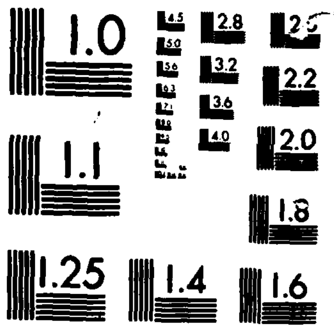
AD-A167 803 REQUIREMENTS FOR THE NEXT GENERATION PACKET SWITCH(U) 1/1
SPARTA INC MCLEAN VA 22 APR 86 URW-6-DCA-053
DCA100-84-C-0085

UNCLASSIFIED F/G 17/2 NL



MICROCOP' CHART

Requirements For The Next Generation Packet Switch

2

Requirements For The Next Generation Packet Switch

Prepared By

SPARTA, Inc.
7926 Jones Branch Drive
Suite 1070
McLean, VA 22102

April 22, 1986

Prepared For

Defense Communications Engineering Center
1860 Wiehle Avenue
Reston, VA 22090

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188
Exp Date: Jun 30, 1986

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS: NONE
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: UNLIMITED
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): UNN-6-DCA-053	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: SPARTA, INC. — 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: DEFENSE COMMUNICATION ENGINEERING CENTER
6c. ADDRESS (City, State, and ZIP Code): 7926 Jones Branch Drive, Suite 1070, McLean, VA 22102	7b. ADDRESS (City, State, and ZIP Code): 1860 Wiehle Ave., Reston, VA 22090
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: DCEC — 8b. OFFICE SYMBOL (If applicable): R640	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: DCA100-84-C-0086
8c. ADDRESS (City, State, and ZIP Code): 1860 Wiehle Ave., Reston, VA 22090	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Requirements for the Next Generation Packet Switch

12. PERSONAL AUTHOR(S) SPARTA, INC.

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM 10/22/84 TO 4/22/86	860422	88

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Computer Networks, Machine Architecture, Packet Switches,

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report discusses the requirements for the next generation packet switch, to be used in the Defense Data Network. These requirements are based upon examination of the current and anticipated trends in the Defense Data Network and upon packet switch hardware and software technology trends. Specific requirements are then developed for the next generation packet switch performance, reliability, maintainability and configuration. Finally, procurement strategies and schedules are discussed.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL: Mr. Ed Cain	22b. TELEPHONE (Include Area Code): (703) 437-2578	22c. OFFICE SYMBOL: R640

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

NEXT GENERATION PACKET SWITCH

TABLE OF CONTENTS

APPENDIX A: DISCUSSION OF PACKET SWITCH ARCHITECTURE AND PERFORMANCE

APPENDIX B - ALTERNATIVE NETWORK ENVIRONMENTS

1.0 INTRODUCTION

1.1 Purpose of the Report

This report has been produced in response to two paragraphs in the Statement of Work (for Contract #DCA100-84-C-0085, "Analysis and Resolution of Packet Switching Issues") which read as follows:

> 2.3 Next Generation Packet Switch
>
> As higher bandwidth long-haul (satellites, optical fiber) and short-haul (broad-band cable, microwave) transmission media become cost-effective for DDN trunking, and as high bandwidth host applications become more widespread, a radically different packet switch will be needed for the DDN functions performed by hardware or firmware. A genuine need for a new switch must be forecast far enough in advance to permit development and testing.
>
> 4.3 Identify the Requirements for the Next Generation DDN Packet Switch
>
> The contractor shall perform the research and necessary analyses, and prepare a report recommending requirements for the next generation DDN packet switch, including functional, reliability, survivability, throughput, maintainability, and security requirements. The report shall predict dates by which the C/30 will begin to fall seriously short of the DDN requirements, estimate when the new packet switch should be made available, and estimate the length of time required for system development and testing. The requirements developed shall be based on: the potential use of alternative trunking facilities in the DDN, such as satellite circuits and local wideband distribution systems; the increasing use of high bandwidth host-to-host applications; and the role of the DDN in an internetworking system.

The intended audience and users of this report is the Defense Data Network (DDN) and those involved in the development of packet switched networks for the Department of Defense (DoD). The reader is assumed to have some knowledge of networking, of packet technology, of the DoD internetwork structure and familiarity with computer/communications hardware and with real-time software for communications. This report is expected to be used by:

the Government to include with a procurement specification for the design of the next generation packet switch

the Government to include with a procurement specification for the development of the next generation packet switch

the contractors who bid on these procurements in order to correctly size the scope of the work

the contractors who perform on these procurements as guidance for their more detailed design and development

As noted above, this report assumes a general understanding of the concepts of packet switching. This understanding can be obtained from a collection of papers such as the DARPA compendium [DAR 81] or the November 1978 issue of the IEEE Proceedings [IEE 78]. The fundamental concepts of packet switching data communications can be summarized as:

messages are broken up into packets (originally typically about 1000 bits) which individually have a high probability of traversing a link without errors

the nodes of the network are computers; they can perform "intelligent" functions and they can provide storage for messages/packets until their receipt is acknowledged

overall reliability is achieved by making use of error detection and requesting retransmission

As the DDN evolves and grows during a time of rapidly changing needs and technology, careful planning and development will produce the Next Generation Packet Switch (NGPS).

1.2 Alternative DDN Environments

In preparing this report attention was paid to the evolving technologies and to commercial developments in data and voice communication. The purpose of this was to determine how, and to what extent, these considerations would affect the Next generation Packet Switch (NGPS) and/or its environment. The conclusion, discussed in Section 5, was that the NGPS environment would be an upward evolution of the present environment and not a radical departure. Alternative DDN environments were considered briefly, however; these are described in Appendix B since we feel that some of the alternatives may have a role to play in a following generation of the evolving DDN.

1.3 Report Organization

This report is organized into eleven sections and two appendices; the Introduction is Section 1. Section 2 summarizes the current, near-term future status, and limits to growth of the DDN. Section 3 discusses the expected growth and changing requirements for the DDN. Section 4 examines technologies which impinge on the future DDN. Section 5 postulates the environment for the NGPS. Section 6 is a discussion of the requirements for the NGPS. Section 7 reviews some commercial packet switch developments and assesses their potential for meeting the NGPS requirements. Section 8 presents some strategies for acquiring the NGPS. A suggested schedule for the NGPS is given in Section 9. Finally a set of recommendations for the NGPS is provided in Section 10. References for the report are in Section 11. Appendix A provides a more detailed discussion of how the requirements for the NGPS were developed, including some discussion of multiprocessor architectures as they fit into the NGPS application. Appendix B, as noted in Section 1.2 above, provides a brief discussion of some alternative network environments to the environment chosen in Section 5.

2.0 WHERE DDN IS AND NEAR TERM FUTURE

2.1 ARPANET-DDN Evolution

We shall examine the current DDN and its past and future evolution as we forecast the major changes that will arise over the next decade. These changes are of two types: those due to specific DoD growth in needs and usage; and those driven by growth in data communication which are largely a result of the changes in computing and communication technology. The current and near term projected DDN is a direct evolution of the ARPANET (described in [DAR 81]). It incorporates part of the original ARPANET as MILNET and additions built out of the ARPANET technology. DDN plans through FY 86 are described in the DDN Program Plan [DDN 82] and in a paper by Heiden and Duffield [HEI 82]. Major changes face DDN as it moves away from the original class of research applications into changing network technology and rapidly growing operational applications for the Department of Defense.

2.2 The ARPANET Technology

A brief discussion of the historical evolution of the ARPANET is useful in understanding how the current DDN was arrived at. The ARPANET was initially a research project intended to develop the advantages of a distributed packet switched network for reliability and efficiency of operation. Baran [BAR 64] proposed that reliability and survivability could be achieved by having multiple source-destination routes through a distributed, highly-connected network of nodes (switches) and trunks. Data communication efficiency was to be obtained by using packets to provide a specialized time-division multiplexing for "bursty" traffic on high speed data links.

2.2.1 The Packet Switches

The first generation packet switches were produced by Bolt Beranek and Newman (BBN) as specially programmed versions of the Honeywell 516 minicomputer; they are described by Heart [HEA 70] and were called Interface Message Processors (IMPs). The network was provided by connecting hosts to IMPs with access lines and IMPs to IMPs over trunks making up the distributed network. Individual terminals were originally connected through their local hosts. Later terminals were connected directly through Terminal Access Controllers (TACs); which used the Honeywell 316 minicomputer and served as a combination of a local host and an IMP [ORN 72]. The IMPs were programmed to carry out the functions of communication with hosts and each other using a set of electrical interfaces and communications protocols which have continued to evolve over time. Both the IMP and the TAC were uniprocessors.

In the late 1970s a new multiprocessor packet switch, the PLURIBUS-IMP, commonly called PLURIBUS [KAT 78] was designed and constructed by BBN in relatively limited numbers; the individual processors were Lockheed SUE minicomputers. As a multiprocessor, the PLURIBUS had higher overall throughput; it was also designed to have fault-tolerance and to be capable of fail-soft operation. PLURIBUS switches have been used on specific, primarily satellite, high data rate links [LIN 79]. Study of the special requirements of satellite links led to a proposed upgrade of the PLURIBUS using a reengineered and faster version of the SUE [NEL 81]. This packet switch was never implemented.

More recently BBN has built an IMP-upward-compatible packet switch called the C/30 [HAV 82]. This uniprocessor is microcoded to implement the basic Honeywell x16 instructions and some common sequences of those instructions. The C/30 also accommodates a larger address space, more communication ports, and utilizes a much higher level of integration in both logic and memories. For high reliability of network service a host may be connected to two different switches; this connection is referred to as "dual homing".

2.2.2 The Communication Links

The original ARPANET internode trunks are nominal 50 kilobit per second (kbps) terrestrial lines leased from commercial communication carriers as are many of the current DDN communication trunks . Other lines are generally at 9.6 kbps. These lines have an error rate of about 1 in 10,000,000, thus most packets of lengths on the order of 1000 bits will transit a concatenation of links and be error free. The time for a packet to completely traverse such a terrestrial link is almost totally made up of the time required by the duration of the packet at the link speed. For example, on a link 1000 miles long the time required for an individual bit to traverse the link is about 1 msec while the transit time of a 1000 bit packet at 50 kbps is 20 msec. Thus transit times dominate propagation delays for terrestrial links of 1000 or 2000 miles.

2.2.3 The PSN Protocols

ARPANET packet switches provided logical host access protocol on three different electrical interfaces. These were the Local Host (LH), Distant Host (DH), and Very Distant HOST (VDH) interfaces. A fourth interface, HDLC Distant Host (HDH) was recently added which allows use of a standard line protocol. ARPANET's original Host Access protocol was known as 1822; it is now known as the ARPANET Host Interface Protocol (AHIP). The DDN now also supports an alternate host protocol, X.25. The links between PSNs originally used the Binary Synchronous Communication protocol (BSC). This has been upgraded to a bit-synchronous protocol, HDLC.

2.2.4 Security

ARPANET originally had no provision for security. At a later time experimental and limited operational communication security was applied to selected data transmission over the ARPANET using the Private Line Interface (PLI). The PLI consisted of two minicomputers and a cryptographic device. The two minicomputers were used to provide separate processing for the "Red" or clear text information and the "Black" or encrypted information. A minimum of destination indicative information is passed between the two minicomputers and the remainder of the information must pass through the cryptographic device [WAL 82].

2.2.5 Summary

ARPANET used a combination of programmable computers with communication ports and high quality leased lines to build a packet switching network. The suite of communication protocols and the functionality of the packet switches evolved with experience and need. As the network grew, both in size and traffic demand, switches were upgraded and communication links were added. The basic switch architecture was enhanced and the software has grown in an upward compatible way. In the next section we examine DDN and its near term growth potential.

2.3 DDN Now and Near Term

The 1986 DDN is described in the (Draft) White Paper on DDN Capacity [PRI 85] from which some of the following material is drawn. At this time, the plans are for two separate networks, MILNET, which is unclassified and DISNET, which will be classified. The 1986 MILNET will have 174 Packet Switching Nodes (PSNs) and 300 trunks.

2.3.1 The Packet Switches

The C/30 PSNs are being upgraded to the C/30E which can logically support 44 connections. A host represents one connection and a trunk represents two connections; software limits the number of trunks to 14. There are some current logical limitations which can be overcome. These deal with addressing of nodes and the number of ports per node. Originally the number of packet switch nodes (PSNs) was restricted to 253. The new IMP End-to-end Protocol (see 2.2.3 below) will raise that to 1024. In order to accommodate individual terminals TACs have been used; each can accommodate 63 terminals. A new device, the Network Access Controller (NAC) [ELD 83], will go into service this year; one of the operational modes of the NAC is as a "mini-TAC" which can act as a concentrator for 16 terminals.

2.3.2 Trunks and Access Lines

There have been no major changes in the trunks and access lines as ARPANET has evolved into DDN. The majority of the trunks are high speed (50 or 56 kbps); the remainder are at 9.6 kbps. Access lines provide speeds from 1.2 kbps to 56 kbps.

2.3.3 The PSN Protocols

The communication protocols within the DDN have had a continuing evolution since the DDN was initiated. The latest version (Release 7.0) will contain a major change to the PSN software, the new End-to-End (EE) protocol for the current packet switches [MAL 84a, MAL 84b, MAL 86]. This protocol establishes duplex connections between endpoint PSNs (the PSNs connected to the hosts which are communicating). This change will improve PSN and network throughput by reducing the number of purely "administrative" messages that support reliable operation. Other enhancements in Release 7.0 include support for satellite links by providing adjustable windows, support for precedence and preemption, and interoperable AHIP-DDN X.25 service.

In release 8.0 the EE protocol will provide even more service; they will include fragmentation and reassembly of datagrams. This will be done with less overhead, but without total reliability which, if needed, would have to be provided by a higher level protocol.

2.3.4 Security

The current security architecture for the DDN is documented for the subscriber in the DDN Subscriber Security Guide

[SHI 83] and its current new draft [SHI 86]. It is also described in the new draft of DDN's future security architecture [DDN 85]. Briefly, the DDN is divided into two segments, one classified and the other unclassified. Both segments use the C/3x PSNs and trunk security will be provided by link encryption with KG-84As for both segments; the link keying material on DISNET will be protected as SECRET. Switches at DISNET sites will be physically protected at the SECRET level. DISNET access lines will also be protected by KG-84As and MILNET access lines will be protected by Low-Cost-Encryption Authentication Devices (LEADs). Subscribers will be able to have additional end-to-end encryption (E3) protection by using BLACKER Front Ends [BFEs] which can provide flexible and dynamic E3 protection.

2.4 Prospective Changes to the DDN

There are a number of prospective changes to the DDN which bear on the capacity and throughput of the network and the PSNs. One such change deals with implementing a congestion control mechanism [EIS 85]. Such a mechanism will improve throughput when the network is heavily loaded. In the past, especially on ARPANET which was always lightly loaded, there were no major congestion problems. With congestion control procedures in place, one can feel more certain about throughput calculations.

Another prospective change deals with Type-of-Service (TOS) routing [GAR 85]. TOS routing takes advantage of the fact that it is appropriate to route different sorts of transmissions over different paths. For example, file transfers are suitable for transmission over satellite trunks with high bandwidth and (relatively) long delay. Single packet transmissions, on the other hand, benefit from low delay paths. TOS routing, of course, interacts with other protocol features and changes.

Still another prospective change which can affect DDN capacity is the expected availability of the BBN C/300 packet switch at the end of 1986. This PSN will have twice the memory of the C/30E, the ability to support 64 (vice 44) attached devices, and increased throughput. It can operate in a network with C/30s and C/30Es. The White Paper on DDN Capacity [PRI 85] makes a number of its calculations under the assumption that the C/300 will be placed into service for DDN. Since the situation regarding the C/300 is unclear, we shall develop the date by which the NGPS is required under both conditions: with and without the C/300.

2.5 Summary and Conclusion

The DDN and its PSNs represent an upward evolution of the ARPANET technology. Architecturally, the PSNs (except for the PLURIBUS and successor satellite nodes) are still uniprocessors. Reliability of PSNs is achieved by the inherent reliability of the computers themselves, by the redundancy provided in the topology of the DDN, and by dual homing of critical hosts. The network protocols have undergone and will continue to undergo change. DDN is growing at about 20% per year. Some growth projections and a discussion of the limits to continuing growth of the present DDN are presented in Section 3.

3.0 WHERE DDN NEEDS TO GO

In this Section we discuss the forthcoming growth of DDN in terms of the needs of the user community and the limits to a DDN using currently planned trunking and PSNs. The material in this Section draws heavily upon Bolt Beranek and Newman's Future Network Technology Study (FNTS) for DCA (the Interim Report is [HER 84], the Final Report is [HER 85a], and portions of the report are summarized in [HER 85b]) and the Draft DCA White Paper On DDN Capacity (WP) [PRI 85].

3.1 Future DDN Requirements

Both the FNTS and the WP note problems in forecasting the growth of DDN. Hard information to support such a forecast is not available. The FNTS produced conclusions that indicated that the maximum number of hosts per backbone network could rise to 25,000 in 1988; other numbers in their reports are very general. Two approaches were used to obtain estimates: top-down - based on planned computer installations, and bottom-up - based on the User Requirements Data Base (URDB). The first method potentially overestimates connections to DDN; the second method potentially underestimates connections to DDN. In either case there are unknown factors which affect the actual number of computing systems which are candidates for connection to DDN.

The WP used an approach which examined the limiting factors to the growth of DDN under its current plans (those summarized in Section 2.4). One factor is the maximum throughput which can be provided by the trunking. This is based on assumptions that the topology and average hop length will remain about the same. Another factor is the throughput of the switches. Both of these estimates result in a maximum network capacity of 6 to 7 Mbps. Trunking is the limiting factor and limits the number of hosts to about 2100 under the assumption that the hosts are connected by 9.6 kbps access lines which are 30% utilized.

There are two user community factors which will have a major impact that cannot be quantified at this time. The first is the extent to which Local Area Networks (LANs) will keep local traffic off of the backbone network and concentrate distant traffic through a single gateway connection. The second is the changes which will occur in the types of traffic; this reflects the difference as traffic makeup goes from heavily interactive to general operational use. Intelligent terminals and PCs will greatly reduce the number of single small-packet messages.

There are also current logical limitations to addressing which can, in principle, limit growth. These limitations can be overcome. The limitations arise in AHIP and in the DDN Internet Protocol (IP) which both limit the number of nodes to 256 (for technical reasons this is actually 253), and the number of hosts per node to 64 (AHIP) or 256 (IP). These limitations could be greatly eased if the DDN X.25 addressing is adopted allowing 999 PSNs with 99 hosts per PSN. Adoption of the ISO Internet Protocol would remove all IP addressing limits.

Over and above the question of addressing, PSN capacity limits are discussed in both the FNTS and the WP. These limits deal with memory limits for tables, overhead for routing updates and routing computation overhead, and packet handling. It is necessary to discriminate between three types of traffic:

> that originating or terminating at a PSN and going to or from a trunk
>
> that going to or from another host connected to the same PSN
>
> that passing from one trunk to another through a PSN (tandem traffic

Both reports conclude that a uniform network of 253 C/300 PSNs could support approximately 7000 hosts which are connected to the PSNs, on the average, by 30% utilized, 9.6 kbps access lines.

3.2 New Technologies in DoD

The advances and spread of computing technology has a potential impact on DDN. There are now many users of personal computers (PCs) who have the potential desire or need to access resources across a data communication network. These PCs are no longer the "dumb" terminals which are accommodated by TACs. They could perform the functions of hosts; however, DDN can not be expected to provide host status to each individual PC. PC's can be expected to exist on Local Area Networks (LANs) within appropriately grouped physical complexes. Each LAN would have a gateway to connect it to DDN; the gateway appears as a host to the NGPS. Any transaction between a PC and another PC or a host on the DDN would be accomplished by passing traffic through the gateway and across the network to the appropriate host or gateway on the LAN for the other PC.

3.3 New Application Level Requirements

The spread of computing resources will encourage the implementation of distributed applications for users who are administratively related, although not physically collocated. These users will perform file transfers of relatively large volumes of data representing one of the contributing factors to the growth of user requirements for DDN.

At the present time, a major source of short messages is traffic from dumb terminals to hosts; in much of this the actual data per single packet message is a single byte (character). This traffic occurs when a terminal is carrying out an interactive application on a remote host. It represents a major inefficiency in network usage. As more intelligent terminals (i.e. PCs) come into use an effort needs to be made to deal with these transactions at the level of screen lines. Further, we can expect that some of the applications may be accomplished on/at the intelligent terminal.

As users gain access to the net from their "own" computers, new applications which generate relatively large volumes of traffic such as mail will spread rapidly; however, the exact effects of application changes are unknown at this time. Computer interaction, like computer utilization, seems to obey a Parkinson's law of expanding to utilize all of the available resources. We have chosen to use the prediction that average host traffic will multiply by 4 for purposes of sizing the NGPS network environment as used in Section 3.6

3.4 Security Considerations

Current DDN policy [DDN 85, SHI 86, LAN 85] calls for separate subnetworks for each security level with the subnetworks themselves, including the PSNs, being operated at system high for that security level. This situation will be ameliorated when BLACKER provides E3 for multi-level security between end users.

There is also a requirement that network control messages be authenticated. This can best be accomplished by using a cryptographically based checksum. The messages should also be protected by link encryption while on the trunks.

The general requirements for network security are still in the process of evolution. The National Computer Security Center has a draft "Department of Defense Trusted Network Evaluation Criteria" (DoDTNEC) [CSC 85] at the present time; the date of completion of DoDTNEC is uncertain. This document is intended to be analogous to the "Department of Defense Trusted Computer Evaluation Criteria" (DoDTCEC) [CSC 83], the so-called "orange book". DoDTCEC provides standards for assessing uniprocessors and their operating systems. Current DDN plans call for the computers within PSNs to meet the C2 criteria, or better, of DoDTCEC. This requirement is moderately stringent. Further, application software for PSNs should be written with computer security validation as a goal.

3.5 Other Features

3.5.1 Precedence

The DDN requires that there be a precedence and preemption capability within DDN to allow for efficient response to critical users while the network is under various levels of stress conditions (p 129 of the DDN Program Plan [DDN 82]). Precedence information must be acted upon by the PSN software to provide the required service.

3.5.2 Fairness

Fairness implies that all users of equal status in terms of precedence and preemption deserve an equal chance at network resources. Implementation of congestion control can, if done in a simplistic way, deny service to users in an unfair manner. There are proposals within the data communication community and within the plan for congestion control experiments to provide fair service to users, all other factors being equal. This is a desirable feature for DDN.

3.5.3 Type of Service

Type of service routing can improve the performance of a network with heterogeneous trunking; the situation we postulate for the NGPS environment in Section 5. High bandwidth trunks with moderate delay are very appropriate for file transfers and for mail. Low delay trunks can be used preferentially for short messages and interactive applications.

3.6 The NGPS Network

The limit of 7000 to 8000 hosts discussed in Section 3.1 is unrealistic insofar as it assumes an even distribution of hosts to PSNs and even average usage by the hosts at a given PSN. Thus, it is an upper bound to the number of hosts that can be supported by the straightforward evolution/expansion of the current technology.

If we use the projection of hosts in the FNTS, we can envision for 1990 a network which has from 2000 to 30,000 hosts. At the low end, this load could be accommodated by a subset of their projection of the 7000 host network of 253 C/300s. We can presume this to be an unevenly populated network with PSNs being either a mix of C/300s and C/30Es or all C/30s. At the high end, the requirements clearly exceed the projected capacity. In any event, we can expect that there will be a time, as discussed in the next section, in which the C/3x based network with 64 kbps trunks will become inadequate.

We can postulate an idealized network which supports 20,000 hosts as follows:

> Assume a 20,000 host network with 400 nodes and, an average of 50 hosts per node, and an average hop length of 4. In keeping with the discussion of Section 3.3 above, we shall assume that the average host generates 4 times the traffic of current hosts, 30% loading on a 9.6 kbps line with 750 bit packets, or almost 4 packets per second. Based on this number we find that the average traffic entering and leaving a node is 0.6 Mbps (16 packets times 50 hosts or 800 750-bit packets per second). If we now assume that 20% of this traffic is intranode (the current value) then 80% of the traffic will be going to the trunks. With our average hop length of 4, another 320% of the 0.6 Mbps must be accommodated as tandem traffic. Thus, the net trunk capacity of the NGPS must be 2.0 Mbps. This increase

in capacity of a PSN calls for the total throughput for the three types of traffic through the switch to be as follows:

Type of Traffic	Throughput (packets/sec)
Host - Trunk	640
Trunk - Trunk	2,560
Host - Host	160

The result of this example calls for the NGPS to have an aggregate throughput of 3360 packets per second; this is about 6 times the capacity of the C/300 (p 44 of the WP) The significance of this requirement is discussed in Appendix A. The NGPS requires about 2 Mbps of trunk capacity. Some of these trunks may have significantly higher rates than 64 kbps. For an average transit length of 4 hops an evenly distributed network should have 5 trunk connections per node. In order to allow for uneven distributions of connectivity and traffic, we propose that the maximal switch have provision for more than three times this number, i.e. 20 trunk connections.

Section 6 presents requirements for the NGPS based upon the above material (Sections 3.1 through 3.6).

3.7 The Date at Which the NGPS is Required

We will make two estimates: one is the date by which a network employing C/30Es and C/300s as PSNs will no longer suffice; the second is the date by which a network of C/30Es will no longer suffice. The first estimate, based on discussion earlier in Section 3, will be the date at which the number of hosts being served exceeds about 7000. If we take the middle ground of the five scenarios of the FNTS Interim Report (p 46 [HER 84]), this date will be about the middle of 1990.

For making the second estimate we will assume a network of C/30Es with some doubling up at high throughput nodes. This suggests that a 4000 host network will begin to exceed the capacity of the C/30E based network. Again, taking the middle ground from the FNTS Interim report we find a date for 4000 hosts to be about the middle of 1987.

Using the first estimate we propose that a date of January 1991 for fielding the first NGPSs would be satisfactory. Such a date would allow sufficient time for a full development cycle for an all new NGPS if that is selected as the means of acquisition (see Section 8). Using the second estimate we propose that a date of January 1988 for fielding the first NGPS would be satisfactory. This would not give time for a full development cycle; it could accommodate the strategy of modifying a commercial switch to be the NGPS (see Section 8). Schedules for both cases are presented in Section 9.

4.0 WHERE TECHNOLOGY IS GOING

Section 3 showed that the DDN will need to handle both a greater number of hosts and a much larger volume of traffic. This growth is one of the major drivers for the Next Generation Packet Switch (NGPS). In addition, technological developments are occurring which will have an impact on the NGPS. Discussions of these developments will be found in [HER 84, HER 85a, ANA 84, FRA 76, and FRA 79]. Increased use of computing resources will be the result of widespread access to computers. Other technological developments will support both expansion and growth; they will also lead to changes in the network and the switches. Changes in trunk and access line transmission technologies, in computing technology, and in commercial communication systems architecture, together with special DDN and DoD requirements will all contribute to the requirements for the NGPS. These are discussed below.

4.1 Transmission Technology

Packet switched networks will have a wider range of technologies for the communication links with both higher bandwidths and a mix of delays available. Telecommunications transmission technology is in a period of rapid change. For high bandwidth transmission, microwave and coaxial cable transmission systems have been overtaken to a considerable extent by satellite and more recently fiber optics development. For lower data rates, improvements in modems make rates up to 19.2 kbps readily achievable over classical voice grade lines. At the same time, voice transmission techniques are becoming all digital as discussed in Section 4.3 below. All of these developments will have an impact on the future DDN and the NGPS.

4.1.1 Satellite Transmission

Satellite transmission affects a packet switched network in two ways. The delay over a single hop, assuming a geosynchronous satellite is significant, it is on the order of half a second. This delay has an immediate effect on the window size, the number of messages which are allowed to be in flight without an acknowledgement being received. The old network protocols provided for a maximum window size of 8; the new EE protocol changes that to 128. In addition much higher bandwidth/increased data rates are available. This higher bandwidth can be exploited in either of two ways; it can be utilized as a number of multiplexed low-speed channels or as one, or a few, high-speed channels. Satellite channels normally incorporate some "built-in" error correction and thus are more reliable than the original channels used on ARPANET. As a result, packet sizes can be larger than those used for ARPANET and a packet size of 8000 bits would improve use of satellite bandwidth [NEL 81]. Larger packets are also more effective when transactions are insensitive to satellite delays.

4.1.2 High Speed Terrestrial Transmission

High speed terrestrial transmission is growing and becoming less expensive as a result of the growth of fiber optics systems and the related growth of T1 carrier services. Bandwidth and error rate will be similar to that discussed above for satellites. Commercially in North America fiber optics transmission systems will be divided into T1 carriers their multiples, and submultiples. (For example, 384 kbps will be a common subset of both T1 and the European 2.048 Mbps lines). Again, because of the lower error rate longer packets are practical if the data transmission needs can support them. The lower delay time represents an improvement over satellite channels for interactive applications.

Another role of fiber optics is its forthcoming use in undersea cable systems. This usage eliminates most of the delay occurring in satellite systems referred above; a trans-Atlantic delay would be less than 10 msec.

4.2 Computing Technology

Computing technology is both the driving force in the expansion of computer use and the support for packet switches. This technology continues to progress rapidly; it improves by a factor of 2 in cost per unit of computation every two years. Important areas for the NGPS, architecture and engineering, devices, and computer security, are discussed below.

4.2.1 Architecture and Engineering

Development of computing systems with very high computational throughput at minimal cost is a current major trend. As speed boundaries are approached (e.g. 1 nanosecond for a electrical impulse to traverse about 7" of wire), it is no longer economical to force higher throughput by sheer serial speed up. As a result, especially where problems can be partitioned into tasks which can run concurrently with little or no inter-task communication required, parallel or concurrent processing becomes very attractive.

Many new computer architectures are providing multiple processors for parallel processing. With the advent of LSI and VLSI, assemblies of identical processors become economically attractive as well. These multiprocessors have the potential to provide redundant, fail-soft configurations. Although not all problems can be partitioned into parallel tasks, packet switch requirements are well suited to such partitioning. On the other hand, operating systems for multiprocessors are not yet at a mature stage.

Hardware and Software Engineering, particularly with computer aids, are making steady advances. Thus, it is possible to build systems of chip level devices in a relatively straightforward manner. Similarly, software applications can be built in an orderly, modular, and self-documenting way which is easy to understand, make operational, and modify.

4.2.2 Logic and Memory

LSI and VLSI technology continue to make advances by factors of about two every 2 years. These advances can be in any of three dimensions: higher complexity/size; higher speed; or lower cost. As a result, devices such as microprocessors, memories, and specialized microcontrollers which can be produced on a single chip attain great advantage. Memory size, especially, is no longer a cost problem and error-correcting memory systems produce more reliable memories than ever before. For packet switches, specialized microcontrollers include communication controllers and hardware implementations of cryptographically based or other checksum algorithms.

4.2.3 Computer Security

Computer security is another area which is receiving significant attention at this time. The National Computer Security Center has developed criteria for trusted computer systems [CSC 83] and is developing criteria for trusted networks [CSC 85]. These activities are beginning to influence the commercial computer environment although much remains to be done.

4.3 Commercial Communication Technology

Analog telecommunications are becoming obsolete. The "telephone industry" is rapidly moving toward providing a digital technology based service. As a result, new digital connections to subscribers are becoming available and there is a strong impetus to provide both voice and packet switched data service using as much common plant as possible. Some of this is exemplified by the developments in the Integrated Services Digital Network (ISDN) and in the development of switches that provide for both circuit and packet switching. These two topics are discussed below.

4.3.1 Integrated Service Digital Network (ISDN)

The ISDN is an architecture and a philosophy that is currently being developed at the standards level by the CCITT. The official description of ISDN is in the I-Series of CCITT recommendations [CCI 85]; MITRE has prepared a set of papers for DCA discussing ISDN and its potential impact on DCA [SAK 83a, SAK 83b, SWE 83a, SWE 83b]. IEEE Communications Magazine has also devoted a special issue to ISDN [IEE 86]. ISDN calls for the use of all digital techniques for switching, any requisite intermediate storage, and transmission from one subscriber's device to that of another subscriber. ISDN will use a contention-avoidance local bus for subscriber premises and digital (i.e. computer or computer technology based) interfaces and switches throughout the overall ISDN. This loop will have the "2B+D" capacity, 2 B channels at a 64 kbps rate and one D channel at a 16 kbps rate. The B channels were normally intended to be used for digitized voice and the D channel was to be used for data and signalling for all channels. The B channels can, in fact, also be used for data.

A major difference between ISDN and much of the current telecommunications technology is the use of "common channel signaling". That is, for circuit switched applications, the control information (call set-up, etc.) is transmitted on a separate all digital channel. Extensive use is to be made of high bandwidth channels (and their major subdivisions) such as North American T1 at 1.544 Mbps or the similar European service of 2.048 Mbps. ISDN for T1 will normally be "23-B+D", 23 B channels and one D channel for signalling. The B channels can be made availble to a switch in a flexible way; the D channel will be used for signalling and administration. An extension to the ISDN Standards is under way to specify how packets are sent over B channels, using the LAP B protocol, and how the D channel will be used for high level control.

The advent of ISDN and its supporting transmission technology will increase the availability and reduce the cost of high bandwidth communication links.

4.3.2 Hybrid Switches

Although there has been extensive experimentation with packet voice, digital voice which is further segmented and sent over a packet switching system, there are a number of drawbacks to its widespread use. Two-way voice communication is not tolerant of highly variable delay; this militates against error correction by requesting retransmission.

In fact, most digitized voice (say, at 9.6 kbps or above) is tolerant of occasional errors and does not require error correction. Finally, although speech is also "bursty", it is desirable to handle full "talkspurts" which are a second or more in length. Such talkspurts represent from tens to hundreds of packets, depending upon the voice coding techniques. There are proposals to do a form of fast circuit switching for talkspurts; this is called "burst switching" [HAU 83].

As a result of these voice/data distinctions, within the commercial telecommunications area there has been increasing attention paid to the concept of "hybrid switches". These are computer based switches which provide both circuit switching for digital voice and packet switching for data. For example, Budrikis and Netravali have proposed a design for a hybrid switch [BUD 84] in which circuit switching is provided by repeatedly allocating memory/ transmission slots for a given call and packet switching is achieved by competition for the unused slots. The AT&T 5ESS described in Section 7.1 below represents another hybrid switch.

4.4 Summary and Conclusions

The current, rapidly changing, technology is both a driver for DDN growth while at the same time it can provide support for a much larger network with higher traffic volumes for users. The requirements for the NGPS developed in Section 6 are completely in accord with the current and very near term technology.

5.0 THE NGPS ENVIRONMENT

5.1 Introduction

In Section 1.2 we mentioned that this study has concluded that the most likely network environment for the NGPS is an extension of the present network topology but that greater throughput is needed for both PSNs and for trunks. More flexibility will also need to be provided as discussed in Section 3. These ideas are elaborated in Section 5.2 below. Some of the possible, but "rejected" environments are discussed in Appendix B in more detail; we have concluded that they are unlikely environments in the anticipated time frame for the NGPS.

5.2 Global Strategy

An extension of the current topology and connectivity appears to be desirable for two primary reasons:

the network can continue to evolve without drastic changes,

no compelling reasons appeared to make the other alternatives stand out as attractive.

This conclusion is also one that is expressed in the Draft White Paper on DDN Capacity [PRI 85] which calls for modular evolutionary growth without major disruptions.

We note here that one of the more interesting possibilities presented in Appendix B is that of a hybrid switch which shares trunking facilities with a voice circuit switch. Such an arrangement allows for maximum flexibility in using DoD owned or leased transmission resources; it is also technologically in line with current voice/data integration as exemplified by ISDN and discussed in Section 4.3. The use of hybrid switches requires investigation by, and coordination with, the voice planning process of DCA.

As a result, we believe that the NGPS environment will be that characterized in Figure 5-1. This is an illustrative diagram which indicates the kind of connectivity expected between nodes. The intent is to show something similar to the present DDN. Three major differences from the current DDN are expected:

There will be differing types of trunks between a given pair of nodes (for example, trunks of differing speeds and delays),

Trunk speeds significantly greater than 64 kbps will need to be accommodated,

A greater number of hosts will need to be accommodated at some high traffic density nodes.

We believe that these differences are justified by the extrapolations developed in Section 3. Further, the NGPS will need to be able to have most of the new functionality which is being or will be placed into service for DDN. These are such features as the new end-to-end protocol, congestion control, and type of service routing. It will also have to have the requisite security features and such DoD features as precedence and priority. The requirements developed in Section 6 and summarized in the recommendations of Section 10 are based on this environment.

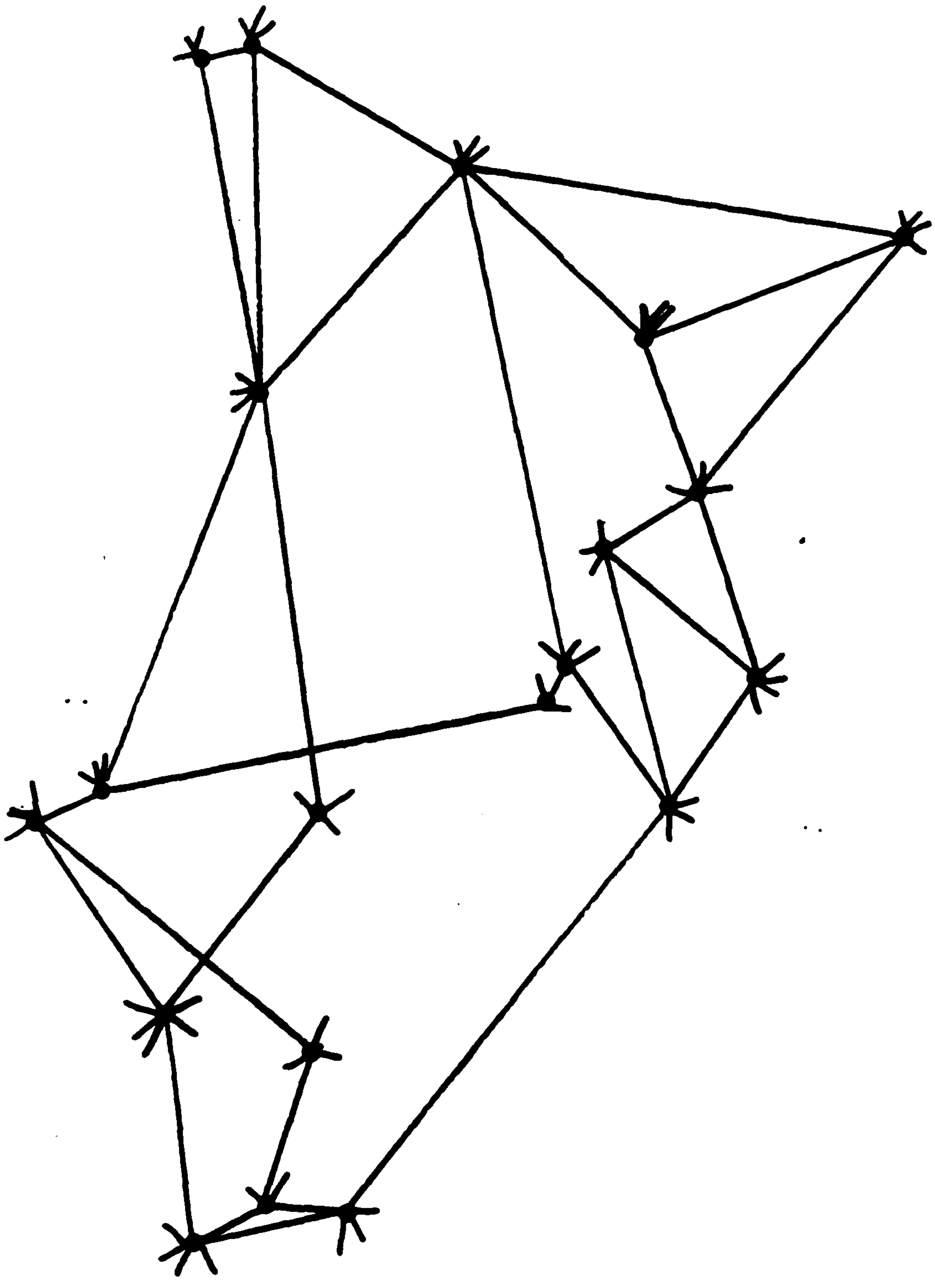


Figure 5-1. Present Day DDN Expanded.

5.3 Switch Families

In Section 8 there is a discussion of two ways of acquiring the NGPS: development of a new, DDN specific, switch; or use of a switch based on commercial developments. Many of the commercial switches described in Section 7 are available in "families" which provide the same service but with different throughputs, number of ports, etc. The NGPS can also be usefully provided as a family with a range of capacities. Since we will have a range of load requirements for PSNs, a family of NGPSs will provide a more cost effective fit to those conditions.

6.0 NGPS REQUIREMENTS DISCUSSION

This Report, and particularly this Section, develop a proposed set of requirements for the NGPS. These NGPS requirements will have to be interpreted in the light of an overall DDN program plan which will be aimed at meeting the service requirements of the 1990-1995 time frame. These service requirements have been estimated in Section 3 and expect the DDN program plan to be in general accordance with the discussion of the NGPS environment in Section 5. Some of the requirements presented here may have to be waived or modified (for example, Section 6.2, Implementation, below) if a commercial offering is adopted or leased.

6.1 Switch Specifications

6.1.1 Quantity

Following the argument of Section 3 we estimate that MILNET, the largest component of the DDN, will require about 400 NGPSs. This suggests that a maximum of 1000 NGPSs will suffice for DoD/DCA generally. Given the state of hardware and software engineering this quantity is a satisfactory size base for developing a unique architecture. As discussed in Section 5.2, the architecture should allow for a family of switches; that is the NGPS needs to be modular in its hardware design.

6.1.2 Capacity

The throughput performance of the maximal switch can be specified by the requirements developed in Section 3.6.

These are:

Type of Traffic	Throughput (packets/sec)
Host - Trunk	640
Trunk - Trunk	2,560
Host - Host	160

When specifying a family of switches, other members of the family could be specified at some fraction (e.g. 3/4, 1/2, and 1/4) of this maximum capacity.

6.1.3 Interfaces

Trunk interfaces should be provided at multiple data rates; these should be standard line rates (e.g. 9.6 and 19.2 kbps, and perhaps 56 kbps) below 64 kbps. 64 kbps trunks should be accomodated. A question arises regarding trunk rates above 64 kbps. There are standard, inexpensive, and reliable ways of using a T1 as many 64 kbps channels. Using a T1 or a major subdivision of a T1 as a single high speed trunk will require new transmitters and receivers at the electrical level. At this time we propose that the external interfaces to trunks be modular and accomodate 64 kbps trunks; provisions should be made to substitute faster interfaces up to at least 384 kbps. That is, internal to the PSN, 384 kbps rates to and from a trunk interface can be supported. The interface protocols should be DDN X.25.

The access line interfaces should accommodate line speeds up to 64 kbps; this will allow for use of the ISDN B channels as access lines. Access line speeds below 64 kbps should be standard multiples of 1.2 kbps. AHIP will no longer be supported (except for transitioning - see Section 6.9 below). DDN X.25 should be adopted for host access lines.

6.2 Implementation

The discussion of implementation presented here is predicated on the assumption that a new NGPS will be developed under contract to DCA . Therefore, acquisition will proceed through the usual phases and appropriate regulations concerning development of computer systems will have to be followed. As noted at the beginning of Section 6, some of these requirements will not apply if commercial packet switches are acquired or employed.

6.2.1 Hardware

The NGPS should be both TEMPEST and HEMP protected. Except for those specifics it should be constructed to the best commercial standards. The NGPS should be capable of graceful degradation under conditions of hardware failure (see also Reliability/Maintainability in Section 6.3 below). This implies that the NGPS is at minimum a dual processor system (as discussed in Section 4). A preferred physical implementation would be one in which link encryptors, and any level 1 and 2 interfaces, such as modems, are housed within the same cabinetry as the NGPS, facilitating the TEMPEST and HEMP protection.

6.2.2 Software

Software for the NGPS should be designed and implemented using a high level language. This should be done for all of the reasons associated with good software engineering practice including simplicity of change and ease of maintenance. The choice of the higher level language should be made at the time at which the formal (A-level) NGPS specification is developed. One obvious candidate is ADA.

At the present time there are some objections to ADA based on the beliefs that ADA is not "mature" enough, does not generate efficient object code, and is not a good basis for secure systems. The first two of those beliefs may be alleviated by time; the last is countered by arguments presented by Anderson [AND 85] in which he makes a case that ADA is a sound basis for secure programs.

6.3 Reliability/Maintainability

The NGPS should have provision for "fail-soft" operation. That is, the failure of a major component such as a processor module, a memory module, or a power supply should not render the NGPS inoperative, although it would reduce the throughput capacity. Although some major component failures might deny service to a particular host or the use of a particular trunk, they cannot be permitted to bring the whole switch down. Software within the NGPS should be provided to monitor the switch status on a continuing basis; to report maintenance requirements which can be satisfied by replacing major modules. It is within the state-of-the-art for a combination of hardware and software to disconnect faulty modules and connect spare modules. Such remote maintenance should be presented as an option when asking for bids.

Overall, it should be possible to specify that the NGPS shall provide 99.995% availability for at least reduced operation, a mean time between (partial) failures of 1000 hours, and a mean time to repair of 0.5 hours

6.3.1 Survivability

DDN survivability for catastrophic events is ensured by the overall redundancy of the network; there are no special survivability requirements for the NGPS.

6.4 Security

The switch will be expected to meet at least the Class C2 criteria of the DoD Trusted Computer Security Evaluation Criteria [CSC 83]. The Class C2 criteria specify that the computer system be required and trusted to maintain discretionary access control. As noted in Section 3.4, it is not clear at this time how the DoD Trusted Network Evaluation Criteria (the draft is [CSC 85]) will apply to the switch. Further, it is not clear exactly how either set of criteria apply to a multiprocessor switch. In any even, the switch should be designed and documented using a formal development methodology. The NGPS (and the new Network Control Center when it is specified and developed) should provide for authentication of control traffic on the network. There will be a large number of link encryptors at each switch. These link encryptors should be integrated with the switch as closely as possible to minimize cabling and space for the switch.

The DDN is continuing to improve security over the whole network. The plan is described in the draft document Defense Data Network Evolution of Security Services [DDN 85] These security plans, some of which are already in being, will place requirements on the network and the NGPS.

6.5 Unattended Operation

The NGPS should be capable of unattended operation and of performing certain configuration changes by either local or Network Control Center direction. These changes would be those discussed under maintainability/reliability above, those changes necessary to accommodate line failures, and those changes dealing with revisions in the composition (names/addresses) of the net. Changes initiated by the NGPS will be reported to the Network Control Center.

6.6 Network Control and Accounting

At the time that the new DDN program plan is developed it will be necessary to develop a set of requirements for a Network Control Center. There will, of course, be multiple Centers for reliability/survivability considerations. If detailed accounting is to be used (as currently proposed) for charging purposes there will also need to be a method of collecting information to provide that accounting. Accounting is commonly performed at the Network Control Center in commercial applications. It might be desirable that these functions not be combined in DDN in order to leave the Network Control Center with maximum capacity for handling stress situations. More detailed discussion of these requirements is outside the scope of this Report.

6.7 Transitioning

The NGPS will have to be brought into service in a phased manner. Nodes containing the NGPS will have to serve trunks communicating with nodes containing C/3x's and serve links communicating with local hosts. As a result, when a node becomes an NGPS node there will be a number of possible strategies. Three possible strategies and a brief discussion of their pros and cons follow:

The NGPS can have a mode in which it emulates the C/3x.

The NGPS can be provided with software, firmware, or a mixture to emulate the C/3x and permit its operation as a C/3x within the existing net. At a time when a grouping of closely connected NGPSs exist, that group will be switched over to the new programs. This strategy implies that the transition will take place in large steps.

The NGPS can cooperate over a trunk with a collocated C/3x, sharing functionality during a gradual switchover of trunks and access lines between the two switches.

A communications channel will be provided between the NGPS and the C/3x ; "transition" software will be written for both the NGPS and the C/3x to provide for movement of packets, control information, routing information, etc. between the two switches. As access lines and trunks are moved from the C3/x to the NGPS or as new access lines and new trunks are installed on the NGPs, more and more of the load will shift from the C/3x to the NGPS. This is a more gradual strategy; it requires keeping dual switches in operation for a transition period. A variation of this strategy would be the provision of a duplicate network having both old and new switches at the nodes. the duplicate networks would be interconnected at many of the nodes by gateways.

The NGPS can provide non-optimized C/3x functionality as well as full NGPS functionality.

The NGPS will be programmed to perform the the minimal set of functions to permit working with users and the C/3x's. This will be done using the protocols which are current at the time of installation. The NGPS will also be programmed to perform the new functions and protocols specified for the NGPS. Performance of the "old" protocols need not be at maximum efficiency since this is only needed for the transition period and the NGPS will have greater capability than the C/3x being replaced. As is the case with current PSN software releases, the Network Control Center can instruct the node as to which release applies for various operations. The Network Control Center thus is in control of the evolving use of the NGPSs. A dual mode operation, similar to that used by AT&T in the 1PSS (see Section 7.1.2) would facilitate transitioning.

All of these strategies require additional programming beyond that required for a network using only the NGPS. The last strategy is the most straightforward and is recommended as a requirement. That strategy would be facilitated if future releases of C/3x software are specified and designed in higher level languages.

7.0 EXAMPLES OF COMMERCIAL PACKET SWITCH CAPABILITIES

In this section we examine some of the commercial packet switches for two reasons: they represent possible ideas which should be considered; they represent possible bases for packet switches to be procured. The discussion of current offerings is not an exhaustive survey; it represents information that was conveniently available. With regard to future offerings only Northern Telecomm and AT&T Bell Laboratories had very much to say.

7.1 Current Commercial Switches

7.1.1 Amnet

Amnet makes the Nucleus 6000, a multiprocessor packet switch using up to 10 Intel 80286's connected together by a proprietary bus technology. A maximum of 256 front end processors handle up to 1024 ports. Trunk speeds of up to 64 kbps are accommodated; X.25 protocol is used. Within the switch, duplicate paths can be provided for hardware redundancy and assurance of higher availability. A network management function is provided on the same hardware base. Amnet states that throughputs of up to 1000 packets (of unstated size) per second can be accommodated on their largest configuration.

7.1.2 AT&T

AT&T currently has their No. 1 Packet Switching System (1PSS) in operation. Like their No. 5 Electronic Switching System (5ESS) which is described in [CAR 85], the 1PSS is based on the ATT 3B20D [BEC 83] computer, a dual processor used for the primary control and supervisory functions. The dual processors of the 3B20D are not used as amultiprocessor system; they are the control processor and a hot standby. When installing a new software release one processor implements the old release and one the new release to facilitate debugging. In the overall switch, Individual microprocessors handle 8 ports (access lines or trunks) each; up to 60 of these microprocessors can be used providing for a total of 480 ports. In Release 3 of the 1PSS all message data is passed through the main memory of the 3B20Ds using DMA techniques. The capacity of the switch is quoted at 1200 packets per second. This is realistically 500 or 600 true data packets of 128 Bytes each; the other packets are used for administration and acknowledgements. Release 4 of the 1PSS (apparently early in 1987) will provide interprocessor communication via a 32 Mbit duplex ring. Up to 960 ports can be accomodated and the throughput is to be over 4000 packets per second.

The 1PSS, using additional software, can operate as a Network Control Center System (NCCS). The switches are intended for unattended operation and extensive centralized diagnostics and display are provided; these can be run at any 1PSS. Up to now the AT&T packet network service has had relatively few switches with high throughput and networks with few hops. The 1PSS will have software for tandem operation by the third quarter of this year. The 1PSS has a four level congestion control strategy which is essentially a source quench; it operates on virtual circuits, not datagrams.

The 5ESS, although primarily a voice circuit switch, can provide packet switching service. Thus, it is an operational hybrid switch. Within the 5ESS two microprocessors, Intel 8086s and Motorola 68000s are used for intermediate control and line handling respectively.

7.1.3 M/A-COM CP9000 Series II

The M/A-COM CP9000 Series II [MAC 85] is also a multiprocessor system. It is based on Intel 80286s; Intel 80186's are used as port processors. Clusters of up to 8 processors are interconnected on a high speed LAN within a node. Hardware redundancy is provided for reliability and the X.25 protocols are adhered to. Data encryption with integral DES devices is available; the software support includes a form of TOS. M/A-COM believe that their 80286 and specialized operating system can receive B-1 or B-2 Computer Security Center certification. Network control is provided on a VAX 11/750.

7.1.4 Northern Telecomm

Northern Telecomm has just introduced their DPN series packet switch; this is an upward compatible version of their SL-10 packet switch [NOR 85]. The switch architecture is based on mutiprocessors connected via high speed buses. A full redundancy configuration is available in which every user has access lines connected to two independent segments of the switch. These segments are in turn connected to two nodes of a trunk network. The high end DPN-50 can handle up to 2880 access lines at 9.6 kbps, or 96 access lines at 64 kbps, or a mixture of those categories. Up to 30 trunks at 192 kbps can be accommodated. The DPN-50 can handle up to 3000 packets per second (of unstated size). Other members of the DPN family consist of smaller packet switches , PADs, and a switch especially configured for trunk and gateway switching. The DPN series has provision for high speed interconnection with a colocated digital voice switching system, thus permitting a hybrid switch configuration.

7.1.5 BBN

In addition to the C/3x series of packet switches BBN has recently introduced the Butterfly multiprocessor system [BBN 85a, BBN 85b]. This multiprocessor system contains a number (4, 16, 64 or 256) of processor nodes which are connected to the same number of common memory modules through an interconnection network which allows any processor node to be connected to any memory module. Processor node-memory module transfers are themselves serial packet transmissions at 32 Mbps. Each processor node contains a Motorola 68000 series processor, local memory of up to 4 Mbytes, and a custom, microprogrammed, coprocessor which acts as the internal packet transmitter and receiver.

The Butterfly processor arose out of a design for a high capacity voice talkspurt multiplexer. It is also being promoted as one of the parallel processing alternatives to supercomputers. Small Butterfly configurations are being developed for use as data communication gateways [MBS 84]. Butterfly-based packet switches will replace the PLURIBUS on DARPA's experimental satellite network.

7.2 Future Commercial Packet Switching Developments

Northern Telecomm has plans for a new DPN-100 packet switch with twice the throughput of the DPN-50. Their research arm, Bell Northern Research (BNR), has been experimenting with more flexible protocols aimed at satellite trunking including larger window sizes and provision for multi-packet retransmission [DRY 85]. They have also been experimenting with mixtures of satellite and terrestrial trunks between the same node-pairs and provision of TOS routing to optimize service over such trunks [CHU 85].

Northern Telecomm believes that the future of digital communications will be dominated, especially for long hauls, by cheap bandwidth as a result of fiber optics developments. As a result, they see hybrid packet-circuit switches sharing internode trunks as principal network components.

AT&T see the requirements for packet switch throughput growing by an order of magnitude every five or six years. thus they foresee requirements in the data world for packet switches to handle 100,000 packets per second before the end of the century. Even so, they see voice as the major driver in their telecommunications world. As a result, they seek maximum commonality between systems; for example, the 1PSS and 5PSS have a lot in common. There is some indication that AT&T thinks that in the future all communications will be digital and packet in nature. Two recent papers by Turner [TUR 85, TUR 86] exemplify this concept. The first of these papers discusses the use of switches with 1.5 Gbps throughput to handle both voice and data/ The second paper adds in broadcast and video for both television and conferencing. In this paper a packet switch terminates 63 fiber optic links operating at 100 Mbps.

7.3 Summary and Conclusions

Commercial packet switches are approaching the capacities and trunk and line speeds required for the NGPS. They also have reliability and maintainability features. They tend not to have some of the DoD/DCA specific requirements such as precedence. Some switches have integrated encryption.

At least one switch is being proposed for computer security certification.

The technical questions concerning the use of a modified commercial switch for the NGPS is the extent and cost of modifications. This can only be addressed in a satisfactory way by the vendors.

8.0 STRATEGIES FOR ACQUIRING THE NGPS

There are only two possibilities for acquiring the NGPS from the viewpoint of the requirements. Only the second one has an impact on the requirements for the NGPS.

a new DDN specific packet switch (family) can be designed, developed, and produced.

commercial packet switches can be procured with modifications to meet NGPS requirements,

Alternative ways of procuring the next generation DDN (including the NGPS) such as leasing versus buying do not affect the requirements for the NGPS. The technological pros and cons of the two acquisition methods are discussed briefly below.

8.1 Design, Develop, and Procure a DDN Specific NGPS

From a technological viewpoint this method will guarantee that all of the specified requirements are met. It will result in receiving exactly what is wanted. The principal question will be one of cost effectiveness for a procurement of the order of 500 to 1000 NGPSs.

8.2 Buy Modified Commercial Packet Switches

As noted in Section 7.3, off-the-shelf commercial packet switches will not meet all of the requirements for the NGPS. At the physical level they are not likely to meet the TEMPEST requirements, nor will they have the desired extent of COMSEC equipment integration. On the protocol level they will not have the set of DDN protocols needed for transition. If DDN is to make the transition to commercially based, ISO standard protocols, those protocols will almost certainly have to be modified to allow for such DDN requirements as precedence and preemption. If the protocol implementations have not been made with computer security in mind, it will be difficult, if not impossible, to validate them at the appropriate level.

If there is interest in considering the acquisition of commercially based switches, it would be valuable to publish the NGPS requirements and ask for comments on those requirements. At that time, some of the difficulties in using commercial switches could be identified and a more detailed examination could be made of ways to cope with them.

9.0 SCHEDULE

9.1 Date NGPS Is Needed

Based on the discussion in Section 3 it appears that the initial installations of the NGPS will be needed by either the beginning of 1988 (no C/300s used) or the beginning of 1991 C/300s used). That is, by the latter date, some switches will require greater throughput than can be provided by the C/300. Installation of operational NGPSs implies both a development schedule and a group of additional requirements. Some of the steps which will need to be taken in making the NGPS available by the desired date(s) are discussed below. Three schedules for the NGPS are prsented in Schedules 9-1 through 9-3 at the end of this Section.

9.2 Development or Adaptation of a Commercial Switch

The schedule presented for the 1991 date is based on the development of a DDN specific device, the alternative described in Section 8.1. The schedule presented for the 1988 date assumes that adaptation of a commercial switch, discusses in section 8.2 is the procurement method; this is the only one that would fit that time frame. If C/300s are brought into the DDN and the 1991 date applies it would be possible to use a less intensive schedule for adapting a commercial switch.

9.3 Testing

Concurrently with the initiation of NGPS development, planning for an appropriate test bed must soon begin. It is likely that the present test bed at DCEC can be the basis for the NGPS test bed. A schedule for the test bed program is included in each of Schedules 9-1 through 9-3.

9.4 Other Requirements

The "Next Generation Network Control Center" has not been examined in the present study. Both the Future Network Technology Report [HER 85] and the Draft White Paper on DDN Capacity state that such a new center will be needed. The requirements for this center must be generated and the necessary centers acquired by the time that NGPS installation takes place.

To assist in making the decision, "develop or adapt" an early call for comments should be made. The current study provides a base to be used in the call for comments.

9.5 Transition

Transition of the NGPS into DDN will require an installation schedule which is coordinated with production and with the installation of new NOCs. Candidate sites for NGPS would be either new sites, those which are heavily loaded with traffic, or those which are candidates for high bandwidth trunks.

Schedule 9-1 Development of a New NGPS

```
NGPS:                       86    87    88    89    90    91
New DDN Plan                 |  ---|     |     |     |     |
Call for Commercial Comments|      *     |     |     |     |
Final Requirements           |     +--   |     |     |     |
A Specification              |     |  ---|     |     |     |
B Specification              |     |  ---|     |     |     |
Publish RFP                  |     |     *     |     |     |
Evaluate Proposals           |     |     | -   |     |     |
Develop Prototypes           |     |     |  ---+--   |     |
Test Prototypes              |     |     |     |  ---+--   |
Begin Production             |     |     |     |     |  ---+-
Develop Transition Plan      |     |     |  ---+--   |     |
Introduce NGPS into DDN      |     |     |     |     |     +-

NGPS Test Bed:              86    87    88    89    90    91
Develop Requirements         |     +--   |     |     |     |
A Specification              |     |  ---|     |     |     |
B Specification              |     |  ---|     |     |     |
Publish RFP                  |     |     *     |     |     |
Evaluate Proposals           |     |     | -   |     |     |
Build Test Bed               |     |     |  ---+--   |     |

Other Requirements:         86    87    88    89    90    91
Plan and acquire new NOC     |  ---+-----+-----+-----+-----+-
```

Schedule 9-2 Adaptation of a Commercial Switch for 1988

```
NGPS:                       86    87    88    89    90    91
New DDN Plan                 |   --|     |     |     |     |
Call for Commercial Comments |  *  |     |     |     |     |
Final Requirements           |   - |     |     |     |     |
A Specification              |    -|     |     |     |     |
B Specification              |    -|     |     |     |     |
Publish RFP                  |    *|     |     |     |     |
Evaluate Proposals           |     |-    |     |     |     |
Test Adapted Switch          |     |    -+     |     |     |
Develop Transition Plan      |   --+-----|     |     |     |
Introduce NGPS into DDN      |     |     |-    |     |     |

NGPS Test Bed:              86    87    88    89    90    91
Develop Requirements         |   --|     |     |     |     |
A Specification              |    -|     |     |     |     |
B Specification              |    -|     |     |     |     |
Publish RFP                  |    *|     |     |     |     |
Evaluate Proposals           |     +     |     |     |     |
Build Test Bed               |     |---- |     |     |     |

Other Requirements:         86    87    88    89    90    91

Plan and acquire new NOC     |   --+-----+--   |     |     |
```

Schedule 9-3 Adaptation of a Commercial Switch for 1991

```
NGPS:                       86    87    88    89    90    91
New DDN Plan                |  ---|     |     |     |     |
Call for Commercial Comments|     *     |     |     |     |
Final Requirements          |     +--   |     |     |     |
A Specification             |     |  ---|     |     |     |
B Specification             |     |  ---|     |     |     |
Publish RFP                 |     |     *     |     |     |
Evaluate Proposals          |     |     | -   |     |     |
Test Adapted Switch         |     |     |     |  ---+--   |
Develop Transition Plan     |     |     |  ---+--   |     |
Introduce NGPS into DDN     |     |     |     |     |     +-

NGPS Test Bed:              86    87    88    89    90    91
Develop Requirements        |     +--   |     |     |     |
A Specification             |     |  ---|     |     |     |
B Specification             |     |  ---|     |     |     |
Publish RFP                 |     |     *     |     |     |
Evaluate Proposals          |     |     | -   |     |     |
Build Test Bed              |     |     |  ---+--   |     |

Other Requirements:         86    87    88    89    90    91

Plan and acquire new NOC    |  ---+-----+-----+-----+-----+-
```

10.0 RECOMMENDATIONS

This Section summarizes the requirements for the NGPS developed in Section 6. The requirements assume a deployment date of January 1991. If an earlier date is chosen as regards schedule 9.2, then the NGPS might not have to meet all the performance requirements if it can be upgraded by 1991. Other supporting information is presented in Section 3.6 and Appendix A.

Date: Production models available January 1991

Quantity: From 500 to 1000

Performance: Throughput of 3360 750-bit packets per second of which 2560 are for tandem traffic, 640 for originating and terminating traffic, and 160 for intranode traffic.

Ports: 20 for trunks to be configured for line speeds as required up to 384 kbps.

99 for hosts to be configured for line speeds up to 64 kbps

DoD Features: Precedence, Preemption

Security: Integral link encryptors, cryptographic checksumming of control messages, switch certified at C-2 level, preferably at B-2.

Compatibility: Must be able to interoperate with DDN PSNs at the initial release.

User Interfaces: DDN X.25, X.75

Operational Features: Congestion Control, Type of Service Routing, Unattended Operation, Remote Maintenance Diagnosis, Fail-soft Architecture

Reliability: 99.995% uptime, 1000 hours MTBF

11.0 REFERENCES

AMN 84 _____, Product Literature for Nucleus 6000 System, Amnet, Inc., Watertown, MA, December 1984.

ANA 84 _____, Networks of the 1990s and Beyond, Report for The National Security Agency, Analytics, (Jul 84).

AND 85 Anderson, Eric R., Ada's Suitability for Trusted Computer Systems, Proc. Symp. on Security and Privacy, Oakland, CA, April 22-34, 1985, 184-188.

BAR 64 Baran, Paul, et. al., On Distributed Communications, Series of 11 reports, The Rand Corporation, Santa Monica, CA, 1964.

BBN 85a _____, Butterfly (TM) parallel Processor Overview, Version 1, (DRAFT), Laboratories, Inc., June 13, 1985.

BBN 85b _____, The Butterfly Gateway, BBN Communications Corporation, September 1985.

BEC 83 Becker, J. O. (Ed.), The 3B20 Processor and DMERT Operating Systems, Special Issue, Bell System Tech. Jour., 62, 1, Part 2, (Jan 83).

BUD 85 Budrikis, Z.L., et. al., A Packet/Circuit Switch, AT&T Bell Laboratories Technical Journal, 63, pp. 1499-1520 (Oct 84).

CAR 85 Carney, D.L., et. al., The 5ESS Switchning System: Architectural Overview, ATT Tech. Jour., 64, 6, 1339-1356, (July - August 1985).

CCI 85 _____, Integrated Services Digital Network (ISDN), Recommendations of the Series I, Red Book, Volume III, Fascicle III.5, VIIIth Plenary Assembly, Malaga-Torremolinos, 8-19 October, 1984, Geneva, CCITT 1985.

CSC 83 Department of Defense Trusted Computer Evaluation Criteria, CSC--STD-001, DoD Computer Security Center, 1983.

CSC 85 Department of Defense Trusted Network Evaluation Criteria, Draft, National Computer Security Center, July, 1985.

CHU 85 Chu, M. et. al., Integrating Satellite Links into a Land-Based Packet Network, Proc. Ninth Data Comms. Symp., Whistler Mountain, BC, 10-13 Sep 1985, 100-104.

DAR 81 DARPA: A History of the ARPANET, The First Decade, DARPA, April 1981 (AD-A115 440).

DDN 82 Defense Data Network Program Plan, Defense Communications Agency, revised May 1982.

DDN 85 Defense Data Network Evolution of Security Services: 1986-1992, DDN PMO, DCA Code B610, Draft of 6 August 85 (with corrections, 30 Aug 85).

DRY 85 Drynan, D. and D. Baker, An Internode Protocol for Packet Switched Data Networks, Proc. Ninth Data Comms. Symp., Whistler Mountain, BC, 10-13, Sep 85, 17-21.

EIS 85 Eisner, S. et. al., Congestion Control Software Preliminary Functional Description, Report No. 6080, BBN Communications Corporation, Draft, November 1985.

ELD 84 Eldridge, C.A., et. al., Network Access Component Specification, MITRE (Sep 84).

FRA 76 Frazer, A.G., The Present Status and Future Trends in Computer/Communication Technology, IEEE Communicaions Society Magazine, Vol 14, pp. 10-19 and 27 (Sep 76).

FRA 79 Frazer, W.D., Potential Technology Implications for Computers and Telecommunications in the 1980s, IBM Systems Journal, No. 2, p. 333 (1979).

GAR 85 Gardner, M. L., Type of Service Routing Requirements, Report No. 6096, BBN Communications Corporation, Draft, November 1985.

HAU 83 Haughney, J., Application of the Burst Switching Technology to the Defense Communications System, GTE Laboratories Inc., 1983 IEEE Military Communications Conference, MILCOM '83, Washington, DC.

HAV 82 Haverty, J., C/30 Information, Internet Protocol Transition Workbook (Mar 82).

HEA 70 Heart F.E., et. al., The Interface Message Processor for the ARPA Computer Network, Proc. Spring Joint Computer Conference, pp. 551-567 (1970).

HEI 82 Heiden, Heidi and Howard Duffield, Defense Data Network, Conference Record EASCON 82, Washington, DC, 20-22 Sep 82, 61-75.

HER 84 Herman, J. et. al., Future Network Technology Study: Interim Report, BBN Communications Corporation, Report No. 5697 (July 1984).

HER 85a Herman, J. et. al., Future Network Technology Study: Final Report, BBN Communications Corporation, Report No. 5894 (Mar 85).

HER 85b Herman, J. G., Architectural Directions for the Defense Data Network: Transmission Service Alternatives, Signal, August 1985, 107-109.

IEE 78 _____, Proceedings of the IEEE, 66, 11, November 1978.

IEE 86 _____, IEEE Communications Magazine, 24, 3, March 1986.

KAT 78 Katsuki, D. et. al., Pluribus -- An Operational Fault-Tolerant Multiprocessor, Proc. IEEE, 66, 1146-1159 (Oct 78).

LIN 79 _____, Experiment Plan for the Wideband Integrated Network - Supplement I, MIT Lincoln Laboratory (Dec 79).

MAC 85 M/A-COM DCC CP9000 Series II Literature, M/A-COM DCC, Gaithersburg, MD, 1985.

MAL 84a Malis, A. G. and W. Gillet, End-to-End Functional Specification, BBN Communications Corporation Report 5575, (Oct 84).

MAL 84b Malis, A. G. and W. Gillett, End-to-End Design Specification, BBN Communications Corporation Report 5576, (Oct 84).

MAL 86 Malis, A. G., PSN End-to-End Functional Specification, RFC 979, March 1986.

MES 84 Mesrobian, W., et. al., High-Speed Gateway Study Final Report, BBN Communications Corporation Report No. 5812 (Dec 84).

NEL 81 Nelson, Ruth, et. al., Pluribus Study Final Report, BBN Report No. 4585 (Jun 81).

NOR 85 Northern Telecomm Literature for DPN Series, Northern Telecom Data Networks, Richardson, TX, 1985.

ORN 72 Ornstein, S.M. et. al., The Terminal IMP for the ARPA Computer Network, Proc. Spring Joint Computer Conference, 243-254 (1972).

PRI 85 Prishivalko, R., White PAPER ON DDN Capacity, Draft, DCA, November 1985.

REI 82 Reiser, M. Performance Evaluation of Data Communication Systems, Proceedings of the IEEE, Vol 70, pp. 171-196 (Feb 82).

SAK 83a Sakamoto, R. D., CCITT Standards Activitiy: The Integrated Services Digital Network, MITRE Report MTR-82W00169, Rev. 1, September 1983.

SAK 83a Sakamoto, R. D., CCITT Standards Activitiy: Ttansmission Systems, MITRE Report MTR-83W00144, September 1983.

SHI 83a Shirey, R.W., Defense Data Network Subscriber Security Guide, MITRE Report, MTR83-W00137, August 1983.

SHI 83b Shirey, R.W., Security Architectures for Long-Haul Packet-Switching Networks, MITRE Report, MTR83-W00163, December 1983.

SHI 86 Shirey, R.W., Defense Data Network Subscriber Guide to Security Services, 1986-1992, Draft, DCA, Code B600, 18 January 1986.

SWE 83a Swetnam, G. F., CCITT Standards Activitiy: Switching, MITRE Report MTR-83W00101, September 1983.

SWE 83a Swetnam, G. F., CCITT Standards Activitiy: Network Management and System Control, MITRE Report MTR-82W00168, Rev. 1, September 1983.

TUR 85 Turner, J. S., Design of an Integrated Services Packet Network, Proc. Ninth Data Comms. Symp., Whistler Mountain, BC, 10-13 Sep 1985, 124-133.

TUR 86 Turner, J. S., Design of a Broadcast Packet Network, Proc. IEEE Infocom 86, Miami, FL, 8-10 Apr 1986, 667-675.

WAL 82 Walker, Stephen T., Department of Defense Data Networks, Signal, October 1982, 42-47.

WIL 81 Wilson, Andrew W., Increasing Speed, Reducing Costs in a Data Network Processor, Computer Design, 143-150 (Sep 81).

WUC 85 Wu, C. and T. Feng, Interconnection Networks for Parallel and Distributed Processing, IEEE Tutorial, 1985.

ZOR 83 Zornig, J.G., IMP Performance Measurements Preliminary Results, BBN Communications, Report No. 5423 (Oct 83).

APPENDIX A

DISCUSSION OF PACKET SWITCH ARCHITECTURE AND PERFORMANCE

A.0 INTRODUCTION

This Appendix is to discuss how salient features of packet switch architecture affect the performance of that packet switch, how packet switch throughput is modeled, and presents an example of a multiprocessor architecture for the NGPS.

A.1 Processor Architecture and Packet Switch Performance

A simple model of the required packet switch processing is developed and used to evaluate the switch architectural features.

A.1.1 Architectural Features

Salient architectural features for a packet switch processor include the following:

- processor instruction set
- processor serial instruction speed
- processing element complement (single, dual, multiple)
- primary memory organization
- memory-processor interconnection
- primary memory bandwidth

An assumption is also made that individual communication link ports are handled directly by specialized processing subsystems. These subsystems contain local storage which can be block transferred to/from global memory, and specialized processors for synchronization, performance of error checks, and for removal or addition of header information.

From an economic standpoint it is desirable if the architecture utilizes standard "building blocks" such as microprocessor, USARTS, CRC generator/checkers, and memory devices. The performance of any architecture is a function of the application which is to be run upon it. In order to focus the analysis of packet switch architecture for this Appendix the packet switch application has been modeled simply in a form which appears to be very useful in assessing the importance of specific architectural features. The numbers proposed in the model are believed to represent realistic nominal values. Additionally, they provide a basis for assessing architecture/performance relationship in a way which is not strongly affected by changes in the nominal values used in the model.

A.1.2 Packet Switch Performance

The packet switch receives and transmits an approximately equal number of bits over its store-and-forward channels. These channels are both those to and from directly connected hosts and those to and from other switches. (The packet switch does not accumulate information that must be stored in a secondary memory). The packet switch successfully uses queues to smooth out variations in packet arrival times so that its operation can be viewed as being in a steady state.

Given that the bandwidth of the communication links is sufficiently high, the rate limiting aspect of packet switch operation is the rate at which decisions and memory management can be performed by the packet switch. Packets are assumed to be of uniform, 1000-bit size; these may represent message segments, control traffic, aggregate acknowledgements, etc.

The model for processing the 1000-bit packets is simply that:

a decision must be made as to how to route and otherwise process the packet (e.g., forward it, hold it for reassembly, etc.);

some memory management must be performed to clear and/or rearrange memory space.

The decision process is assumed to require an average of 500 elementary instructions and 1000 bits of memory must be accessed for the packet memory management. With this processing model the time required for the packet switch to handle a single packet is the sum of:

the time to handle 2 block data transfers between local and global memory;

the time to execute 500 instructions;

the time to perform memory management (taken to be 1 instruction per word).

Memory bandwidth is prominent in these expressions for both data transfers and for the rate at which the decision instructions are fetched and performed.

The second term above is related to both the processor serial instruction rate and the instruction set. The serial instruction rate can be expressed in instructions per unit time (e.g., millions of operations per second or Mips). The cycle rate of the processor affects its speed but so does the processor's internal architecture. For example, a processor with a large register set can concentrate on fast register-to-register operations. Thus more of the instruction mix can be in these fast operations. Some processors have complex operations particularly suited for packet switching functions such as enqueue and dequeue instructions. These instruction types can significantly reduce the average number of instructions required for the packet switching operations.

The maximum throughput for this model packet switch can be improved by supplying additional processors to perform decision calculations in parallel for a number of packets at one time. The rationale is as follows: Memory/ data bus width will have an impact on the first and third terms above. There will be less memory access required to transfer the 1000 bit packets if wide memories and data busses (e.g., 32-bit) are available. Assuming a 32-bit word the first and third terms will require about 100 memory access instructions and the processor time for step 2 instructions will predominate.

In the situation just examined the ratio of decision instructions to data movement instructions was 5:1. Using 5 processors would bring the decision processing time down to the data movement time representing a more balanced situation. However, each processor would require its own instruction stream and thus global memory access and data paths could become the new bottleneck.

In summary, a specific packet switch architecture can be examined with the following steps:

identify a packet size;

postulate the number of processing steps for one packet;

postulate requirements for data movement and memory management;

use the processor specifications for memory bandwidth and instruction time to find the time required to process one packet; invert to find maximum packet rate;

examine effects of increased memory bandwidth and faster instruction execution;

examine ways to balance decision instruction time with data movement time

Two important factors which are not related to packet switch throughput were not taken into account in this discussion. *The first factor is* the ability of a multiprocessor system to provide graceful degradation if one or more processors fail. This ability can only be achieved wit an appropriate operating system (note that we are not referring to a general purpose operating system but a limited purpose operating system for the specific packet switch application). The second factor is the problem of providing a certain level of computer security in the packet switch. Both of these factors are addressed in the discussion of overall packet switch requirements. They are likely to be important in the selection of specific processors for a multiprocessor system.

A.1.3 Multiprocessing/Parallel Architectures

The discussion in the previous section has made a strong implication that multiprocessor architectures are a desirable direction for the next generation packet switch. At the same time, it is clear that the interconnection of processors and memories is a key issue if bottlenecks are to be avoided. In the example of the previous section the use of 5 processors removed the bottleneck having to do with the ratio of decision instructions to data movement instructions. This was accomplished at the expense of a potential bottleneck in the global memory and data paths.

There are at least two ways in which this potential bottleneck can be avoided. A current topic of major interest in parallel computation is the provision of multiple, high bandwidth, paths between processors and memories. (A recent collection of papers on this topic will be found in an IEEE Tutorial volume [WUC 85]). A less conventional solution can be envisioned in a system whose data storage consists of shift registers and FIFO buffers, resulting in a special purpose processing system.

A.1.4 Summary of the Model

This simplified discussion of a model for packet switch performance has highlighted the following steps:

1. Identify a basic model of the processing performed by a packet switch; here a 1000 bit data segment was used.

2. Postulate the steps required to process basic event(s); here 500 elementary instructions were assumed necessary for decisions.

3. Postulate requirements for data movement and memory management.

4. Use processor specifications for memory bandwidth and instruction timing to assess time required to process a segment of data; invert to obtain the maximum data rate.

5. Examine the effect of increased memory bandwidth, either through faster access times or wider data paths.

6. Examine the effect of faster instruction execution.

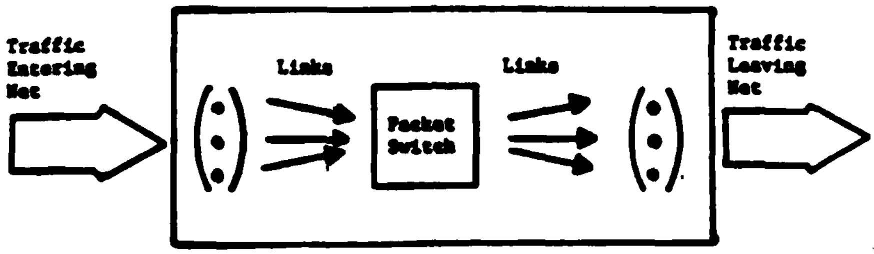
7. Examine the ratio of decision instruction execution time to data movement time.

8. The steps above explore the domain in which processing is the limiting factor; also explore the internal limiting data bandwidth as defined by memory and data path bandwidth.

A.2 Network Throughput and Packet Switch Throughput

A.2.1 Top Level Model

The underlying packet switch model views the switch as a node in a network of servers. The switch's outgoing telecommunication links are its own network servers. Incoming telecommunications links are the paths for introducing new service requests. This model is illustrated below.



The load experienced by a packet switch in a network will depend upon the following network characteristics:

1. The set of traffic flow demands between all possible network entry and exit points (i.e., nodes that have hosts attached). The heavier these flows are, the heavier the load experienced by the packet switches.

2. When traffic flow demands are very unbalanced and the allocation techniques are based upon minimizing delay, some network paths may deviate significantly from shortest hop count paths to avoid overloading.

3. The capacities of the communication links interconnecting the node limit the number of bits that can flow through a packet switch.

4. The topology of the interconnections of packet switches. A packet switch with very few links is apt to experience a lighter load than one with many incoming and outgoing 'inks.

A.2.2 Estimation Techniques

Given the wealth of parameters that determine the load on a packet switch, an accurate yet simple load-predicting model is not obtainable. Instead, one of several estimation techniques may be used to predict packet switch loading. Computer system engineers agree upon a hierarchy of estimation techniques for loads and performance of computer systems subject to multiple job streams of variable intensity (such as the nodes of a computer network).

The most accurate estimates come from benchmarks in which the actual computer systems are run under actual loads while measurements are being made. This also allows distributions rather than single estimates to be made for node loading and performance.

Accurate estimates of system performance are frequently made from discrete event simulations. These are computer programs often constructed using specialized languages representing processes and series of random events treated by those processes. Simulations model a large number of system features in such a way as to represent their interactions in real time. (Real time itself is simulated by a clock that is advanced only under control of the simulation program.) Discrete event simulation programs can model a large number of events to report statistics about loading and performance, rather than single estimate answers.

Satisfactory estimates can be made from queuing theory calculations based upon the arrival frequencies of jobs (i.e., frequency of arrival of packets over a line) and upon the probability distribution of job service times (i.e., the distribution of packet lengths divided by the link's bit rate). In computer networks there is (usually) an interaction between the path allocation technique and the queuing delays estimated at each node. The allocation technique dynamically adapts to the delays experienced at each node, rerouting flows away from congested nodes. Algorithms for evaluating network node load and performance can take into account the allocation technique at the expense of computational complexity. (Section B.3 below describes steps in estimating detailed network loading from topological descriptions and flow requirements. Loading and performance of individual nodes can be derived).

Adequate estimates can often be made using "rules of thumb". In the case of networks, such rules of thumb include judgments about the average number of hops in a network path and the implicit assumption that every path constrains this number of hops, and about the ratio of store-and-forward traffic (from other packet switches) to newly entering traffic (from attached hosts). These rules of thumb are derived from experience with benchmarks and from more detailed analytical or discrete event simulations. As such they represent "average" or "typical" values. They do fail to represent the variance beyond "typical" cases and as a result the possible consequences of this variance. (In contrast, analytic queueing theory calculations do calculate the consequences of variance in flow demands, packet lengths, etc.)

For the purpose of illustrating sample packet switch performance requirements we have used rule-of-thumb calculations. These are easier to follow, yet they are supported by experience with the more detailed calculations.

A.2.3 Iterative Techniques for Network Load Leveling

This section describes an algorithm for assigning flows (i.e. of data) over a network of packet switch nodes and communication links. The procedure is briefly presented here.

First, an initial flow assignment is attempted by seeking the shortest path (in terms of number of hops) for each flow demand. The technique for balancing flows in the initial assignment attempt (and in subsequent refinements) is to use per link delay estimates, based upon queueing theory, rather than number of hops in path length calculations. The same link may be used for more than one flow and may be over utilized. In such a case, an attempt must be made to balance the flow so that no link is over utilized.

Next, the success of the initial assignment is evaluated according to link utilization. Whenever a flow assignment over utilizes a link, the flows are hypothetically rescaled to a utilization arbitrarily just under 100% (e.g. 99%). This has an effect upon the queueing theory calculations in a next iteration -- very large delays are calculated for the link in question. Consequently, minimum-delay paths will avoid the link in question, thus re-balancing the load.

Finally, repeated attempts are made to optimally balance the load so as to minimize a grand delay average. Solution convergence is used to halt the iterations. However, this does not guarantee a globally optimum flow assignment.

This technique involves iterative, compute-intensive optimization and only partially represents how a distributed routing algorithm performs its own flow assignments. (A distributed routing algorithm performs at each node a shortest path calculation based upon per-link delay estimates that are received at each node). It also does not provide a way of evaluating dynamic aspects, such as the time required to adapt to network configuration changes. It does provide a hypothetical "performance envelope" of network delays, against which the performance of a routing algorithm can be judged. It can also be used to evaluate the maximum traffic handling capacities of hypothetical network configurations.

A.3 An Example Architecture for the NGPS

In this Section we establish that an economical and elegant architecture can be found to satisfy the requirements of the Next Generation Packet Switch. Factors that need to be taken into account for the architecture include: throughput, reliability, security, flexibility.

These factors are discussed in general terms below.

A.3.1 Throughput

A multiprocessor architecture is appropriate for the NGPS as the following discussion makes clear. A requirement was developed in Section 3.6 that approximately 3360 packets per second be handled. The majority of these are tandem traffic and we can use the figure of 500 instructions per packet, developed in A.1 as the dominant factor. Therefore, we need to execute over 1.6 million instructions per second (MIPs). A good goal would be 2 MIPs. The nature of the instructions and of the data handling operations must also be taken into account. The way in which current computing architectures provide increased power at lower cost is the use of parallel processing when the computing tasks can be partitioned. The computing tasks of a packet switch are just such a case; an aggregate of processors whose total computing power will be about 2 MIPs can handle the throughput. We might do well to double this requirement as a safety factor and to allow for additional features to be provided.

A.3.2 Reliability

Although computing components continue to become individually more reliable, the NGPS requires a very high degree of reliability. Further, failure of one component should not disable the switch. That is, the system should be fault-tolerant or fail-soft. As in the example of the PLURIBUS and the Butterfly, a multiprocessor architecture with appropriate interconnections is well suited to such a requirement.

A.3.3 Flexibility

The next generation packet switch will have to support a dynamic communication world. The current generation of switches has evolved steadily in terms of its interfaces and internal operations, but always on the same basic architecture. The use of a stored program computer has permitted such flexibility; this facility can be realized in a multiprocessor switch. Multiprocessor operating systems are not as mature as those for well established uniprocessors; however, orderly development methods for such operating systems are currently well in hand.

A.3.4 Security

The multiprocessor system postulated here for the NGPS is a distributed computer system or a network of computers. Assuring computer security for distributed systems or for networks of computers is still in a primitive stage. The National Computer Security Center has not yet released their final version of the DoD Trusted Network Security Criteria (the draft is [CSC 85]). Nevertheless, an orderly and formal approach to computer security can be taken in developing a multiprocessor based NGPS.

A.3.5 Summing Up

This Section has shown that a multiprocessor-based solution to the hardware architecture of the NGPS is realistic and attractive. The critical point for throughput of that solution is how the processors work with the aggregate of memory; that is, the nature of the processor-memory bus structure and performance. This solution is also attractive to commercial vendors as the examples in Section 7.1 have shown.

APPENDIX B - ALTERNATIVE NETWORK ENVIRONMENTS

B.0 Introduction

In Section 5 we presented an environment for the NGPS which can best be described as a scaling up of the DDN as it exists today. Topologically, the network would not be changed very much and the average hop-length would be about the same. Differences would be mainly in volumes of traffic carried, in accommodation to new communication technologies, and in provision of more specialization such as Type of Service.

In this Appendix we discuss briefly some alternatives to that topology and environment for the NGPS. These possibilities received a very modest amount of attention during the development of this report. The Introduction to Section 5 presents the rationale for electing continuation of the "classical" topology. Nevertheless, for completeness, we describe here some alternatives which differ from that "classical" topology. These alternatives are described in Section B.1 below. The probable effects of the alternatives on NGPS requirements are discussed in Section B.2. .

The Future Network Technology Study identified two ways of using high-speed trunks in DDN (pp 236-37 of [HER 85a]). One of them is the hierarchical network in B.1.2 below. The other multiplexes T1 circuits into 64 kbps trunks and limits individual trunks to that speed. We have rejected the second option in Section 5 of this report.

We believe that the alternatives discussed below, as well as some which have not yet been "invented", represent possibilities which will need to be considered in the generation which follows the NGPS.

B.1 Alternative Strategies for the DDN

In this section a number of alternative architectures/environments/strategies for the DDN of 1990+ are briefly presented. They all allow for the substantial growth in user support requirements developed in Section 3, though they represent substantially different ways of getting there. Each description of an alternative also contains a very concise discussion of the "pros" and "cons" for such a system.

B.1.1 DDN as an Internet

Section 5 developed a concept of DDN as a continuation of the current DDN in which the net is, at least topologically, a uniform whole. That is the net can be conceptually diagramed as shown in Figure B-1 (a duplicate of Figure 5-1, but shown here for convenience). An alternative architecture for the DDN would be the provision of many independent networks provided or maintained by groupings of users which fall together logically. The majority of the traffic in these independent networks would be intra-network. The networks would be interconnected by a DDN which provided only internet service. Figure B-2 is a diagram of that configuration.

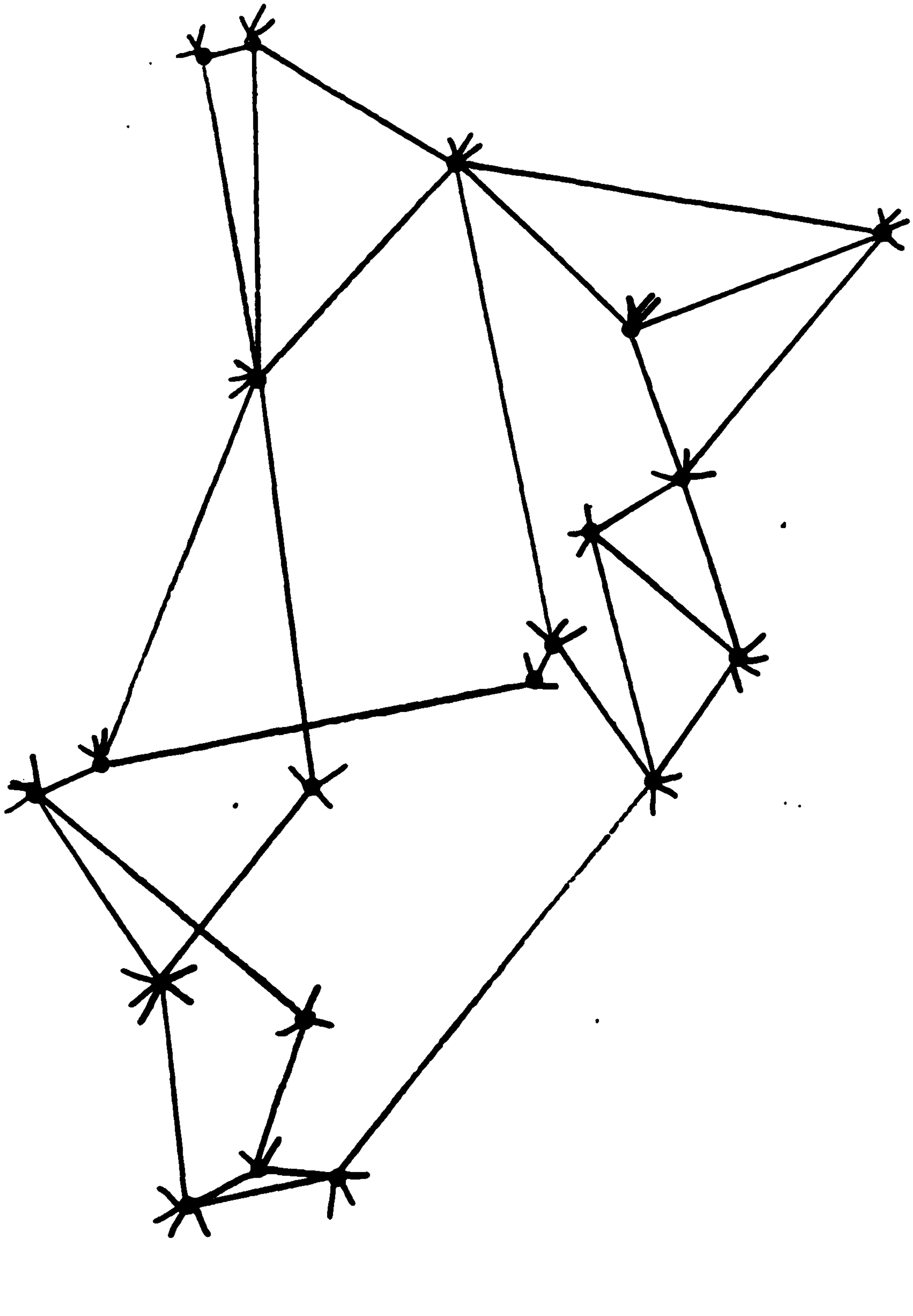


FIGURE B-1 PRESENT DAY DDN EXPANDED

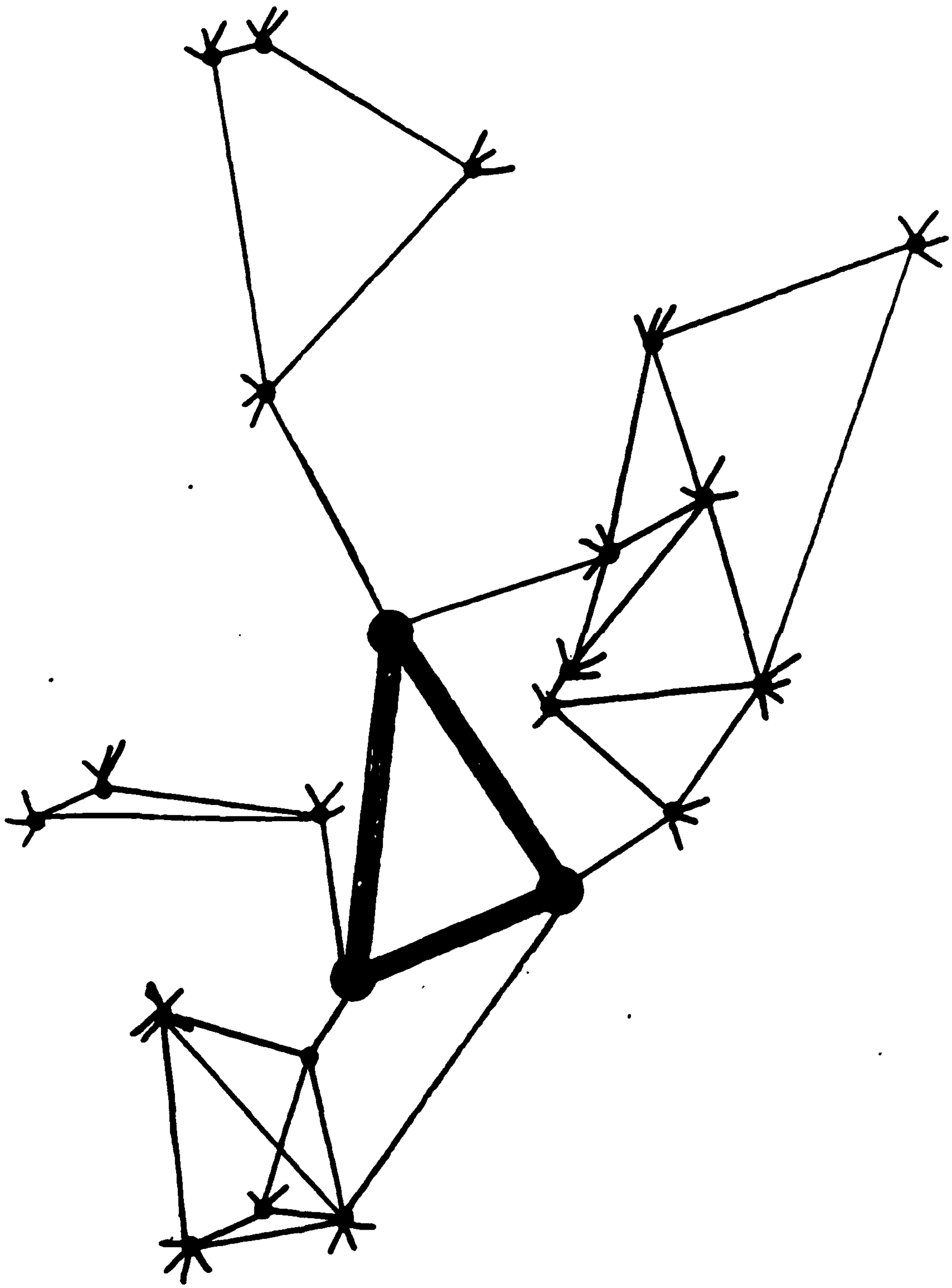


FIGURE B-2 DDN AS INTERNET

B.1.2 DDN as an Hierarchical Network

This version of a DDN can be considered as a physical implementation of the concept of area routing for DDN. Such a network is illustrated conceptually in Figure B-3. As far as groupings of users are concerned, it is somewhat like the diagram of Figure B-2. There are two major differences in the hierarchical concept: the groupings are not connected to each other through a gateway but through a somewhat specially configured packet switch; the nodes are all supported by DDN.

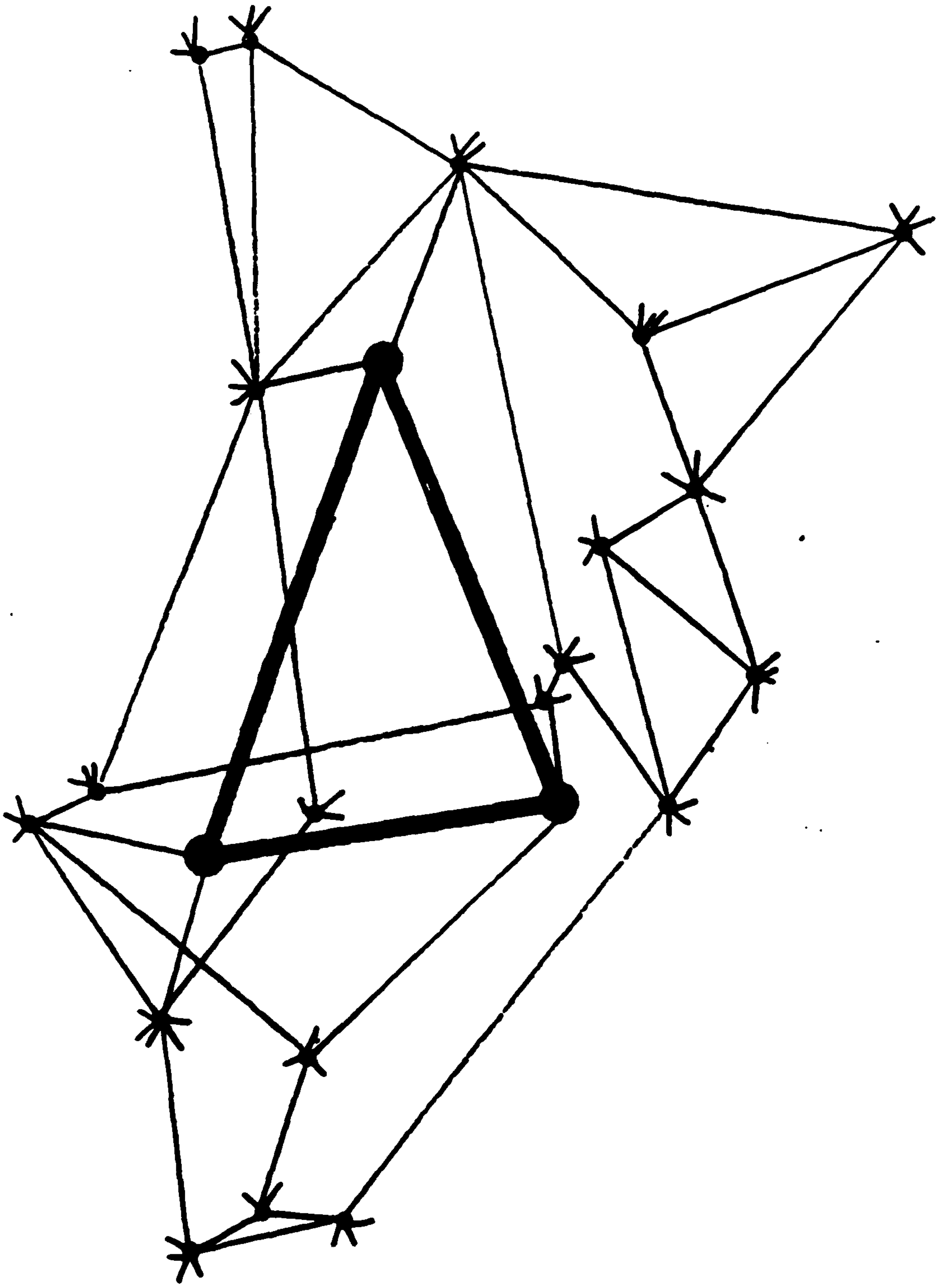


FIGURE B-3 DDN HIERARCHICAL NET WITH INTERIOR (SUPER NODES.)

B.1.3 DDN as a Multi-Vendor Network

This alternative assumes that it will be possible to rely on that, at most, minor adaptations of commercial packet switches will be necessary to provide the functionality needed by DDN. It further supposes that different vendors will be chosen for successive acquisitions and further assumes that standardization will continue to take place to the extent that switches from different vendors will be able to be interconnected without difficulty. (If not, networks of different vendors can always be connected through an internet.) Under these conditions it will be possible to add new switches which represent the best value at the time of acquisition and older equipment will be able to coexist with newer. Figure B-4 represents such a network.

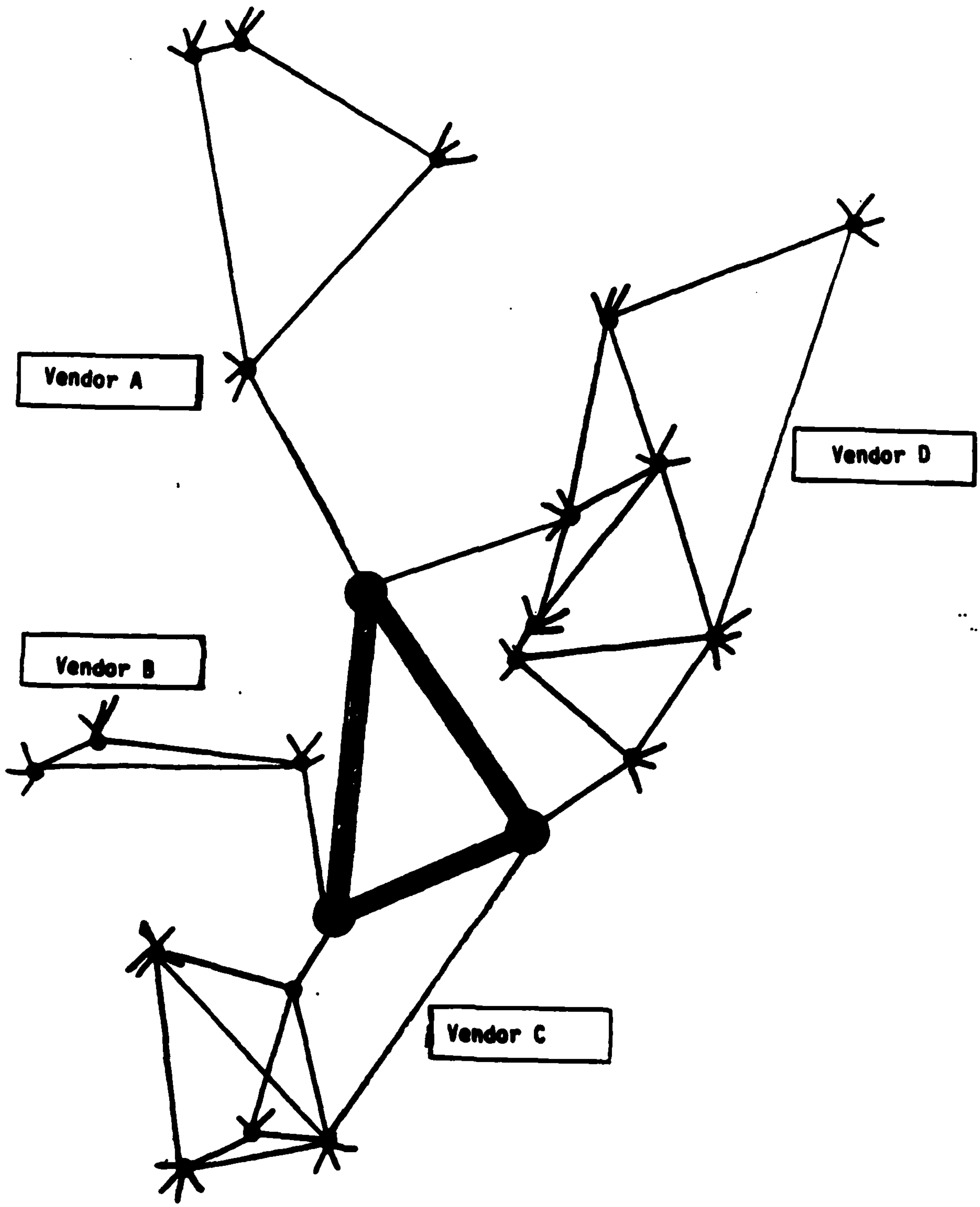


FIGURE B-4 DDN AS A MULTI-VENDOR NETWORK

As a practical matter, the adaptations to commercial switches which will need to be made will represent a substantial software effort. Since the standards met by such switches will be at an interface level, use of multiple vendors will increase the software adaptations by the number of vendors.

B.1.4 DDN as a Maximal Backbone

This alternative, diagramed in Figure B-5, would provide a network with relatively few nodes. Each node would serve many more subscribers than present nodes. This maximal backbone is equivalent to the top tier in the hierarchy discussed in B.1.2. It is also the same as the top tier in the high speed backbone of the Future Network Technology Study (Figure 4.12, p 237 of [HER 85b]. Such nodes would represent a more significant target for physical denial of service. Therefore, for reliability and survivability reasons it would be necessary for subscribers to be homed to two different nodes. The nodes themselves would be interconnected by high data rate communication links. From the standpoint of the network, this configuration would be similar to an upscaling of the present DDN; the nodes would have greatly increased throughput and many more user ports.

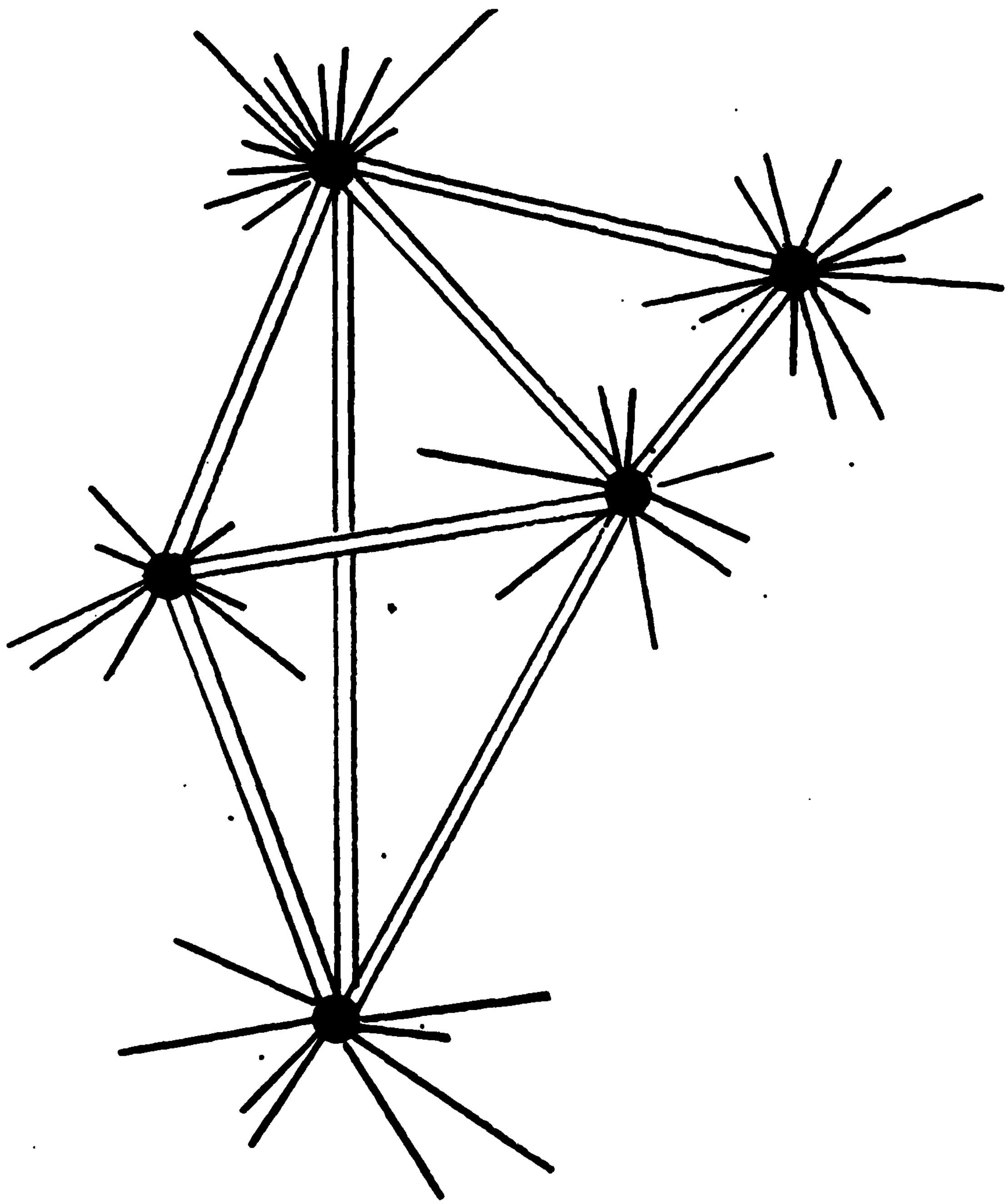


FIGURE B-5 MAXIMAL BACKBONE

B.1.5 DDN as a Hybrid Network

This alternative is rather different from those described above. The others (Sections B1.1 through B1.4) are all data networks and this architecture represents a network for data, voice, and perhaps other transmissions as well. Such a network, shown in Figure B-6, would provide a flexible use of inter-node trunks. These trunks would be allocated to data or to voice in a time-varying mixture. Technologically, this would be in accordance with the intended development of ISDN. The packet switch would have to take on new functionality as it reallocated trunk capacity. In principle, this would be a powerful and flexible way to maximize communication trunk resources. In practice we probably do not know enough at this time to make intelligent plans for such usage.

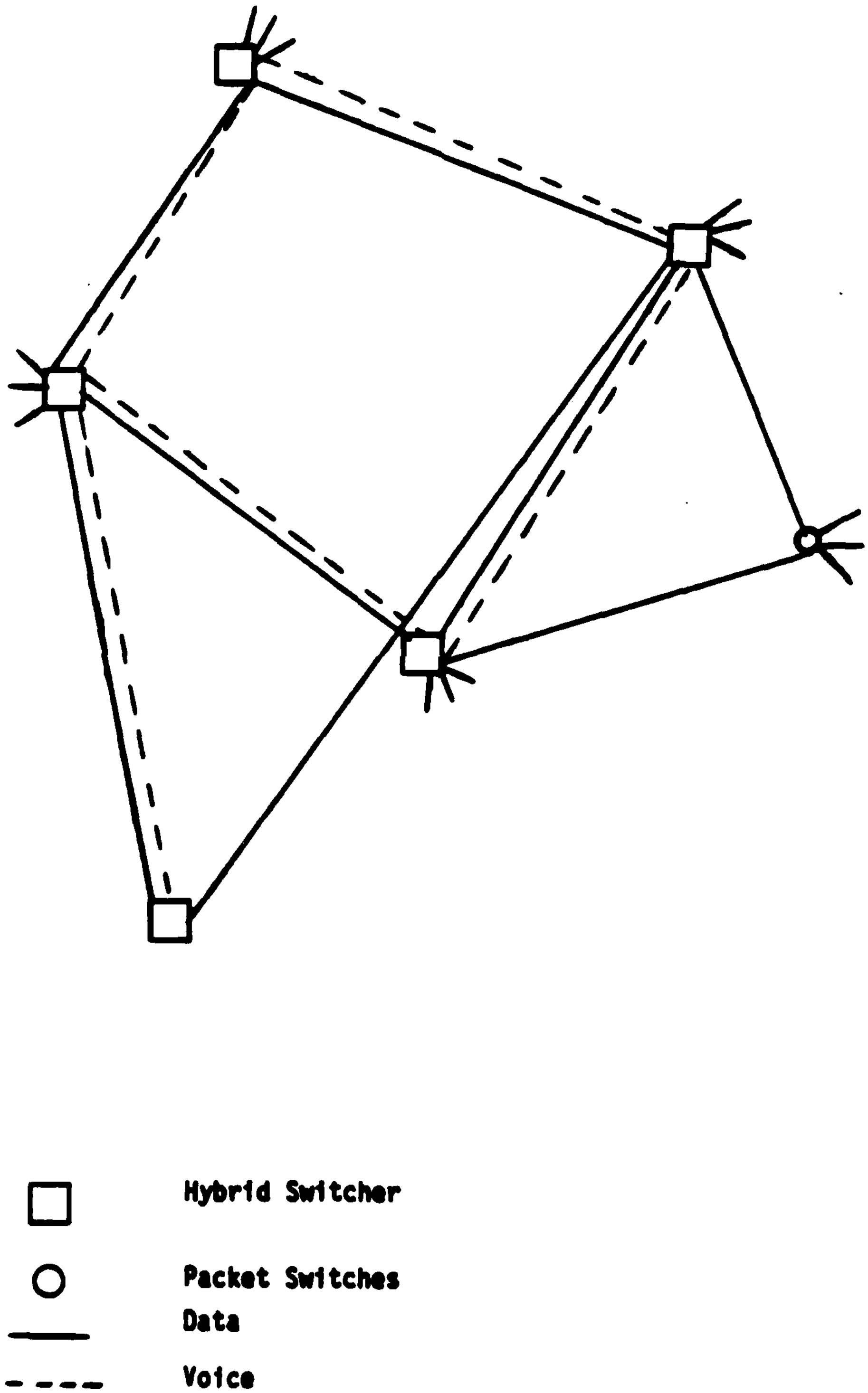


FIGURE B-6 A CONFIGURATION WITH A MIXTURE OF HYBRID SWITCHES

There are a number of problems with such a system for DoD/DCA usage at this time. Although there are commercial hybrid voice/data packet switches (for example the No. 5 ESS described in section 7.1), these switches utilize separate voice signaling and they do not necessarily have any provision for DoD quality communication security of either end-to-end or traffic flow security.

Burst switching, another alternative to packet or circuit switching has been proposed recently for voice transmission [HAU 83]. Burst switching is a form of very fast circuit switching. Each time a talkspurt starts a new circuit allocation is made. This lasts for the length of the talkspurt which is typically a few seconds.

B.2 The Effects of the Alternatives on NGPS Requirements

The conversion of DDN to an internet would require less switches. These might not need to be of higher performance than the C/300s. They would require new software, a major cost in any PSN procurement. The hierarchical and maximal backbone environments would require a moderate number of very high performance switches. A multi-vendor network would mean accommodating DDN to modifications of commercial packet-switching practice. The hybrid network could possibly capitalize on commercial offerings which may arise out of ISDN.

END

DTIC

6-86