3.7 Arpanet Control Algorithms

Jim Herman
Bolt, Beranek and Newman

NETWORK CONTROL ALGORITHMS

- Routing

  Ensure Delivery, Purge Packets for Unreachable Destinations

- Flow Control

  Match the Rate of the Sender to That of the Receiver

- Congestion Control

  Throttle Senders That Are Over-Loading Network Resources

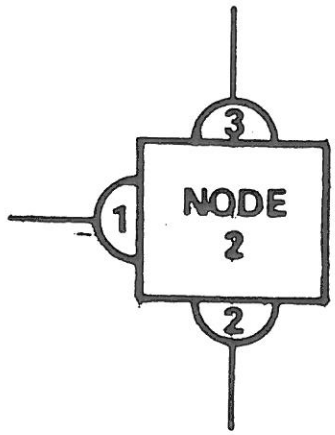
THE IMP PROGRAM
ROUTING

- FORWARDING PACKETS
- MEASUREMENT
- UPDATING
- ROUTE CALCULATION

ROUTING
FORWARDING PACKETS

- PACKETS FORWARDED THROUGH TANDEM NODES
- NEXT HOP DETERMINED BY SIMPLE TABLE LOOKUP
- PACKETS FOR UNREACHABLE DESTINATIONS PURGED

ROUTING
FORWARDING PACKETS



PACKET

1	INPUT MODEM
6	DESTINATION

ROUTING TABLE

2	1
0	2
-1	3
3	4
1	5
3	6
⋮	

ROUTING MEASUREMENT

- CONTINUOUS MONITORING OF LINE ERROR RATE
- LINES DECLARED UP OR DOWN AUTOMATICALLY
- CONTINUOUS MEASUREMENT OF PACKET DELAYS
- DELAY = RESIDENCY + TRANSMISSION TIME + PROPAGATION DELAY
- AVERAGED OVER 10 SECONDS

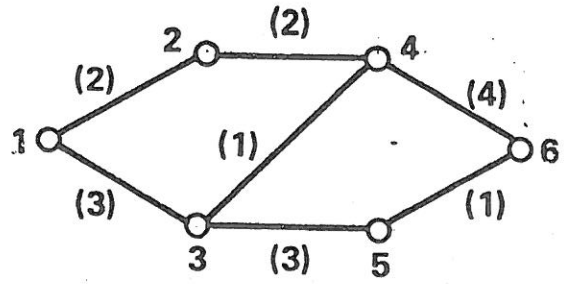
ROUTING UPDATING

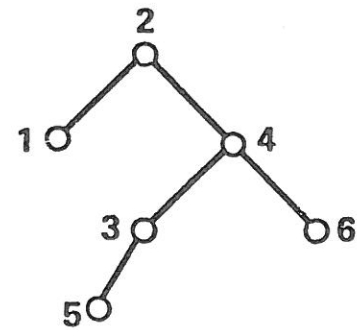
- CHANGE IN LINE STATE OR SIGNIFICANT CHANGE IN DELAY TRIGGERS UPDATE
- UPDATES PROPAGATE AT HIGHEST PRIORITY LEVEL
- UPDATES PROPAGATE WITHOUT ROUTING-FLOODING
- UPDATE SENT OUT ALL LINES
- UPDATE ALREADY RECEIVED IS QUENCHED
- NETWORK UPDATED IN LESS THAN 100 ms

ROUTING
ROUTE CALCULATION

- DATABASE OF LINKS, ENDPOINTS AND DELAYS
- SHORTEST PATH FIRST ALGORITHM
- CONSISTENT DATABASES = CONSISTENT ROUTING TABLES

SPF TREE FOR NODE 2





Unreachable Nodes Do Not Appear on Tree
Not All Lines Appear in Tree (e.g. Line From 1 to 3)

SPF CALCULATION

For Each Entry in An Update for Which There Is a Change in Delay

- Delay Worse But Line Not in Tree — (No Action)
- Delay Worse For Line in Tree — Consider Subtree
- Delay Better for Line in Tree — Consider All but Subtree
- Delay Better for Line Not in Tree — Consider Attaching

Complexity Bound-

- Average Subtree Size = Average Path Length
- Probability That a Line Is in the Tree = $\frac{1}{\text{Connectivity}}$
- Expected Number of Routing Changes = $\frac{\text{Average Path Length}}{\text{Connectivity}}$

END-TO-END FLOW CONTROL

- Operates on a Message Basis
- Sequence Number on Each Message
- Limit of Eight Outstanding Messages to Any Destination Host
- Ninth Message Will Block the Host Interface
- Acknowledgement (RFNM) Advances Sequence Number
- Explicit Reservation of Buffer Space in Receiving IMP

POSSIBLE IMPROVEMENTS TO END-TO-END FLOW CONTROL MECHANISMS

- Variable Sized Windows Depending on Delay to Destination
- Uncouple Sequencing, Error Control and Flow Control
- Non-Blocking (Refusable) Host-to-IMP Protocols

CONGESTION CONTROL

- Measurement
- Updating
- Determination of Congested Destinations
- Metering

CONGESTION CONTROL MEASUREMENT

Global Node Congestion — Length of Task Queue

- Exceeds Threshold

Link Congestion — Numberof Packet Refusals

Must Be Smoothed Over Suitable Interval

Link Can Be In One of Three States

- Underloaded — No Congestion Events
- Maximally Loaded — Congestion Events Below Threshold
- Congested — Number of Congestion Events Exceeds Threshold for Several Successive Intervals

CONGESTION CONTROL UPDATING

- Two Bits of Information Per Line
- Carried With SPF Updates
- Change in Congestion State Triggers Update Also
- Should Cause Minimal Increase in Number of Updates
- Should Not Contribute to Congestion

DETERMINATION OF CONGESTED DESTINATIONS

- Destination Is "Congested" if Path to It Has At Least One Congested Link
- Destination Is "Maximally Loaded" if Path to It Has No Congested Links, But at Least One Maximally Loaded Link
- Destination Is "Underloaded" if Path to It Has Only Underloaded Links
- SPF Algorithm Modified to Propagate Congestion State Down Subtrees

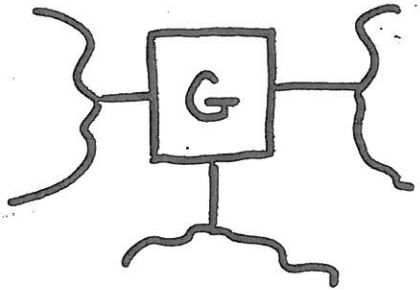
CONGESTION CONTROL METERING

- Controls Applied When Destination Enters Congested State
- Controls Tightened Until Destination Leaves Congested State
- Controls Loosened Gradually When Destination Is In Underloaded State
- Simple Control: Limit Number of Outgoing Channels That Can Be Filled With Packets for a Given Destination
- Complex Control: Measure Flows to Each Destination and Explicitly Limit Flows Per Interval

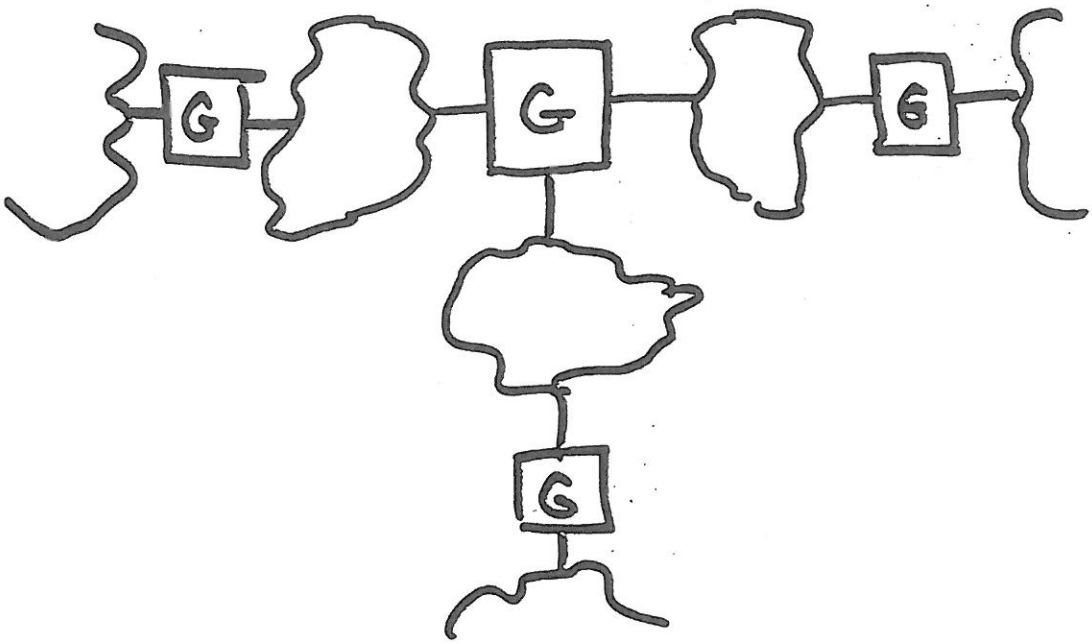
3.8 Internet Management and Control

Jack Haverty
Bolt, Beranek and Newman

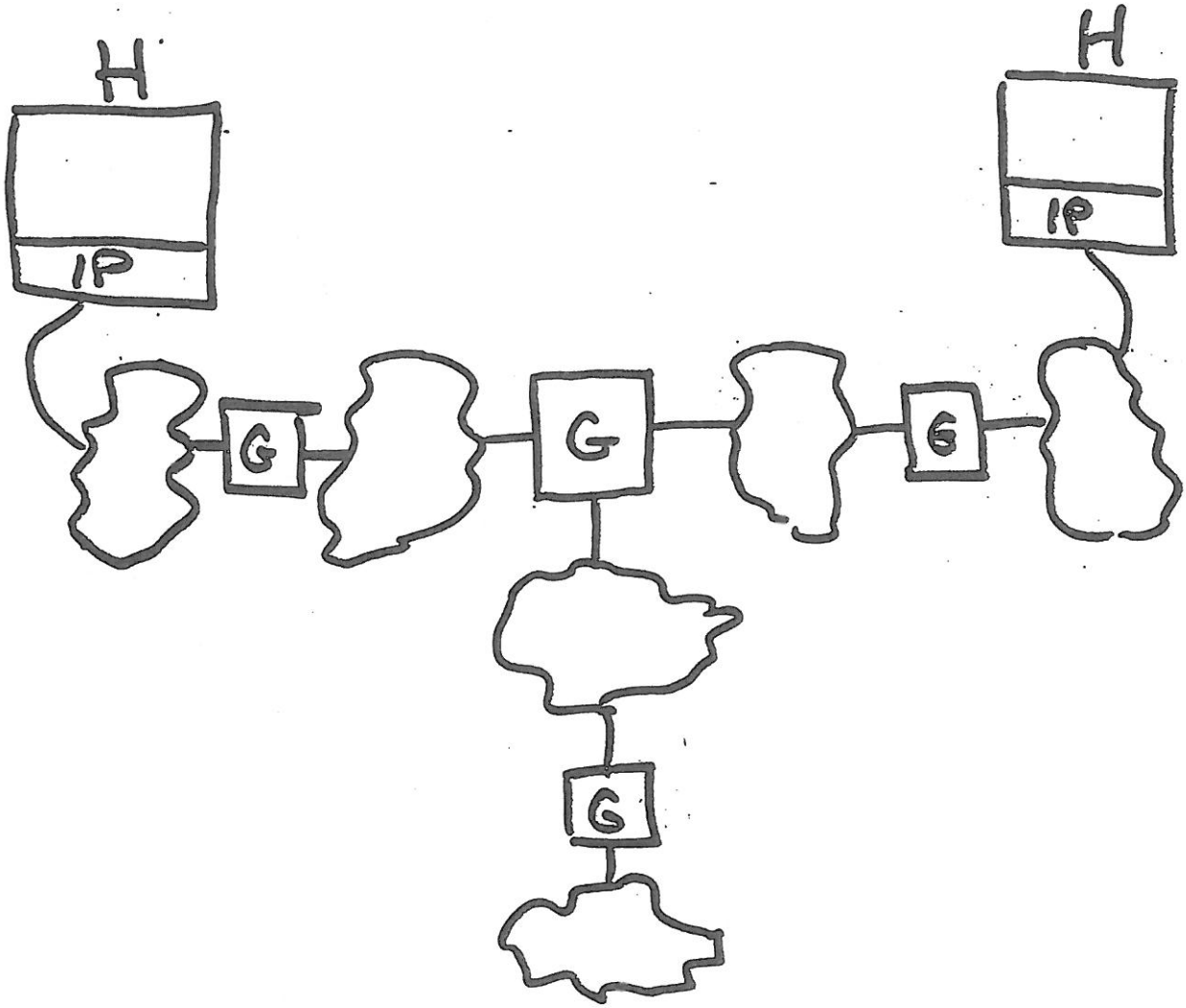
INTERNET NETWORK



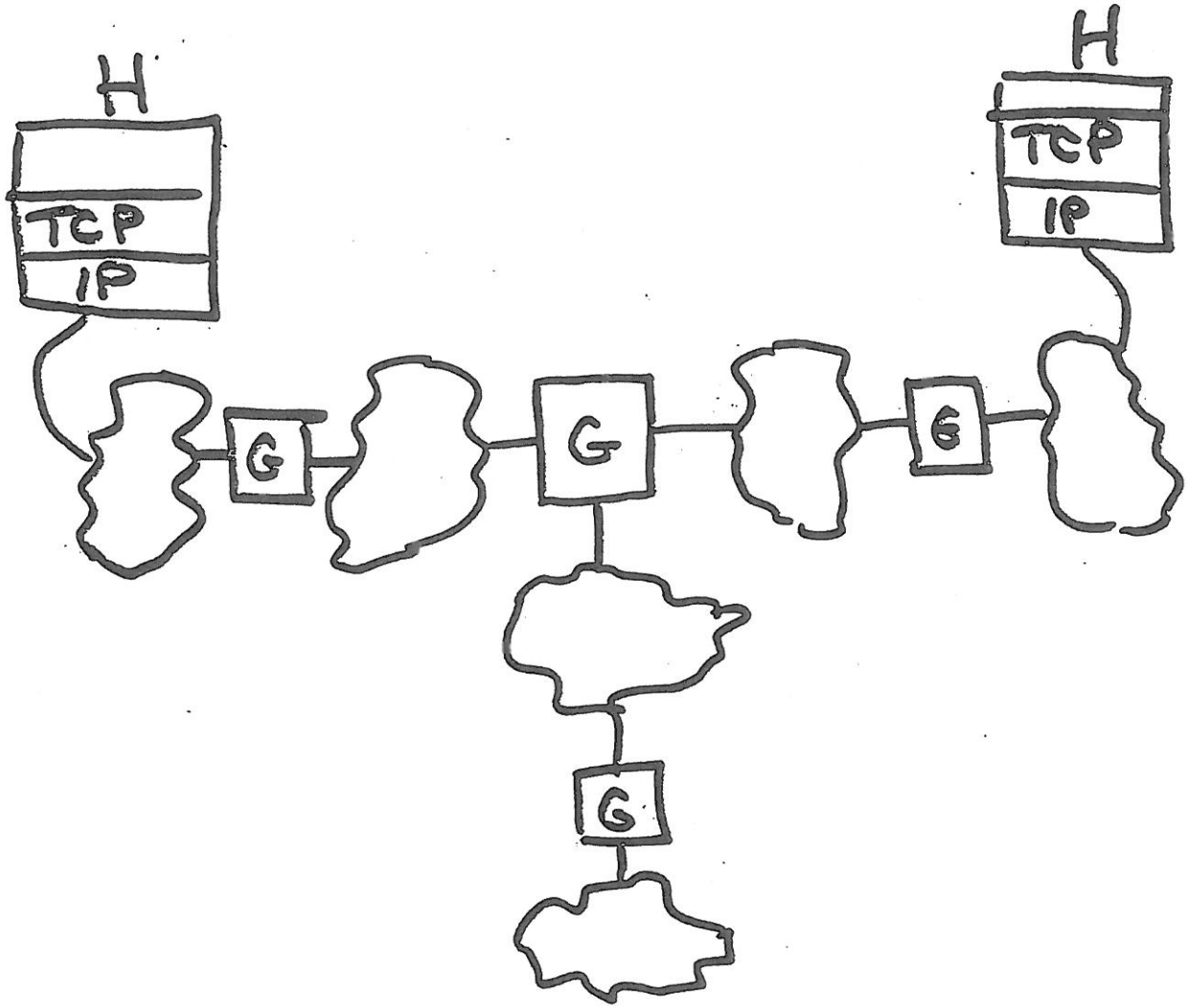
INTERNET NETWORK



INTERNET NETWORK



INTERNET NETWORK



ROUTING

- MIN HOP (GGP)
- SOURCE ROUTING
- "DUMB" GATEWAYS (FIXED ROUTING)
- VARIOUS UP/DOWN PROTOCOLS

FLOW CONTROL

- TCP LEVEL (H - H CONTROL ONLY)
- DOES NOT CONTROL DG FLOWS

CONGESTION CONTROL

- SOURCE QUENCH
- RATE CONTROL ONLY, DOES NOT INTERACT WITH ROUTING

- ROUTING → SPF
  - DELAY-BASED
  - MULTIPLE TOPOLOGY SUPPORT
    - TYPE OF SERVICE
    - RESTRICTED ROUTING

- CONGESTION CONTROL
  - MODEL AFTER ARPANET
  - AFFECT ROUTING AND FLOW CONTROL

- FLOW CONTROL
  - INTRODUCT STATE INFO
  - GW ↔ GW INFO (ON ENTRY, SELECT EXIT GW)
  - HOST INTERFACE ENHANCED TO SUPPLY FLOW INFO

- G - G INTERACTIONS
  - NOW GGP AND IP
  - CHANGE TO CONTROL PROTOCOL AND DATA PROTOCOL
  - CONTROL PROTOCOL
    - ROUTING, RESOURCE MGMT, FAIRNESS
  - DATA PROTOCOL
    - ALLOWS AGGREGATION, INTRA_NET REPACKETIZATION, ETC.
  - EGP

1) FIGURE OUT WHAT YOU ARE BUILDING

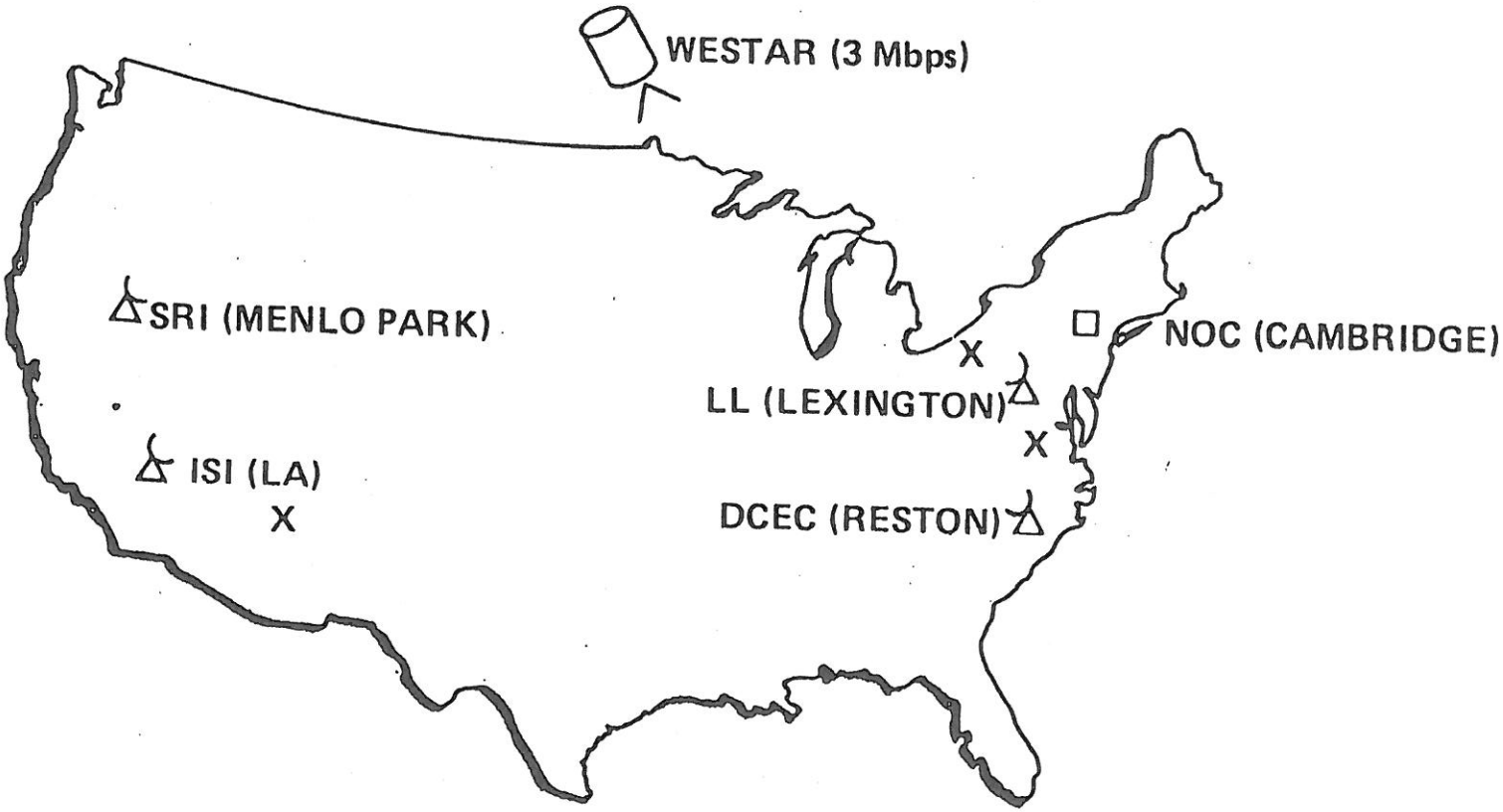
2) BE PARANOID

3) EXPECT CHANGES EVERYWHERE

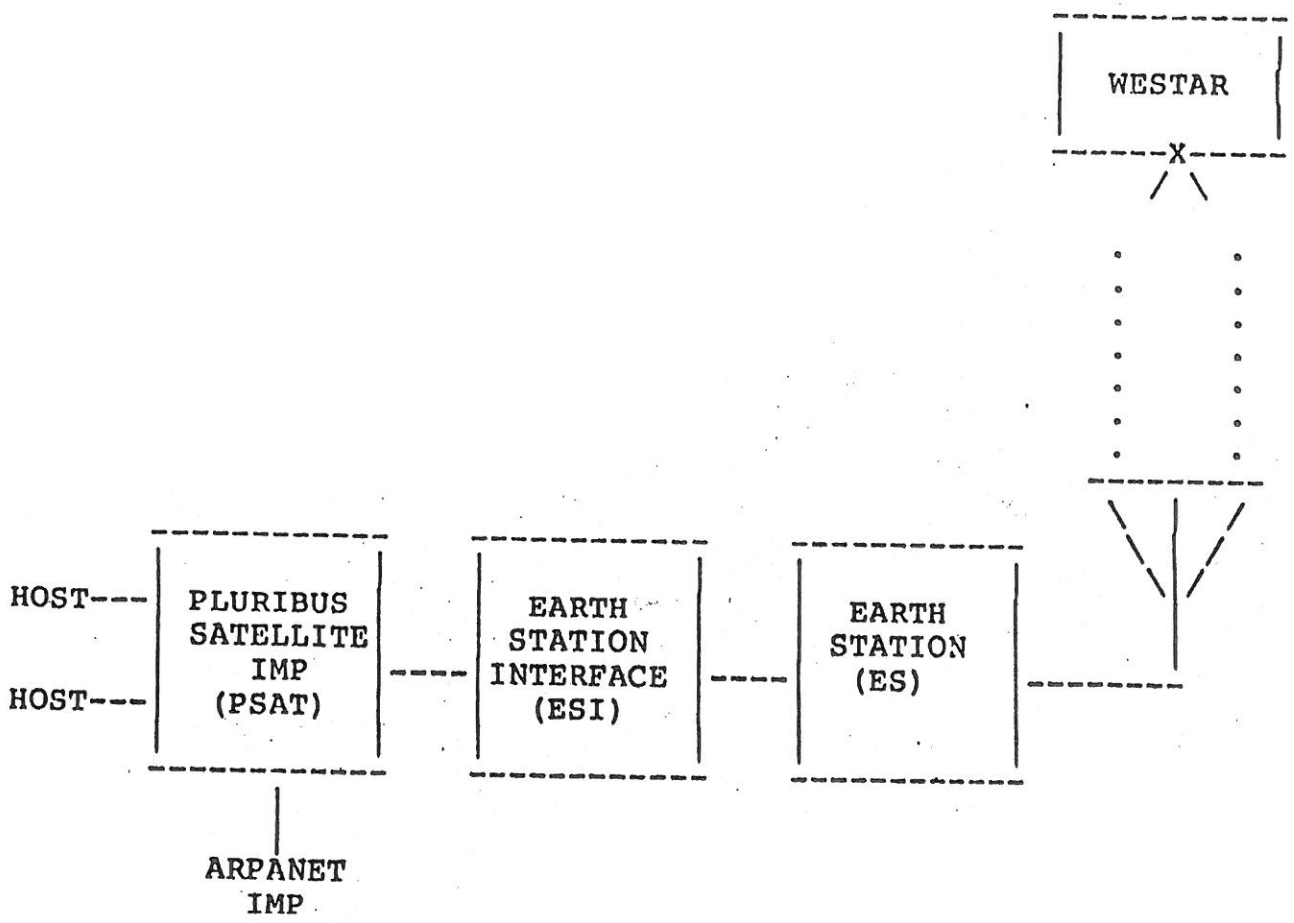
3.9 Wideband Satellite Network

Gil Falk
Bolt, Beranek and Newman

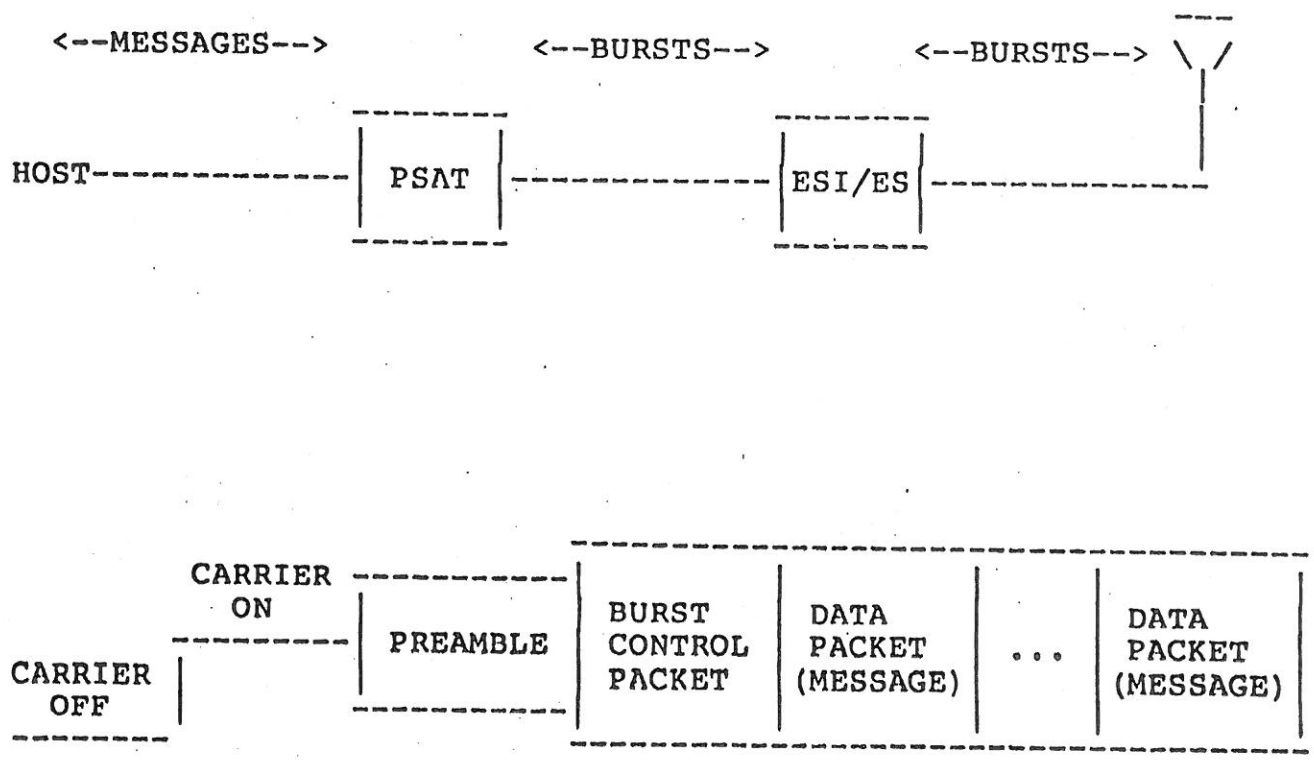
EISN/WIDEBAND NETWORK SATELLITE SUBSYSTEM



SITE CONFIGURATION

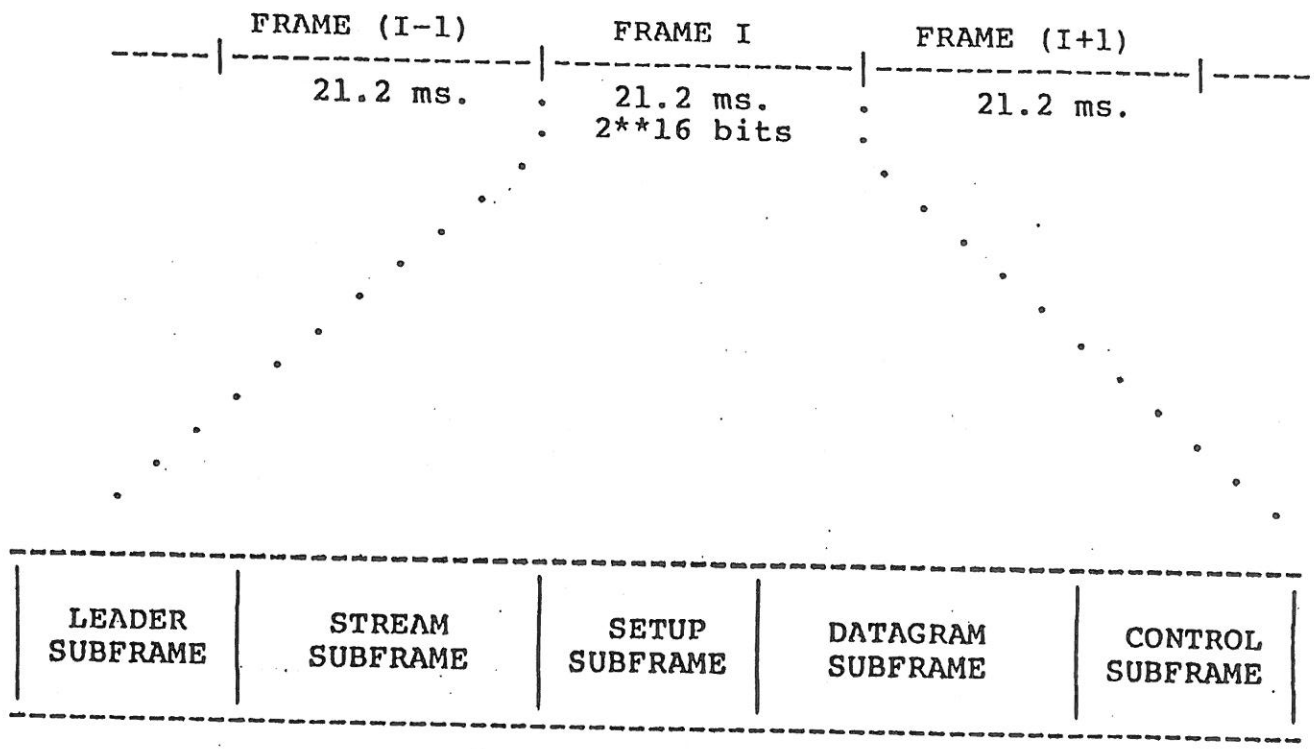


BURST STRUCTURE



MESSAGE = DATAGRAM MESSAGE OR STREAM MESSAGE

CHANNEL FRAMING



PODA STREAM SUBFRAME

```
|L|  Si  | D | C |L| Si+1 | D | C |L| Si+2 | D | C |
+-----------------+----------------+----------------+
|     FRAME i     |    FRAME i+1   |    FRAME i+2   |
```

L = LEADER SUBFRAME
S = STREAM SUBFRAME
D = DATAGRAM SUBFRAME
C = CONTROL SUBFRAME

```
        | CS1=(HS1)   CS2=(HS2,HS3)   CS3=(HS4)  CS4=(HS5)  |
        |+---------+ +-----+--------+            +---------+|
Si      || HS1     || |HS2 | HS3    |            | HS5     ||
        ----------------------------------------------------

        |+---------+ +-------+-------+            |
Si+1    || HS1     || | HS2   | HS3   |            |
        ----------------------------------------------

        |+---------+ +---------+            |            +---------+|
Si+2    || HS1     || | HS3     |            |            | HS5     ||
        ----------------------------------------------------
```

A COMPARISON OF SATNET AND WIDEBAND NETWORK MECHANISMS

	SATNET	WIDEBAND NETWORK
SATELLITE	INTELSAT IV-A (C-BAND)	WESTAR (C-BAND)
ANTENNA SIZE	30 M	5 M
IDENTICAL STATIONS	YES TESTED WITH UET 11 DB DISADVANTAGE	YES
CHANNEL RATE (MAX)	64 KBPS 128 KBPS (FUTURE)	3.088 MBPS
CHANNEL SYMBOL RATES	32 KSYM/SEC	1544 KSYM/SEC 772 KSYM/SEC 386 KSYM/SEC 193 KSYM/SEC
MODULATION TYPES	BPSK, QPSK	BPSK, QPSK
FEC RATES	1/2, 1	1/2, 3/4, 7/8, 1
HOST ACCESS	VDH HDH (FUTURE) HOST/SATNET PROTOCOL	VDH (FRAMING ONLY) 1822 HAP
CURRENT CONTROLLER IMPLEMENTATION	H316 C/30	PLURIBUS

A COMPARISON OF SATNET AND WIDEBAND NETWORK MECHANISMS (CONTINUED)

	SATNET	WIDEBAND NETWORK
FRAME LENGTH	317 MS 20.3 KBIT	21.2 MS 64 KBIT
FRAME SUBSTRUCTURE	INFORMATION CONTROL	LEADER STREAM SETUP DATAGRAM CONTROL
PODA OPTIONS	FPODA CPODA	FPODA CPODA
SUPERFRAME	HELLO FRAME EVERY 4 PODA FRAMES	NONE
MESSAGE TYPES	DATAGRAM STREAM	DATAGRAM STREAM
BURST STRUCTURE	1 MESSAGE AT SINGLE PRIORITY	AGGREGATED MESSAGES AT DIFFERENT PRIORITIE
MAXIMUM LENGTH (MESSAGE/BURST)	2 KBITS/4 KBITS	16 KBITS/32 KBITS
SCHEDULING UNIT	16 BITS	1 CHANNEL SYMBOL

A COMPARISON OF SATNET AND WIDEBAND NETWORK MECHANISMS (CONTINUED)

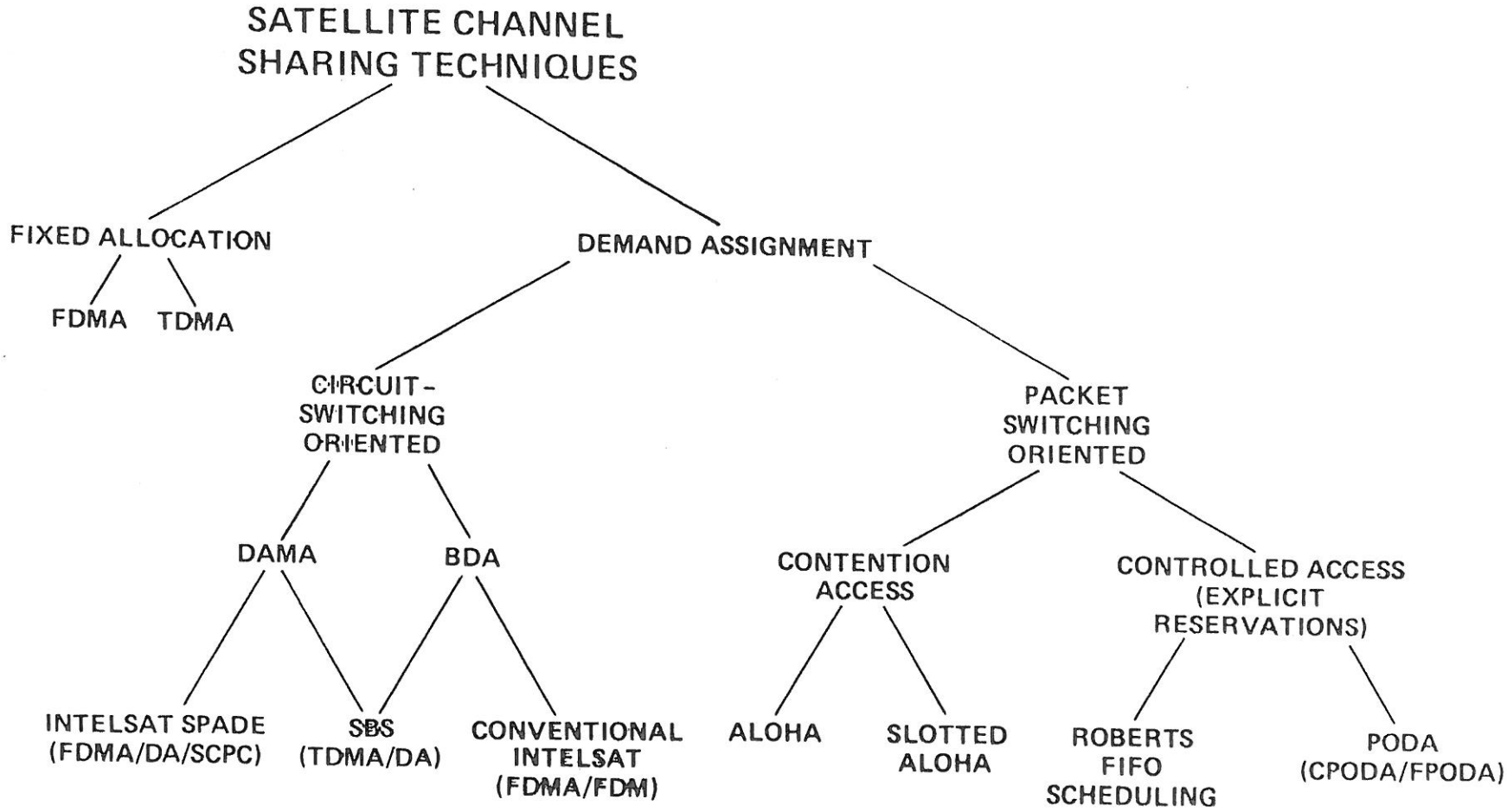
	SATNET	WIDEBAND NETWORK
BASIS FOR HANDLING MESSAGES	URGENCY	PRIORITY
BASIS FOR ORDERING RESERVATION TRANSMISSIONS	URGENCY	PRIORITY
BASIS FOR SCHEDULING RECEIVED RESERVATIONS	URGENCY	FIFO
RECOVERY FROM LOSS OF DATAGRAM SYNC	COMPUTATIONALLY INTENSIVE	ADDITION OF OFFSET ONLY
MULTIFRAME BURST FRAGMENTATION	NO	DATAGRAMS
RESERVATION TRANSMISSSION	CONTROL BURST OR PIGGYBACKED	CONTROL BURST OR PIGGYBACKED
STREAM RATES SUPPORTED	ARBITRARY	EVERY 1,2,4, OR 8 FRAMES
STREAMS SHARED BY MULTIPLE SITES?	YES	NO
STREAM SCHEDULING	INTERMIXED WITH DATAGRAMS USING PSEUDO RESERVATIONS	SEPARATE SUBFRAME

A COMPARISON OF SATNET AND WIDEBAND NETWORK MECHANISMS (CONTINUED)

	SATNET	WIDEBAND NETWORK
ADDRESSING	LOGICAL STATIC GROUP	LOGICAL DYNAMIC GROUP
UNCODED BIT ERROR PERFORMANCE AT MAX CHANNEL RATE	10**-6	10**-3
ERROR CONTROL	CRC ARQ (OPTIONAL)	FEC CRC
OPTIONALLY PASS ON OR DISCARD MESSAGES CONTAINING ERRORS	YES	YES
CONGESTION CONTROL	LIMITED	LIMITED
END-TO-END FLOW CONTROL	LIMITED	NO

3.10 Packet Satellite Network

Dale McNeill
Bolt, Beranek and Newman


SATELLITE CHANNEL SHARING TECHNIQUES
FIXED ALLOCATION
FDMA
TDMA
DEMAND ASSIGNMENT
CIRCUIT-SWITCHING ORIENTED
PACKET SWITCHING ORIENTED
DAMA
BDA
CONTENTION ACCESS
CONTROLLED ACCESS (EXPLICIT RESERVATIONS)
INTELSAT SPADE (FDMA/DA/SCPC)
SBS (TDMA/DA)
CONVENTIONAL INTELSAT (FDMA/FDM)
ALOHA
SLOTTED ALOHA
ROBERTS FIFO SCHEDULING
PODA (CPODA/FPODA)

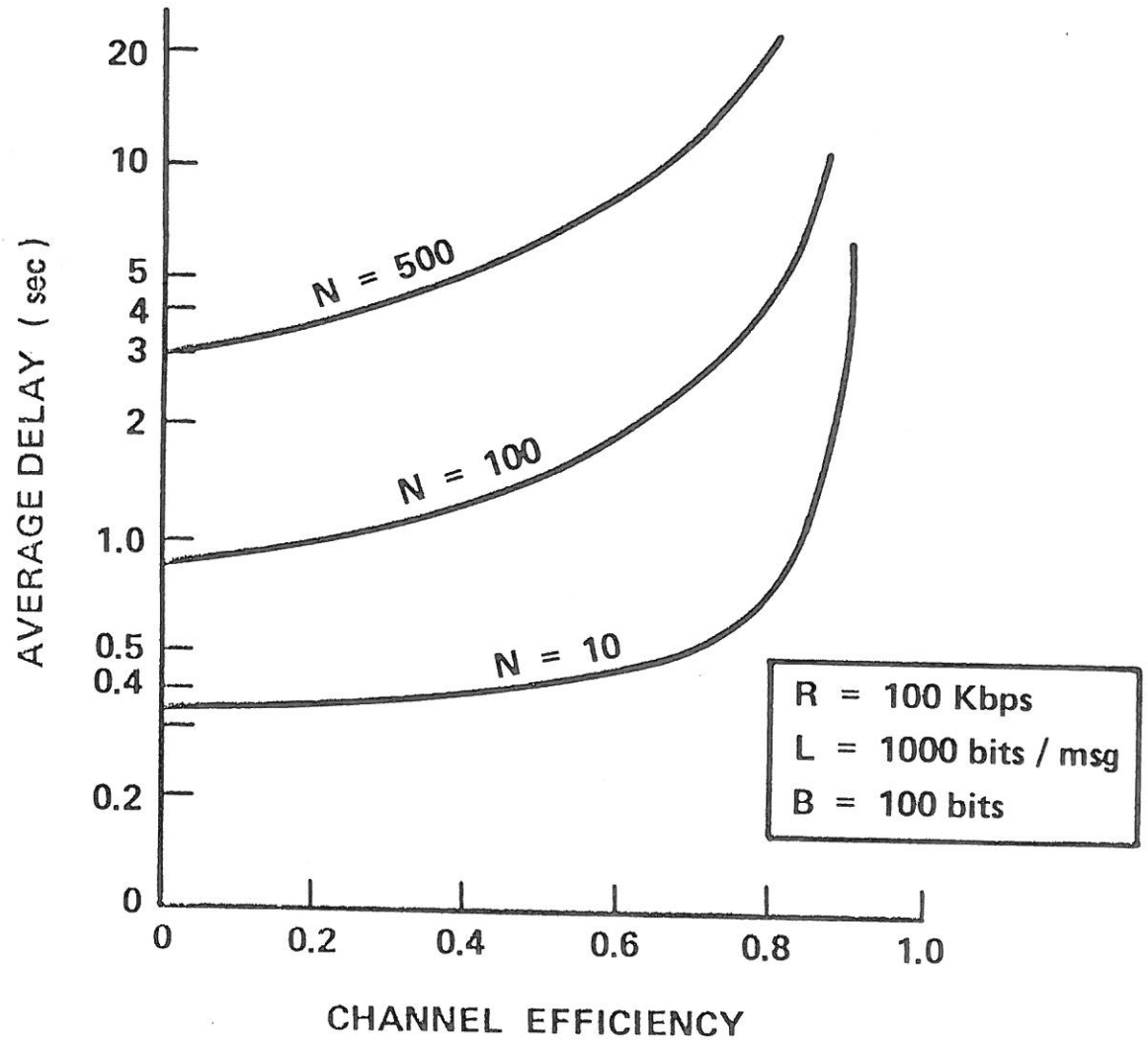
PACKET SATELLITE TECHNOLOGY

- A Single Channel is Shared by All Stations, Providing Full Connectivity with Single-Channel Hardware at Each Station
- Both Bursty and Constant Data Traffic can be Handled Efficiently
- Each Transmission From a Station can be Received by Multiple Destination Stations
- New Stations may be Added Without Changing Hardware at the Other Stations
- Changes to Station Bandwidth Requirements are Automatically Accommodated within the Limits of the Total Available Channel Capacity

FIXED ASSIGNMENT

- Simple Implementation
- Very Robust
- Good Performance for High Traffic Concentration Per Station and Non–Varying Traffic Distribution Among Stations
- Variable– Length Problem
- Does Not Allow Takeover of Capacity by Stations with High–Priority Traffic
- Stream Traffic Not Accommodated at its Required Intervals

FIXED ASSIGNMENT
DELAY – EFFICIENCY PERFORMANCE



ALOHA

- Good Average Performance (but High Variance)
  If :
  - Very Low Traffic Concentration Per Station
  - Short Messages
  - Large Number of Stations
- Requires Stability Control
- Priority Access Achieved at Expense of Efficiency
- Slotted ALOHA has Variable – Length (Fixed Assignment) Problem
- Stream Traffic Not Effectively Accommodated

PODA - PRIORITY ORIENTED DEMAND ASSIGNMENT

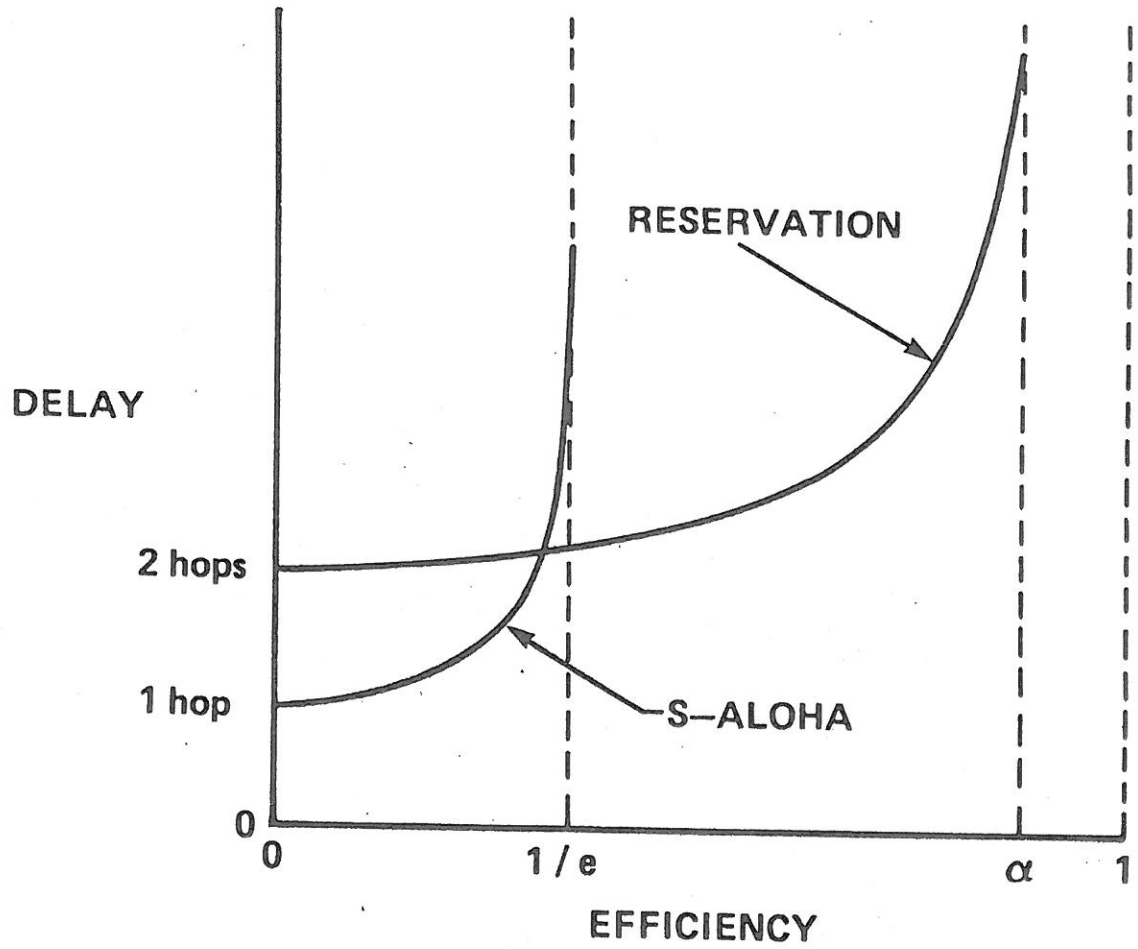
A Reservation System Providing Maximum Flexibility for Channel Allocation:

- Arbitrary Number of Stations Can Share Channel
- High Availability Achieved Through Use of Distributed Control Techniques
- Datagram and Stream Service Provided

PODA DATAGRAM SERVICE

- Messages Sent as Individual Packets
- Per Packet Reservation
- Channel Scheduling Incorporates:
  - Priority
  - Delay Class
  - Holding Time
  - Reliability
- Traffic Adaptive
- Distributed Control Results in Two - Hop Service

COMPARISON OF EXPLICIT RESERVATION AND S–ALOHA

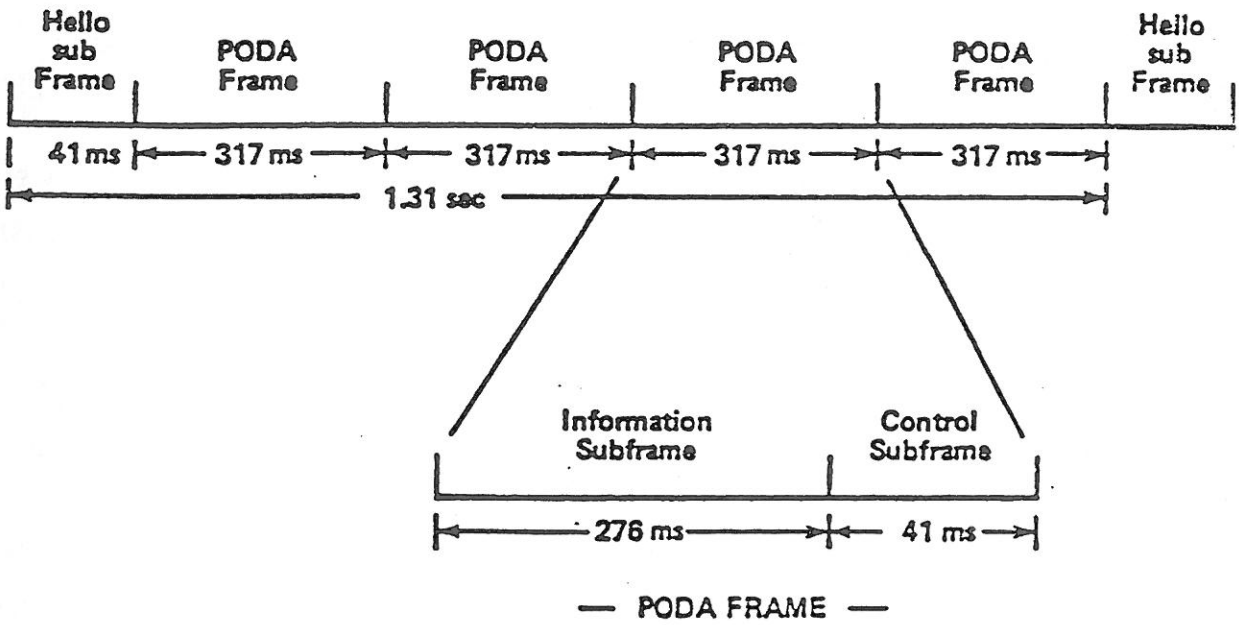


CENTRALIZED ASSIGNMENT

- 3 – Hop Datagram Delay
- Channel Overhead Penalty
- No Synchronization Required
- Any 'Large' Station can Automatically Take Over as Central Controller to Maintain High Availability

DISTRIBUTED ASSIGNMENT

- 2 – Hop Datagram Delay
- Synchronization Required
- Performance Limited by Channel Data Rate and Local Error Rates
- High Availability



SATNET Frame Structure
Figure 2