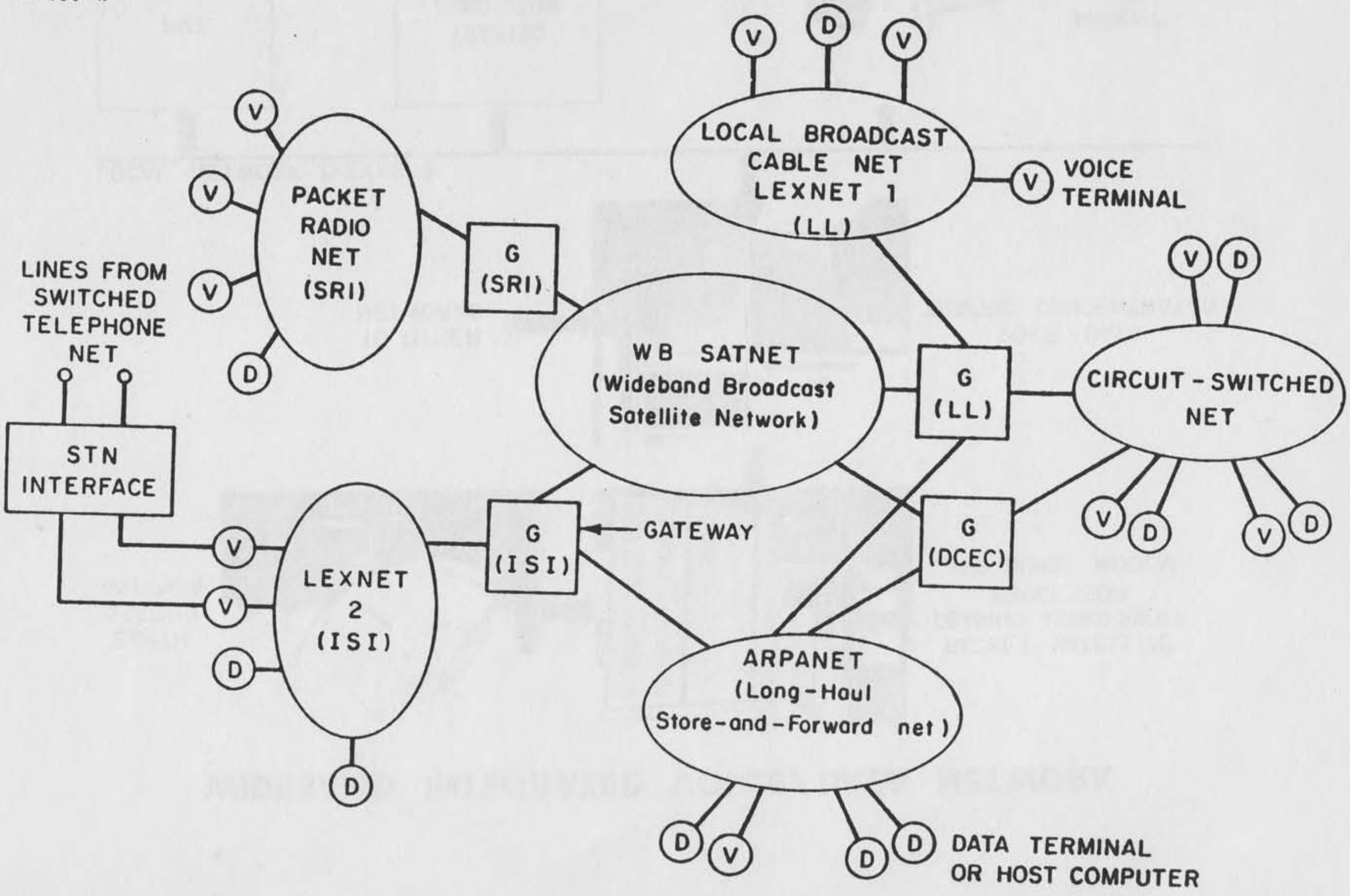


LEXNET = LOCAL PACKET VOICE NETWORK
PRNET = PACKET RADIO NET
TOE = TELEPHONE OFFICE EMULATOR

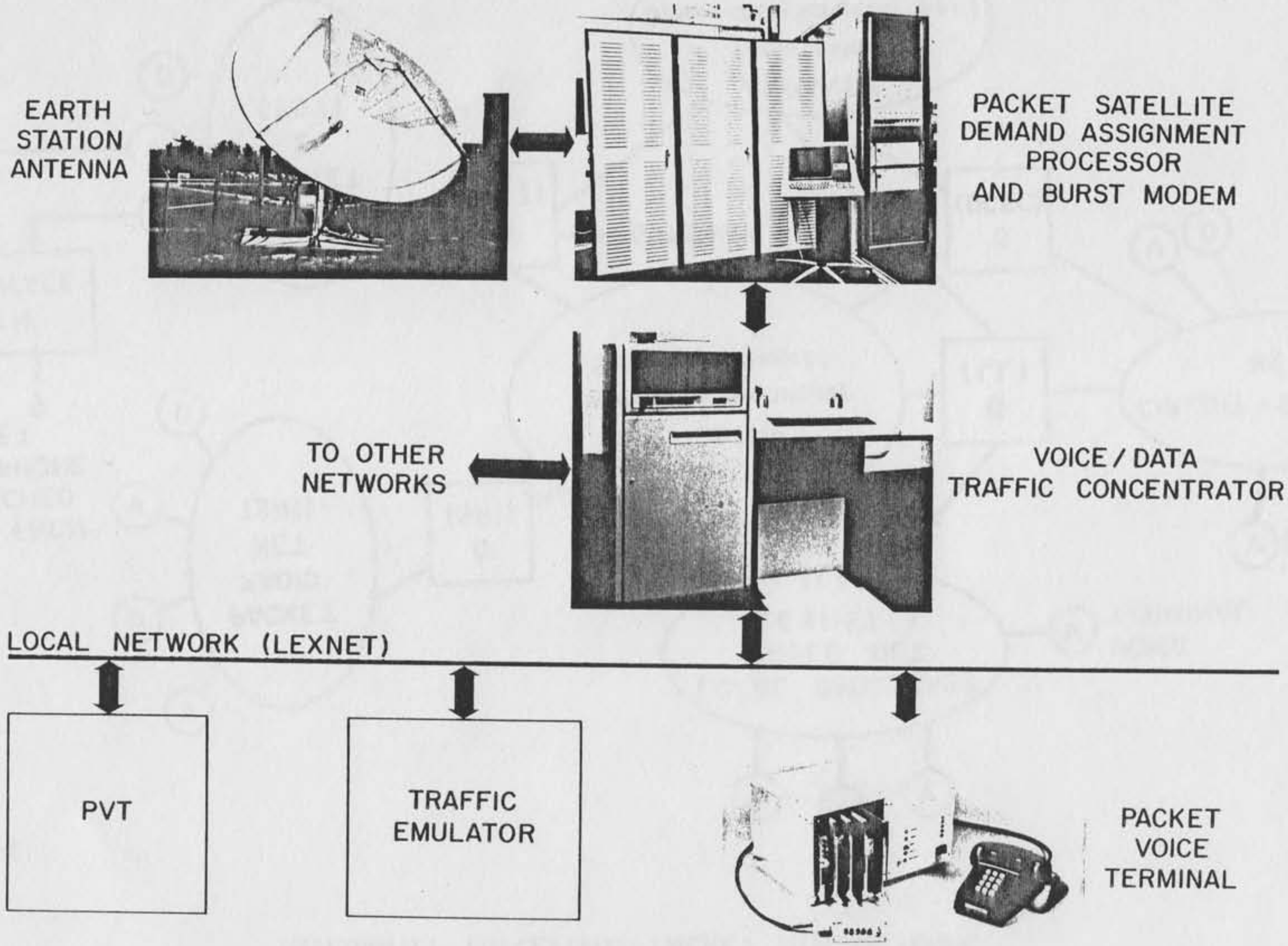
114814-N-01

INTEGRATED VOICE/DATA PACKET INTERNETWORK

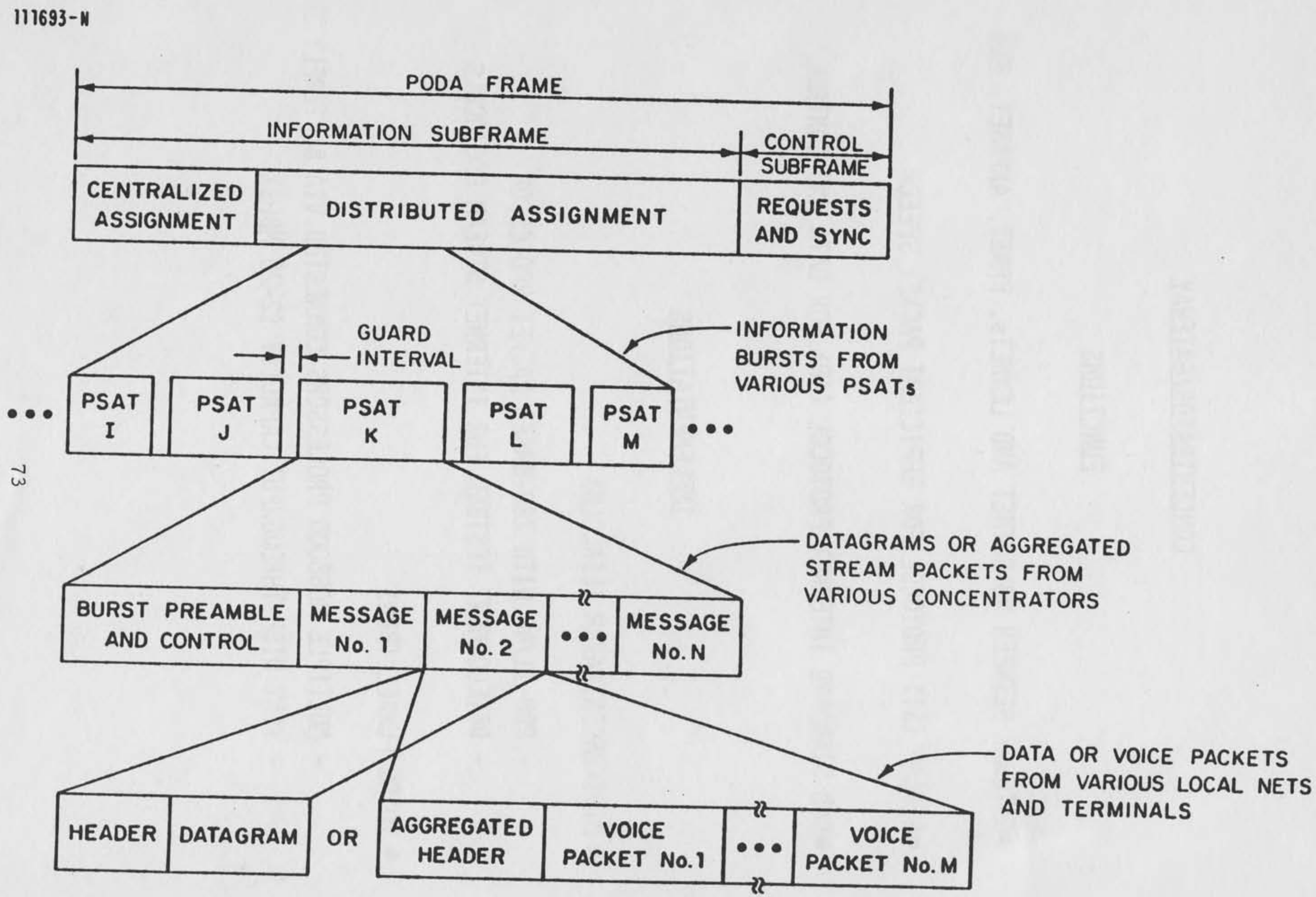
111686-N



WIDEBAND INTEGRATED VOICE / DATA NETWORK



PACKET MULTIPLEXING HIERARCHY



CONCENTRATOR/GATEWAY

FUNCTIONS

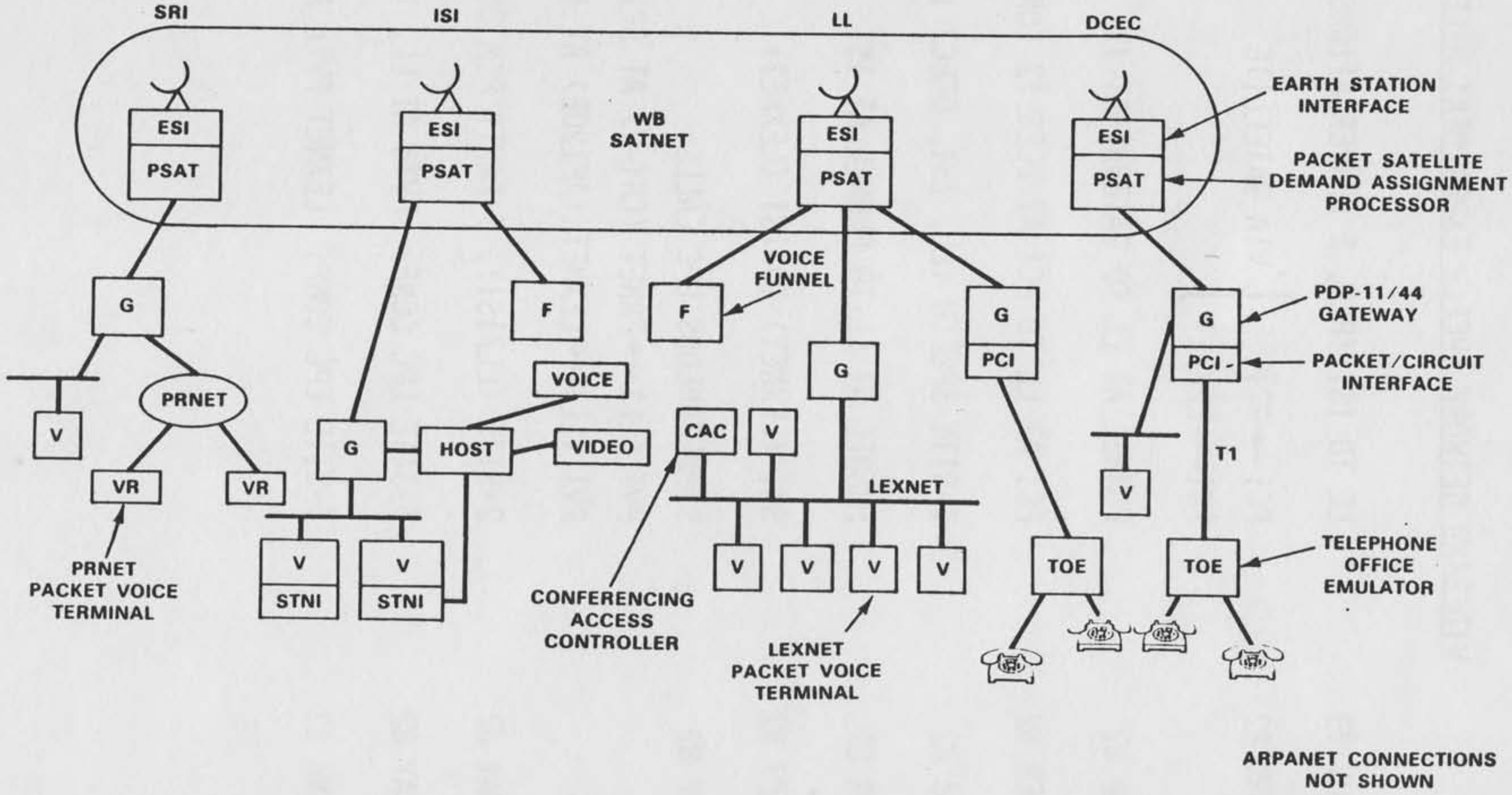
- GATEWAY BETWEEN WB SATNET AND LEXNETs, PRNET, ARPANET, EDN
- STREAM (ST) PROTOCOL FOR EFFICIENT PACKET SPEECH
- DoD STANDARD INTERNET PROTOCOL (IP) FOR DATA AND CONTROL

IMPLEMENTATIONS

- MINICONCENTRATOR (LINCOLN)
  - PDP-11/44 WITH Z80-BASED PACKET PROCESSOR
  - DEVELOPMENT TESTBED FOR INTERNET STREAM PROTOCOLS
- VOICE FUNNEL (BBN)
  - MULTIPLE MC68000 PROCESSORS CONNECTED VIA BUTTERFLY SWITCH
  - VERY HIGH THROUGHPUT CAPACITY (2-20 MBPS)

117750-N

PACKET SPEECH EXPERIMENT STATUS — JUNE 1982



WIDEBAND NETWORK SPEECH EXPERIMENT MILESTONES

18 NOV 81	LL TO ISI (PCM, 2 CONVERSATIONS, ONE USING STNI AT ISI)
22 JAN 82	PCI → PCI, PCI → LEXNET } VIA SATELLITE { 2 HOSTS ON LL PSAT, 2 CONVERSATIONS
3 FEB 82	FUNNEL AT LL TO GATEWAY AT ISI (PCM)
22 FEB 82	PCI AT LL TO PCI AT DCEC (2 CONVERSATIONS)
4 MAR 82	3-SITE SPEECH (LL, ISI, DCEC) IN 2-PARTY COMBINATIONS
9 APR 82	FUNNEL AT LL TO FUNNEL AT ISI
22 APR 82	SRI (LEXNET) TO ISI (LEXNET)
7 MAY 82	SIMULTANEOUS LPC CALLS PVT (LL) → PRNET (CHI-V) AT SRI PVT (LL) → LEXNET (AP120B) AT ISI
13 MAY 82	2-SITE (LL/ISI), 3-PARTY PCM CONF
21 MAY 82	3-SITE LPC CONF: PVTs AT LL, ISI, SRI
1 JUNE 82	3-SITE LPC CONF: LEXNET PVTs AT LL, ISI; PRNET AT SRI

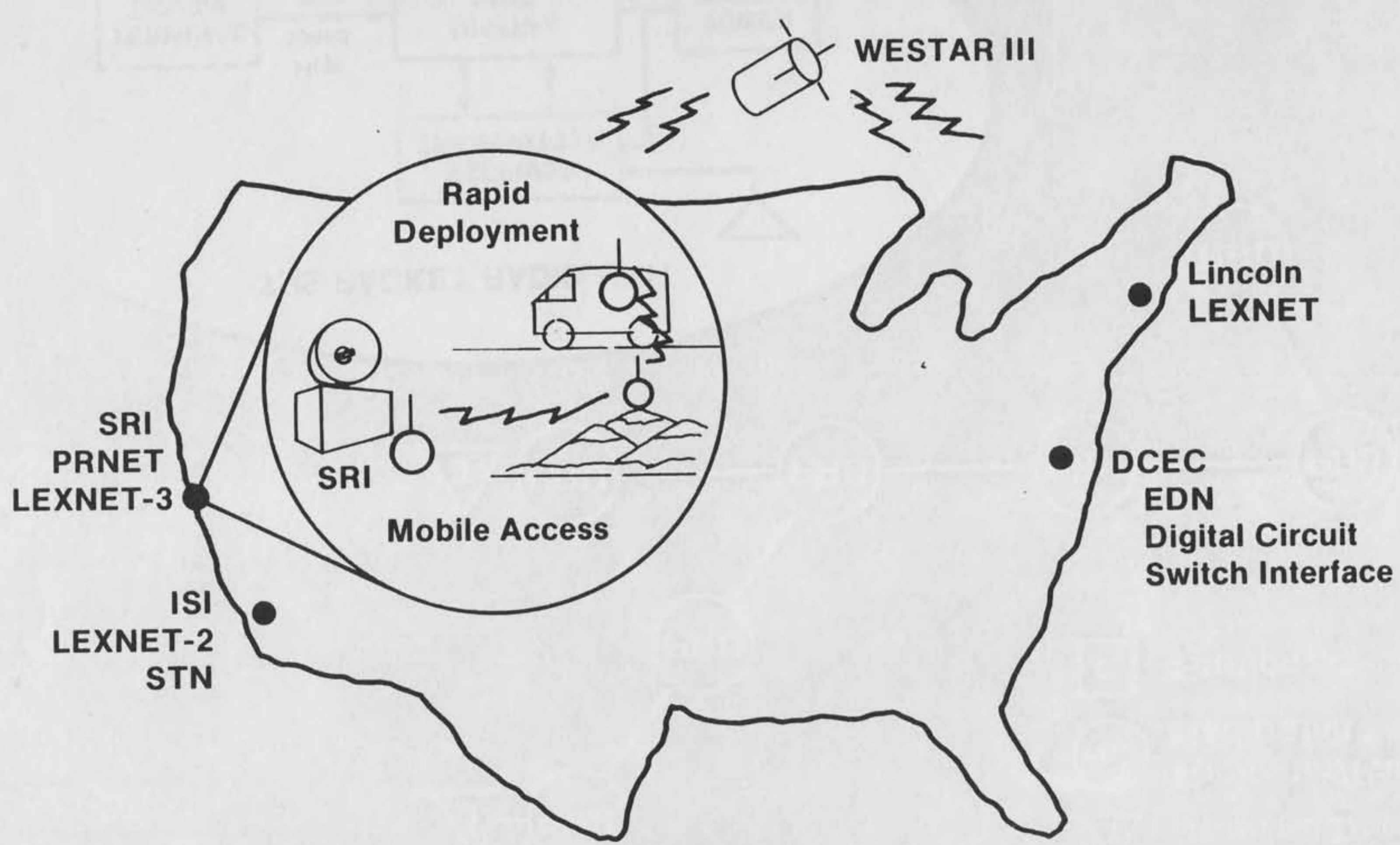
FUTURE DIRECTIONS

- MULTI-USER, MULTI-SITE EXPERIMENTS
  - MULTIPLEXING TO TEST SYSTEM LIMITS
  - ADAPTIVE TRAFFIC CONTROL OF VOICE AND DATA
  - INTERNET VOICE CONFERENCING
  - VOICE-CONTROLLED SYSTEMS
- PACKET VIDEO DEVELOPMENT/EXPERIMENTS
- SECURE PACKET SPEECH SYSTEM DEVELOPMENT/EXPERIMENTS

PACKET RADIO SPEECH

EARL J. CRAIGHILL
SRI INTERNATIONAL

EXPERIMENTAL WIDEBAND VOICE/DATA SATELLITE NETWORK

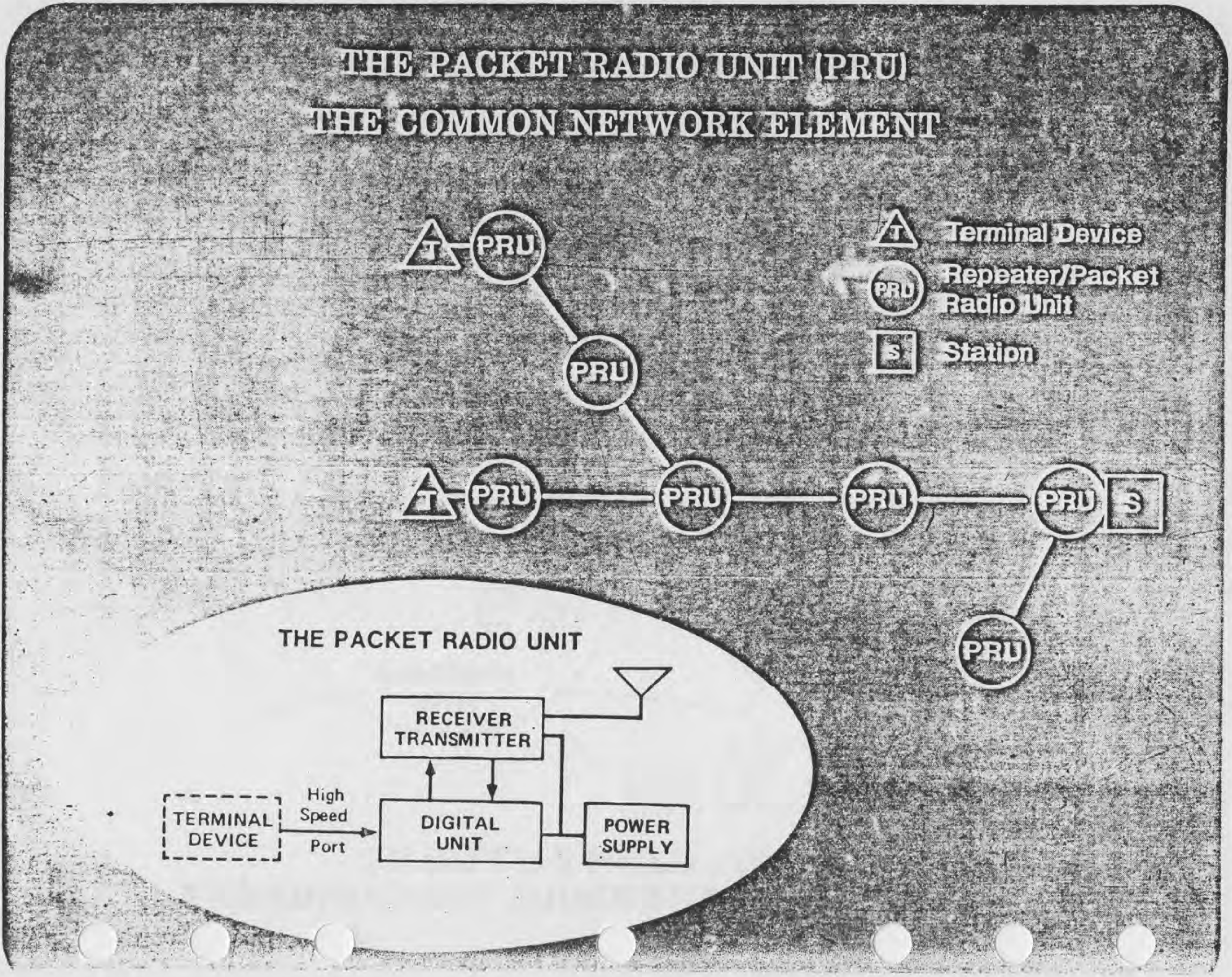


Subsystems at Each Satellite Net Node

— Earth Station (RF)
— Flexible Burst Modem
— Demand Assignment Processor

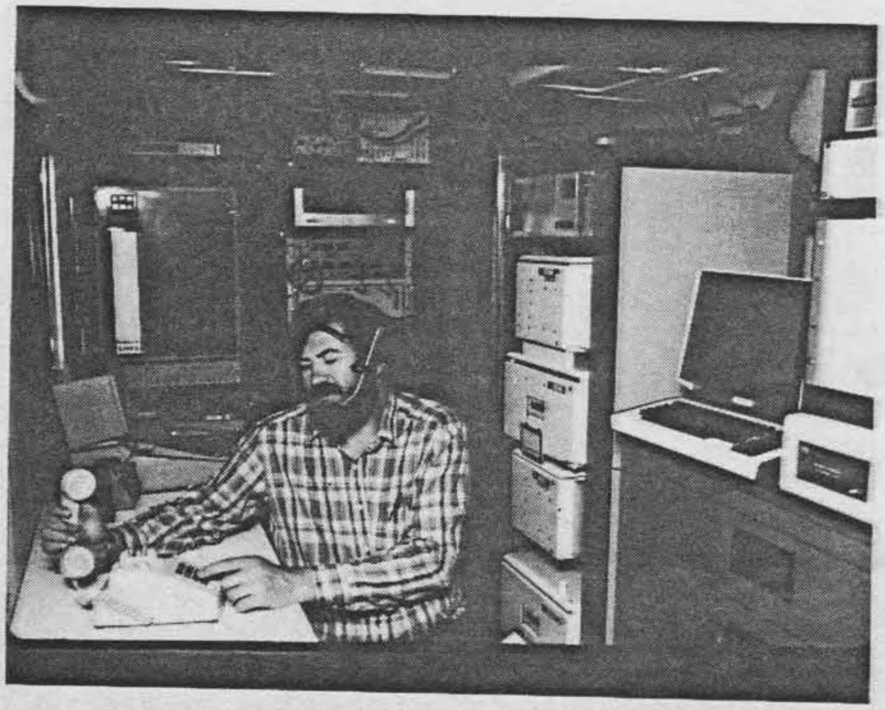
FEATURES

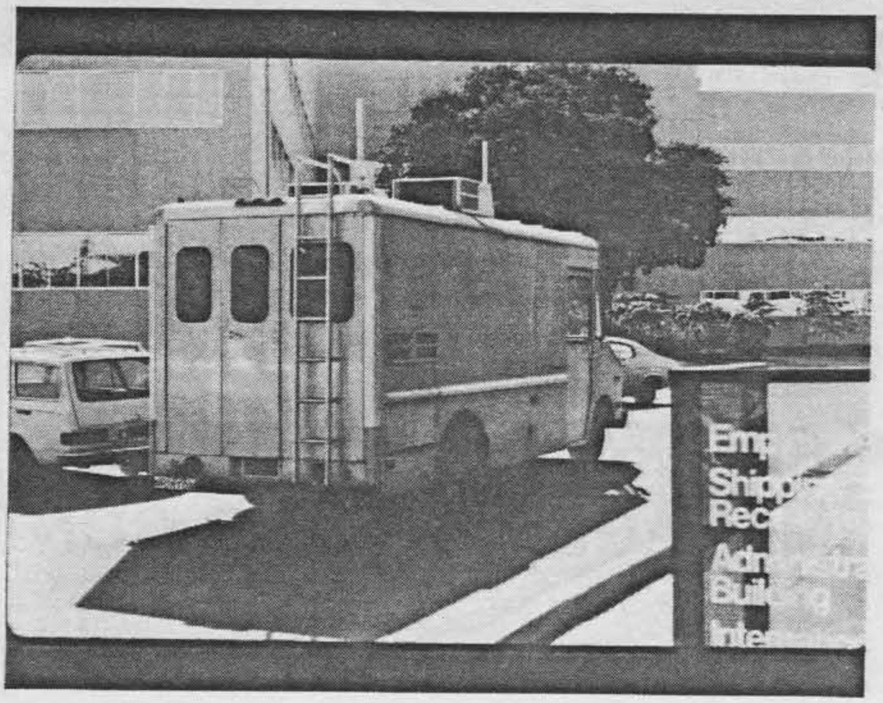
— Multiuser Voice/Data
— Satellite/Terrestrial
— Distrubuted Control
— Internet/Security Protocols
— Mobile Access for Voice
— Flexibility for Experiments


THE PACKET RADIO UNIT (PRU)
THE COMMON NETWORK ELEMENT
T
PRU
S
Terminal Device
Repeater/Packet Radio Unit
Station
THE PACKET RADIO UNIT
RECEIVER TRANSMITTER
TERMINAL DEVICE
High Speed Port
DIGITAL UNIT
POWER SUPPLY

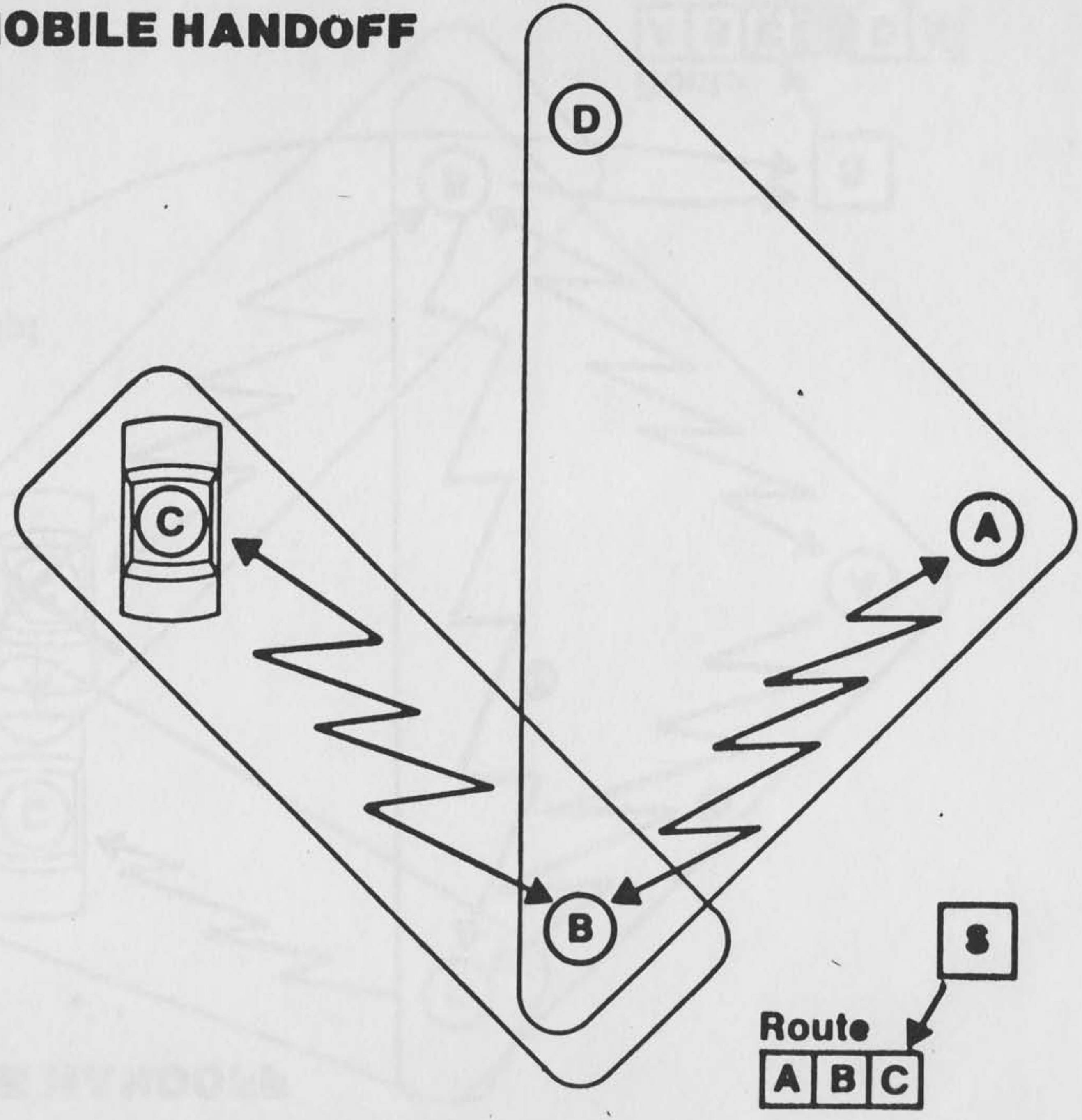
PACKET RADIO CONCEPTS

- BROADCAST RADIO
- PACKET-SWITCHED
- MULTIPLE ACCESS
- MULTIPLE-HOP
- MOBILE OPERATION
- AUTOMATED NETWORK MANAGEMENT AND CONTROL

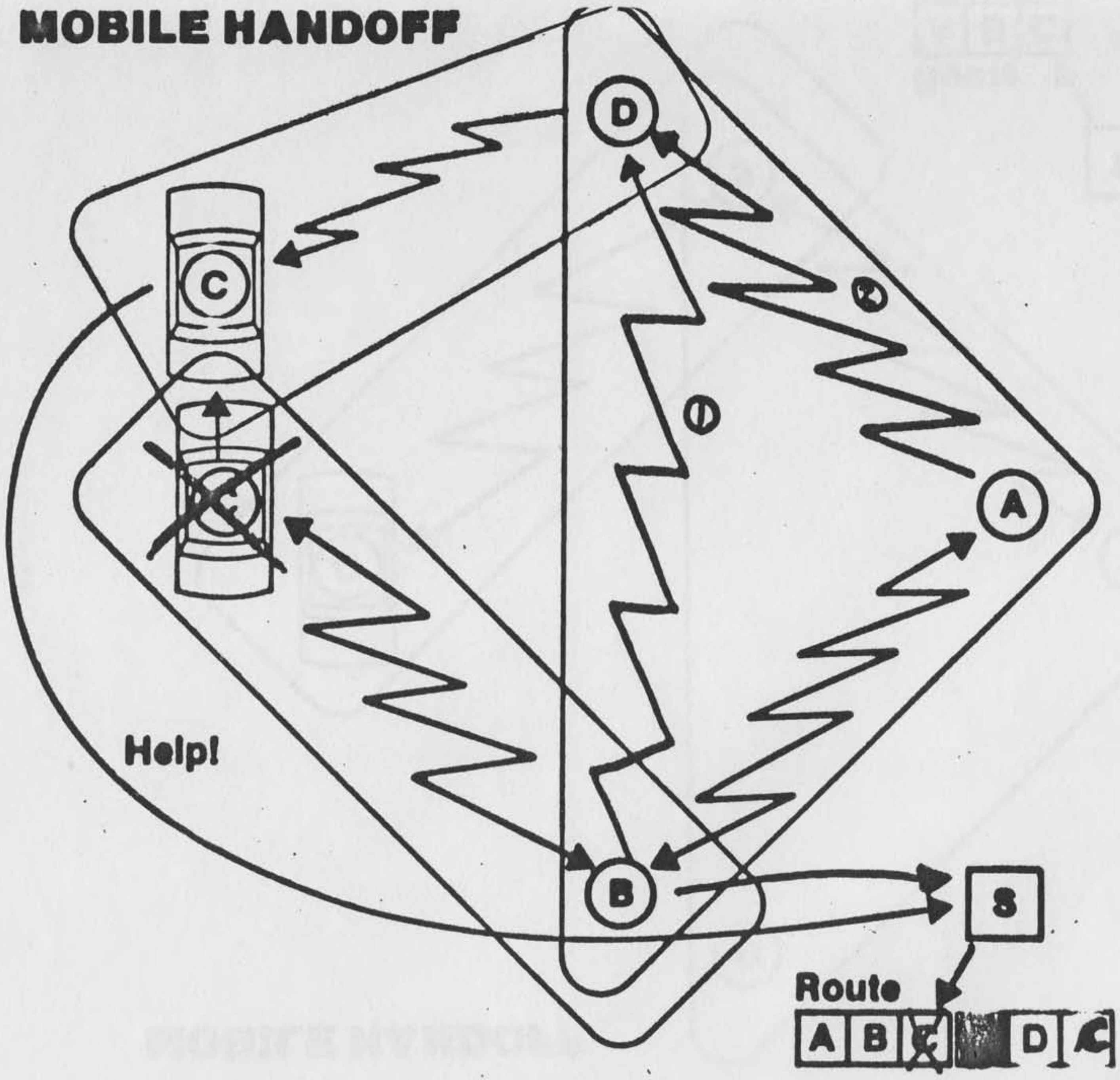




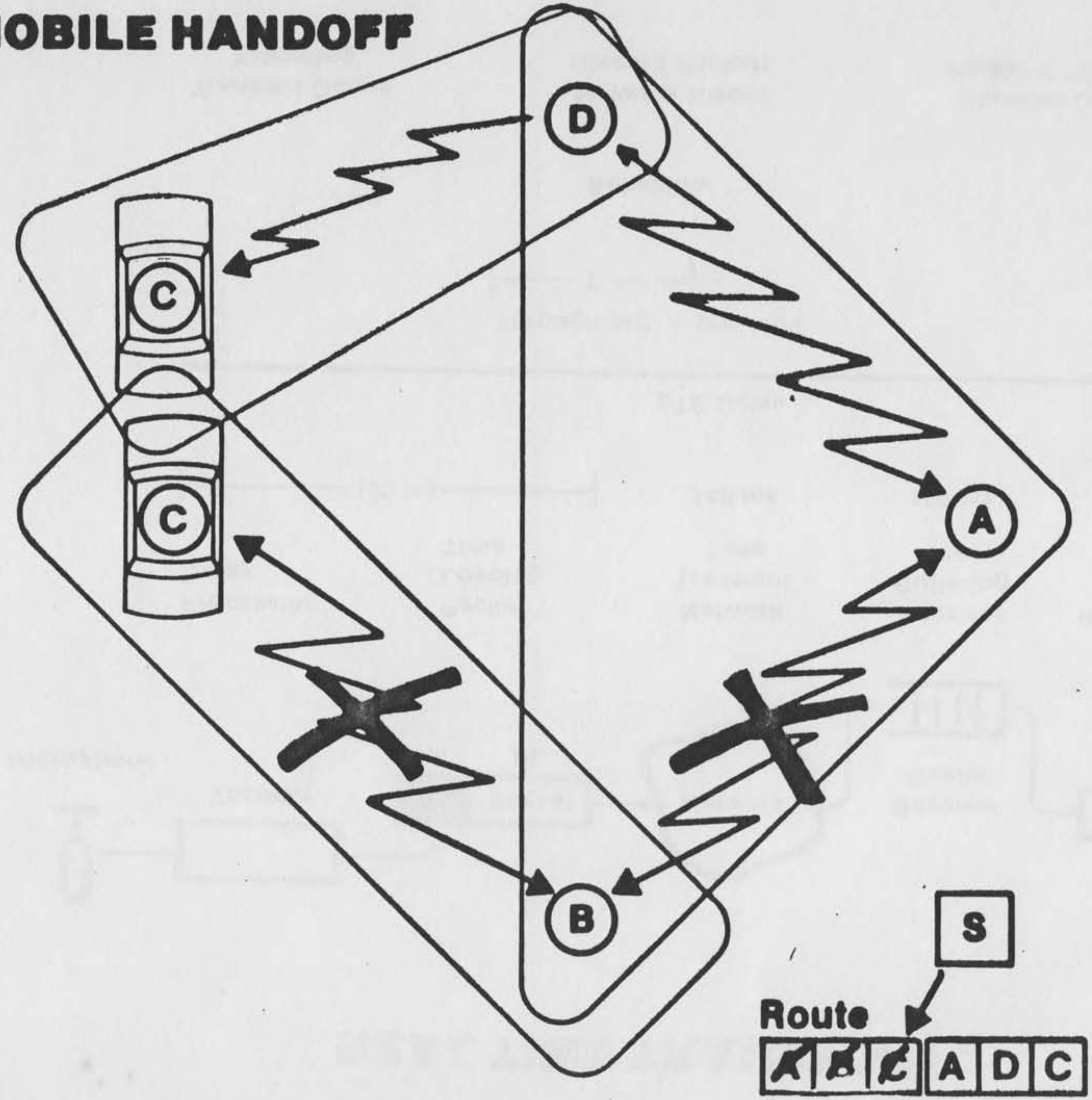
MOBILE HANDOFF



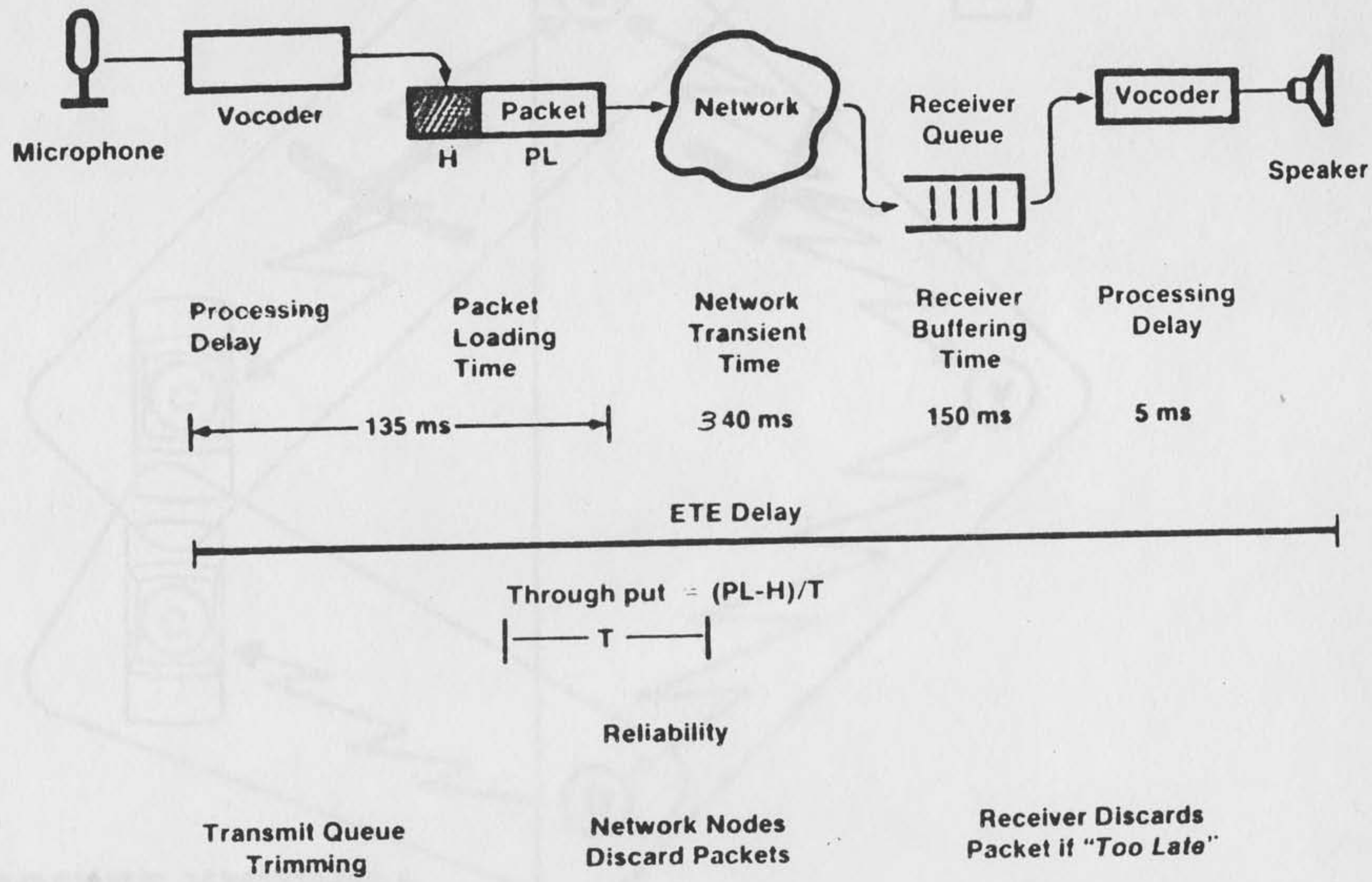
MOBILE HANDOFF



MOBILE HANDOFF

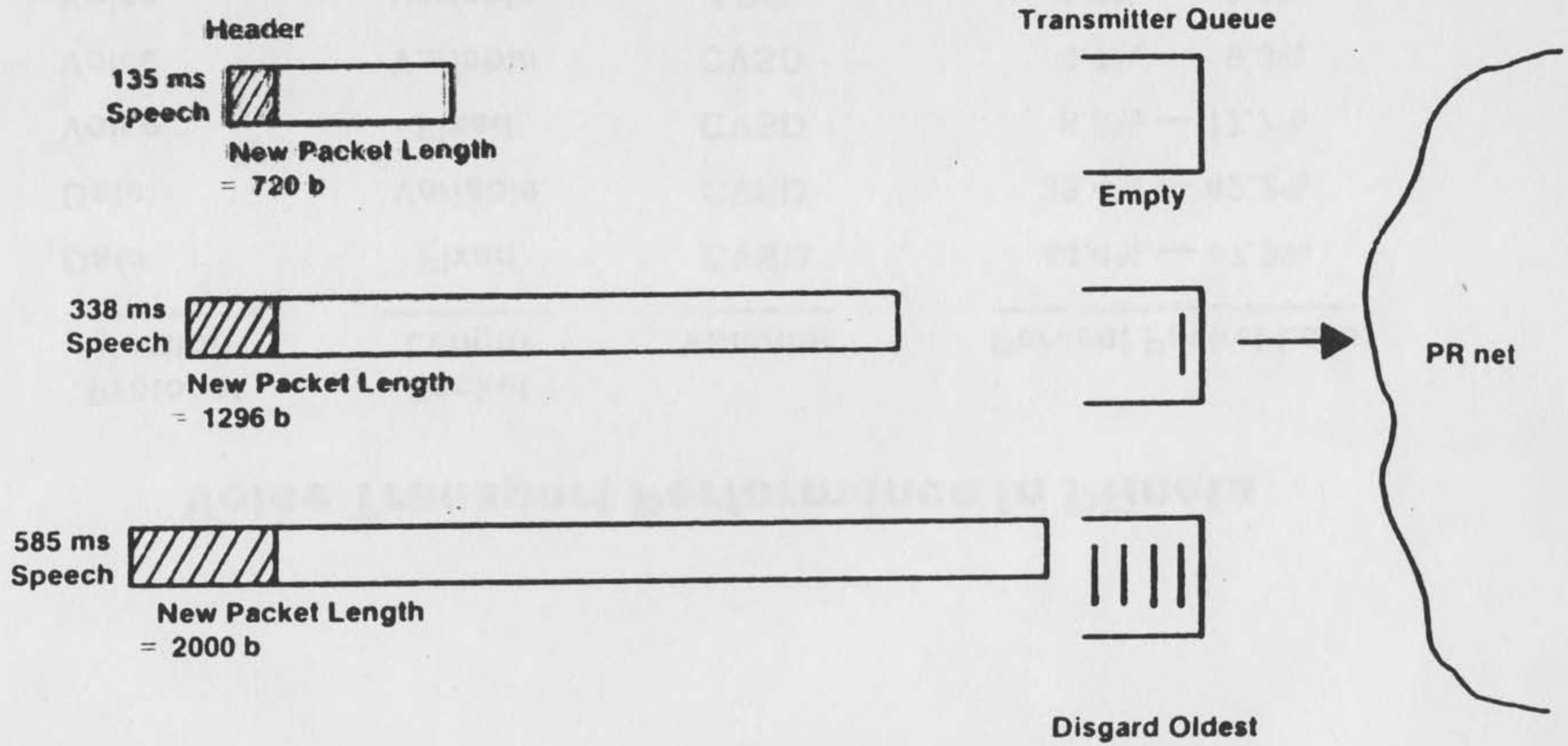


REAL TIME TRANSPORT



TRANSMITTER FLOW CONTROL

Variable Packet Length
Controls the offering rate of packets to the network



Voice Transport Performance in PRnets

Protocol Version	Packet Length	Vocoder	Percent Packet Loss
Data	Fixed	CVSD	44.4% — 47.9%
Data	Variable	CVSD	35.4% — 42.2%
Voice	Fixed	CVSD	8.6% — 17.7%
Voice	Variable	CVSD	4.7% — 9.3%
Voice	Variable	LPC	1.8% — 3.1%

Need Type of Service to Support Data/Voice on the Same Network

— Different Requirements for Data and Voice
— Network Transport Parameters Changed on a Packet-by-Packet Basis
— Adjustments for Dynamic Network Conditions

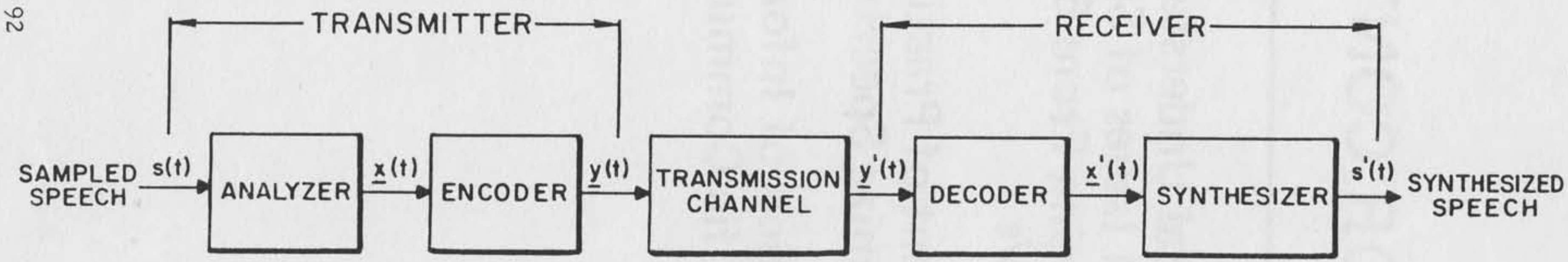
ALGORITHMIC ACHIEVEMENTS

JOHN MAKHOUL,
VISHU VISWANATHAN
BOLT BERANEK AND NEWMAN INC.

MAJOR CONTRIBUTIONS

- Fundamental Understanding of the Basic Theoretical Issues of Speech Coding at Data Rates Ranging From 64,000 bits/s Down to 100 bits/s
- Development of Practical Coding Algorithms That Maximize Speech Quality and Intelligibility
- Dissemination of Information to DoD Agencies and Scientific Community at Large

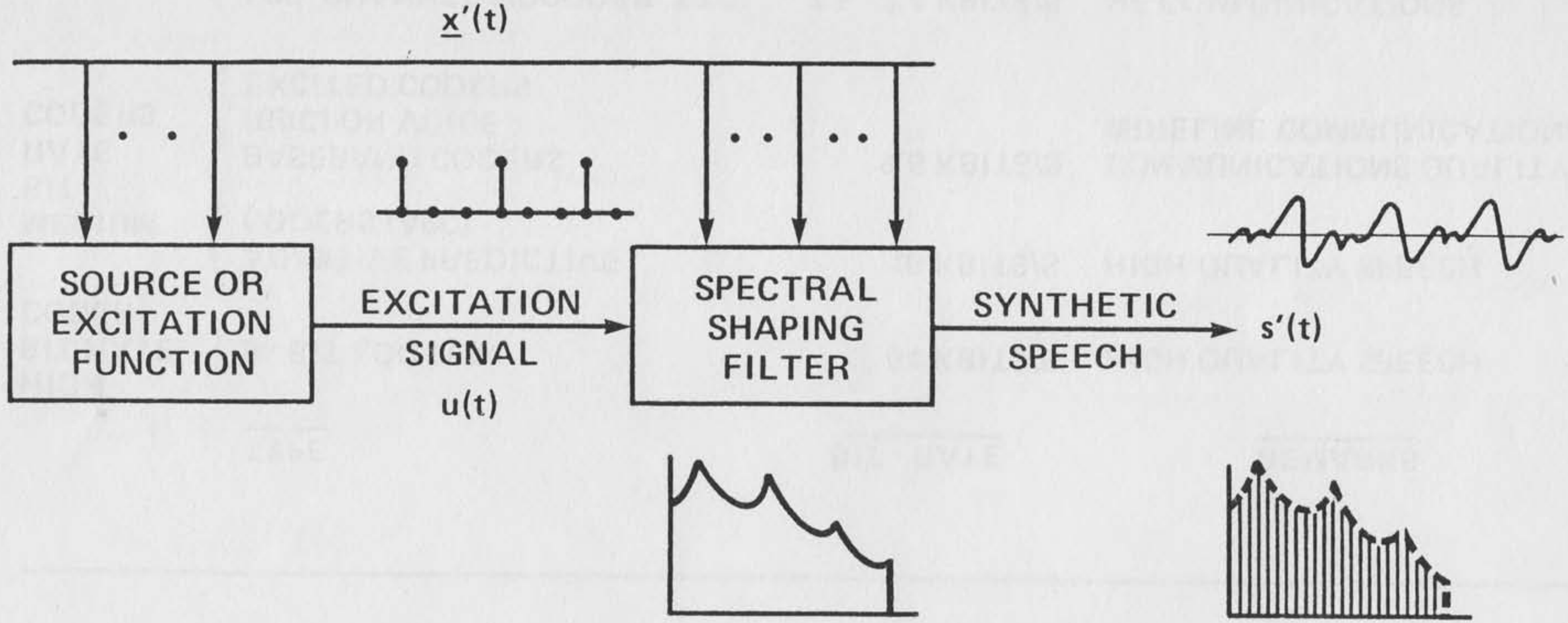
COMPONENTS OF A SPEECH COMPRESSION SYSTEM



TYPES OF SPEECH CODERS

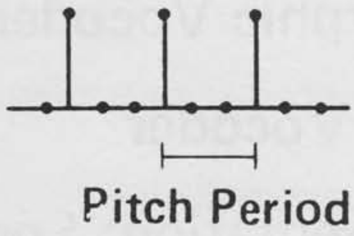
	TYPE	BIT - RATE	REMARKS
HIGH BIT RATE CODER	8 - BIT LOG PCM	64 KBITS/S	HIGH QUALITY SPEECH
MEDIUM BIT RATE CODERS	ADAPTIVE PREDICTIVE CODERS (APC)	16 KBITS/S	HIGH QUALITY SPEECH
	BASEBAND CODERS (BBC) OR VOICE - EXCITED CODERS	9.6 KBITS/S	COMMUNICATIONS QUALITY, WIRELINE COMMUNICATIONS
LOW BIT RATE CODERS	LPC, CHANNEL VOCODER, ETC.	1.5 - 2.4 KBITS/S	HF COMMUNICATIONS
	VECTOR QUANTIZATION	400 - 800 BITS/S	NAVY AND MINING SYSTEM APPLICATIONS
	SEGMENT VOCODER	150 - 300 BITS/S	LOW TRANSMISSION POWER.

A GENERAL SYNTHESIS MODEL

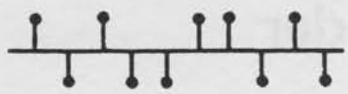


SOURCE MODELS FOR LOW BIT RATE VOCODERS

1. Pulse/Noise Model



Pulse Source for Voiced Sounds



Random Noise for Unvoiced Sounds

- Developed Robust Pitch and Voicing Extraction Algorithms
- Model too Idealized for Certain Sounds — Buzzy Speech Quality

2. Mixed-Source Model

- Uses Pulse and Noise Sources Simultaneously
- Eliminates Buzzy Quality and Produces More Natural-Sounding Speech
- Application in Phonetic Synthesis and Text-to-Speech Systems

SPECTRAL MODELS

- Linear Prediction (All-Pole) Model: LPC Vocoder
- Filter-Bank Model: Channel Vocoder
- Cepstral Model: Homomorphic Vocoder
- Formant Model: Formant Vocoder
- Spectral Amplitude Representation: Spectral Envelope Estimation Vocoder
- Pole-Zero Model

WHAT WE LEARNED:

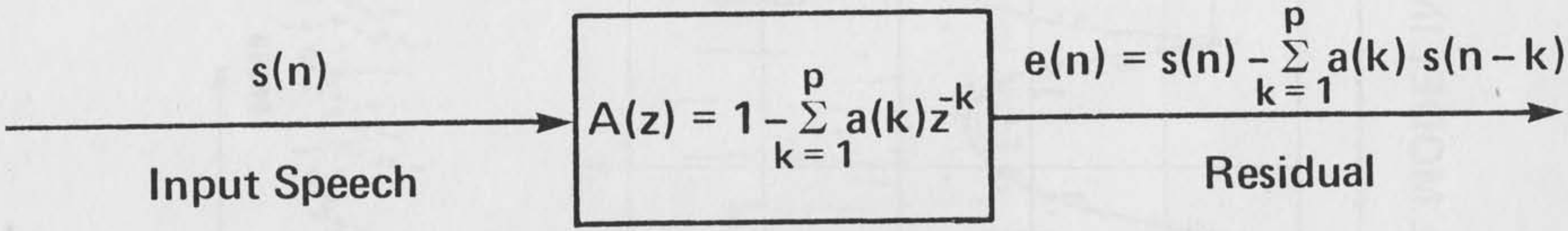
An Accurate Representation of the Short-Term Speech Spectral Envelope, Especially at Spectral Energy Peaks, Is Necessary for Good-Quality Speech Synthesis

LINEAR PREDICTION

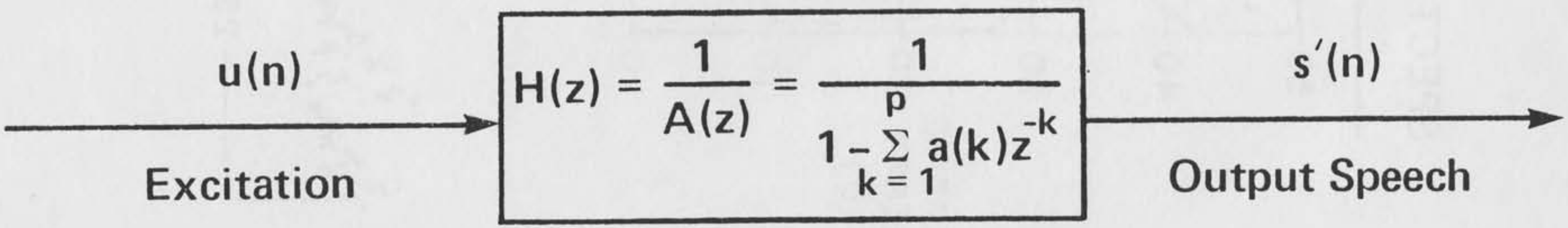
Each Speech Sample is Approximated by a Weighted Linear Summation of Previous Speech Samples:

$$s(n) \cong \sum_{k=1}^{p} a(k)\, s(n-k).$$

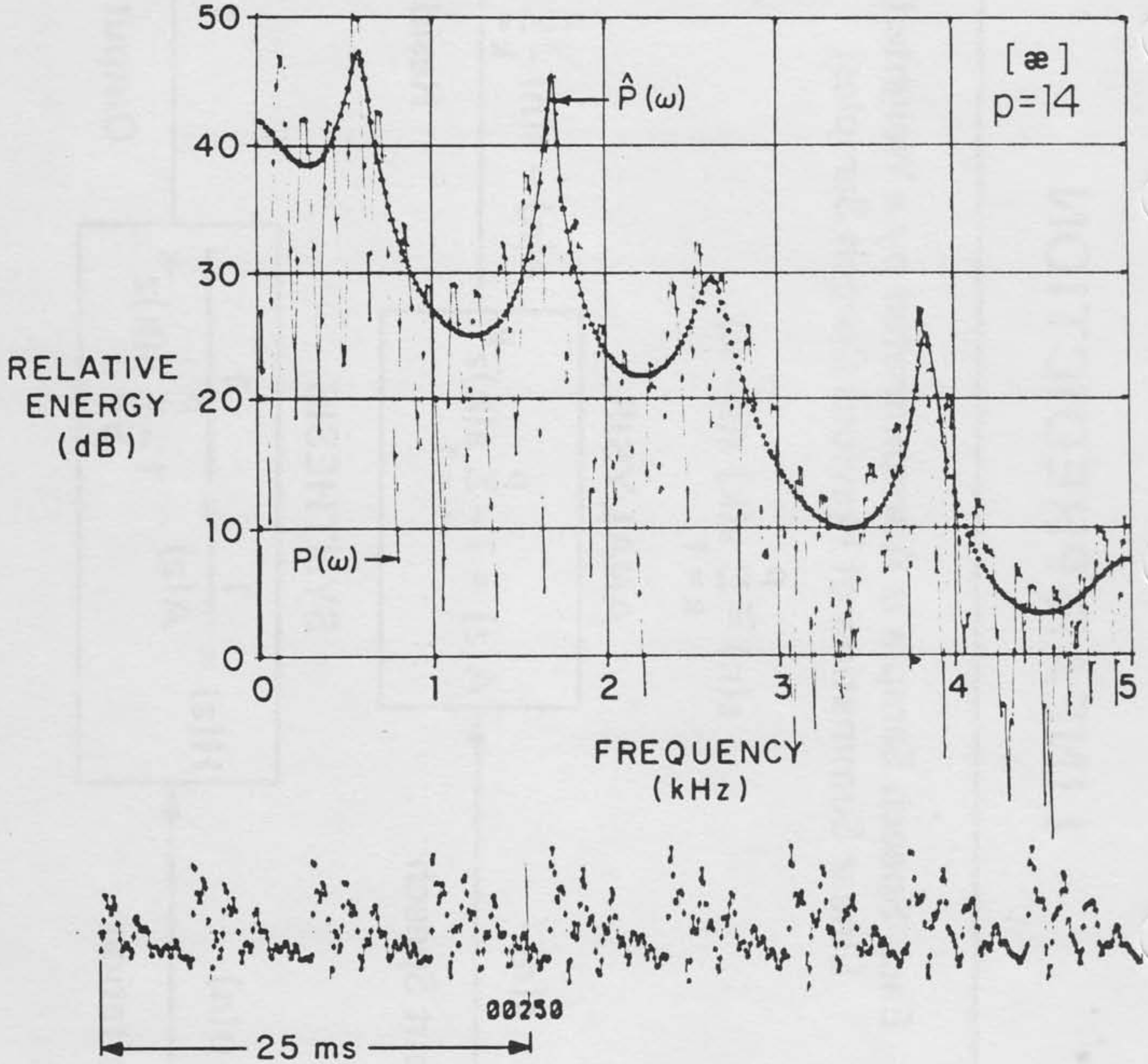
ANALYSIS



SYNTHESIS



SPECTRAL MODELING BY LINEAR PREDICTION

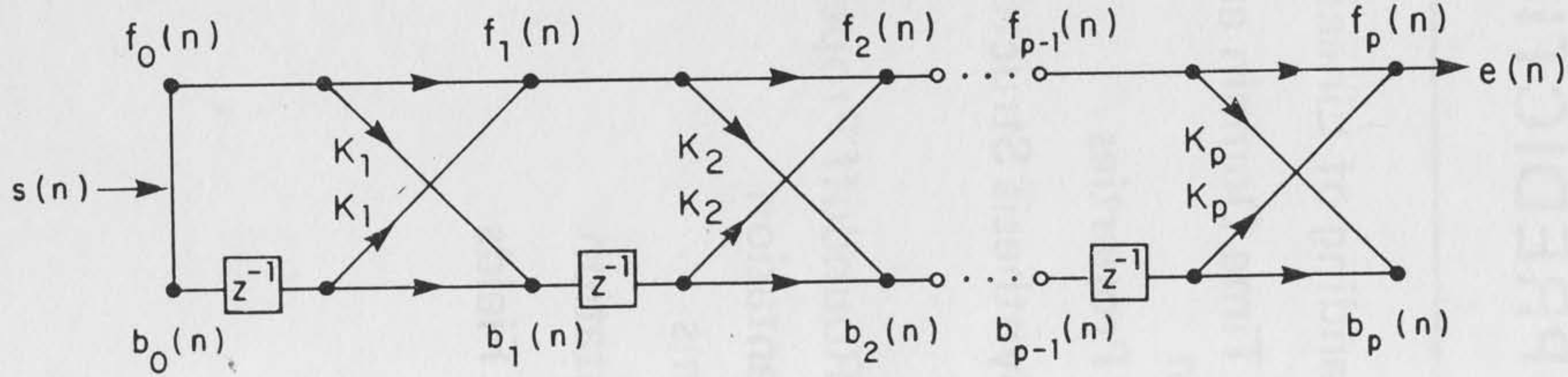


CONTRIBUTIONS IN LINEAR PREDICTION

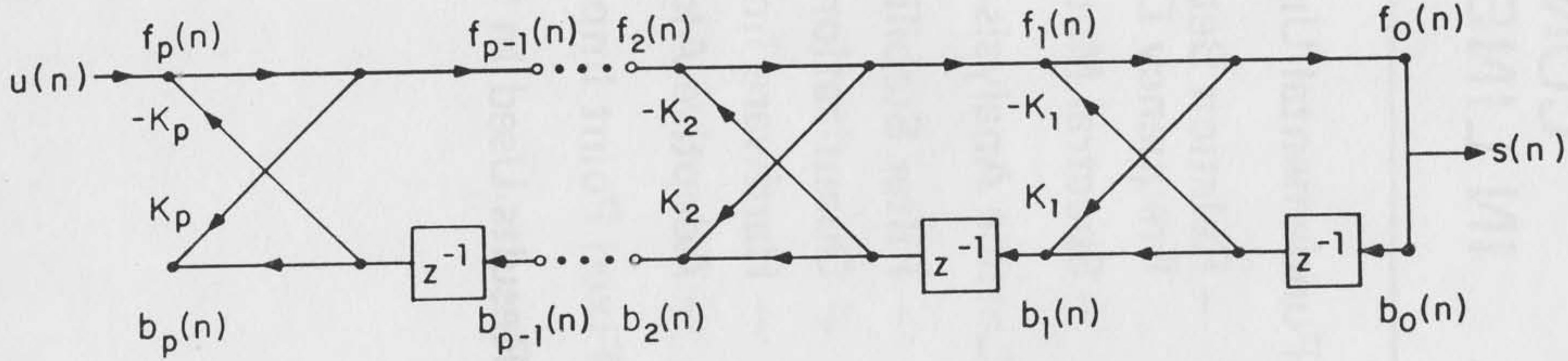
1. Fundamental Understanding of Linear Prediction
   - Relation Between Time Domain and Frequency Domain
   - Spectral Matching Properties
2. Lattice Analysis and Synthesis Structures
   - Filter Stability
   - Quantization and Roundoff Properties
   - Hardware Implementation
   - Adaptive Algorithms
3. Fixed-Point Implementation
4. Results Used In Other Fields

LATTICE METHODS

ANALYSIS

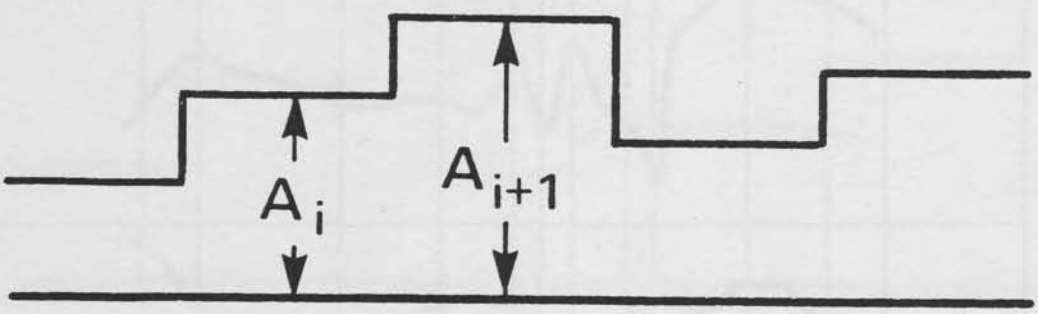


SYNTHESIS



QUANTIZATION OF SPECTRAL PARAMETERS

- Quantization Properties of Different Parametric Representations
  - Reflection Coefficients $\{K_i\}$
- The Notion of Spectral Sensitivity for Optimal Quantization
  - Log Area Ratios $\{L_i\}$
  - Arc Sine Parameters



VOCAL TRACT MODEL

$$L_i = \log \frac{A_{i+1}}{A_i} = \log \frac{1+K_i}{1-K_i}$$

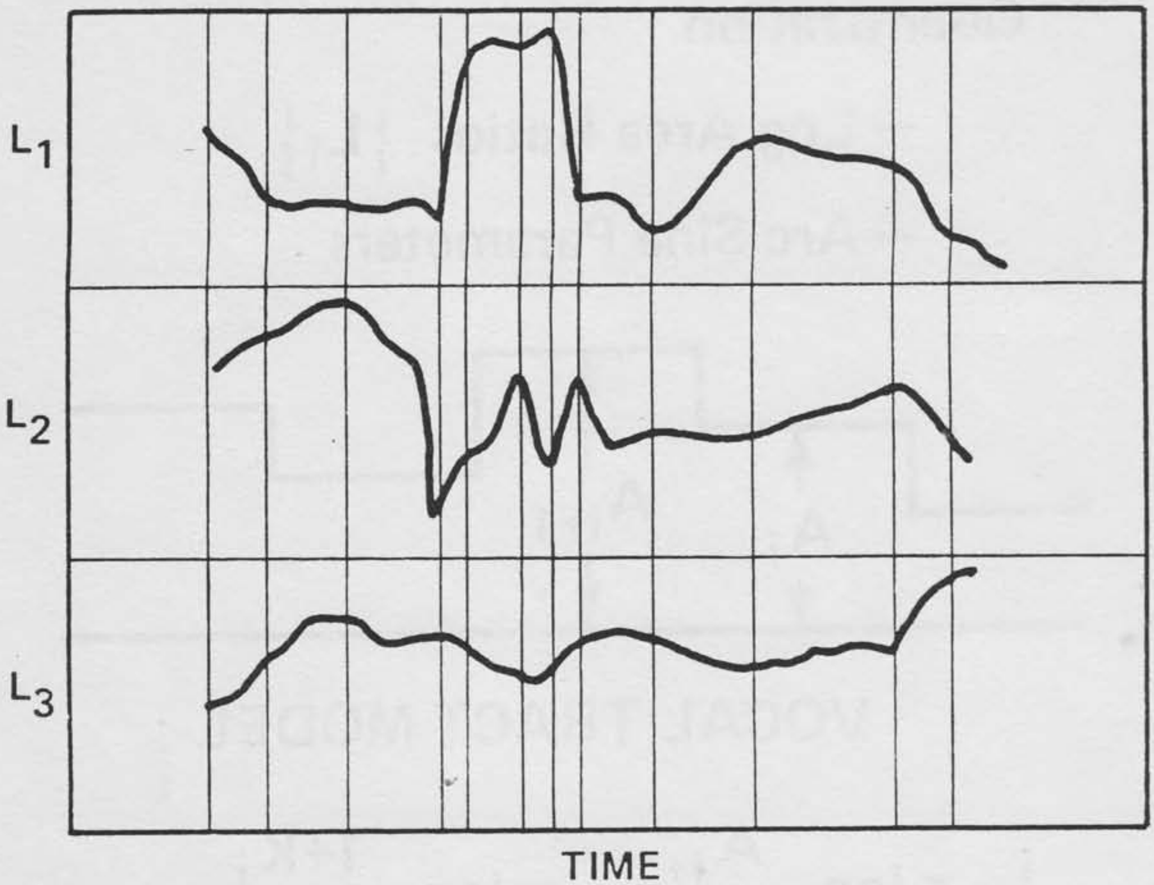
VARIABLE FRAME RATE TRANSMISSION

IDEA:

- Coder Parameters Are Transmitted Only if They Have Changed Sufficiently Since the Last Transmission

METHOD:

- Transmit Selected Frames of Parameters, with the Selection Made Independently for Pitch, Gain and Spectral Parameters
- Reconstruct the Untransmitted Parameters by Linear Interpolation



VARIABLE FRAME RATE TRANSMISSION-RESULTS

	Bits/s
1. Fixed Frame Rate	
100 Frames/s	5600
2. Variable Frame Rate	
Average: 31 Frames/s	2075

RESULT:

Systems 1 and 2 Produce the Same Speech Quality

ROBUST PERFORMANCE UNDER ACOUSTIC BACKGROUND NOISE

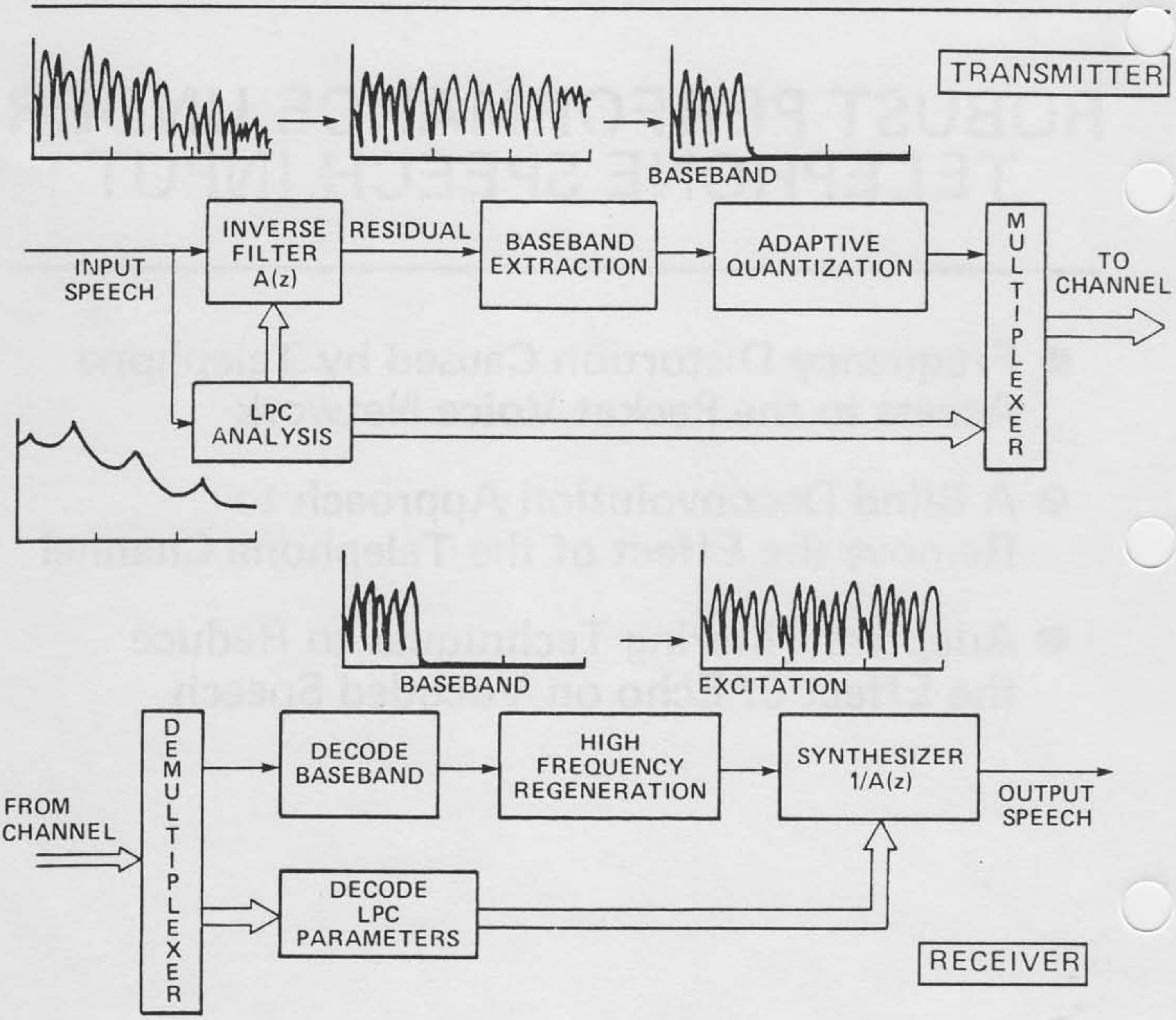
THREE APPROACHES:

1. Noise-Cancelling Microphones (Sponsored by Other DoD Agencies)
2. Two-Microphone Input for Adaptive Noise Cancelling
3. Single Microphone Input
   - Preprocessing by Spectral Subtraction Method
   - Preprocessing Enhances the Quality and Intelligibility of Vocoded Speech in Acoustic Background Noise

ROBUST PERFORMANCE UNDER TELEPHONE SPEECH INPUT

- Frequency Distortion Caused by Telephone Access to the Packet Voice Network
- A Blind Deconvolution Approach to Remove the Effect of the Telephone Channel
- Adaptive Filtering Techniques to Reduce the Effect of Echo on Vocoded Speech

MEDIUMBAND SPEECH CODING



MULTIRATE SPEECH CODING

- Transmission at a Wide Range of Bit Rates to Accommodate Varying Traffic Loads on the Packet Network
- Embedded-Code Systems
- Specific Designs Considered:
  - CVSD/PCM (16 to 64 kbits/s)
  - LPC Vocoder/Adaptive Transform Coder (2.4 to 16 kbits/s)
  - Channel Vocoder/Sub-Band Coder (2.4 to 9.6 kbits/s)

LOW BIT RATE VOCODERS

1. Scalar Quantization of Spectral Parameters (1.5-2.4 kbits/s)
   - Independent Quantization of Each Parameter
2. Vector Quantization of Spectral Parameters (400-800 bits/s)
   - Parameters Quantized as a Vector in Multi-Dimensional Space
   - Use Clustering to Determine Quantization Templates
3. Segment Quantization (150-300 bits/s)
   - Quantize a Sequence of Spectra as One Unit
4. Phonetic Vocoder (100 bits/s)
   - Use Speech Recognition Techniques

STATE-OF-THE-ART IN SPEECH COMPRESSION

- Virtually Transparent Speech Quality at 16 kbits/s
- Quality and Intelligibility Degrade With Decreasing Bit-Rate, Especially
  - Below 4.8 kbits/s
  - For Female Speech
  - Under Acoustic Noise and Channel Distortion Conditions
- Substantially Improved Intelligibility and Quality of 2.4 kbits/s Vocoders

AREAS OF FURTHER RESEARCH

PROBLEM: How to Maximize Speech Quality and Intelligibility for Every bit That We Use in Transmission?

- Source Modeling
  - Pitch and Voiced/Unvoiced Decision
- Spectral Modeling
  - Female Speech
- Pattern Recognition and Information Theoretic Coding Techniques
- Understanding and Modeling of the Effects of Coding Techniques on Human Speech Perception

VERY LOW RATE VOCODER

RICHARD SCHWARTZ
BOLT BERANEK AND NEWMAN INC.

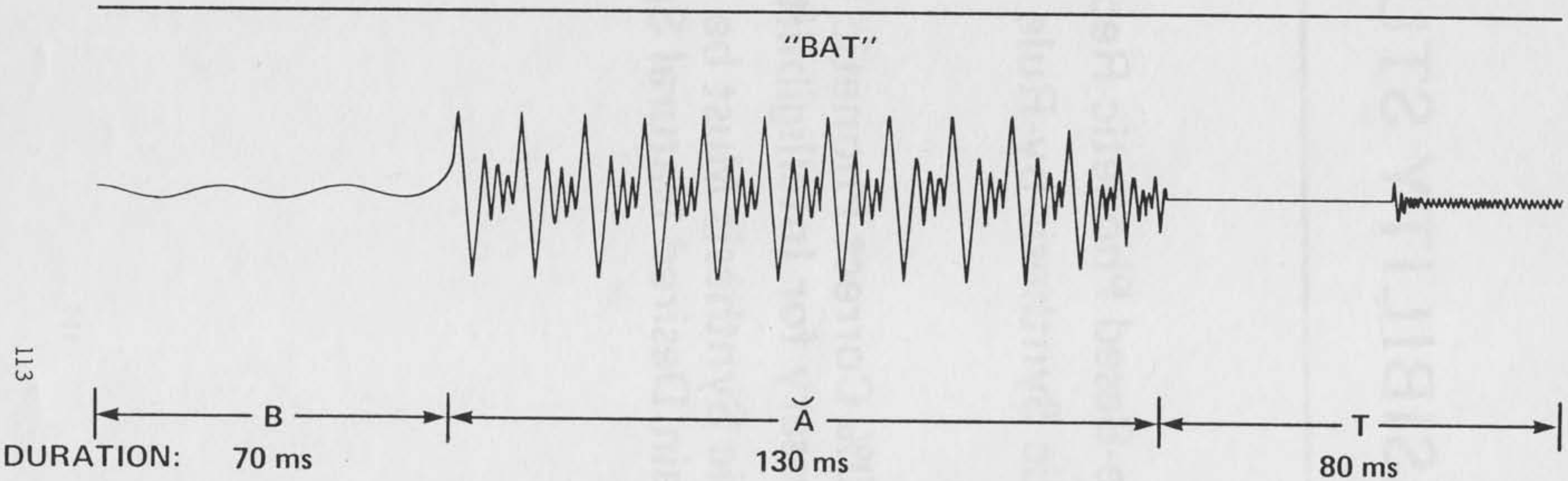
VERY LOW RATE VOCODER

- Requirements
  - 100-200 bits/s
  - Intelligibility in a Conversation
  - Naturalness
- Applications
  - Low Power Communication
  - Degraded Channel Conditions
  - "Burn Through" Jamming Networks
  - Speech Storage

VERY LOW RATE VOCODERS

1. Phonetic Vocoder ~100 bits/s
2. Segment Vocoder ~150-300 bits/s

PHONETIC VOCODER

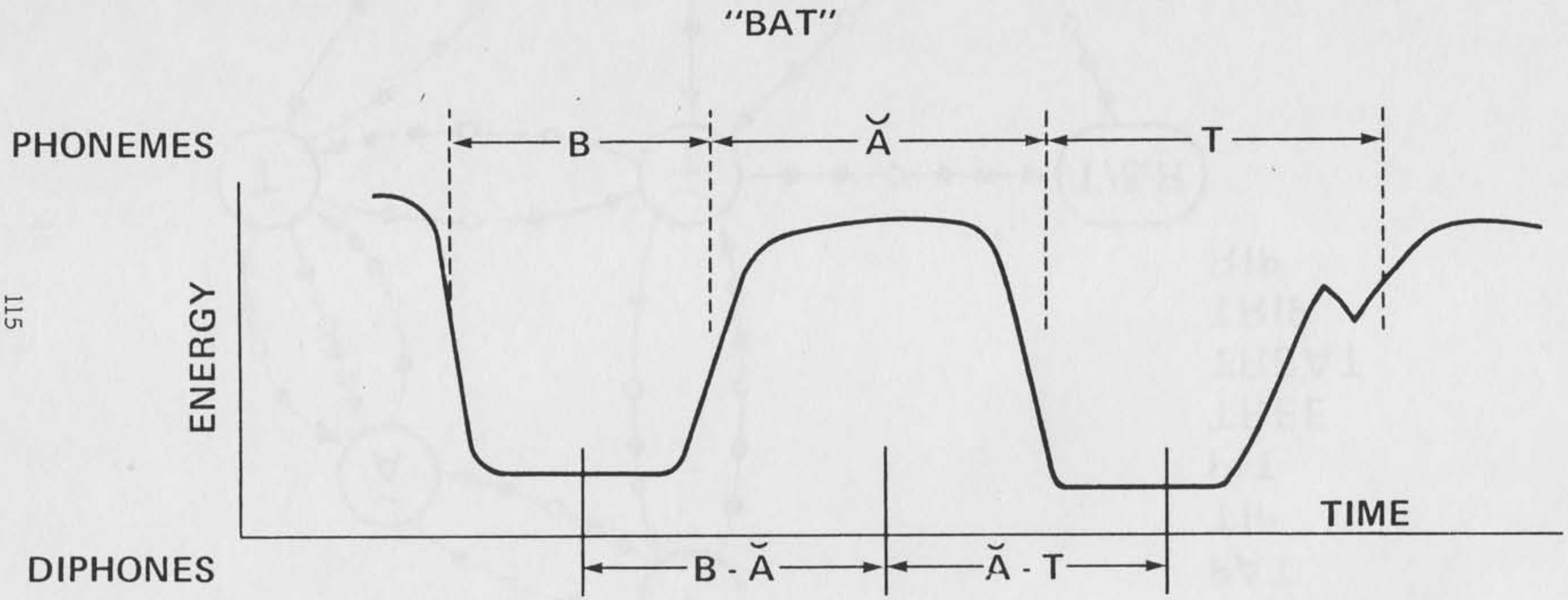


Number of Phonemes in English: 42
Typical Speaking Rate: 12 Phonemes/sec.

FEASIBILITY STUDY

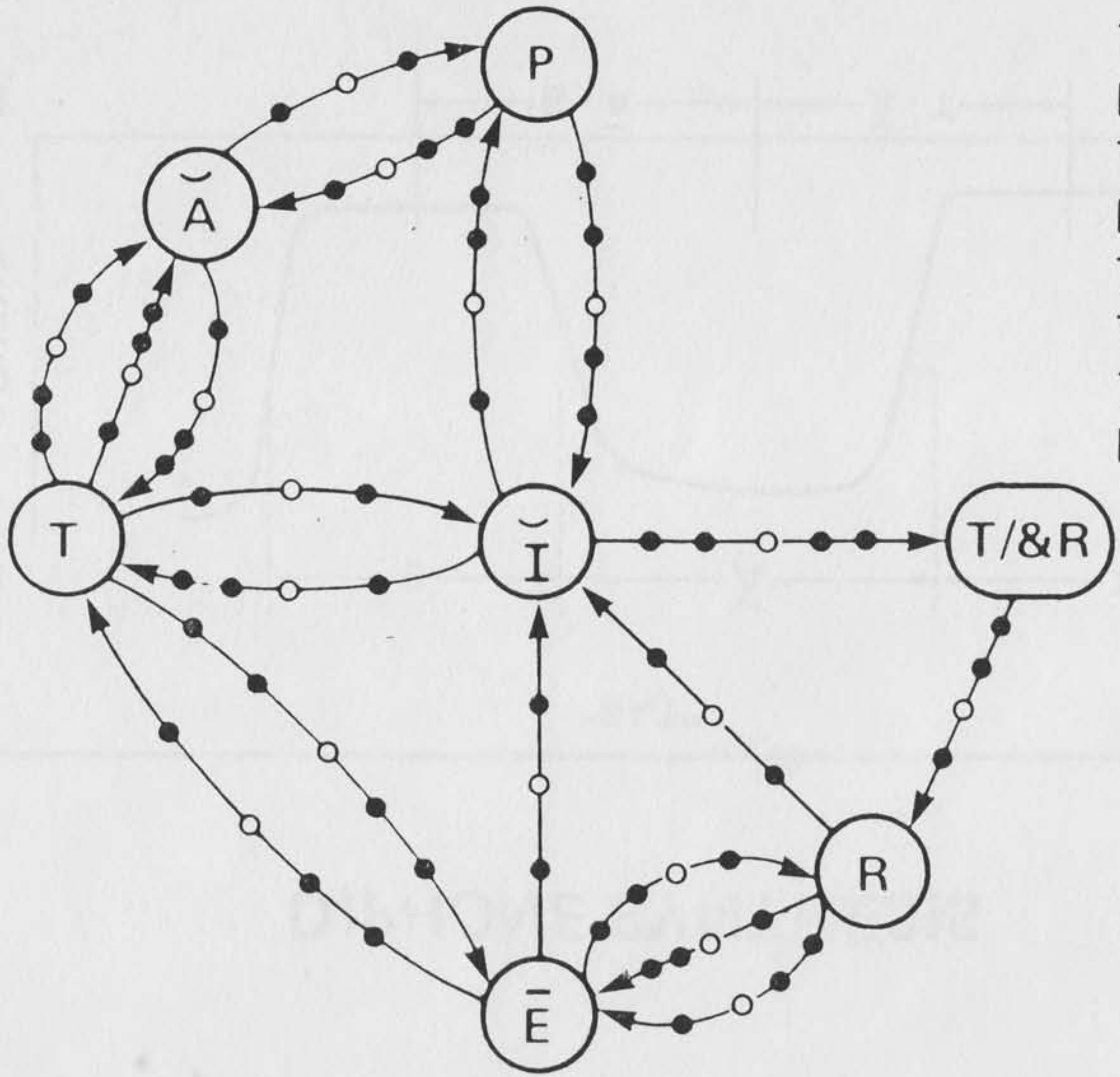
- Approach
  - Feature-Based Phonetic Recognition
  - Phonetic Synthesis-by-Rule
- Conclusions
  - Over 80% Correct Phonetic Recognition Is Necessary for Intelligibility in Context
  - Phonetic Synthesis Must be Improved to Obtain Desired Natural Speech Quality

DIPHONE SYNTHESIS



SPEECH IS SYNTHESIZED BY CONCATENATING DIPHONE TEMPLATES

DIPHONE RECOGNITION NETWORK



TAP = T Ă P
PAT
TIP
PIT
TREE
TREAT
TRIP
RIP

PHONETIC VOCODER

- Requires Phonetic Recognition
  - Still a Difficult Problem
- Requires Manual Labeling of Diphones
  - Labor Intensive

SEGMENT VOCODER

- A Segment Is a "Diphone-Like" Sequence of Spectra Between Two Steady State Sounds
- Segment Templates Are Selected Automatically (Unsupervised)
  - There Is No Phonetic Labeling
- Result: Segment Vocoder Speech Is Intelligible at About 150 bits/s

AUTOMATIC TEMPLATE SELECTION

- Training Data: 15 Minutes of Unconstrained Speech
- Segmentation Algorithm:
  - Divides Training Speech Into "Diphone-Like" Segments (15 Minutes Contains About 8,000 or 2^{13} Segments)
- Random Quantizer
  - All Segments are Used as Templates
- Template Network:
  - Determines Constraints on Template Sequences by Requiring Spectral Continuity Between Successive Templates
  - Only 1024 or 2^{10} Templates Can Follow Any Given Template

BIT ALLOCATION FOR SEGMENT VOCODER

	Bits/Segment
Segment Template Code	10
Segment Duration	1½
Pitch	1
Energy	1½
Total Bits/Segment	14
Average Segment Rate	11 per sec
Average Bit Rate = 11x14 =	154 bits/s

PACKET VOICE TERMINALS
AND
THE LEXNET LOCAL NETWORK

GERALD C. O'LEARY
MIT LINCOLN LABORATORY

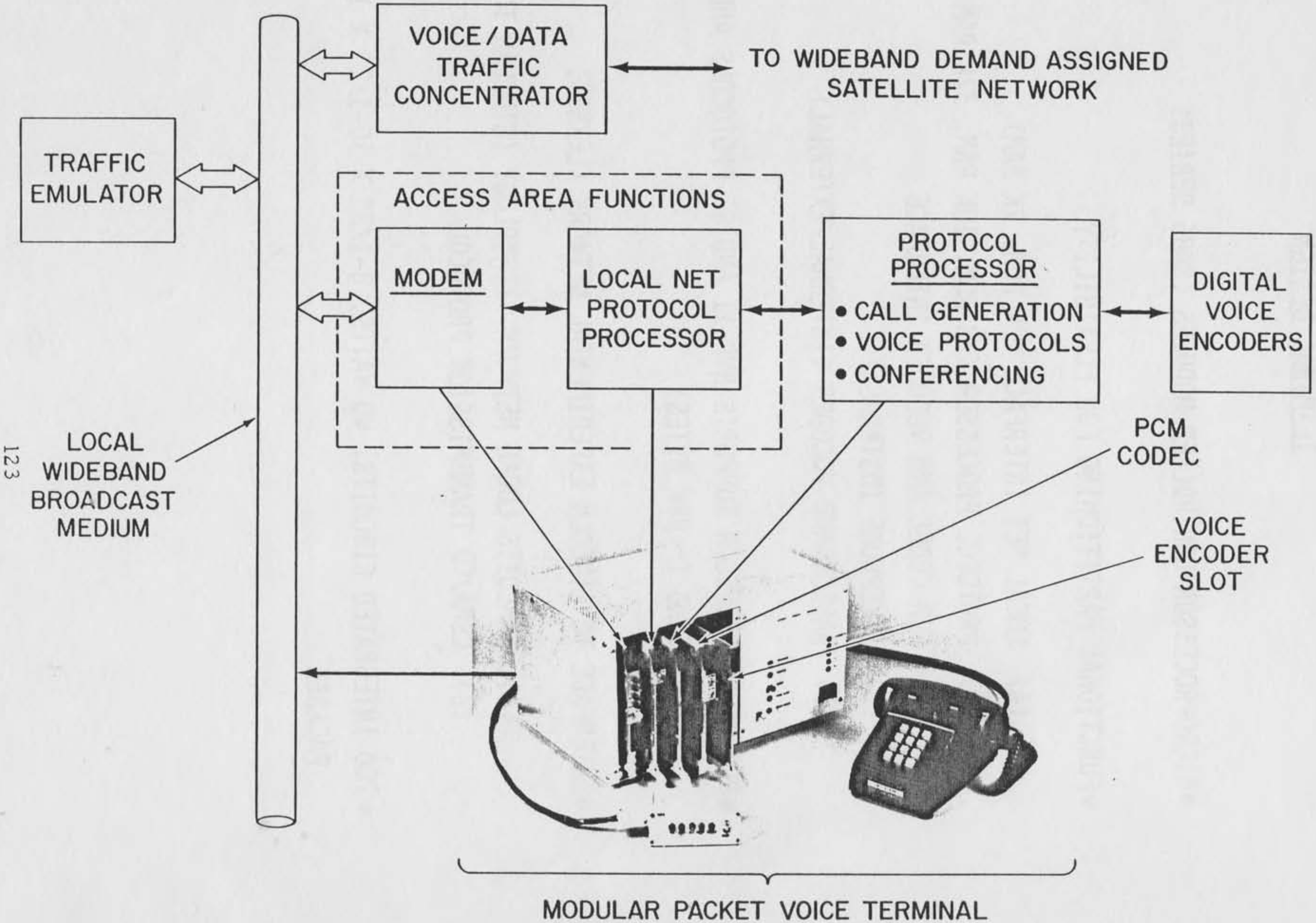
PACKET VOICE TERMINAL AND LEXNET

- PVT REQUIREMENTS
- ARCHITECTURE
- SPEECH PROCESSOR OPTIONS
- LEXNET
- CURRENT STATUS
- FUTURE DIRECTIONS

PACKET VOICE TERMINAL REQUIREMENTS

- ACCOMMODATE PCM AND NARROWBAND SPEECH PROCESSORS
- IMPLEMENT PACKET VOICE PROTOCOLS
- SUPPORT CONNECTION TO DIFFERENT PACKET NETS
- TELEPHONE-LIKE USER INTERFACE
- SOURCE OF PACKET SPEECH TRAFFIC FOR WIDEBAND PACKET VOICE EXPERIMENTS

LINCOLN EXPERIMENTAL PACKET VOICE NETWORK



TERMINAL DESIGN

- MICROPROCESSOR CONTROLLED MODULES (8085 SERIES)
- FUNCTIONAL PARTITIONING FOR FLEXIBILITY
  - (A) LOCAL NET INTERFACE (2K ROM, 2K RAM)
  - (B) PROTOCOL PROCESSOR (2K ROM, 10K RAM, 32K RAM/ROM)
  - (C) PCM CODEC AND VOCODER INTERFACE
  - (D) TELEPHONE INSTRUMENT
  - (E) NARROWBAND VOCODER (INTERNAL/EXTERNAL)
- PROTOCOL PROCESSOR SUPPORTS NVP-11 AND ST PROTOCOLS AND CONFERENCING (~ 28K BYTES)
- INTERFACE TO LINCOLN EXPERIMENTAL NETWORK (LEXNET)
  - (A) EXPLOITS LOCAL NETWORK TECHNOLOGY (SIMILAR TO ETHERNET)
  - (B) CSMA/CD TRANSMISSION PROTOCOL
- 200 INTEGRATED CIRCUITS, 40 WATTS, 8-1/2" X 10-1/2" X 16-1/2" PACKAGE

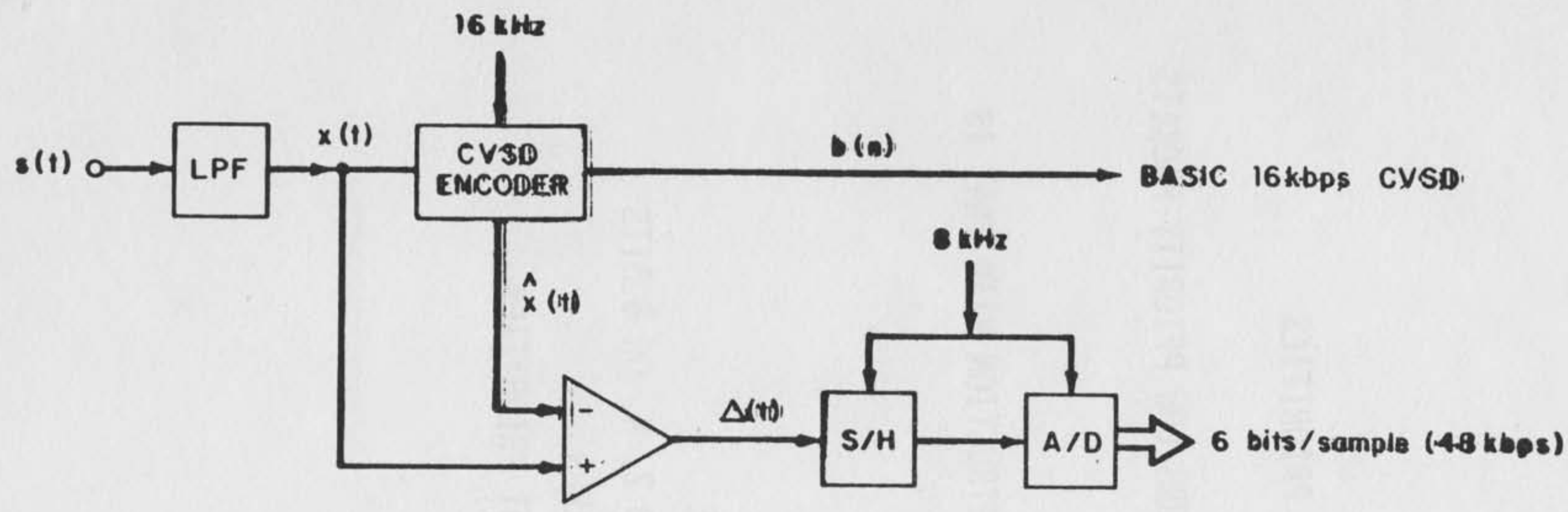
SPEECH PROCESSOR OPTIONS

- 64 KBIT/SEC PCM
- 16 - 64 KBIT/SEC EMBEDDED CVSD
- 2400 BIT/SEC LPC
- SWITCHED TELEPHONE NETWORK INTERFACE CARD
- NETWORK TIMING MEASUREMENTS CARD

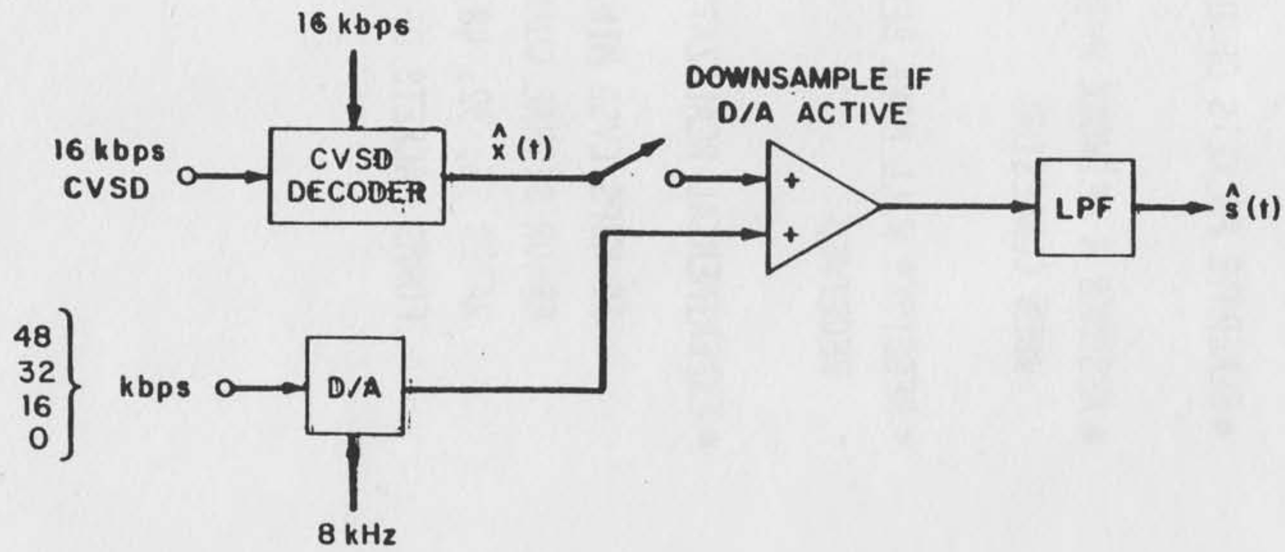
EMBEDDED VOCODER

- VARIABLE RATE VOCODER
- RATE-QUALITY TRADE-OFF
- GENERATE PACKETS OF DIFFERENT PRIORITIES
- ASSUMES A NETWORK WHICH DISCARDS LOW PRIORITY PACKETS WHEN CONGESTED
- RECEIVER WILL MAKE BEST RECONSTRUCTION WITH WHAT IS RECEIVED
- EXPERIMENTAL REALIZATION

  16 KBPS CVSD MINIMUM RATE
  ERROR SIGNAL QUANTIZED TO 2, 4, OR 6 BITS
  RATES 16, 32, 48, OR 64 KBPS
  FORMS PACKETS OF DIFFERENT PRIORITIES


16 kHz
s(t)
LPF
x(t)
CVSD ENCODER
b(n)
BASIC 16kbps CVSD
8 kHz
x̂(t)
Δ(t)
S/H
A/D
6 bits/sample (48 kbps)

(a)


16 kbps
16 kbps CVSD
CVSD DECODER
x̂(t)
DOWNSAMPLE IF D/A ACTIVE
LPF
ŝ(t)
48
32
16
0
kbps
D/A
8 kHz

(b)

LINCOLN EXPERIMENTAL PACKET VOICE NETWORK (LEXNET)

- CONNECTS PVTs TO FORM DISTRIBUTED TELEPHONE SYSTEM
- CONCENTRATOR INTERFACE BOX PROVIDES HIGH RATE CONNECTION TO SATELLITE CHANNEL
- CONFERENCE ACCESS CONTROLLER
- LINCOLN'S NET SPANS ABOUT 1000 FEET
- 1 MBIT/SEC ON RG-59/U COAXIAL CABLE

SUMMARY

- CURRENT STATUS

  NETWORKS AT LINCOLN (2),

  ISI

  SRI

  DCEC

  10 TERMINALS

- FUTURE DIRECTIONS

  INTEGRATE PACKET VOICE CAPABILITY TO OTHER DoD DATA NETWORKS

  SECURE SPEECH TERMINALS

  TECHNOLOGY TRANSFER

  VOICE CONTROLLED CONFERENCE SET UP

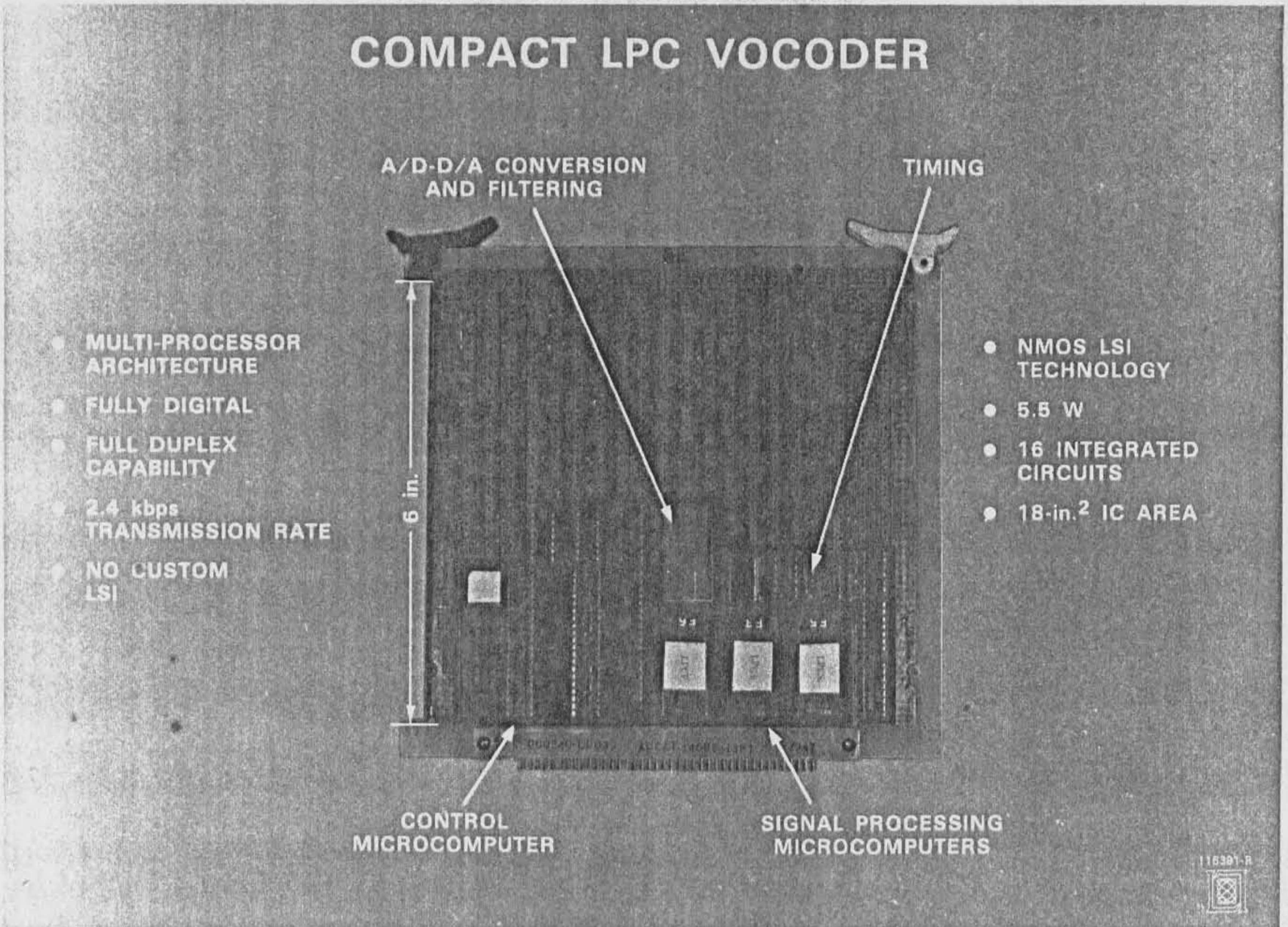
  NETWORK TIMING MEASUREMENTS

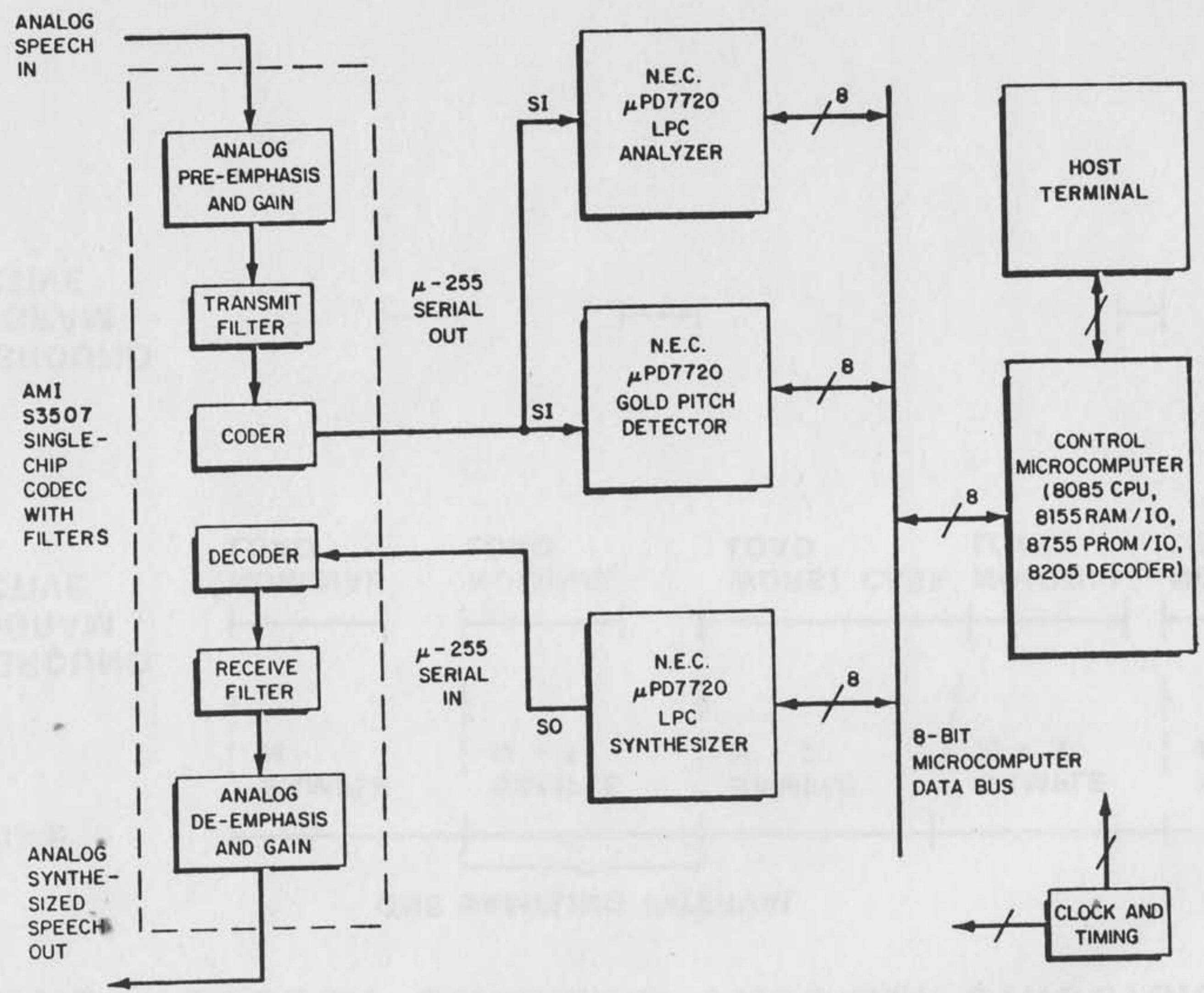
COMPACT LPC VOCODER

JOEL A. FELDMAN
MIT LINCOLN LABORATORY

PVT SINGLE BOARD 2.4 KBPS LPC VOCODER

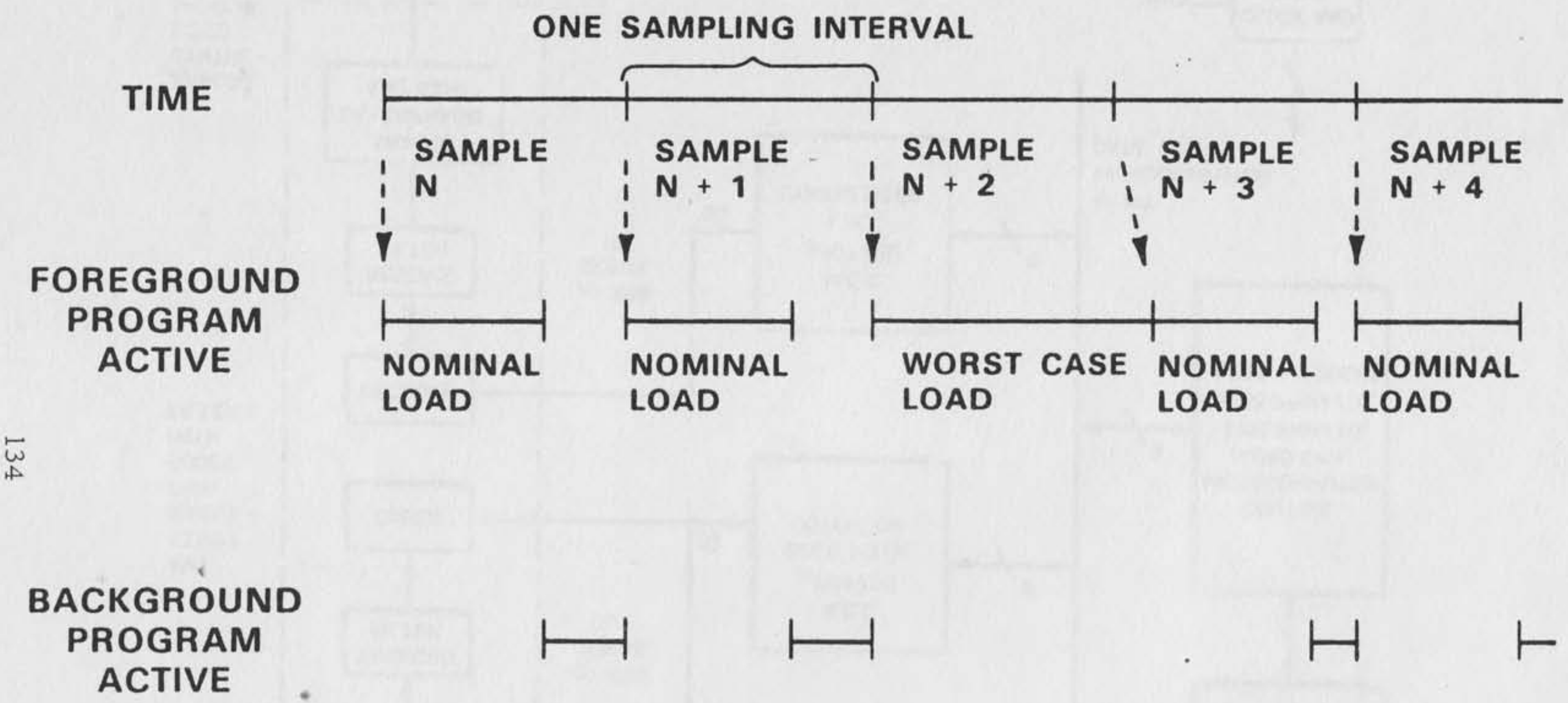
- HARDWARE
- ARCHITECTURE
- PROCESSORS' TASKS
- PROCESSORS' REAL-TIME AND MEMORY USAGE
- INITIALIZATION OPTIONS


COMPACT LPC VOCODER
A/D-D/A CONVERSION AND FILTERING
TIMING
MULTI-PROCESSOR ARCHITECTURE
FULLY DIGITAL
FULL DUPLEX CAPABILITY
2.4 kbps TRANSMISSION RATE
NO CUSTOM LSI
6 in.
NMOS LSI TECHNOLOGY
5.5 W
16 INTEGRATED CIRCUITS
18-in.² IC AREA
CONTROL MICROCOMPUTER
SIGNAL PROCESSING MICROCOMPUTERS
116391-R


ANALOG SPEECH IN
ANALOG PRE-EMPHASIS AND GAIN
TRANSMIT FILTER
CODER
AMI S3507 SINGLE-CHIP CODEC WITH FILTERS
DECODER
RECEIVE FILTER
ANALOG DE-EMPHASIS AND GAIN
ANALOG SYNTHE-SIZED SPEECH OUT
μ-255 SERIAL OUT
μ-255 SERIAL IN
SI
SI
SO
N.E.C. μPD7720 LPC ANALYZER
N.E.C. μPD7720 GOLD PITCH DETECTOR
N.E.C. μPD7720 LPC SYNTHESIZER
8
8
8
8
8-BIT MICROCOMPUTER DATA BUS
HOST TERMINAL
CONTROL MICROCOMPUTER (8085 CPU, 8155 RAM/IO, 8755 PROM/IO, 8205 DECODER)
CLOCK AND TIMING

112777-N-01

COMPACT VOCODER 'STREAMED' PROGRAM STRUCTURE



116365-N

LPC ANALYZER TASKS (1 N.E.C. μPD7720)

- EXECUTED EACH SAMPLE (Interrupt Driven Foreground)

  (a) INPUT SAMPLE, MU-LAW TO LINEAR CONVERSION, HAMMING WINDOW, DOWNSCALE, UPDATE CORRELATION COEFFICIENTS

  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR

  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE AND SWAP OUTPUT PARAMETER BUFFERS

- EXECUTED EACH FRAME (Background)

  (a) BLOCK FLOATING POINT CORRELATION COEFFICIENTS AND COMPUTE REFLECTION COEFFICIENTS

  (b) TRANSFER REFLECTION COEFFICIENTS AND ENERGY ESTIMATE TO CONTROL PROCESSOR

116356-N

GOLD PITCH DETECTOR TASKS (1 N.E.C. μPD7720)

- EXECUTED EACH SAMPLE (Interrupt Driven Foreground)

  (a) INPUT SAMPLE, MU-LAW TO LINEAR CONVERSION, LOW-PASS PRE-FILTER, PEAK-DETECTION, UPDATE SIX PARALLEL PITCH PERIOD ESTIMATES

  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR

  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE

- EXECUTED EACH FRAME (Background)

  (a) PITCH PERIOD SCORING AND FINAL VOICING DECISION AND PITCH ESTIMATION

  (b) OUTPUT VOICING DECISION AND PITCH ESTIMATE TO CONTROL PROCESSOR

116357-N

LPC SYNTHESIZER TASKS (1 N.E.C. μPD7720)

- **EXECUTED EACH SAMPLE** (Interrupt Driven Foreground)

  (a) UPDATE EXCITATION GENERATOR, UPDATE LATTICE FILTER, OUTPUT SAMPLE

  (b) CHECK FOR FRAME MARK FROM CONTROL PROCESSOR

  (c) IF FRAME TIME, ACTIVATE BACKGROUND ROUTINE AND SWAP INPUT PARAMETER BUFFERS

- **EXECUTED EACH FRAME** (Background)

  (a) CONVERT ENERGY ESTIMATE TO PITCH PULSE AMPLITUDES (Voiced Frames) OR PSEUDO-RANDOM NOISE AMPLITUDES (Unvoiced Frames)

  (b) INPUT PITCH/VOICING PARAMETER, ENERGY ESTIMATE AND REFLECTION COEFFICIENTS FROM CONTROL PROCESSOR

116358-N

CONTROL MICROCOMPUTER TASKS (1 INTEL 8085)

- EXECUTED EACH FRAME

  (a) GIVE FRAME MARK TO ANALYZER, PITCH DETECTOR AND SYNTHESIZER

  (b) INPUT ANALYSIS PARAMETERS FROM ANALYZER AND PITCH DETECTOR, CODE TO 2400 bps, OUTPUT TO HOST TERMINAL

  (c) INPUT SYNTHESIS PARAMETERS FROM HOST TERMINAL, DECODE TO 16-bit VALUES, OUTPUT TO SYNTHESIZER

116359-N

LPC VOCODER MEMORY REQUIREMENTS
(10th Order Model)

	Percentage of:		
	Data RAM (128 x 16)	Program ROM (512 x 23)	Data ROM (512 x 13)
Gold Pitch Detector	94%	95%	10%
Linear Predictive Analysis	70%	80%	3%
Synthesis	63%	67%	0%

116361-N

LPC VOCODER REAL-TIME USAGE

(a) 10th ORDER LINEAR PREDICTIVE MODEL
(b) 8-kHz SAMPLING FREQUENCY
(c) 22.5-ms SPEECH FRAMES

PERCENTAGE OF REAL-TIME USED

GOLD PITCH DETECTOR	≈35%
LINEAR PREDICTIVE ANALYSIS	63%
SYNTHESIS	46%

116360-N

CONTROL MICROCOMPUTER REAL-TIME AND MEMORY REQUIREMENTS

- 67% REAL-TIME (22.5 ms Frames, 8 MHz X-TAL)
- 2013 BYTES ROM
- 139 BYTES RAM

116362-N

COMPACT LPC VOCODER INITIALIZATION OPTIONS

- LINEAR PREDICTIVE MODEL ORDER (up to 15)
- ANALYSIS AND SYNTHESIS FRAME SIZE
- SPEECH SAMPLING FREQUENCY
- SPEECH INPUT/OUTPUT CODING FORMAT
  - (a) 16-bit LINEAR
  - (b) 8-bit MU-255 LAW
- PERFORMANCE OPTIMIZATION FOR INPUT SPEECH BACKGROUND NOISE LEVEL

116363-N

SUMMARY

- VERY FLEXIBLE, COMPACT LPC VOCODER BASED ON A COMMERCIAL SIGNAL PROCESSING MICROCOMPUTER

- IMPLEMENTATION ACHIEVES ORDER OF MAGNITUDE DECREASE IN

  (a) POWER DISSIPATION

  (b) INTEGRATED CIRCUIT AREA

  (c) PRODUCTION COST

- VOCODER SUCCESSFULLY IMPLEMENTED AND DEMONSTRATED USING EPROM VERSIONS OF THE N.E.C. μPD7720

116364-N

FLEXIBLE ARRAY PROCESSORS

GLEN CULLER
CHI SYSTEMS, INC.

DEVELOPMENT OF FLEXIBLE ARRAY PROCESSORS

1. ARRAY PROCESSOR OVERVIEW
2. SIGNAL PROCESSING FRO ARC
3. THE FIRST LPC SPEECH ON ARPANET
4. THE FPS-120B COMMERCIALIZATION
5. THE PLASMA SIMULATION SPINOFF
6. THE LPCAP
7. THE CHI-5, A GENERAL PURPOSE ARRAY PROCESSOR
8. CHI-5 ARCHITECTURE ONTO VLSI CHIPS

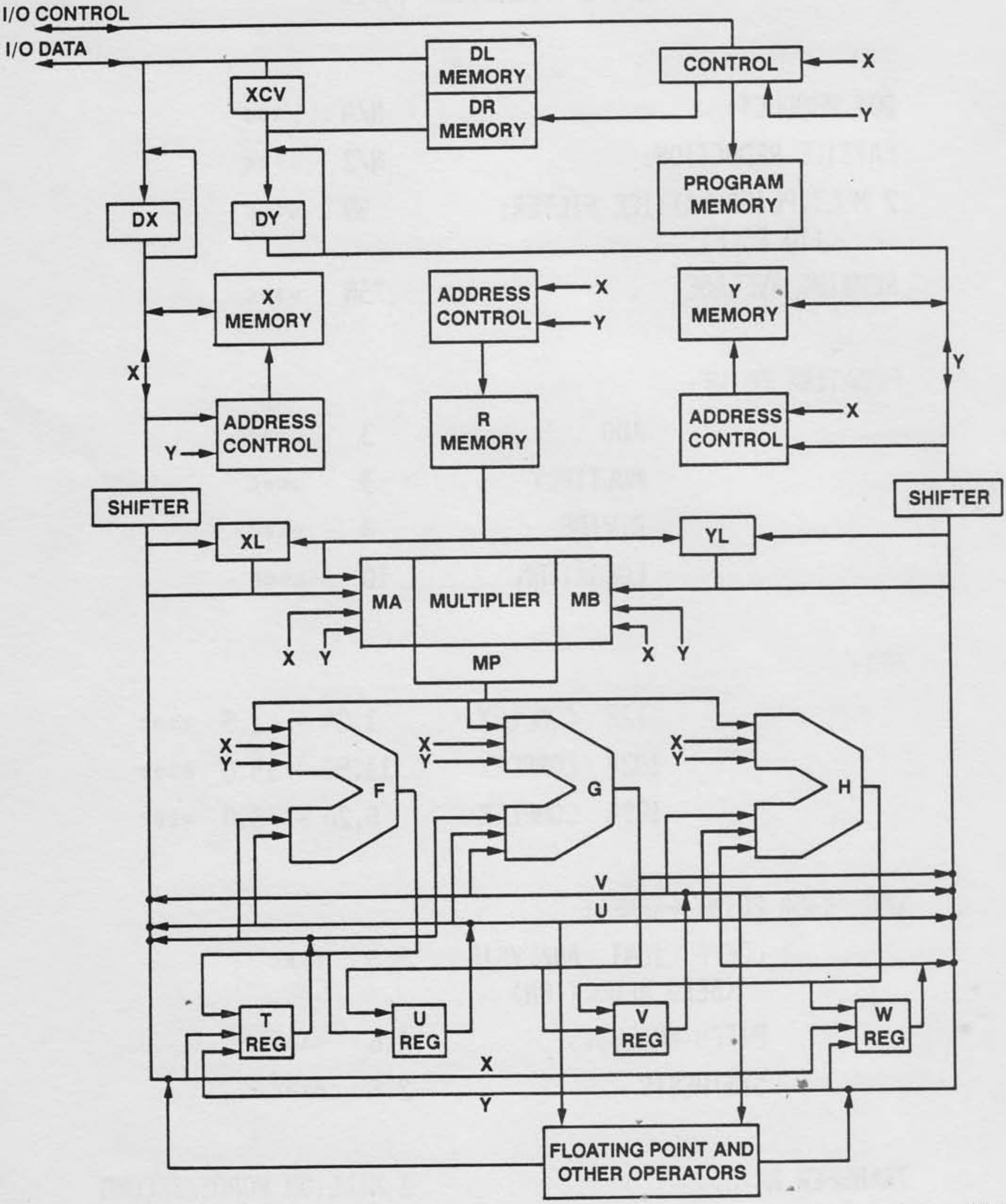
CHI-5 CHARACTERISTICS

1. GENERAL PURPOSE COMPUTER WITH ARRAY PROCESSOR INTERIOR
2. TWO ARRAY MEMORIES 1024 WORDS EACH
3. MATRIX ADDRESSING FOR ARRAY MEMORIES
4. FOUR STAGE PROGRAMMABLE PIPELINE
5. 16 PORT MEMORY
   - 8 PROCESS PORTS
   - 8 IO PORTS
6. Q-BUS AND UNIBUS INTERFACES
7. TWO CHANNEL RS-232 INTERFACE

CHI-5 EXECUTION TIMES

DOT PRODUCT:		N/4	µsec	
LATTICE REDUCTION:		N/2	µsec	
2 MULTIPLIER LATTICE FILTER: (10 POLE)		9N	µsec	
RUNNING AVERAGE:		.75N	µsec	
FLOATING POINT:				
	ADD	3	µsec	
	MULTIPLY	3	µsec	
	DIVIDE	8	µsec	
	LOGARITHM	10	µsec	
FFT:				
	128 COMPLEX	1.24 - 1.5	msec	
	1024 COMPLEX	11.50 - 15.0	msec	
	1024 COMPLEX	6.20 - 8.0	msec	
LPC (164 POINT FRAME):				
	COEFFICIENT ANALYSIS (BERG REDUCTION)	2.3	msec	
	PITCH ANALYSIS	1.6	msec	
	SYNTHESIS	2.5	msec	
TRANSFER RATE:		1 MILLION WORDS/SECOND		

ARRAY PROCESSOR ARCHITECTURE



8477-1

VLSI ARRAY PROCESSOR

EDWARD GREENWOOD
MOTOROLA - GOVERNMENT
ELECTRONICS GROUP

PROGRAM OBJECTIVES

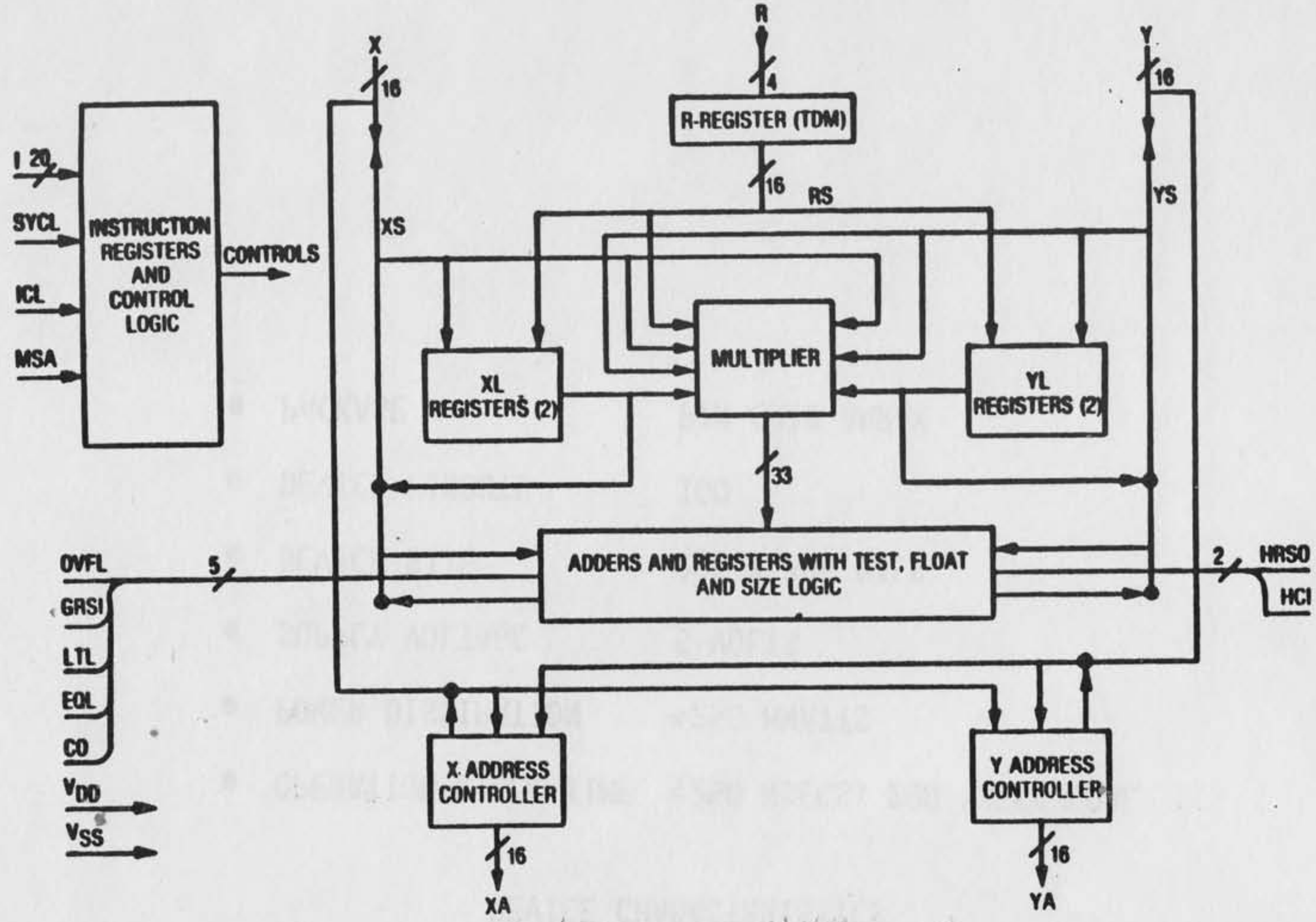
- DELIVER VLSI ARITHMETIC PROCESSOR CHIPS
- DELIVER TWO DEMONSTRATION SYSTEMS
- DEVELOP SINGLE MODULE ARRAY PROCESSOR CONCEPT
- DELIVER SOFTWARE DEVELOPMENT TOOLS
  - TRANSPORTABLE - FORTRAN 77
  - LAYERED - ARITHMETIC PROCESSOR
    - DEMONSTRATION SYSTEM
  - SIMULATORS AND ASSEMBLERS

ALGORITHM EXECUTION TIMES

ALGORITHM	TIME[(1)]
1024 POINT COMPLEX FFT	5632 µSECS
512 POINT REAL FFT	1152 µSECS
1024 POINT REAL VECTOR MULTIPLICATION	260 µSECS
1024 POINT COMPLEX MULTIPLY	520 µSECS
DARPA 2400 BIT FULL DUPLEX VOICE PROCESSING	≤5263 µSECS
COMPLEX CONJUGATE POLE-ZERO PAIR FILTER WITH GAIN	1.25 µSECS/POINT
32 BIT MANTISSA FLOATING POINT ADDITION	1.5 TO 2.5 µSECS
32 BIT MANTISSA FLOATING POINT MULTIPLICATION	.75, 1.0 OR 1.75 µSECS
1 MULTIPLICATION AND 3 DATA ADDITIONS	250 NSECS

(1) BASED ON 4 MHz SYSTEM CLOCK FREQUENCY AND BLOCK FLOATING POINT ARITHMETIC FOR PROCESSES

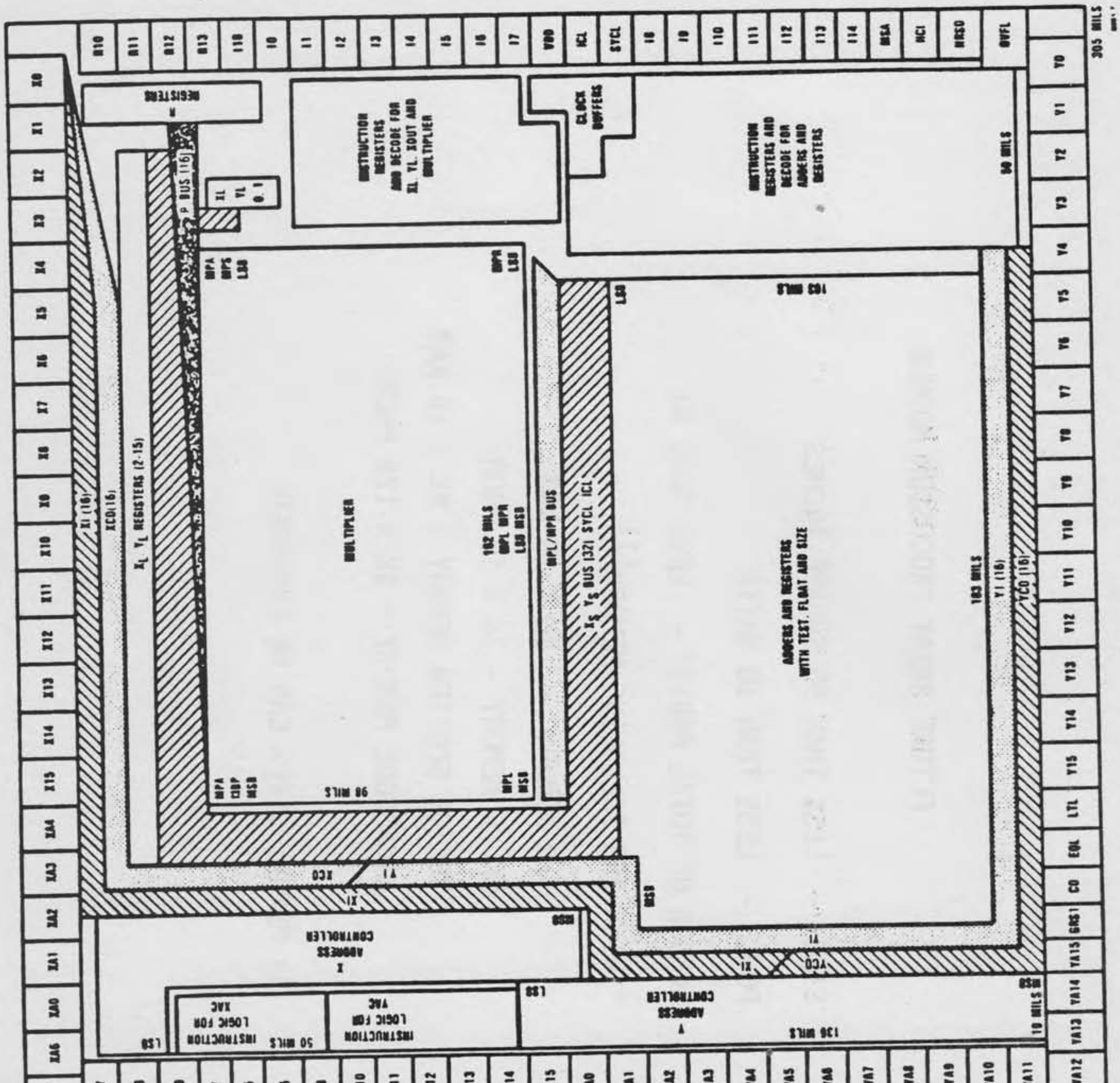
ARITHMETIC PROCESSOR UNIT



DEVICE CHARACTERISTICS

• OPERATION CYCLE TIME	≤250 NSECS; 200 NSECS GOAL
• POWER DISSIPATION	<250 MWATTS
• SUPPLY VOLTAGE	5 VOLTS
• DEVICE SIZE	298 X 305 MILS
• DEVICE PINOUTS	100
• PACKAGE	PIN GRID ARRAY

APU CHIP PLAN



FUTURE ARRAY PROCESSOR MODULE

SIZE - LESS THAN 50 SQUARE INCHES

POWER - LESS THAN 10 WATTS

NUMBER OF TOTAL PARTS* - LESS THAN 30

* WITH THE FOLLOWING MEMORY[(1)]

I/O MEMORY - 8K X 16 RAM

MACRO MEMORY - 1K X 16 PROM

VECTOR SCRATCH MEMORY - 4K X 16 RAM

MICROCODE MEMORY - 2K X 128 PROM

(1) MEMORY SIZES CAN BE EXPANDED

PROGRAM STATUS

- CHIP IN MASK SHOP
- DEMONSTRATION SYSTEM IN TEST
- SOFTWARE TASKS COMPLETED

SINGLE CHIP LPC

ROBERT BRODERSEN
UNIV. OF CALIFORNIA/BERKELEY

SINGLE CHIP LPC

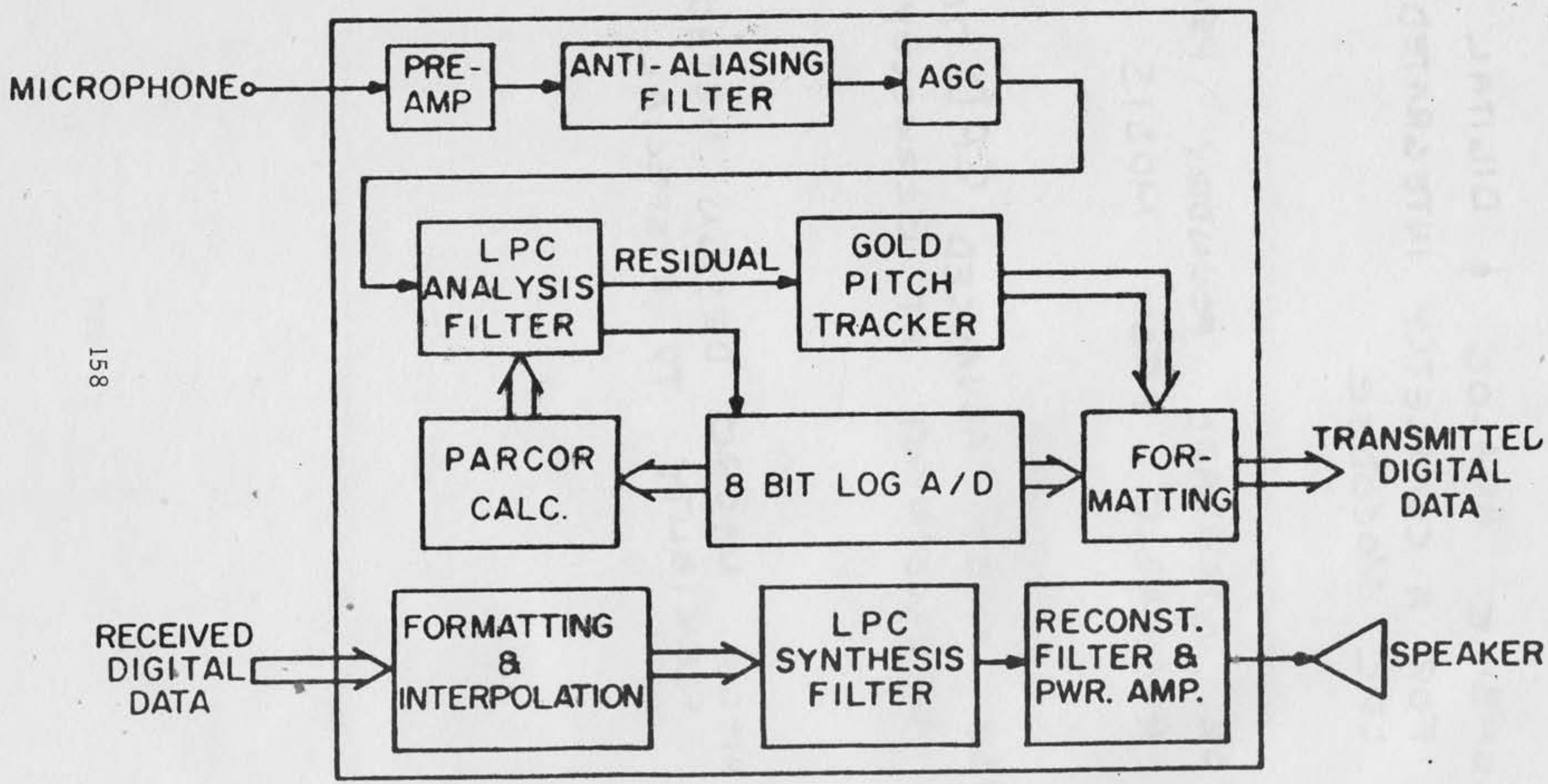
- COMBINE ANALOG & DIGITAL FUNCTIONS FOR A COMPLETLY INTEGRATED, FULL DUPLEX LPC VOCODER

- USE STANDARD FOUNDRY PROCESSING AVAILABLE FROM MOSIS

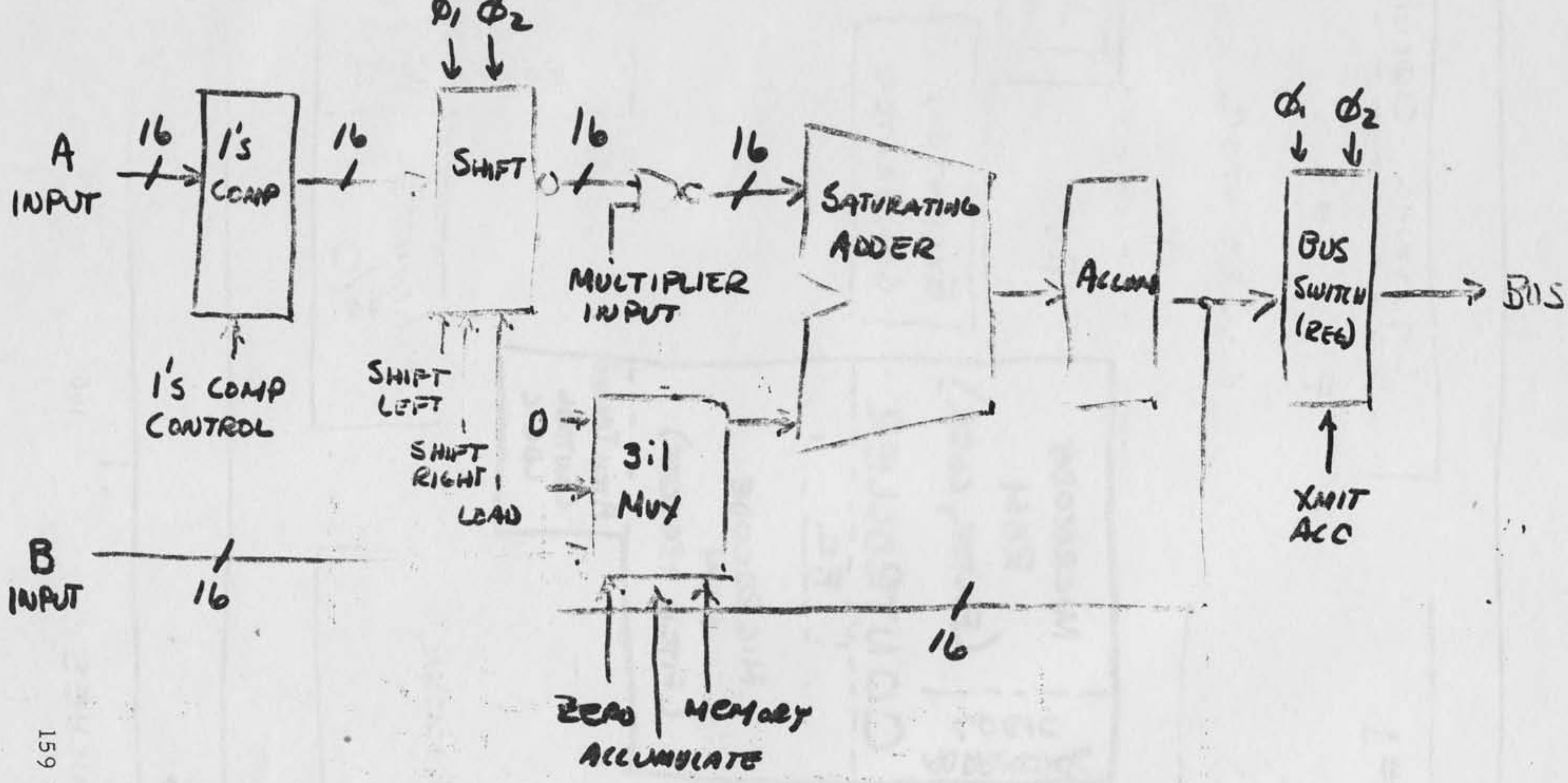
- USE MOST ADVANCED CAD TOOLS UNDER DEVELOPMENT AT BERKELEY

- LARGE MACRO DESIGN METHODOLOGY TO GIVE FLEXIBILITY TO A SPECIAL PURPOSE DESIGN

PROPOSED LPC VOCODER CHIP

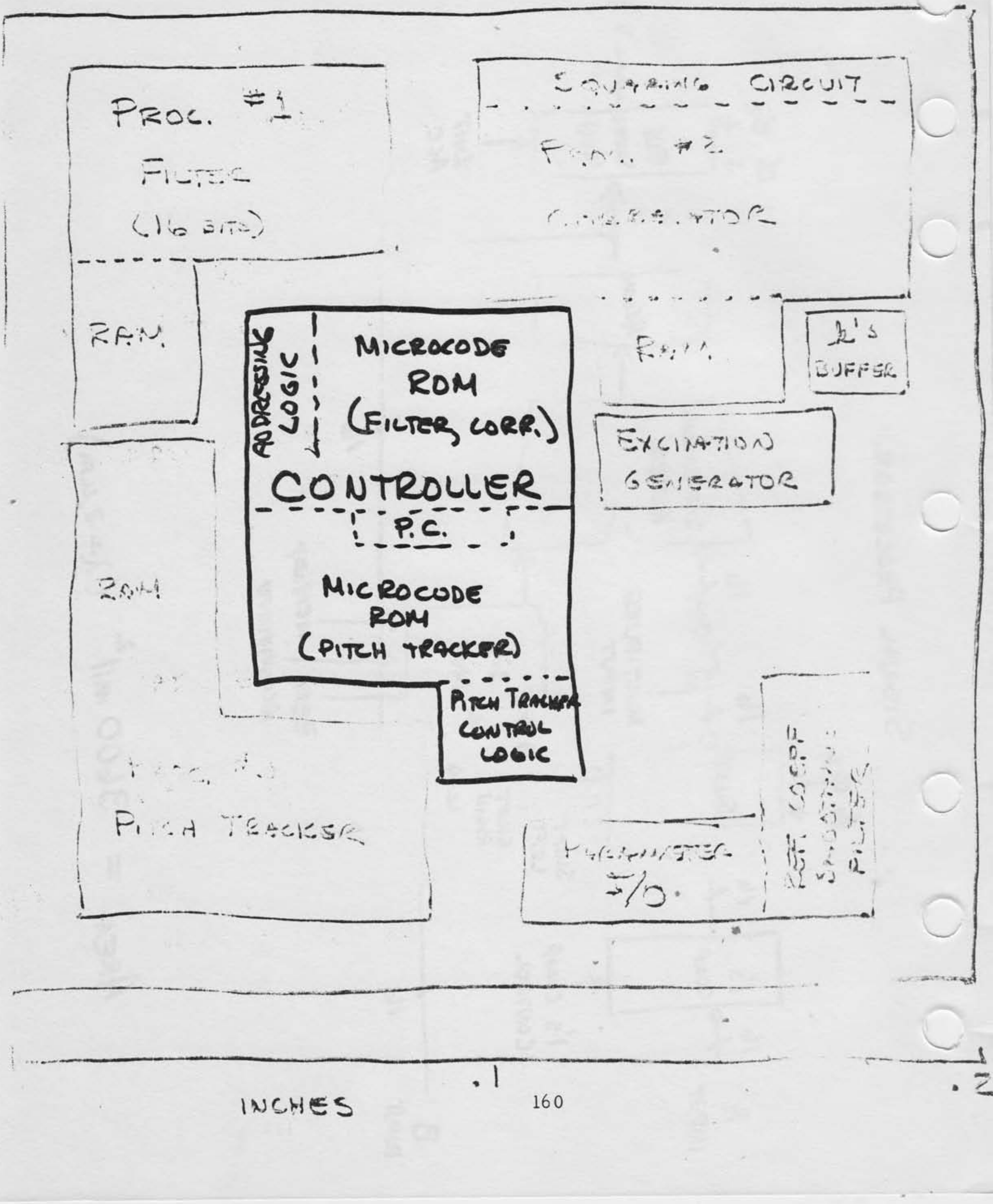


SIGNAL PROCESSOR



AREA = 3600 mil^2 ($\lambda = 2\mu m$)

Chip Layout



INCHES

PARALLEL PROCESSORS

PROCESSOR 1: ANALYSIS FILTER,
SYNTHESIS FILTER, PRE-EMPHASIS, DE-EMPHASIS
PITCH TRACKER FILTER

PROCESSOR 2: REFLECTION COEFFICIENT
CALCULATION AND RESIDUAL ENERGY
ESTIMATE

PROCESSOR 3: GOLD-RABINER ALGORITHM
AND VOICING DECISION

MISC. BLOCKS

EXCITATION GENERATOR: FORMS EXCITATION
WAVEFORM

PARAMETER I/O: COMMUNICATES WITH μP

CONTROLLER: PROGRAM COUNTER, MICROCODE ROM,
ADDRESSING ARITHMETIC, GENERATES
ALL CONTROL SIGNALS

LPC Vocoder chip summary

	Area (sq. mils)	Power (milliwatts)
Pre-amp	– 400	5
Anti-aliasing filter	– 2400	8
AGC	– 2000	10
LPC filters	(5000) 2x5000 =10,000	2x10 =20
LPC digital correlator	(7500) 10,000	40
A/D converter	– 10,000	150
Pitch tracker	(10,000)	
filters	4200	15
estimators	4800	15
scoring	7800	20
Formatting		
Interpolation		
Clocks	(2000) 8000	32
Power amp & output filt.	– 2000	15
Total	61,600	330

SIGNAL PROCESSOR EXECUTION CYCLE

PREAMBLE : PRE-EMPHASIS, DE-EMPHASIS AND LOW PASS FILTER FOR PITCH TRACKER; NOMINALLY 32 INSTRUCTIONS

ANALYSIS/
SYNTHESIS : 10 CYCLES OF 32 INSTRUCTIONS = 320 ; 16 INSTRUCTIONS FOR EACH FILTER

CLK. RATE = 352 × 8 kHz = 2.816 MHz

SUMMARY

- 20,000 TRANSISTORS
- .85 WATT POWER DISSIPATION (NMOS)
- 40,000 mil^2 – 4 μm CHANNEL LENGTH
- 6K CONTROL ROM
- 2K DYNAMIC RAM
- 3 PARALLEL PROCESSORS
-