Precedence and Priority Access Implementation For Department of Defense Data Networks

William W. Barns

July 1991

MTR-91W00029

SPONSOR:
Defense Information Systems Agency
Contract No.:
DAAB07-91-C-N751

The MITRE Corporation
Washington C^3 Center
7525 Colshire Drive
McLean, Virginia 22102-3481

Department Approval: _____

Gladys A. Reichlen
Associate Manager
Networking Center

MITRE Project Approval: _____

Adel I. Messeh
Associate Department Head
National Telecommunications
and Data Systems

ABSTRACT

Multilevel Precedence and Preemption (MLPP) functions are used in military telephone and message systems to help assure that important communications activities are not blocked or unduly delayed by less important calls or messages. The Defense Data Network (DDN) provides MLPP features for general data traffic, but additional implementation actions in other networks and in computer systems connected to these networks are necessary to make MLPP features perform usefully in the total communications environment. This report presents background information and a methodology for designing the MLPP implementations in network and host components, and identifies areas of the DDN implementation that should be enhanced to make MLPP perform well in the current internetwork environment.

Suggested Keywords: Defense Data Network (DDN), Communication networks, Packet switching, Multilevel Precedence and Preemption (MLPP), Precedence, Priority.

ACKNOWLEDGMENTS

Sections 2, 3, and 4 of this work are a synthesis of past experience of many government organizations and contractors in the development and use of military communications systems that include precedence features. Many of the design and evaluation documents that include this information are not available to the general public because they contain sensitive military information. Although these sources cannot be cited as references in this report, the author wishes to note that many unnamed people and organizations have contributed indirectly to this report through their work on these military systems.

TABLE OF CONTENTS

SECTION **PAGE**

SECTION **PAGE**

LIST OF TABLES

EXECUTIVE SUMMARY

The Defense Data Network (DDN) is the primary common-user, long-haul data communications facility of the Department of Defense (DOD). The DDN now consists of four packet-switched data networks: Military Network (MILNET), which carries unclassified traffic, and three similar but physically separate Defense Secure Networks (DSNETs) 1, 2, and 3, which carry classified traffic. In the future, the three DSNETs are planned to merge into a single Defense Integrated Secure Network (DISNET). The Defense Information Systems Agency (DISA), formerly the Defense Communications Agency (DCA), is responsible for engineering and operation of the DDN.

A multitude of military systems are connected directly or indirectly to the DDN. The rapid growth in numbers of computer systems used by military organizations has led to widespread implementation of local area networks (LANs), which provide relatively high speed interconnection of (potentially) large numbers of systems in a limited geographical area such as a building or a base. These LANs are often connected to the DDN using *local distribution systems* or *concentrators*, known more technically as *routers*. (The term *gateway* was formerly synonymous with *router*, but this usage is obsolete.) Many varieties of routers are available commercially from a multitude of sources.

A communication network intended to support military organizations must provide certain levels of service under various adverse conditions to which military systems are exposed. In the case of the DDN, the DDN Program Plan (DCA, 1982), currently being updated, calls for various security and survivability provisions to meet these needs. Many of these provisions have already been implemented in full or in selected portions of the network. Others are in the process of being implemented, or are awaiting planning or design work.

One of the special military features planned for the DDN is a Multilevel Precedence and Preemption (MLPP) capability. This name describes a general class of features that has been provided in many military communication systems in the past, such as the DOD Automatic Voice Network (AUTOVON) and successor Defense Switched Network (DSN) telephone systems, and the Automatic Digital Network (AUTODIN) used for formal message traffic. The reference to Preemption in MLPP is an artifact of implementation decisions in pre-existing systems; the term is sometimes used in a more general sense to include systems achieving an equivalent result without using preemption as the means of providing the service.

The general notion of MLPP is that any traffic submitted to the communication system is assigned to one of several precedence levels. These levels are administratively predefined and they have an ordered relationship from highest (most urgent or important) to lowest (least urgent or important). Elements of an MLPP system are expected to make some arrangement to ensure that the service provided to low precedence traffic does not adversely affect the service provided to high precedence traffic. In practice, this means that when the communication system cannot completely satisfy all demands that its users attempt to place upon it, degradation or denial of service ought to be forced upon the low precedence traffic rather than the high precedence traffic.

PURPOSE AND SCOPE OF THIS REPORT

The 1982 DDN Program Plan includes a fairly detailed technical description of how the formal requirement for precedence in the DDN was to be implemented in the network components. This implementation has been substantially but not totally carried out in the course of DDN software implementation and evolution of the original packet-switched network backbone. There has been no guidance of comparable depth for DISA's implementation of the router-based high speed backbone network or for subscriber networks and end systems. Also, the DDN approach to MLPP requires updating to provide a technical approach for extending the precedence function to the increasingly prevalent environment of end users being connected to LANs, which are in turn connected to the DDN through some type of gateway device. As a result of this lack of guidance, the network components and end application systems being acquired and used by subscribers often do not have the necessary features to employ MLPP service, or would be prevented from using MLPP service because their access to DDN goes through some other component that does not support the service.

The original DDN planning also did not address the router-based backbone network now being implemented by DISA.

This document provides technical recommendations for the implementation of precedence and preemption features in DOD data networks, with primary emphasis on completing the implementation of MLPP in the DDN packet switches and implementing the corresponding mechanisms in subscriber networks and systems that connect to the DDN. The technical discussion is generally applicable to any layered protocol structure, but the details are described in terms of the Transmission Control Protocol and Internet Protocol (TCP/IP) protocol family and the Open Systems Interconnection (OSI) protocols since TCP/IP is prevalent on DOD networks now and OSI is the Government standard for the future. To support the recommendations and assist in detailed planning for specific system implementations, information is provided in the following areas:

- Explanation of the benefits of MLPP and, conversely, the quality of service features that are not within the scope of MLPP to provide. This information should be used by planners of both networks and user systems to determine whether including MLPP features would improve the performance or functional capabilities of a component or system.

- Description of the types of mechanisms used to implement precedence functions–queueing, preemption, and blocking. This is background information for use in planing an implementation of MLPP features or analyzing those that have already been designed or implemented.

- Design methodology for determining an appropriate technical means of providing MLPP service in a particular component or set of components. This can be used to develop a detailed design specification for a component or system, to help guide the actual implementation process, or to analyze an existing design or implementation.

- Technical issues to consider in designing end user software components for use in the MLPP environment. These issues are relevant to specification and design of end user

systems, and some of them also affect the network provider if the network is intended to apply restrictions to the user system's choice of precedence level.

- An interpretation of MLPP services in internetworks, and specific recommendations for MLPP implementation in internetwork routers. This will assist military network service providers, including operators of both long-haul and local networks, in specifying appropriate features for such components as Internet Protocol (IP) routers. This information can also be used by vendors to develop such products with MLPP capabilities.

- An approach to using the Priority features of OSI protocols to support precedence signalling in networks and systems conforming to the Government Open Systems Interconnection Profile (GOSIP) standard. This is a strawman for implementing MLPP in an OSI environment, but refinement and coordination with standards organizations are likely to be needed to obtain commercial OSI products that support this approach (or any alternative approach).

The treatment of general quality of service (QOS) issues is beyond the scope of this report. These are more complex problems and the focus of this report is on the simpler problem of ensuring that the available resources are provided to the most important traffic.

OBJECTIVES OF PRECEDENCE FEATURES

Precedence and *Priority* both refer to an administratively assigned attribute of a communication instance such as a telephone call or a text message. *Precedence* is the DOD term for a standardized designation of importance of timely delivery and handling. *Priority* is a generic term used in computer and communications engineering for such notions as precedence.

Preemption is the premature suspension or termination of an activity in order to permit some other activity to proceed. It is one of several actions that a system element may take when the demand for a system resource exceeds the available supply. Preemption is employed in many existing systems that respond to precedence level selection. However, mechanisms other than preemption can also support a precedence feature in many situations.

From a user's point of view, the goal of a precedence feature is to ensure that important communications are carried out in a timely way. At a more technical level, the primary objective of a precedence feature is to allow users or authorities superior to the users to guide the allocation of communication-related resources when the demand for a resource exceeds the available supply of that resource.

Precedence mechanisms do not provide guaranteed quality of service (QOS) in any sense. They may be used in networks that guarantee a certain QOS, but the overall design of the network determines the QOS it can provide. Precedence mechanisms change the order in which operations happen. Assigning appropriate precedence levels to different uses of the network or system allows the most important uses to have the greatest likelihood of receiving the QOS that the network is designed to

provide in an unstressed environment. This result can also be described as shifting delays from high precedence users to low precedence users.

MLPP features are therefore useful in situations where resource exhaustion is a reasonable possibility and the impact of resulting service delays is more serious for some users than for others. If it is sufficiently certain that resource exhaustion will not occur, there is no need for MLPP features. If the real-world impact of resource exhaustion is about the same for all users, MLPP features cannot improve the situation. The types of situations in which MLPP may be useful include:

- Large, unpredictable (or unsupportable) increases in traffic in a crisis situation
- Systems subject to battle damage, sabotage, etc., which could reduce capacity suddenly and unpredictably
- Major traffic pattern shifts due to redeployment into configurations for which the system was not engineered
- Gateways interconnecting networks with greatly differing bandwidth capabilities

It is sometimes suggested that the motivations for providing MLPP features in older communications technology do not apply in a modern packet-switched network environment. This assertion is overly broad. If enough capacity is certain to exist, MLPP is not needed in any environment, packet-switched or otherwise. Packet switching changes the delay characteristics of the network in a way that often makes it possible for more users to receive timely service with an equivalent investment in network capacity. However, the capacity of any network is still a fixed and finite amount. If the load offered to a packet-switched network grows large enough, delays in the network will grow to a point that makes the service unsatisfactory to the user. Under these conditions, MLPP can provide a useful service by shifting these delays to the least important users, ensuring that the most important users are still supported.

Likewise, the expected widespread use of integrated communications services such as the Integrated Services Digital Network (ISDN) technology and fiber-based high bandwidth transmission will change the specifics of the issues, but not the principles. If the available bandwidth and switching capacity are sufficient to support all the traffic that users will ever try to send, then MLPP is not needed. However, the capacity of the network may be limited by factors other than the bandwidth of the transmission technology, such as the capacity of the switching components of the network or the capacity of the end systems. Also, it is likely that for some time to come, some network users will still be using lower-bandwidth technology at some point in the communication path. These considerations suggest that in the integrated services era, MLPP features will be most important in switching elements and gateways that connect together dissimilar (or simply disjoint) subnetworks, and in the end systems themselves.

RESOURCE IDENTIFICATION, PRECEDENCE SIGNALLING, AND PRECEDENCE PROCESSING

Two major concepts—(1) resource identification and management, and (2) precedence signalling and processing—are used throughout the technical discussion to break down the analysis problem into manageable pieces.

The entire communication system is considered as a set of resources. System components (the network components and the communicating devices attached to the networks) use some process to allocate resources to individual instances of communication that require those particular resources. The need for MLPP and the specific means of providing MLPP features are both determined by identifying the resource that make up the communication system and deciding on a management process for allocating each one. For this purpose, a resource is any discrete ‘‘thing’’ that exists in limited quantity. Examples include the transmission capacity of individual transmission links of a network, data buffers in processing elements, or connection-related data structures or physical resources that must persist for the life of a connection.

The use of MLPP features requires two kinds of actions to occur, which we describe as signalling and processing. Signalling is the process by which a system component becomes aware of the relative importance (precedence) of a particular demand for use of a resource. Processing is the action taken by a system component based on its knowledge of the precedence of demands made upon it. A correctly functioning signalling path must exist from the user, who chooses the precedence, to each point in the network where precedence-related processing is expected to occur. Therefore, some components may be required to provide signalling functions even though they perform no precedence processing, because they are part of the signalling path from the user to some other component that does perform such processing.

Precedence processing actions are considered to be of three types, with many implementation variations possible for each type. These types are *queueing*, *preemption*, and *blocking*. When precedence processing is required at some point in the system, queueing is usually the preferred approach. Preemption would be used if the expected waiting time of a queueing approach is unacceptably large (what is acceptable depends on the usage context). Blocking or its extreme form, discarding traffic, is needed as a fallback strategy for sustained excess loads in a queueing system, and is a natural side-effect of a preemption system.

PRECEDENCE IMPLEMENTATION DESIGN METHODOLOGY

When designing a communication system or a component of such a system, a five-part design methodology is recommended. The five parts are:

(1) Identification of resources subject to exhaustion

(2) Analysis of expected and required waiting times for resources

(3) Selection of precedence processing mechanisms for management of resources

(4) Selection of protocols to support the precedence signalling needed to control the selected precedence processing mechanisms

(5) Validation of the resulting design

Some of the steps may have to be repeated to take into account the effects of revisions in the design. For example, when the protocols to be used are preselected, the available signalling capability restricts the points at which processing mechanisms can be used.

The waiting time analysis is the basis for selection of processing mechanisms, and in fact it also shows whether MLPP is needed at all. In this analysis, the characteristics of the hardware, software, and protocols are used to formulate qualitative statements about the probability of having to wait more than a certain interval of time to gain some resource. This is described as *expected waiting time*, but it is really a probability distribution rather than a single number. In practice, there is rarely enough information to calculate the exact expected waiting time, but some estimate can be made. This is compared with the user's requirement to gain a resource within a certain time interval in order to provide timely service. If the expected waiting time is longer than the required waiting time, a more powerful precedence processing mechanism is put into the design and the analysis is repeated.

PRECEDENCE CAPABILITIES AND LIMITATIONS OF DATA SERVICES USERS

Section 5 of this report reviews a number of technical topics related to the use of precedence on end user systems. Although this mainly involves implementation and configuration of software on such end systems, one topic of global significance needs to be determined by service providers: the restrictions, if any, on what precedence levels a user may select. The phenomenon of ‘‘precedence creep’’ (users increasing their precedence level to whatever level is necessary to get the service they want) makes it desirable to have some means of preventing users from inappropriately selecting a high precedence level. However, mechanisms to restrict the precedence level that a user can successfully assert may have undesirable side effects, possibly serious enough to make it undesirable to actually prevent a user from selecting any precedence desired. There are several options and some tradeoffs between difficulty of implementation and assurance of control; the broad choices are enforcement that dynamically detects and prevents usage not explicitly authorized, out-of-band procedures that create some penalty for using high precedence (such as cost and accountability), or no restrictions at all.

The main implementation requirements for MLPP on end systems are in the area of signalling. Even if the local system has no resource shortage, if any user of that system needs to communicate over a path that may pass through a precedence processing mechanism, the signalling mechanisms must be implemented correctly in the local system so that the other component can use its precedence processing mechanism correctly.

Within certain applications, it is often desirable to implement some internal processing mechanisms based on precedence. This is notably true of message handling systems. They perform many local

actions which are sometimes time-consuming, so it is possible that some precedence processing may be needed in the message handling software to meet the timeliness requirements of the system's users.

It is also very important that derivative communication actions triggered by an initial communication event use the appropriate precedence level on the derived communications. This has been neglected in the TCP/IP environment; specifically, many interactions involving the Domain Name System (DNS) should be handled on a precedence basis since other applications that use precedence may be dependent on the DNS to enable the application to work.

PRECEDENCE IN INTERNETWORKS

Internetworking technology is now being used widely in DOD and its use is likely to grow still more. The most common application is the interconnection of local area networks with each other and with the DDN. Another rapidly growing application is the construction of router-based networks such as DISA's T1 Router Network. The 1982 DDN Program Plan did not take into account this evolution. As a result, the MLPP services expected to be available to users as a result of the actions outlined in that Program Plan are not necessarily available to users accessing the DDN through routers. Some provisions are made in the IP protocol and also in its OSI counterpart, the Connectionless Network Protocol (CLNP), to allow precedence signalling to work in this environment, but implementation of these capabilities is not mandated either by the specification or by policy guidance. To date, few router implementations have made any provision for using these capabilities. This is partly due to a lack of market demand and partly due to a lack of specific guidance to vendors on what actions such devices should take in response to a precedence or priority indication.

Since internetwork communication is now routinely used, the precedence feature requirements that apply to the original DDN packet switch hardware should logically apply to the internetwork components and attached networks. There may be some cases where it is unnecessary to provide MLPP support, but it should be generally assumed that a shared facility may have users with such a requirement.

In order to assist potential vendors in developing products with MLPP capabilities, implementation guidelines are presented in section 6. The implementation described is considered to be fairly simple to provide in an internetwork router. It is suggested that such requirements should be specified in procurements of routers by DISA and by other DOD organizations. Although there is some cost associated with the initial implementation of these features, it could be very small when amortized over a large installed base.

The precedence of management information is very important in internetworks. This is especially true when network management requires the managed system and the management system to communicate across subnetwork boundaries. It would be very undesirable to create a situation where management control of the network is disrupted by high user traffic loads. Likewise, if routing information exchange within the internetwork environment is disrupted by high user traffic loads, a ''collapse'' situation could develop in which user traffic fails because the routers do not have up-to-date routing information. To prevent problems of this kind, appropriate (usually high) precedence

levels should be used in management and routing communications in the internetwork. Section 6 identifies recommended precedence levels for various types of management and routing traffic.

It may be desirable to monitor and restrict the transit of high precedence traffic between environments that control the use of high precedence (presumably including DOD networks) and those that do not have such controls (typically external networks such as the National Science Foundation Network, NSFNET). This capability would not be needed in most routers, but it may be desired in components such as the DDN Mailbridges. There are some technical problems in building this type of filtering mechanism, but there are promising approaches to mitigating these problems. The appropriate future actions in this area depend on policy decisions regarding DOD needs for this type of filtering.

PRECEDENCE IN THE DDN—REVIEW AND RECOMMENDATIONS

The analysis of precedence service in the DDN components and the recommended actions are combined in this section for convenient reference.

Although a full analysis of the requirements for precedence features in each DDN component (or adequacy of existing features) would be useful to have, carrying out the entire design process presented in this report for every DDN component would have required an increased level of effort which would be beyond the scope of this task. Therefore, this analysis takes the form of a review of the qualitative adequacy and operational effectiveness of the mechanisms currently implemented, without any mathematical analysis.

Packet Switch Node (PSN): A precedence signalling and processing capability has already been implemented in the PSN. The basic design of these PSN features is generaly aligned with the technical approach described in this study. Some deficiencies were noted in three areas: (1) precedence of network management traffic, (2) signalling of precedence in incoming X.25 calls, and (3) lack of precedence-based allocation of X.25 logical channels. The management traffic precedence issue is discussed further below. The PSN software should be changed to resolve (1) and (2). For problem (3), ongoing quantitative analysis of logical channel usage in the network is recommended in order to determine whether there is a risk of logical channel exhaustion. If such a risk is found to exist, the PSN software should be changed to perform X.25 logical channel reservation and preemption.

Terminal Access Controller (TAC) and Network Access Component (NAC): For the TAC and NAC, user requirements for precedence access are somewhat unclear. The TAC has no precedence setting capability. The NAC allows precedence to be configured for a particular physical port, but the user of that port cannot choose precedence on a per-connection basis. This is considered to be a poor design that limits the usefulness of the feature. However, allowing precedence selection by the user raises the policy question of whether some dynamic authorization checking mechanism is considered necessary to restrict users from selecting inappropriate high precedence. The alternative would be to rely on the usage billing system or another similar management product to disclose such usage to the responsible managers.

Therefore, the recommended approach to providing precedence in the NAC is to determine (1) whether precedence selection on other than a semi-permanent basis is desirable to support any user requirements, and (2) whether dynamic verification of a user's authorization to select a precedence level is a required function. If (1) is not required, no further action is needed. If (1) is required but (2) is not, then a NAC software upgrade is necessary to add the user command to select precedence. If both (1) and (2) are necessary, then both the NAC user command and the authorization checking software are needed, which would require changes both to the NAC and to some external authorization system such as the proposed Login Host.

Mailbridges: The existing DDN Mailbridges (or any replacements for them) require a software modification to select appropriate precedence levels for management and control traffic (similar to those required for other DDN components as described below). It is also important to implement mapping of IP precedence to DDN X.25 subnetwork precedence. The necessity for adding precedence queueing at this time is uncertain, but this feature should at least be specified in any new Mailbridge procurement.

Routers: Precedence support should be included in routers used to provide the backbone service, such as the IP and CLNP routers that are expected to gradually take over much of the current MILNET traffic. This support should include (1) use of appropriate mappings between IP and CLNP precedence/priority and subnetwork precedence or priority features, (2) precedence-ordered queueing at entry to the forwarding function and at queues for transmission links, (3) selection of appropriate precedence values for management and routing traffic to and from these components. It should be possible for the network manager to reconfigure the specific precedence values assigned to different management and routing functions, and to alter the mappings of precedence values between layers.

Information Services: The lack of proper handling of precedence in Domain Name System (DNS) implementations is a serious problem which would globally impair the use of precedence by all network users. This problem applies both to DNS software operated by the DDN and similar software used by subscribers. The needed software changes are not technically complex (provided that the underlying system supports precedence as recommended in this report), but they probably will not occur without clear direction and encouragement from the management of the DDN and other networks and systems. A promising approach to alleviating this problem would be to modify the most widespread implementation of DNS software and make the result publicly available.

Network Management and Control: The Monitoring Center (MC), as well as all of the other components discussed above, needs to use appropriate precedence levels on management and routing traffic. A complete and coordinated review of the MC network usage in conjunction with the corresponding functions in the PSNs, TACs, and NACs should be made to develop a list of all functions that pass traffic through the network for monitoring and control purposes. This should be construed broadly to include usage accounting, access control, and other management functions. The appropriate precedence level should be determined for each function, and the software in the corresponding components should be modified to use the levels determined. It is recommended that the precedence levels used be separately configurable for each function in each device so that future changes can be made by management control if the need arises.

OTHER RECOMMENDATIONS

Subscriber network components should provide similar features to those just described for the DDN. This is especially important for routers, since they now occupy a critical role in carrying a large proportion of long-haul data traffic. However, the need for priority features in the internal subnetwork technology must be separately determined for each subnetwork, based on the projected usage environment. Internetworking components should have the capability to use these precedence and priority features even if subnetwork priority is not being used.

End (application) systems should use protocol implementations that include the basic mechanisms needed to support applications that use precedence. The choice of whether to acquire application software with precedence selection capabilities should be based on the expected usage environment. Large-scale acquisitions of generic systems should consider the likelihood that some users of these systems may require or derive benefit from the availability of these capabilities.

The proposals in this report for implementing precedence functions in an OSI environment using the Priority features of OSI protocols need to be reviewed in coordination with appropriate organizations involved in OSI standardization and the decisions made will need to be published.

SECTION 1

INTRODUCTION

1.1 BACKGROUND

The Defense Data Network (DDN) is the primary common-user, long-haul data communications facility of the Department of Defense (DOD). The DDN uses packet-switching technology together with leased terrestrial lines and satellite links to interconnect over 1500 computer systems and local area networks used by DOD services and agencies and defense contractors. These systems and networks are located at hundreds of government and non-government facilities in the United States, its allied nations, and their territories.

The DDN now consists of the Military Network (MILNET), which carries unclassified traffic, and three similar but physically separate Defense Secure Networks (DSNETs) 1, 2, and 3, which carry classified traffic. In the future, the three DSNETs are planned to merge into a single Defense Integrated Secure Network (DISNET). The Defense Information Systems Agency (DISA), formerly known as the Defense Communications Agency (DCA), is responsible for engineering and operation of the DDN.

Since the DDN is intended to support military organizations, it must provide certain levels of service under various adverse conditions to which military systems are exposed. The steps taken to meet these special requirements can be divided into broad categories of security and survivability measures. Security measures such as encryption and physical protections are used to protect network components and communications traffic against hostile actions. Since it is not possible to provide complete security protection against every possible threat, survivability measures are also needed to allow service to be provided under adverse conditions. The DDN Program Plan (DCA, 1982), currently being updated, calls for various security and survivability provisions to meet these needs. Many of these provisions have already been implemented in full or in selected portions of the network. Others are in the process of being implemented, or are awaiting planning or design work.

One of the special military features planned for the DDN is a Multilevel Precedence and Preemption (MLPP) capability. This type of feature has been provided in many military communications systems in the past, and the terminology has been in use for decades. As a result, there is a widespread notion of what it generally means, but it does not have a precise definition that applies to all systems. The reference to Preemption in MLPP is an artifact of older circuit-switched voice systems, in which the nature of the technology provided no other practical way of implementing selective degradation or denial of service. Alternative techniques are feasible in many current technologies, but the term MLPP is often used in a generic sense for any of these techniques, even though it may not be technically precise.

The general notion of MLPP is that all traffic submitted to the communications system is assigned to one of several precedence levels. These levels are administratively predefined and they have an

ordered relationship from highest (most urgent or important) to lowest (least urgent or important). Elements of an MLPP system are expected to make some arrangement to ensure that the service provided to low precedence traffic does not adversely affect the service provided to high precedence traffic. In practice, this means that when the communications system cannot completely satisfy all demands that its users attempt to place upon it, degradation or denial of service ought to be forced upon the low precedence traffic rather than the high precedence traffic.

The formal requirement for precedence features in the DDN exists mainly because the DDN is expected to support the functions of earlier systems that provided precedence features. The most notable of these systems were the Automatic Digital Network (AUTODIN), which is still operational today, and its intended successor, AUTODIN II, which never reached fruition. The AUTODIN II program was terminated in April 1982, and the DDN program was implemented in its place. AUTODIN is essentially a store-and-forward message handling system in which several thousand user terminals are connected to a small number of AUTODIN Switching Centers (ASCs) that perform the store-and-forward function.

Since the AUTODIN precedence system is based on the precedence system used in the Defense Switched Network (DSN) and its earlier stage, the Automatic Voice Network (AUTOVON), an explanation of AUTOVON precedence may be helpful. AUTOVON and DSN provide worldwide long-haul telephony capabilities to DOD units. In these systems, each call is associated with one of five precedence levels (ROUTINE, PRIORITY, IMMEDIATE, FLASH, FLASH OVERRIDE) specified by the Joint Uniform Telephone Communications Precedence System. The maximum precedence level allowed for a given telephone user is administratively predetermined. The appropriate precedence level for a particular call is usually signalled by prepending a special code to the number being dialed, or by pushing a special button on the telephone instrument. Operator assistance is needed to place precedence calls on a line not previously authorized for calls at that precedence level.

In many respects, the AUTODIN precedence system emulates the behavior of the AUTOVON precedence system. Even though AUTODIN is a data network and AUTOVON is a voice network, the use of precedence is similar because AUTODIN processing is essentially message-oriented. Only one message is transmitted on a particular line (either between an AUTODIN terminal and an ASC, or between ASCs) at a time. Therefore, if a low precedence message is being transmitted on a line and a higher precedence message needs to be sent on the same line, either the transmission of the low precedence message must be aborted or the high precedence message must wait for it to be completed. Given the worst allowable case, a 75 baud line and a 40,000 character message, a message could occupy a line for approximately one hour. This is called the *holding time*. When AUTODIN was implemented in the 1960s, low speed lines were very common. With a line speed of 4800 bits per second (bps), more typical of current AUTODIN usage, the holding time would be slightly over one minute. The range of holding times found in AUTODIN is similar to that in AUTOVON for voice calls, and they share the property that it is not possible to have more than one communication in progress at a time on a single line.

1.2 PURPOSE

This document provides technical recommendations for the implementation of precedence and preemption features in DOD data networks, with primary emphasis on the DDN and subscriber networks and systems that connect to the DDN. It includes a technical overview of the objectives of precedence and preemption features in the DOD context, the general implementation mechanisms that are used to implement precedence services, and the technical factors to be considered in planning the implementation of precedence in a network component or user system.

1.3 SCOPE

This document considers the full range of technical issues that arise in trying to design, implement, and use a precedence system for the purpose of providing resource allocation on the basis of application importance as defined by some external authority. This corresponds to what queueing theory describes as a head-of-line queueing discipline. The discussion is generally applicable to any protocol stack, but the specific presentation is naturally framed in terms of Transmission Control Protocol and Internet Protocol (TCP/IP) and Open Systems Interconnection (OSI) protocols since TCP/IP is prevalent on DOD networks now and OSI is the Government standard for the future.

The treatment of general quality of service (QOS) issues is not within the scope of this report. These are more complex problems and the focus of this report is on the simpler problem of ensuring that the available resources are provided to the most important traffic.

1.4 ORGANIZATION

Section 2 discusses the functional objectives that precedence and priority features can help meet, and distinguishes between precedence and other quality of service features.

Section 3 presents an overview of certain concepts employed in the precedence analysis and design methodology, including the notion of the network environment as a collection of resources and the general categories of actions that components can take to produce precedence-based behavior.

Section 4 presents a precedence analysis and design methodology in some detail.

Section 5 describes and analyzes the precedence features and mechanisms that might be useful or necessary in end user equipment, including general-purpose networking software for hosts and specific application implementations. This includes a discussion of interactions between the subscriber and network, such as validation of authority to use a precedence level.

Section 6 describes the issues specific to implementation of precedence in an internetwork environment.

Section 7 describes and analyzes the precedence features and mechanisms used in the DDN backbone, including ancillary functions furnished by or through the network administration.

Section 8 summarizes the current situation with regard to the provision of precedence features to DOD users and describes the steps needed to provide precedence features where they are needed but not currently available.

The appendix to this report discusses the appropriate mapping of precedence signaling and processing features to the seven layers of the OSI Reference Model, and the precedence or priority mechanisms that are available in various data link layer and network layer technologies.

SECTION 2

OBJECTIVES OF A PRECEDENCE/PRIORITY FEATURE

2.1 TERMS DEFINED

The following basic definitions are assumed in this document:

- *Precedence* and *Priority* both refer to an administratively assigned attribute of a communication instance, denoting the relative importance of early processing of the communication. In military communications, *Precedence* is the common term for a standardized designation of importance of timely delivery. *Priority* is the more commonly used term in general computer applications, in queueing theory, and in the specification of communication protocols.

- *Preemption* is the premature suspension or termination of an activity in order to permit some other activity to proceed. Preemption has several variations.

  – *Preemption with resume*, also known as a *preemptive resume discipline*, means that the preempted activity is suspended in such a way that only the uncompleted portion needs to be performed upon resumption.

  – *Preemption with repeat*, or a *preemptive repeat discipline*, requires that the portion of the preempted activity which had been completed at the time of preemption must be repeated when the activity is resumed.

  – *Preemptive repeat with resampling* describes a system in which the amount of work required to complete the preempted activity can change between the time it is preempted and the time it is resumed.

Note that there is no necessary connection between precedence and preemption. Either could exist without the other. Precedence is a characteristic of the information being communicated; preemption is an action taken by some processing agent. Although there are many situations in which precedence indications exist in order to trigger preemption actions, they could be used to trigger other actions instead. Section 3 describes other potentially useful actions that might be used to respond to precedence indications.

2.2 THE PRIMARY OBJECTIVE: ALLOCATION OF OVERSUBSCRIBED RESOURCES

The purpose of precedence, as seen by the user, is to ensure that important communications are carried out in a timely way. In order to formulate and analyze the problem of when and how to implement precedence, though, a more specific objective must be defined. This analysis takes as an axiom the notion that the immediate objective of precedence mechanisms is to deal with the situation

where demand for a communication resource exceeds the supply of that resource available. It is important to realize that this is not a matter of making a communication process happen faster. Digital communication system elements have intrinsic speeds of performing the actions they are designed to perform, and those speeds are determined by design or configuration decisions. Precedence features do not change the speed with which those actions are performed. Precedence features can change the order in which actions are performed, and can suppress some of the offered traffic to bring the resource demand into balance with the resource supply.

In any system with multiple inputs, some process is used to allocate resources to each input. In queueing theory terms, there is an ordered queue of customers for each resource. The customers become ordered in a certain way as a result of the *scheduling discipline* applied to the queue. One of the simplest and most common scheduling disciplines is the first-come-first-served (FCFS) process. When one speaks of a communication system without precedence features, an FCFS scheduling discipline for all resource queues is usually implicit. When a requirement for priority or precedence features is stated, this typically means that FCFS resource allocation is considered unacceptable. It often does not imply any specific notion of what discipline should be used instead of the FCFS discipline. However, it is historically interpreted to mean a discipline in which a fixed number of usage classes are predefined to have an ordered relationship of importance. Within a class (precedence level), the allocation of resources is probabilistic using some scheme such as FCFS, but more important classes always have their resource demands met before resources are made available to less important classes.

2.3 DISTINCTION BETWEEN PRECEDENCE AND OTHER QUALITY OF SERVICE FEATURES

Precedence does not attempt to guarantee that particular quality of service (QOS) parameters are met for a communication application. In other words, the precedence level does not correspond to quantitative metrics such as round trip delay, aggregate throughput, setup time, probability of failure, or any other specific metrics of performance in the communication system.

Given this definition, one might wonder whether precedence is a useful feature to try to build into a communication system. It might be argued that it is more reasonable to manage the quantitative quality of service metrics directly. This argument is well founded in principle; if all of the end user's absolute needs are being met, the relative standing of that user's importance in the resource allocation discipline is of no practical importance to anyone. However, there are two reasons why the relative parameter is often preferred in practice:

(1) To provide quantitative quality of service assurance requires some global knowledge about the entire network environment. This knowledge would have to be gathered and processed in some way to determine whether the requested QOS can be provided. In a very large network environment, this type of calculation may be fairly expensive to perform. In some respects, it isn't clear whether the state of the art is advanced enough yet to perform resource allocation at this level of sophistication in a large and complex environment.

(2) Even assuming that the optimal resource allocations can be calculated, there would have to be enough actual resources to meet the demand. When the available resources are fundamentally inadequate for the demand, some traffic must be removed from the system. Knowing the resources required by each user to do their job doesn't give any guide to which jobs should be removed in this situation. However, the relative metric, precedence, does give information of this kind.

The relative precedence parameter makes it possible to build independent implementations of scheduling disciplines in individual components that can provide assurance of adequate service to important users without requiring any component to have global knowledge of network resource availability and usage. It is not guaranteed that all of the resources in the network will be allocated in the way that would optimally support the user demand to the extent that capacity allows (this would require global knowledge), but most of the effects of non-optimality will fall on the least important classes of users (lowest precedence levels). Later sections of this paper describe the specific precedence handling actions that can be taken in network components to help produce this result.

2.4 CRITERIA FOR EVALUATING NEED FOR PRECEDENCE

From the foregoing discussion it should be evident that precedence features are potentially useful whenever there is a risk that resource demand and supply will be far enough out of balance to make it impossible to meet the QOS requirements of the network users. Conversely, if it is clear that enough resources are available to meet all demands, then there is no reason to implement precedence features.

In the military communications environment, there are at least three interesting types of situations where major excess demand is a realistic possibility:

(1) Traffic surges in a crisis situation. It is well known that large surges in communications activity can reasonably be expected in crisis situations. Since major crisis situations are relatively infrequent, there is some degree of uncertainty about any estimates of the maximum surge level to which a network or network element might be exposed. This suggests that a great deal of surge capacity should be provided. However, cost considerations strongly discourage this. In practice, it is common for systems not to be able to support the maximum possible surge load, either because of intentional economizing or unintentional error in estimating the load when the system was designed.

(2) In the event of actual hostilities, the availability of communications facilities may be reduced by hostile actions such as battle damage or sabotage. This reduces the capacity of the system to handle user demand.

(3) Redeployment of military units to respond to actual or potential acts of the adversary changes the traffic patterns within a network. While it is possible to plan for some deployment scenarios in advance, the progress of events may result in some unplanned deployments and thus produce different resource needs in the communications systems that support the deployed forces.

Besides the special risks of the military environment, the risks of network disruption due to natural disasters or operator errors are present in military as well as civil communication networks.

2.5 SPECIAL CONSIDERATIONS FOR DATA SERVICES

Since precedence features were developed originally for circuit-switched voice networks and message-switched data networks, it is reasonable to ask whether these features are relevant, useful, or cost-effective in current generation data networks, which are based on various forms of packet technology. Some specific concerns are addressed below.

2.5.1 Relevance of Precedence in a Packet-Oriented Network

It has been argued that precedence is not a relevant service in a packet-switched data network. This argument seems to be mainly based on two considerations:

(1) Data input to a packet network tends to be delivered within some time scale which is short compared to the needs of the users. Therefore, there is little point in making distinctions between classes of traffic.

(2) Packet switching eliminates the problem of dedicated resources being tied up by whichever individual user has acquired them because the packetization of data allows the channel to be used by any number of users with an acceptably good illusion of simultaneous support of all of its users.

The technical distinctions between packet switching and circuit switching (such as telephone calls) or message switching (such as AUTODIN) are very much as implied here. However, in both of these arguments, it is assumed that there will be a relatively short delay imposed by the network. This assumption is not valid when a sufficiently large load is imposed on a network, regardless of whether it uses packet switching, circuit switching, or message switching. It is a well known result of queueing theory that queue lengths and waiting times grow without bound as the utilization (imposed load) increases. Dividing the load into smaller pieces changes the probability distribution function of the delays somewhat, but the delays still grow when the utilization increases. At some utilization level, the delay will be unacceptable from the viewpoint of the user. In other words, from the user's viewpoint, there is no practical difference between an unavailable network and a network with an indefinitely long delay. It follows that the two assertions cited above are not true in the general case.

2.5.2 Precedence Processing Overhead

Another objection to precedence features in data services is that the processing actions that implement precedence features consume some of the resources that would otherwise be used to process user traffic. From this, the inference is made that the users would be better served if the network devoted all available resources to handling traffic instead of expending some of them on precedence processing.

This objection has merit, but the inference is too broad. Clearly it is unproductive to implement a precedence feature that consumes more resources than the traffic it drops would have consumed. The tradeoff depends upon the resource consumption of the feature and the amount of excess traffic attempting to use the system. If the amount of excess traffic is large enough, almost any amount of resource consumption by the precedence mechanism is justifiable, since the alternative is total network collapse. On the other hand, if there is significant excess capacity in the network at a given time, the use of resources by the precedence mechanism is not important because it has little effect on users. Between these extreme cases, there is some range of loads in which this tradeoff is worth examining more closely.

The amount of overhead associated with a precedence feature depends on the fundamental nature of the algorithms used and on how frequently the processing algorithms must be used. The frequency of processing is a special concern with connectionless packet processing, which implies that some precedence-related processing is performed on every packet. This reduces the limiting throughput of some types of network elements. It is clearly desirable to minimize the amount of per-packet processing involved in supporting precedence.

2.6 APPLICATION TO INTEGRATED SERVICES

It might also be asked whether precedence is a relevant or useful feature in the context of integrated communications services such as the Integrated Services Digital Network (ISDN). In this context, ‘‘integrated services’’ usually refers to the use of common network components for voice telephony and for data applications, but it can also include other communications applications such as high-quality audio services and video services. Also implicit in this issue is the question of what impact the increasing availability of extremely high capacity transmission facilities will have on the need for precedence features or the way they should be implemented or used.

It would be premature to try to answer this question fully now. However, the general capabilities and limitations already discussed still apply in an integrated environment. Specifically, the need for precedence features is determined by the likelihood of significant imbalances between resource availability and resource demand. The potential for such imbalances needs to be estimated in the context of the special military issues of traffic surge, network disruption, and potential redeployment, described in section 2.4 above. Also, the resources subject to exhaustion need to be identified for this environment. Transmission capacity is only one of these resources; the limitations of switching components and of the information generating and processing systems attached to the network also need to be considered.

This report does not attempt to analyze the integrated services environment in depth, but several observations can be made. The notable features of this environment are a more diverse range of applications competing for the same resources, and the availability of many high bandwidth transmission links. This implies that more of the resource management activities will need to be focused in the internal resources of the switching components and in the devices that provide value-added information processing. Also, more attention will need to be given to the effects of a diversity of applications on the behavior of the network. The relative treatment of communications needs to be

rationalized not only for different users of the same application, but for users of different applications. Finally, it must be remembered that the information-generating components attached to networks are built with the same technologies used to build the networks themselves. It follows that the ability to generate demand grows as fast as the ability to support that demand. Even with enormous bandwidth available, it is never certain that a demand will not arise to consume it—and when there are many ‘‘demanders,’’ to exceed it.

SECTION 3

PRECEDENCE SERVICE CONCEPTS AND TECHNIQUES

This section introduces several concepts useful in determining whether and how to implement precedence in communications system elements. These concepts will be used in section 4, where an analysis process is defined in more depth. However, a brief overview of the design analysis process may help explain why these particular concepts are useful.

The entire communications environment, including the network and the communicating devices attached to it, is viewed as a collection of finite resources. (There is no need to be concerned about infinite resources, so they are just ignored.) Users attempt to use the network to support many individual communication instances. Each instance has some precedence level associated with it (at least conceptually). In order for the communications instance to be supported, some of the resources of the network must be allocated to that instance. The designer makes some determination of which resources are potentially subject to exhaustion, given the set of communication instances that may require use of each resource. For each such resource, the designer estimates the likely impact of exhaustion and chooses some precedence processing mechanism to be used to manage that resource. The designer also considers how the component managing the resource becomes aware of a claim for use of that resource and tries to identify a way of associating the precedence level with that claiming event (putting it another way, the precedence level must be signalled explicitly or implicitly to the point where the allocation is made, so the precedence processing mechanism there can do the right thing).

3.1 IDENTIFICATION OF PRECEDENCE-RELATED RESOURCES

The basic requirement for a precedence processing feature is that it should ensure that a higher precedence communications instance is not prevented from taking place by a lower precedence communications instance. Successful execution of a communications instance requires that a particular set of resources be made available within some time constraints.

Since the precedence mechanism is responsible for assuring availability of resources to higher precedence instances, it is necessary to identify each of the resources that are needed to support the communication instances expected to occur in a network. In practice, it is not important to separately define everything that might be considered a resource. It is only necessary to determine the resources that are potentially subject to being exhausted in the course of network operation. Examples of these resources include:

- Transmission capacity on a link
- Switching capacity in a switching element

- Buffer space in a processing device (for data in transit)
- Memory space for protocol state information (connection state blocks)
- Processor cycles in a processing element
- Identification elements (sequence numbers, logical channel numbers, etc.)

The list of interesting resources will differ from one type of component to another, and it may differ even in similar components (for example, packet switches made by different vendors may use different implementation approaches that use different internal resources).

3.2 PRECEDENCE SIGNALLING AND PRECEDENCE PROCESSING

Two different types of actions must take place in a communication system that provides precedence features to users. We describe these as signalling actions and processing actions.

3.2.1 Signalling Mechanisms

Section 2.1 stated that precedence is an administratively assigned attribute of a communication instance. The network components that are expected to respond to this attribute need to know the precedence that was administratively assigned to the communication instance that generated each unit of work requested of them. The mechanisms used to convey this information from the point of assignment to the processing mechanisms are called signalling mechanisms.

Since one of the main functions of data communication protocols is to convey structured information from place to place, it is logical that most signalling mechanisms are embedded in implementations of protocols. (An important exception will be described below.) Furthermore, since layered protocol architectures are now the norm, the signalling scheme as a whole needs to fit into the layered protocol concept in which entities at a given layer (normally) communicate only with the next higher and next lower layers in the local system, and (in effect) with the peer entity at the same layer in other systems. The practical result (described in terms of the OSI Reference Model) is that the precedence information starts out at the application layer and is passed down to lower layers in the same component, each of which passes the precedence information to its peer entity at the same layer.

In reality, not all protocols provide ways to communicate precedence information between peer entities. This is only a problem if the peer entity is expected to perform a processing action (at that layer), or if it must signal the precedence value to some other entity that will perform a processing action. In principle, it would be best to first determine the user requirements, then choose the precedence processing actions for all the components to best meet the requirements, and finally choose (or develop) protocols with the needed precedence signalling capabilities. In practice, the choice of protocols is often constrained either by external requirements such as interoperability needs, or pragmatic concerns such as a desire to save money by using commercial off-the-shelf (COTS) products. Whatever the reason may be, if there isn’t a signalling path from the point of precedence

assignment to a given point of precedence processing, the precedence processing at that point will not be able to function.

One essential signalling mechanism is not normally embedded in a protocol. This is the mechanism that makes the administratively determined precedence value for a particular communication instance known to the communication system. This mechanism is usually part of the application software used to initiate the activity. Several implementation approaches are possible, such as:

- **User input:** A typical implementation requires the human user to specify the precedence level when the activity is started. Often there is a precedence level assigned by default if the user doesn't explicitly select the precedence level.

- **Fixed value:** In some applications, a fixed precedence value can be predetermined, either separately for each user in some type of database, or uniformly for all users.

- **Derivation from context:** The application may determine the precedence automatically from other information available to it. This is a useful approach for applications in which one communication gives rise to others; the precedence of the initial activity can be applied to the derived activities. Connectionless application servers are likely to use this approach.

3.2.2 Processing Mechanisms: Queueing, Preemption, Blocking

Once the resources have been identified, some method must be developed for allocating (i.e., scheduling) the resource for *customers* (claimants for that specific resource), taking into account a precedence value associated with the customer. Although many different mechanisms can be devised, they fall into three broad categories of procedures: queueing, preemption, and blocking.

3.2.2.1 Queueing Procedures

The most widely useful technique for precedence-related resource allocation is to maintain a precedence-ordered queue for the resource. This means that whenever the resource is available for use, the highest precedence waiting customer acquires the right to use the resource. There are variant techniques for implementing these queues; there may be a single queue into which new customers are added at the point corresponding to that customer's precedence, or there may be one queue per precedence level and processing logic to check the queues in descending order of precedence. The relative efficiency of these two methods depends on whether the number of waiting customers is typically small or large compared to the number of precedence levels defined.

In a strict precedence scheduling discipline, the effect of queueing is to reduce delays for high precedence customers by increasing them for low precedence customers. This necessarily increases the variance of the delay for low precedence requests as well.

Queueing is known to be an inadequate mechanism in two situations:

- If the expected holding time for the resource is longer than the customer can afford to wait,

no benefit can be obtained by queueing. Therefore, some type of preemption or blocking action is needed. This is the usual motivation for implementing preemption mechanisms.

- Queueing necessarily involves some resource consumption in itself. If these resources are unavailable, then some type of blocking is needed. Note the potential recursion in resource allocation: the memory used to hold queued work may itself be allocated in a precedence-dependent way, and this clearly has to use preemption or blocking when the space runs out, since by definition there is no room to queue it.

At a more detailed level, a designer may find that queueing demands for certain resources in a certain component is undesirable because overall performance is better if that component blocks the activity, causing the whole activity to be queued elsewhere in the network. A typical compromise is to restrict the length of the queue to some value thought to be just large enough to ensure that the resource doesn't become idle because its next user is blocked upstream.

3.2.2.2 Preemption Procedures

In section 2.1, three variations of preemption were defined. In each variation, the general notion is to acquire a resource by terminating the use of the resource by some other user. The preemptive resume discipline is the best behaved in theory, but the majority of communications protocols are not capable of supporting it (they require a packet header contiguous with all of the packet body). Therefore, most preemption implementations will use a preemptive repeat discipline. This causes some additional consumption of network resources because the initial portion of the preempted packet is sent more than once. Since this repetition decreases the effective net throughput of the network, it is desirable to avoid using preemptive resume schemes.

Even though there are problems with both of the possible ways of doing preemption, preemption is still sometimes necessary. Section 4.3.1 below discusses how to determine whether such a need exists.

Preemption of a communications transmission is only practical when there is some way of indicating to the receiving party that preemption has occurred. The form of this indication (if it exists) is protocol-dependent. A connection-oriented protocol is almost certain to have a construct for closing a connection; for example, in X.25, a Clear Request or Clear Indication packet would be sent. Connectionless protocols (a category that implicitly includes most physical layer protocols used for data signalling) obviously do not have a similar notion. Either type of protocol may or may not have a means of prematurely terminating a data unit. For example, High-level Data Link Control (HDLC) protocol allows a frame to be aborted at any time by transmitting seven consecutive 1-bits.

3.2.2.3 Blocking Procedures

In any finite system, there is a possibility that the demand for a resource will be so great that queueing is inadequate (the queue space is all used up, or the expected waiting time is so large that queueing is useless). Therefore, all precedence handling schemes must resort to some type of blocking procedure to deal with work that can be neither processed nor queued. If there is a way to

explicitly refuse the request, it would normally be used in this situation. This allows the upstream process or component to queue the work, if it has that capability and decides to use it. If the work cannot be explicitly refused, the only option left is to discard it (and hope that it will be regenerated upstream at a more convenient time).

Blocking is also useful to deal with a phenomenon known as the ‘‘repeated attempts problem.’’ When a communications process fails to complete as the user expected, it is likely that the user will immediately and repeatedly attempt to repeat the operation. In an excess demand situation, this is not desirable because some operating resources are used up attempting to respond to each attempted repetition. It is desirable to block unsupportable traffic as close to its point of origin as possible, so that downstream components do not dissipate resources on unsuccessful repeated attempts. Of course, it is only desirable to block the repeated attempts that will fail, and not the ones that will succeed.

SECTION 4

METHODOLOGY FOR PRECEDENCE FEATURE DESIGN

This section presents a design methodology for analysis and development of precedence features in a commuication system. Its major steps are as follows:

- Identify resources potentially subject to exhaustion
- Compute and compare required and expected availability for each resource
- Select processing mechanisms separately for each resource
- Determine signalling requirements
- Validate results

When the implementation options are constrained, the analysis may need to be performed repeatedly to converge to the best solution possible under the given constraints. For example, if the design initially developed requires precedence signalling to a point that can only be reached by protocols that don't provide precedence signalling and can't be modified to provide it, the designer should try to throttle the demand at some point upstream. This might introduce new signalling requirements or change the type of processing mechanism that should be used elsewhere in the network.

4.1 IDENTIFYING RESOURCES

Section 3.1 above introduced the notion of a network as a collection of resources and explained that it is necessary to identify only those resources in the communications environment that are potentially subject to exhaustion under a traffic load thought to be achievable in the network. These will generally fall into two categories: transmission link capacity and processing component resources. Since precedence features are typically intended for use in situations where unpredicted large utilization is a risk, it is probably best to begin the analysis with an extremely conservative set of assumptions.

Specifically, one can analyze the processing components by assuming that all of the links in and out of the component are operating at full capacity with whatever type of traffic load is thought to place the most stress on the processing components. Conversely, the link analysis can be predicated on attempted full capacity utilization of all access points to the link. In the case of a simple full-duplex point-to-point link, these assumptions are equivalent. They are not equivalent when there are multiple transmitters on a single link, because this configuration allows link congestion to occur even when all transmitters are operating below the rated capacity of their individual interfaces. A half-duplex radio or wireline connection exhibits the same property.

4.1.1 Processing Component Resource Analysis

A component will have some set of transmission units that it recognizes (connections, packets, or both). For a continuous connection (as in a conventional telephone switch), some physical connection path is presumably allocated to each connection, and the number of such paths through a component is a known quantity constrained by hardware design. For a packet handling design, there will be some type of memory buffer allocated for each packet.

If virtual connections are involved (and their existence is known to this particular component), there will also be some memory resource used to retain the connection state information. It should be expected that any nontrivial protocol implemented in the component will have some memory demands for state information.

Within the software that implements a protocol, additional resource exhaustion possibilities may be caused by limited sequence number spaces or by consistency mechanisms. For example, the IP protocol mechanism for reassembly of fragmented datagrams relies on the uniqueness of datagram identification (ID) numbers in all unfragmented datagrams in transit from a given source to a given destination. A 16-bit field is provided for this value, so there can be at most 65,536 uniquely identified datagrams in transit from source to destination. This sets a limit on how rapidly IP datagrams may be injected into a network. Thus, datagram ID values may be thought of as a resource that could be exhausted. The designer should determine for each IP datagram source either that it is definitely not capable of sending more than 65,536 datagrams within the lifetime of an IP datagram in the network, or that some precaution is taken to give the highest precedence traffic preference in acquiring ID values.

Consistency mechanisms necessarily imply a resource that could be exhausted. An example of such a mechanism is a semaphore-based lock used to prevent more than one process from modifying a data structure, or to prevent it from being read while it is being updated. Such a lock might be used to prevent any routing decisions from being made while a new set of routes is being calculated. The designer should determine whether the total demand for the routing table will become a constraint on system performance; if so, it is a resource subject to exhaustion and it may be necessary to prioritize access to it.

An active switching component usually has many functions which are carried out under the control of a central processing unit (CPU). The CPU has an inherent speed of operation and therefore a limited capacity to do work. Therefore, it is a resource that could be exhausted. It has long been recognized by designers of multitasking computer systems that some enforced prioritization of internal operations is necessary if the system is to service its own internal operations in a timely way, and there is normally some combination of hardware and software features to do this. Further prioritization in communication processors based on precedence is usually implemented in software only, and works by sub-allocating the resources allocated to a task. In other words, the design of the basic computing system may result in a certain amount of CPU time being available for interpreting and acting upon protocol constructs contained in incoming packets. The precedence mechanism (if any) doesn’t change the proportion of CPU time used for packet processing; it might, however, change the order of processing packets or the type of processing done on them.

4.1.2 Transmission Link Resource Analysis

The capacity of a transmission link is necessarily limited, and the limit is usually a well known, fixed data rate. This simplifies the analysis greatly, since it is only necessary to determine whether the traffic sources can demand a higher rate than the link provides. The answer is usually yes, so it is normal to consider links as resources subject to exhaustion.

The designer should keep in mind that some links may exhibit variable rate properties. For example, the link may use modems that select a data rate based on the measured error rate of the media. The analysis should take into account the possibility that the modems will use a lower rate and thus make the link more likely to be overloaded.

4.1.3 Coupled Resource Allocations

Different resources may be coupled in the sense that the usage or non-usage of one resource implies the usage or non-usage of another. In situations where the resources have a size or duration associated with them, the coupling may be completely predictable (such as every X.25 logical channel requiring a fixed amount of memory for its protocol state storage) or only loosely predictable (such as the number of buffers holding unacknowledged data). When couplings exist and are well enough understood, it is possible to save some resource management effort by only considering precedence when allocating the ones that might be exhausted first. This is especially relevant when preemption is possible. Given a choice, it is desirable to preempt at the highest possible layer so that all of the related lower-layer resources are released. However, if the lower layer resources are more likely to be exhausted, the next best option is to have the lower layer preemption communicated to the higher layers so that they can shut down also.

4.2 ANALYZING EXPECTED AND REQUIRED WAITING TIMES

The expected waiting time for a resource depends on the demand for the resource, the average holding time (also called service time), the number of servers, and the scheduling discipline. The demand and holding time parameters are probability distribution functions, the number of servers is a simple count, and the scheduling discipline is an algorithm (also capable of being expressed as a mathematical function or functions). Likewise, the holding time is a probability distribution of times. The number of servers is a simple count. The scheduling discipline is an algorithm chosen by the designer and can be described by a mathematical function or set of functions. Each of these is discussed further below.

The acceptable waiting time depends on the required application performance, the cumulative waiting time for the entire application activity, and indirectly on the serial or parallel nature of various resource requirements. Just what delay is acceptable is highly application-dependent. The usual analytic approach is to begin with requirements defined in an external context (e.g., an assertion that a certain application should be able to be performed in 30 seconds), and then break this down into the lower level operations that are used in the application (number of connections, number of packets, size of packets, etc.). From this information, some approximate requirements can be calculated for

the individual resources such as link bandwidth, switching capacity, etc. The total delay can often be divided up in more than one way—using a faster component of one type may make it possible to use a slower component of another type, and so forth. These choices are usually made on pragmatic grounds such as the expected cost and availability of different kinds of resources.

Ideally, the expected and acceptable waiting times would be expressed and compared in terms of probabilities.

4.2.1 Demand Estimation (Arrival Time Distribution)

Demand is normally thought of as a process in which *customers* arrive at various times, so a high demand implies a shorter time between arrivals. Demand is typically expressed as a probability distribution of time between one arrival and the next arrival. The term *customer* is a conventional term in queueing theory and does not imply anything about the type of entity that arrives; in data networks, the customers are often packets, messages, or virtual circuit instances.

Absent any known constraint on the distribution of arrival times, it is customary to treat the absolute arrival times of customers as independent events. This is known as a Markov process and the resulting distribution of interarrival times is known as an exponential distribution. The exponential distribution has a single parameter which controls the absolute scale of the distribution; this parameter is normally expressed as the arithmetic mean of interarrival times. (The upshot of this relationship is that if the customers are known to be independent and the average interarrival time is known, perhaps from a measurement of a comparable system in operation, various probabilities of interest such as the probability of k customers arriving in a given time interval $[t_0, t_1]$ can be calculated to any precision desired using known statistical formulas. This is useful for determining whether the performance of a system is acceptable as judged by a probability-based criterion.)

Unfortunately it is common that the arrivals are not completely independent of each other. Clustering of arrivals is a common artifact of window-based flow control schemes such as those in X.25, TCP, and HDLC. Sometimes it is safe to ignore the effect of clustering, but it is difficult to generalize about when this is safe. The practical solution is to measure the behavior in a similar network to derive an observed mean and variance. A detailed statistical discussion is beyond the scope of this document; it is enough to say that if the designer doesn't have confidence in the demand assumptions and the mathematics used to apply them, an experiment should be performed to see whether the assumptions reflect reality.

Computer simulations are of some help in this type of analysis, but it is important for the simulation to use the same algorithms that would be used in the actual components. In a simulation, it is still necessary to make some assumptions about the demand on the system, but the demand may be input in some form which is easier to validate. For example, it may be easier to determine the distribution of attempts of human users to send electronic mail messages than the distribution of the network layer packets used to carry those messages. If the correct algorithms are used in the simulation, the appropriate clustering of packets (or lack of clustering) will be produced by the simulation software.

4.2.2 Holding Time Estimation

The holding time of a resource (known to theoreticians as the service time) is also typically probabilistic. In some cases, it may be a fixed quantity because an external constraint forces it to be fixed. This type of behavior can be observed in time division multiple access (TDMA) systems. At the other extreme, the holding time may be dependent on human actions. An example is the holding time for physical path resources used for a telephone call; the resources will be held until one of the parties frees them by hanging up their telephone. Between these extremes are cases in which the holding time is variable within certain limits, or is distributed according to some known (or hypothesized) probability distribution. The lengths of packets in packet data networks are often variable within known limits (they obviously cannot be less than zero, and it is common to have a maximum packet size).

As with the arrival time distribution, an exponential distribution of holding times can be assumed when there is independence of customers. Once again, the validity of this assumption is questionable. One source of correlation is the tendency of the human end users to each be performing one of a limited set of applications, each of which has a distinctive pattern of network usage. Fortunately, some pragmatic considerations simplify the estimation of holding time. Many of the holding times in packet network components are dependent on packet lengths. Since the packet size is often constrained by the protocol being used, the maximum holding time for a single packet can often be calculated directly. This is necessarily an upper bound on the actual holding time for a particular packet, and it can be used as a conservative estimate of the average holding time.

For virtual circuit resources, the holding time estimation problem is similar to that for telephone circuits. The exponential distribution is probably not a good assumption because human behaviors, although they are unpredictable, are not random. It is more likely that there will be some set of functions for which the network is used, each of which has an independent (possibly exponential) distribution of holding times. Therefore it is probably better either to measure holding times in real networks performing comparable functions in a comparable way, or to try to determine the specific functions for which the network will be used and the likely holding times for each.

4.2.3 Number of Servers

The notion of servers is perhaps best explained by an example. Suppose that a telephone system consists of two switches, each with many customers, and one transmission cable between the switches. If this cable can carry only one conversation, then the resource has one server. If it can carry 1000 conversations, the resource has 1000 servers.

The number of servers has a major effect on the expected waiting time for a resource. If there is only one server, then the expected waiting time is very much dependent on the expected holding time of the previous user of the resource. If there are many servers, not only is it more likely that one of the busy servers is near the end of its holding time (about to become free), but if there is substantial variation in holding times, it is even more likely that a server will be free soon. The relevance of this will be even more apparent when we consider the analysis of whether to include a preemption mechanism in a component.

4.2.4 Scheduling Disciplines

By definition, a scheduling discipline exists for any shared resource. The choice of discipline is entirely within the control of the component designer. If no conscious choice is made, the resulting implementation is likely to be one of three simple disciplines: (1) FCFS, (2) last-come-first-served (LCFS), or (3) random service (RS). If a queueing mechanism is used to implement precedence, a more complex discipline will result. It is usual to begin by predicting the performance of a simple discipline, and then progress to other disciplines involving precedence or other prioritization schemes.

Other disciplines have been developed to optimally handle certain cases. The Shortest Job First (SJF) discipline has the property of producing the lowest average delay for all traffic taken as a whole (provided that the holding time for each customer is known upon arrival, so that the ‘‘shortest job’’ can be determined). A simple SJF discipline doesn’t have any notion of precedence, but the SJF procedure could be used separately within each precedence level in a head-of-line (HOL) system. Section 4.3.2 below discusses the issues to consider in choosing a scheduling discipline.

4.3 SELECTING PRECEDENCE PROCESSING MECHANISMS

4.3.1 Evaluating the Need for Preemption

Preemption is chiefly useful when the activity being considered ties up a scarce resource for a long period of time. In such a situation, it may be impossible to meet the needs of a higher precedence user if a lower precedence user has control of the resource and can’t be preempted. This situation arises in voice telephony and is the main motivation for preemption in voice systems.

The need for preemption is dependent upon the expected holding time of the resource. If a free resource can be expected to be available in a short time, there is little gain from preemption. Therefore, the designer should try to determine what the expected holding time will be for each resource in the system. Usually the holding times will be relatively large for those resources that exist throughout the life of an application instance, such as the stored state information for a connection-oriented protocol. Therefore, preemption is most often used to manage connections or connection-like resources.

Sometimes it is possible to influence the expected holding time by changing the design or configuration of the system. For example, the holding time for a data transmission link is the time required to transmit one data link service data unit (often called a frame). If the data link layer protocol allows a choice of frame size without significantly affecting other layers, it is probably reasonable to use smaller frame sizes on slower data lines, which reduces the expected holding time of the data link.

4.3.2 Selecting Scheduling Disciplines

For the first attempt at analysis, the FCFS discipline (without precedence) is normally assumed. If FCFS works well enough to provide the required performance, then it is not necessary to explore precedence processing strategies. In some applications, though, any performance enhancement is considered worthwhile. This is a matter of judgment.

If a precedence queueing feature is being considered, the usual first choice is the head-of-line scheduling discipline. This provides the equivalent of separate FCFS queues per precedence level. The HOL scheme is traditionally used in military systems because it guarantees that a higher precedence customer will always be served before a lower precedence customer.

Other priority disciplines provide better performance for certain types of situations, but in their usual forms, they do not guarantee absolute precedence ordering. The SJF discipline is optimal (in the sense of minimizing the average waiting time of all customers for that resource) when each customer's holding time is exactly known upon arrival, but SJF must be modified to maintain the guarantee of absolute precedence ordering. The guarantee is maintained when SJF or any other discipline is used only to reorder customers of equal precedence level.

Although there seems to be a good argument in theory for using SJF scheduling within each precedence level and HOL scheduling of the levels as a whole, there is additional work involved in carrying out a more complicated scheduling algorithm. This increases the consumption of one type of resource, computation effort, to try to more efficiently use other resources. This is a bad idea if the computing capacity is scarcer than the resources being managed by the computing processor.

There has been extensive research in optimal scheduling of tasks when there are known deadlines for completion of the tasks or specified portions of the tasks. These algorithms tend not to be useful in the general-purpose data network components because the tasks performed by the individual components are usually very small portions of the large task whose deadlines are known, and the decomposition of the large task deadline into small task deadlines seems to require either a global optimization throughout the network (which requires gathering a great deal of information and performing a somewhat expensive series of calculations) or signalling of static deadlines in protocol data units (not a feature of most protocols, and probably requires a synchronized time base in network components). It may be possible to use these algorithms effectively in application level scheduling, such as message transmission scheduling in a message transfer agent, but the use of these algorithms in lower layer entities is a research topic and not a practical implementation strategy at this time.

Other considerations related to SJF-type disciplines include:

(1) SJF and many similar disciplines produce more benefits when there is a very large dispersion of service times (e.g., a wide range of packet sizes), and no benefit when service times are identical.

(2) When applied to datagrams, many dynamic priority disciplines (including SJF and schedule-to-deadline schemes) tend to reorder datagrams in transit. Although this is

acceptable by definition for a datagram service, it has some effect on the performance of connection-oriented services built on top of the datagram service. This effect is partially mitigated by the tendency of bulk transfers to use equal sized packets for most of the transfer, thus preventing those packets from being shuffled by the scheduling process.

(3) SJF tends to discriminate between applications and between implementations. Some applications such as remote terminal access naturally tend to use smaller packet sizes. Also, different implementations of the same application may use different packet sizes for implementation-specific reasons. There could also be discrimination on the basis of smallest maximum transmission unit (MTU) size in the complete path; a host connected to a local network with a small MTU would receive quicker service in distant networks than hosts accessing the same distant network through a local network with a larger MTU. It is also imaginable that some users might force the use of smaller packet sizes in order to gain preferential treatment. The existence of an overriding HOL scheme that enforces precedence ordering mitigates these problems but does not completely eliminate them.

4.3.3 Throughput Allocation and Preemption

An interesting problem with many data applications is that at certain times, the instantaneous throughput desired is potentially unbounded. Consider a file transfer application. Although there is some preliminary protocol interaction to set up the parameters of the transfer operation, the bulk of the resource consumption all occurs in one compact burst when the contents of the file are sent through the network. From the user's point of view, it would be desirable for the file transfer to be finished instantaneously, or as nearly so as possible. This implies using the maximum throughput available on the path for whatever time is required to transfer the file. If this transmission has been assigned some precedence level other than the lowest level, this would seem to imply totally blocking out all lower precedence traffic from all the links involved.

In practice, this behavior is rarely desired. It would be appropriate for other traffic to be blocked out if the high precedence file transfer cannot be completed in an acceptable elapsed time otherwise, but the file transfer user usually doesn't have such a stringent requirement. It may be useful to think of a network with precedence features as providing each precedence level with one of three levels of service: good service, poor service, or no service. At a given time, one or at most two precedence levels will be experiencing poor service. Higher levels will receive good service and lower levels will receive no service (or service so poor that it is effectively valueless). Each user hopes to receive good service, but there is relatively little value attached to receiving greater service than whatever is considered good service. Therefore, it is desired to limit the maximum instantaneous resource consumption of each user even for high precedence users.

The problem is to determine an appropriate limiting throughput value and to implement it. This problem decomposes into signalling and processing problems resembling the original precedence problem (except that the processing part is somewhat simpler since it is, by definition, only a blocking problem). Some protocols provide means of selecting the target throughput value for a connection, but available implementations of these protocols rarely act on it, and in any case, most applications do not specify this value. Therefore, some default target throughput value needs to be

assumed and applied in those cases. It might be reasonable to determine this value on a per-subnetwork basis. Implementation would probably be based on manipulation of flow control mechanisms, and thus is relatively straightforward when a connection-oriented protocol is being regulated and somewhat more problematic for connectionless protocols. (This suggests two inferences: first, applications that will send large volumes of data probably should not use pure connectionless protocol stacks; second, it makes sense to implement the throughput throttle in a layer where a connection-oriented protocol is being used.)

When extra flow control is applied to a connection as part of a precedence mechanism, there should be some consideration of the minimum throughput necessary to provide a useful service. If the minimum cannot be met, it may be more efficient to preempt (terminate and block) that connection. As with target throughput, the minimum acceptable throughput logically ought to be determined by the user but usually is assumed by default. There is little practical experience in this area, so any implementations should be designed thoughtfully and carefully validated.

4.4 SIGNALLING REQUIREMENTS: PROTOCOL AND INTERFACE IMPACTS

Once a set of precedence processing mechanisms has been chosen, the means of signalling the precedence level to the point of processing must be determined. For our analysis we employ a strict layering model. This model divides the total processing activity of a component into portions associated with specific layers. The approach is the same regardless of the specific layer reference model being used; it is assumed that each layer entity communicates with higher layer and lower layer entities in the same component through well-defined service interfaces, and with peer (same layer) entities in other components using a well-defined protocol. There may also be local management information (e.g., configuration settings) associated with the layer entity; for this analysis, it can be assumed that management information is static, and how it is acquired is irrelevant.

4.4.1 Protocol and Service Interface Requirements

The analysis task, then, is to determine which service interfaces must convey a precedence value, which parameters at that interface are used to convey it, the mappings between the protocol constructs used by the layer and the service interface parameters, and any management information needed to support local processing or signalling needs. When existing protocols and interfaces are to be used (the most common situation), the analysis should take into account not only the service interface and protocol standards, but also the extent to which actual implementations support the needed features of those standards.

4.4.2 Mapping Precedence Value Encodings Between Layers

Since different protocols and services don’t necessarily use the same representations of precedence levels, the mapping between all of the representations used in a layer must be specified. For example, the International Organization for Standardization (ISO) 8072 connection-oriented transport protocol (COTP or just TP) has a 16-bit integer priority parameter with smaller numerical values representing higher priority (precedence) than larger values, but the ISO 8473 connectionless-mode network

protocol (CLNP) has only four bits (expandable to eight) to represent priority, and ISO 8208 connection-oriented network protocol (CONP) based on X.25 has six different priority parameters, each of four bits potentially expandable to eight. Therefore, transport entities that use COTP will need to map between the 16-bit values used within the transport layer and the 4-bit values used at the network service interface.

It is normally desired that direct peers at a given layer use the same mappings to the next higher layer's representation. However, technology-specific characteristics of lower layer media or different absolute timing characteristics caused by a different workload mix on a particular subnetwork make it more likely that different systems or networks would use different mappings to lower layer representations. It is obviously impossible to make the mappings on a network both consistent and invertible if different downward mappings are used. This suggests that the protocol stack yields the most functionality when peer-to-peer signalling occurs at those layers where downward mappings may differ.

4.5 DESIGN VALIDATION

When a complete design has been defined, it should be reviewed for consistency and correctness. It often happens in practice that the initial design parameters cannot be fully reconciled with each other. Such conflicts should be detected during the validation phase, and the design parameters should be adjusted to resolve the discrepancy. This may make it necessary to re-analyze portions of the design. The cycle of design and validation is repeated until a satisfactory result is reached.

Discrepancies most often result from restrictions on the hardware, software, or protocols that can be employed in the network. These restrictions might be based on pragmatic concerns such as cost, policy concerns such as security, or on requirements for interoperation with systems that are outside of the control of the designer. For example, the analytical design process may determine that certain preemption events should occur in a device that must communicate with its peers using a protocol that doesn't provide a precedence signalling mechanism. Either the protocols (and implementations) must be changed to provide the necessary precedence indication, or the preemption function must be removed from the device. If the preemption function must be removed, the behavior of the network changes, and re-analysis may show that some other component should be changed to cope with the network's behavior.

The validation process should encompass both static and dynamic aspects of the design. Some appropriate validation activities are described below.

4.5.1 Static Validation of Assumptions

Aspects of a precedence service design which can be statically validated include:

- Assumed network configuration and (stressed) load

- Performance criteria
- Continuity of precedence signalling
- Appropriate selection of entities for preemption

The network configuration and maximum load assumptions should be documented. Since the analysis process often brings to light additional knowledge, the initial assumptions should be reviewed when a design has been prepared, in order to verify that they are still considered appropriate design goals.

Any performance objectives established for the design should also be compared with the predicted performance of the system as designed. This is the point at which one considers whether some feature of the design seems to add complexity or cost out of proportion to its benefits. If this seems to be the case, it may be appropriate to revise the performance objectives.

At a technical level, it should also be possible to review the precedence processing mechanisms postulated by the design and the protocols to be used in the system to ensure that there is a signalling path from the point where precedence is administratively determined to each of the points where precedence-based processing occurs. The paths should result in the precedence value for outgoing events being passed from the source application downward through its layer service interfaces to the lowest layer where precedence processing will occur within the source component or its peer entity at that layer. Conversely, the receiving component should pass the signalled values to its higher layers. Components that perform relaying functions (such as packet switches, routers, bridges, and gateways at various layers) should retain the precedence value through the process that operates between the incoming and outgoing protocol stacks.

If the system employs preemption, each point of preemption processing should be checked to ensure that its preemption decision is based on valid information that a particular preemption will actually free the specific resource for use by a higher precedence customer, and will preempt the lowest precedence customer that holds the specific resource. Some designs do not follow this rule; they need to be analyzed dynamically via simulation or actual operation. In either case, the objective of this step is to verify that preemption actually moves the scarce resource from lower precedence customers to higher precedence customers.

4.5.2 Dynamic Validation of Mechanisms

Dynamic validation can be attempted in two senses. One is to validate the performance of the mechanisms selected for a particular component. The other is to validate the operational performance of the components when assembled into a network. Either of these may be carried out using simulation or using actual operational components.

Consider first the validation of individual component designs. Simulation carries the risk that the simulated component or network may not behave the same as the real component or network would behave. Discrepancies can result from simplifying assumptions made in the simulation or from errors

in constructing the simulation. The simulation is not necessarily more subject to implementation flaws than the real component, but when unexpected results arise in a simulation, there is some uncertainty as to whether the flaw is in the design being simulated or in the way it was simulated. As long as the effort is essentially theoretical, this is not a problem. Simulations can be used to compare alternative designs and a later, real implementation can be made to conform to the behavior of the simulation. However, only a real test of the actual component can verify that the design in the component actually performs as the simulation did.

The second type of dynamic validation, network-wide performance validation, tends to be neglected but is at least as important as the first. The local optimization of resource usage in components does not guarantee global optimization in entire networks. In other words, there are some designs in which each individual component manages its local resources in a way that appears to be reasonable, but the network as a whole does not perform as much work as it could. A globally optimized network should not be expected, but it should be established that the actions of individual components are not producing gross inefficiencies.

Undesirable global effects can be induced either by precedence signalling flaws or by side effects of precedence processing mechanisms. Signalling problems are not difficult to recognize in the design process, but the side effects of processing actions are sometimes difficult for the designer to anticipate. This is especially true when preemption is employed. Although preemption is normally done to free a single resource that is known to be critical at the time, the preemption is likely to also free other resources involved in the same application, with some resulting effect on other competitors for those resources.

4.5.3 Effects of Resource Interactions Within Components

The system designer should also consider the interactions of different resources in the system. The discussion here has assumed that the scheduling of different resources is either closely coupled or independent, but this is only a simplifying assumption. Cases of concern include both generalized resource management problems and possible failure modes of the hardware or software.

As an example of a generalized resource management problem, consider a computer system that is used both as an internetwork router and a general-purpose computer. The computing activities on the system are likely to have some effect on the performance of the router functions. If possible, the priority of the communications processing for system-wide resources should be such that there is a reasonable prioritization of all of the activities on the system.

Failure conditions may cause a resource to be unavailable for a longer period than could occur in normal operation. Although a device cannot reasonably be expected to operate correctly in the presence of arbitrary internal failures, the designer should give some thought to detecting and recovering from partial failure modes where possible. For example, if one task in a multitasking software system ceases to operate, it might be possible for another task to detect the failure and reclaim any resources tied up by the failed activity.

SECTION 5

PRECEDENCE CAPABILITIES AND LIMITATIONS OF DATA SERVICES USERS

This section presents a general discussion of the precedence capabilities and limitations of the data services currently provided in the DDN and associated subscriber equipment, as seen from the point of view of a host computer system connected to the DDN either directly or by way of some type of subscriber-controlled network facility such as a local area network (LAN).

Specific implementation issues related to end user systems are covered in this section. Issues related to internetwork gateways are primarily covered in Section 6, and issues related to subnetwork components are primarily covered in section 7. However, it will be necessary to mention some aspects of the internetwork and subnetwork that impinge upon the end user systems.

End system issues include both the implementation of precedence signalling and processing mechanisms in the end systems, and management processes used to regulate (statically or dynamically) the precedence assigned to each end user's communications activities. The purely administrative and policy questions of precedence management, such as the basis and procedures for deciding what precedence level is justified for a specific user or application, are outside the scope of this paper. It is sufficient to state that some competent authority must make these determinations. This following discussion is intended to explain the technical issues to consider and actions to take to give practical effect to the precedence assignments that have been made by the proper authorities.

5.1 GENERAL TECHNICAL PROBLEMS IN USER PRECEDENCE MANAGEMENT

We begin by identifying some technical problems in implementing and using precedence features that apply across the full range of end-user applications.

5.1.1 Connection Management

When a connection-oriented protocol is used in an environment that acts upon a signalled precedence value, each connection needs to have some precedence level associated with the connection as a whole. This is not just a matter of philosophical consistency. Even if the protocols being used allow different portions of the data on the connection to have a different precedence level (which is improbable), the effects of precedence processing are corrupted by such a mixture. The most serious problem (and also the most likely) is that the use of flow control on the connection can't be made to depend on the precedence unless the connection actually has a single precedence level associated with it. Otherwise, flow control might be applied on the basis of a perceived low precedence, hindering delivery of high precedence data that is in transit at the time. Preemption of the connection would also be ambiguous and possibly incorrect if the connection's traffic flow is of mixed precedence.

To prevent such problems, communicating components must generate distinct connections for each precedence level to be used, and they must place only traffic of the corresponding precedence level on a connection.

5.1.2 Consistent Precedence Usage for Related Activities

In some application environments, several different communication activities will occur in the course of a single user application. An example is the use of the domain name system (DNS) in the DDN or Internet to obtain domain name to IP address translations. The DNS is used by many other applications to obtain an IP address for a host specified by the user. If the user's planned activity has a certain precedence level, any DNS query and response messages should have the same precedence level. Otherwise, precedence mechanisms in the network might block or delay the DNS operation, and thus indirectly block or delay the user's application.

Since the DNS itself may generate additional communication activity to consult other domain name servers for information not found locally, the DNS must also retain the precedence of any unresolved queries and set the same precedence in its recursive queries. Likewise, when the DNS response is finally sent, it should have the same precedence set. (Since DNS is most often accessed using the connectionless User Datagram Protocol (UDP), the domain server must make its own provision to retain the precedence of the query, as there will be no connection state block containing the precedence as there would be if TCP were used.)

A similar case comes up in mail-based query/response servers, a number of which are operational on DDN subscriber systems. The precedence of the incoming mail message containing the query should be remembered and applied to the outgoing mail message containing the response.

5.1.3 Assuring Proper Usage of Precedence Levels

Systems that implement any type of prioritization of activities of human users must deal with the natural tendency of the human users to first decide the level of service they want to receive and then assert whatever level of importance they feel is necessary to obtain that level of service. This behavior has been observed in DOD communication systems (notably in AUTOVON during the 1970s), where it is colloquially known as *precedence creep*. The creeping action refers to the use of successively higher precedence levels until the call succeeds. This behavior defeats the purpose of a priority scheme. Therefore, it is usually desirable and sometimes necessary to take some action to prevent or at least minimize this behavior. The preventive actions may include regulatory mechanisms in the communication system or outside of the system. It appears that no system of controls is completely ideal, so we now consider the alternatives and the technical impacts of different possible mechanisms.

5.1.3.1 Authorization and Validation Procedures

In order to restrict the use of high precedence levels, it is current practice in DOD to assign maximum authorized precedence levels to users of communication systems. The question is how to obtain user compliance with these assignments. Some possibilities include:

(1) Validation mechanisms in some network component.

(2) Validation mechanisms in the end system.

(3) Cost-based controls.

(4) Precedence usage reporting to management.

Methods 1 and 2 both imply that some characteristic of the traffic is used to look up the authorized precedence level in a database. In the AUTOVON and DSN telephone systems, each telephone instrument has an authorized precedence level for its outgoing calls (only) and an attempt to use a higher level than authorized results in the call being blocked or diverted to a human operator who will place the call if the caller provides an acceptable authorization code.

The next section discusses one problem (described as the reconnection problem) with using this approach for data network applications. Laying aside the reconnection problem for the moment, these methods require definition of the information elements to which an authorization level will be assigned (the counterpart to the calling telephone number or authorization code in DSN), creation of a precedence authorization database, and provision of software in network components to interpret the arriving traffic, extract appropriate information for validation, consult the database to actually perform the validation, and dispose of unauthorized attempts in some way.

Methods 3 and 4 both imply that some statistical data or logs are compiled within the network. This presumably would associate the traffic level at various precedences with users or applications. The type of information available to identify the user responsible for a given traffic flow depends on where in the network the data is collected,

With either type of approach, there is a tradeoff between performing actions in the end system and performing them at a more remote location. The end system has the most accurate information available to it, but the management process is more trustworthy when the validation mechanisms are out of the immediate control of those who would benefit from subverting them. A reasonable compromise might be to perform all enforced validation in the end system, but collect statistical usage data elsewhere in the network.

When precedence validation is performed at a location where multiple traffic sources have been aggregated into a single flow, it is clearly necessary to set the authorization limit to the highest level appropriate for any of the subsidiary flows.

Internetwork components need to use high precedence communications for routing and management functions. (The reason for this is explained in section 6 of this report.) Therefore, it is not useful to have precedence limits on subnetwork interfaces to internetwork routers. Any necessary enforcement will have to occur in the internetwork entity or closer to the traffic source.

5.1.3.2 Reconnection Problems

One of the useful features of a layered protocol architecture is that service interruptions at a lower layer can be recovered transparently by higher-layer actions. In some situations, these recovery mechanisms need to be initiated by the responding peer rather than the originating peer. If a precedence validation scheme is to be incorporated in the network or end system, it must make some provision for this communication to occur when the responding peer is not authorized to originate traffic at this precedence level.

One way to avoid this type of problem is to perform only pairwise validations of authorizations. In other words, a connection or datagram is considered to have a valid precedence if either the source or the destination is authorized to use that precedence level.

Another approach is to perform validation in end system software and design this software so that validation limits are bypassed when reconnection procedures are in progress. In effect, only the precedence at initiation is validated, and all subsequent traffic on that connection or association is deemed valid.

5.2 TCP/IP PRECEDENCE IMPLEMENTATION GUIDELINES

5.2.1 TCP Precedence Matching

All devices that implement TCP must perform the precedence negotiation as defined by the protocol specification. The effect of the negotiation process is that each party specifies a desired precedence level, and both parties use the higher of the two desired levels for the duration of that TCP connection. This is called ‘‘precedence raising’’ and it should not be prevented from working by any authorization checking mechanisms in the host.

5.2.2 Application Programming Interface for Precedence Service

The transport layer service interface provided to application programs should support the following precedence-related functions:

- Selecting precedence level for an active TCP Open
- Selecting minimum acceptable precedence level for a passive TCP Open
- Determining the precedence level of each TCP connection associated with an application (the negotiation may have raised the precedence from the level selected in the Open primitive action, and there should be a way to find out the precedence level that resulted)
- Selecting the precedence level for each UDP datagram sent by the application
- Determining the precedence level of each UDP datagram delivered to the application

These requirements are clearly without meaning in the case of a system with no programmable interfaces (single function components). However, it is likely that some equivalent interface exists internal to the software of such a component, and some or all of these functions would be used at that interface to provide precedence features to the single application function.

5.2.3 Precedence Mapping Between IP and Subnetwork

5.2.3.1 Devices Using the DDN X.25 Interface

Table 1 shows the defined IP precedence values, which are used in the precedence field of the IP datagram header, and the mapping of IP precedence to the DDN X.25 precedence facility as defined in the DDN X.25 Host Interface Specification (DCA, 1983). All hosts connected to the DDN backbone are required to use this mapping for their IP protocol traffic.

Table 1. IP Precedence Levels and Mapping to DDN X.25 Precedence

IP Precedence Value	DDN X.25 Precedence Facility	Descriptive DOD Name
0	xxxx xx00	Routine
1	xxxx xx01	Priority
2	xxxx xx01	Immediate
3	xxxx xx11	Flash
4	xxxx xx11	Flash Override
5	xxxx xx11	ECP/CRITIC
6	xxxx xx11	Internetwork Control
7	xxxx xx11	Network Control
x = Don't care when received, set to 0 when generated		

5.2.3.2 Devices Connected to LANs

For hosts attached to LANs, the subnetwork treatment of precedence depends on the subnetwork technology. Most standardized LAN technologies employ the ISO 8802 priority structure, which provides eight priority levels at the Logical Link Control (LLC) service interface. The mapping of the eight LLC priorities to actual subnetwork priorities is different for each type of ISO 8802 subnetwork. An important LAN technology not using LLC is the Ethernet. (There is a subtle distinction between an Ethernet and an ISO 8802-3 LAN. They can coexist on a single physical cable, but the bit structure of the frames is different. Ethernet framing is usually used with IP, but

ISO 8802-3 framing is used with OSI protocols.) Ethernet has no notion of priority or precedence, so no useful subnetwork mapping can be performed by a device connected to an Ethernet.

Some research questions the usefulness of LAN priority schemes in ring networks for any situation other than an absolute excess of demand (Peden and Weaver, 1988). According to these studies, users of these networks tend to either receive very good service or very bad service, so the priority mechanism comes close to being a preemption mechanism. If the LAN is not too large, the bandwidth of the LAN may be enough to adequately support all foreseeable user requirements. For these reasons, there is often little use for subnetwork precedence functions on LANs. The decision to use or not to use subnetwork priority to support IP (or CLNP) precedence needs to be made on a per-subnetwork basis. If it is decided to use the subnetwork precedence, the eight IP precedence levels may be mapped directly to the eight LLC priority levels.

5.3 GENERAL APPLICATION PRECEDENCE IMPLEMENTATION GUIDELINES

If an application is intended to be used with different precedence levels by different users, or by the same user on different occasions, the user interface must provide a way to select the precedence level desired. This might be implemented as an optional command, with ROUTINE precedence used if the special command to select another precedence level hasn't been given. Such a capability is easy to implement and should be routinely included in application software.

The application may check this against some authorization database to guard against misuse; the need for such verification must be determined by the managers responsible for the system. For example, there is a policy in DOD that precedence levels higher than PRIORITY should not be used for bulk data transfers without special high-level authorization. It would be reasonable for a bulk data transfer application to restrict users from selecting higher precedence levels.

Some thought should be given to the reconnection problem described in section 5.1.3.2 above. Any validation mechanisms built into the application should not obstruct the resumption of an activity which was set to a high precedence level by the other party to the communication.

Applications that use multiple connections need to ensure that corresponding precedence levels are used on each connection involved in servicing a given instance of usage. For example, a server program for the File Transfer Protocol (FTP) should open data connections with the same precedence level as that used on the control connection.

When the user interface presents many applications or data sources simultaneously, special provisions may be desired to attract the attention of a user to activity on high precedence connections. For example, if a display system with overlapping windows is used, it would probably be undesirable for activity in a high precedence application window to be unseen because it is covered up by a low precedence window. Changing the stacking order of the windows under system control might be too drastic, but it would be desirable to have some way of alerting the user that something in a hidden window may require immediate attention.

5.4 PRECEDENCE IMPLEMENTATION GUIDELINES FOR MESSAGING SYSTEMS

The store-and-forward nature of a messaging system calls for a somewhat different treatment of precedence in the messaging software than would be used in other applications. Each message has a precedence level (or conceivably different precedence levels for different addressees), and this precedence level needs to be stored with the message at each point where it is queued for forwarding or delivery. The message item file format (or equivalent data structure) must provide a way of doing this. The means of specifying the precedence might not be a command, as in other applications, but some type of control field in the message header which the user fills in when composing the message.

Queue management in the message delivery software is one of the most important aspects of overall performance of the messaging application. The message transfer agent (MTA) should maintain its queue of undelivered messages in such a way that it can employ precedence-ordered queueing. Messages not delivered on the initial attempt (due to problems in the network or remote host) should be retried at some interval that depends on the precedence of the message. The retry intervals may be different for different sites, so it is advisable to make this parameter configurable by the system manager without needing to change the software itself.

If the MTA cannot support multiple simultaneous connections for sending mail to other systems, it may also need to implement preemption of outgoing messages.

Conversely, if there are limits on the number of concurrent incoming connections that can be handled, it would be desirable to keep one ‘‘slot’’ free in case a high precedence message comes in. The slot could be kept free by using the TCP application interface feature described in section 5.2 of setting a minimum acceptable precedence for a passive TCP Open call. The minimum should be set to one level higher than the lowest level incoming connection in progress. If the reserved slot is actually used by a high precedence incoming connection, one of the lowest precedence connections should be aborted so that its slot can be reserved for a higher level.

There are some practical differences in the allocation of system resources depending on whether messaging is the primary function of a system or an auxiliary facility of a system mainly used for other functions. Resource consumption limits for the message handling software may be used to avoid undue impact on other system activities, but these limits should be set much higher on a system used primarily for messaging.

5.5 PROPOSED PRECEDENCE IMPLEMENTATION APPROACH FOR GOSIP PROTOCOLS

This section presents a proposed mapping of DOD precedence values and functions into the primary transport and network protocols specified by the Government Open Systems Interconnection Profile (GOSIP). It is intended to represent a functional equivalent to the TCP/IP guidelines presented in section 5.2 above.

Table 2 presents a mapping of DOD precedence levels to CLNP priority values as proposed in (Scott, 1990). This has not been formally adopted but it is now being used as the basis for the Phase II BLACKER Front End interpretation of CLNP Priority, and thus will prevail in practice unless and until a different policy is adopted.

Table 2. Proposed Mapping of DOD Precedence Levels to CLNP Priority Values

CLNP Priority	Encoded Value	Descriptive DOD Name
0	0000 000x	Routine
2	0000 001x	Priority
4	0000 010x	Immediate
6	0000 011x	Flash
8	0000 100x	Flash Override
10	0000 101x	ECP/CRITIC
12	0000 110x	Internetwork Control
14	0000 1110	Network Control
x = Don't care on input, set to 0 on output		

Table 3 shows the corresponding mapping of CLNP priority to the current DDN X.25 Precedence facility. Although no final decision has been made, there is a substantial possibility that the DDN will be upgraded in the future to use the Priority facility from Annex G of International Consultative Committee for Telegraphy and Telephony (CCITT) Recommendation X.25 (1988) and ISO 8208 (second edition). At this time, the DDN X.25 interface allows the CCITT/ISO Priority facility to be passed between data terminal equipments (DTEs), but the network does not interpret the contents. As a matter of correctness, the CCITT/ISO Priority facility should be included in X.25 calls when GOSIP protocols are being used above X.25. For the present, it is recommended that all six subfields of the Priority facility be set identically to the CLNP priority value.

Table 3. Proposed Mapping of CLNP Priority Levels to DDN X.25 Precedence

CLNP Priority Option	DDN X.25 Precedence Facility	Descriptive DOD Name
xxxx 000x	xxxx xx00	Routine
xxxx 001x	xxxx xx01	Priority
xxxx 010x	xxxx xx01	Immediate
xxxx 011x	xxxx xx11	Flash
xxxx 100x	xxxx xx11	Flash Override
xxxx 101x	xxxx xx11	ECP/CRITIC
xxxx 110x	xxxx xx11	Internetwork Control
xxxx 111x	xxxx xx11	Network Control
x = Don't care when received, set to 0 when generated		

SECTION 6

IMPLEMENTATION OF PRECEDENCE SERVICES IN INTERNETWORKS

Internetworks built up from subnetworks of differing technology are an increasingly prominent component of data communications in general, and this is no less true in the Department of Defense than elsewhere. Two major forces are motivating the widespread use of internetworking technology in DOD:

- The enormous increase in the numbers and types of computing devices used in day-to-day work has made the implementation of local area networks almost essential. These LANs may connect hundreds or even thousands of personal computers (PCs) and multiuser systems at a single location. The most economical way to enable these systems to communicate with distant locations is to employ internetworking technology to interconnect the LANs through a wide area network such as the DDN.

- The uses of data communications in lower echelons of operational military organizations are growing, and applications requiring integrated communications among tactical, strategic and sustaining activities are being developed. Since the computer and communications systems deployed with operational military units necessarily have some special characteristics to make them survivable in adverse environments, integrated communications service for these systems and for more traditional systems also involves internetworking technology. The Integrated Tactical-Strategic Data Network (ITDN) program is a prime example.

Although requirements for precedence features have been defined for the DDN and for a number of tactical systems, there has never been a defined requirement specifically for the internetworking service elements. The widespread insertion of internetwork technology into the operating environments of these systems causes their system-specific features to break down at the individual system or network boundaries unless some provisions are made for a corresponding service in the internetwork components. Therefore, it seems logical to include precedence features in the internetwork components.

Internetwork components are also being used as the basis of some subnetworks. The DDN currently has an IP-based subnetwork in an early stage of operational service. It is unclear whether such a network is formally required to provide precedence features, but it is clearly reasonable to consider how it can be made to do so using the internetworking protocols and concepts from which it was built.

This section presents an analysis of how precedence features would be provided and used in the components and protocols that support internetworks. The analysis is predominantly oriented toward TCP/IP internetworks since this is the current predominant technology in DOD. The issues and alternatives for OSI-based internetworking using the connectionless network service are essentially the same as for IP-based networking, so it is sufficient to note the corresponding protocols and

entities in those two types of networks. Currently, there is little usage of connection-oriented (X.75) internetworking mechanisms in DOD, but a brief discussion of the distinctive aspects of precedence support in the connection-oriented internetworking environment is also provided.

6.1 TCP/IP INTERNET SUBSCRIBER SERVICES

The DDN's increasing involvement in internetwork service support to subscribers causes a number of additional issues to arise regarding precedence support at this layer.

6.1.1 IP Precedence Levels

The IP protocol defines an 8-level precedence scheme. It does not specify exactly what processing actions any component should take based on the precedence. Table 1 in section 5 shows the numerical values and descriptive names of the eight levels. The precedence of an IP datagram is set by storing the numerical value in a three-bit field of the IP datagram header. This three-bit field is present in every IP datagram; there is no way to omit it.

6.1.2 Precedence Across Administrative Boundaries

There could be a concern about the interpretation of precedence values when different authorities were responsible for the criteria used to assign the precedence and those used to interpret it. This is a special concern when the origin of the traffic doesn't have any effective control to assure that high precedence levels aren't abused. In a relatively open internetwork environment, it can be expected that some portions of the internetwork will not have the policy and administrative controls on precedence usage that exist in a military organization. Therefore, the possibility should be considered that some traffic marked as high precedence in an uncontrolled environment will enter controlled environments and divert resources from traffic of real importance.

In order to control this problem, it would be desirable to have a means of restricting the precedence levels of traffic across the boundary between the controlled and uncontrolled areas of the internetwork environment. This might be attempted either by changing the marked precedence values on traffic at the boundary, or by blocking such traffic altogether.

Blocking would be the easier solution to implement if not for the problem that there will be some high precedence traffic which should traverse administrative boundaries, such as routing, management, and control traffic. A partial solution would be to allow the traffic through but restrict the rate at which it can be sent from the uncontrolled side, using flow control or other blocking means to limit the flow to the administratively determined value.

Modifying the precedence markings would be preferable in that it doesn't restrict the flow of traffic across the boundary as long as adequate resources are available to handle the traffic. A complication is that modifying the marked precedence disrupts certain protocols, notably TCP, which expects the precedence level to be negotiated at connection startup to the higher of the values desired by the two endpoints. If the precedence of the active synchronizing packet is lowered without the knowledge of

the host that sent it, the receiving host will attempt to synchronize at the lower precedence level, and its low precedence packet would appear to the initiating host as a protocol error, causing the connection to be reset (aborted).

It would be possible to change precedence levels at the boundary if the gateway device recognizes the TCP protocol (and any other protocols of interest), keeps track of the precedence mappings for individual connections, and modifies precedence values symmetrically for each connection. This is in the nature of a transport bridge function rather than a router function, but in principle it could be built. It could also be adapted to include some type of lookup function to allow, disallow, or modify precedence levels differently for different combinations of endpoints (since the relay would have knowledge of individual TCP connections, the validation would happen only once per connection). However, it requires that all segments on the connection be routed through the same relay device, unless some additional protocol is used to share knowledge of connection precedence mappings among all devices at the boundary of the controlled region. This is technically feasible, but not a trivial undertaking.

6.2 INTERNETWORK OPERATION AND MANAGEMENT SERVICES

6.2.1 Routing Protocols and Procedures

Since correct routing of traffic is necessary to provide a usable internetwork service, routing should never be disrupted by the processing of ordinary traffic. Therefore, the precedence of routing traffic should be higher than any user traffic. Routing services in TCP/IP internets commonly employ protocols such as Exterior Gateway Protocol (EGP), Border Gateway Protocol (BGP), Routing Information Protocol (RIP), or Open Shortest Path First (OSPF) protocol. Regardless of the specific protocol being used to circulate network routing information, its traffic should be set to an appropriate high precedence level. Normally this would be precedence level 6, Internetwork Control. If subnetwork routing information is propagated using internetwork routing protocols, it would be more appropriate to use level 7, Network Control.

6.2.2 Monitoring and Control Protocols and Procedures

This discussion will consider only precedence issues pertaining to the internetworking aspects of general management functions and to specific management functions unique to the internetwork environment. Management requirements for internetworking components such as routers are generally similar to those for other components, and that information will not be repeated here.

In a TCP/IP internet, the predominant management protocol is the Simple Network Management Protocol (SNMP). Other vendor-specific management protocols exist and are also used over IP. Whatever protocols are used for monitoring and control should select an appropriate IP precedence value for their traffic. IP precedence levels 6 (Internetwork Control) and 7 (Network Control) are appropriate for essential monitoring and control functions essential to continued acceptable operation of the network. There may be other access to management information which is of a less critical

nature. A lower precedence level would then be appropriate. The following suggestions are offered for selecting precedence levels:

- If the essential management activities of a subnetwork uses IP-based protocols, then that traffic should use IP precedence 7, Network Control. Essential management activities across subnetwork boundaries should use IP precedence 6, Internetwork Control.

- The importance of such management functions as accounting is essentially a policy question. It should have at least the next higher precedence level above that of the most important traffic for which accounting is considered an important function. For example, if it is only considered essential to have accounting data for Routine traffic (IP precedence 0), the accounting reports from network components might be sent at IP precedence 1, which would protect them from loss if there is serious congestion at precedence 0. As long as the accounting traffic is not a major percentage of network burden, it seems reasonable to assign a high precedence (Internetwork Control or Network Control, as appropriate) to accounting reports, the same as is done with other management reports.

- Performance analysis and reporting tools should have a selectable precedence level so that the analyst can separately examine the properties of the network at different precedence levels. For example, ‘‘pinging’’ with precedence 0 may yield very different results from pinging with precedence 7, and both of these results may be useful to the analyst or network manager.

- General-purpose tools used under direct human control should allow the precedence level to be selected at the user interface. The importance of these activities depends on the exact situation. It seems reasonable to allow the user of the tool to decide whether a high precedence is warranted in a specific situation.

6.3 OSI CONNECTIONLESS NETWORK SERVICE (CLNS) INTERNETWORKING

The use of precedence queueing and of subnetwork preemption in a CLNS internetwork router is entirely comparable to the use of the same features in an IP router. The only differences to be dealt with are the specific protocol constructs used to perform the signalling.

The allocation of functionality to specific protocols differs somewhat between the CLNS and the IP protocol family. IP uses the Internet Control Message Protocol (ICMP) and the Address Resolution Protocol (ARP) as auxiliary protocols for error reporting and link layer address resolution. CLNS achieves approximately the same results using the data (DT) and error report (ER) protocol data units (PDUs
in the CLNP protocol and the ISO 9542 connectionless end system to intermediate system (ES-IS) routing information exchange protocol. The approximate correspondences between the two stacks are: CLNP DT generally corresponds to IP, CLNP ER to the error reporting aspects of ICMP, and ISO 9542 ES-IS to ARP and the addressing and routing aspects of ICMP. ICMP is actually a subprotocol of IP and it is carried in IP PDUs, but ISO 9542 is conceptually parallel to CLNP and has

its own distinct PDU formats. As a result, the IP datagram precedence field can be used to set the precedence of an ICMP message, but in OSI, each ES-IS PDU type must have a priority option in order to signal precedence to a peer entity.

Routing protocols for use between intermediate systems have not yet reached full standardization in OSI. The intermediate system to intermediate system (IS-IS) protocol for use between intermediate systems in a single routing domain is ISO Draft International Standard (DIS) 10589. A working draft exists for an inter-domain routing protocol (IDRP). The priority mechanisms currently defined as applicable to the draft standards are briefly discussed below, but it should be recognized that there may be changes in the protocols or their usage environment before the final standards are approved.

6.3.1 CLNP Priority Signalling

CLNP defines a Priority option with the appropriate semantics to enable it to be used to control a precedence feature. Like all CLNP options, the entire option field can be included in or omitted in the CLNP header as the case requires. IP precedence, on the other hand, is signalled in a header field that is present in every IP datagram. Including the Priority option in a CLNP PDU increases the length of the PDU by two octets. ISO 8473 states that if the priority option is not included, the PDU is treated as having priority value zero (lowest priority). The priority option provides an 8-bit field for the value, but the specification currently allows only values 0 through 14 decimal.

There is no formally standardized mapping of DOD precedence to CLNP priority at this time, but the mapping shown in Table 3 is being used by the Phase II BLACKER Front End (BFE) and as such is a *de facto* interim standard which may or may not become official in the future.

6.3.2 ISO 9542 ES-IS Priority Signalling

The ISO 9542 connectionless ES-IS protocol allows a priority option to be included in all of its PDUs. The format of the priority option for ISO 9542 is identical to the format of the CLNP priority option.

Although there is no definitive guidance on what priority value should be used for this protocol, it seems reasonable to use the Network Control priority (14) since the information passed in this protocol may be essential for address resolution on the local subnetwork.

6.3.3 ISO DIS 10589 Connectionless Intra-Domain IS-IS Priority Signalling

There is no priority signalling construct defined as applicable to DIS 10589 PDUs. Like the ISO 9542 ES-IS PDUs, there are many different PDU types, each free-standing and not embedded in ISO 8473 DT PDUs. Unlike ISO 9542, DIS 10589 does not define the optional fields that are defined by ISO 8473, one of these being the Priority option. DIS 10589's conformance requirements call for a specific internal prioritization of packet processing in the router, which has the effect of giving higher priority to DIS 10589 PDUs than to data PDUs. This obviates the requirement for a separate priority indication.

6.3.4 ISO Connectionless Inter-Domain Routing Protocol (IDRP) IS-IS Priority Signalling

This working draft, soon expected to be registered as a Draft Proposed International Standard, embeds the IDRP PDUs in CLNP DT PDUs. This allows the precedence of the routing information exchange traffic to be signalled using the CLNP priority option.

6.4 X.75 INTERNETWORKING

The X.75 protocol is used to interconnect X.25 public data networks and in some cases for private networks as well. X.75 is essentially an adaptation of X.25 used between two half-gateways (one on each network). The DTEs on the X.25 networks involved do not execute X.75 protocol on an end-to-end basis; an X.25 connection is established to the local X.75 gateway, and the remote X.75 gateway establishes an X.25 connection to the next X.75 gateway or to the destination DTE.

Since the X.75 gateway is strictly connection-oriented, it has a fixed maximum number of connections it can support. These connections could conceivably all be used up, so connections (and all the state information and per-connection buffers used to support connections) are resources that might need precedence processing applied to them. Logical channels and protocol state blocks on local or internetwork interfaces are also candidates. Note that the gateway function may support more or fewer connections than the interfaces; an X.75 gateway connected to more than two networks could run out of gatewaying capacity before running out of capacity on the specific interfaces involved, but a gateway connected to only two networks will be limited by only one resource (whichever is fewest in number).

If the X.75 gateway performs packet or window size conversions, it will need to keep more data in buffers, which makes buffer exhaustion more of a concern. A very simple-minded gateway would never need to buffer more than one packet in each direction per connection.

Throughput on either the local network or internetwork interfaces of the gateway might also be subject to exhaustion due to flow control being applied by one network or the other. The processor of the gateway itself could be subject to exhaustion as well.

There is no accepted standard for precedence or priority signaling in an X.75 (or X.25) environment, so provision of this service in an X.75 internetwork would be likely to require custom development. (The CCITT-specified DTE Priority facility is specified for use between the DTEs attached to the network, rather than for interpretation by network components. Although such an interpretation can be made, this cannot be considered as standard or widely accepted practice.)

SECTION 7

PRECEDENCE CAPABILITIES AND LIMITATIONS OF DDN-PROVIDED SERVICES

7.1 PSN SUBNETWORK SUBSCRIBER PRECEDENCE SERVICES

This section discusses the precedence features supported in the services provided by the PSN subnetworks of the DDN (MILNET and DISNET segments).

7.1.1 Requirements

Precedence service was established as a requirement for the design of the DDN by the Office of the Secretary of Defense when the DDN program was established in 1982. In response to this requirement, a design for implementation of precedence within the DDN backbone was prepared by the DDN Project Management Office in consultation with the principal contractor and other parties involved in implementing the DDN. This design is documented in the DDN Program Plan, edition of April 1982 (DCA, 1982).

7.1.2 Operational Status

In practice, no MILNET subscriber systems have been enabled to use non-default precedence. DISNET PSNs have been configured to allow FLASH precedence traffic among components of the BLACKER end-to-end encryption system. FLASH precedence is used by the BLACKER system for administrative traffic such as key placement.

7.1.3 Design As Implemented

Precedence service within the DDN backbone is currently available to subscriber systems connected directly to MILNET and DISNET PSNs using ARPANET Host Interface Protocol (AHIP) or DDN Standard X.25 interfaces. Based on the precedence level signalled by the subscriber equipment, the PSNs give preferential treatment to higher precedence traffic.

The PSN software provides a capability to configure the maximum precedence allowed for outgoing X.25 call requests on each PSN port.

Distinct mechanisms exist for signaling the precedence level on an AHIP host connection and on an X.25 host connection. These procedures make it possible for hosts to simultaneously send traffic of multiple precedence levels to the same or different destinations.

In the case of an AHIP interface, the destination and precedence of each AHIP datagram are signaled explicitly in the AHIP protocol header of every datagram. The method of signaling precedence on an X.25 interface is a DDN-specific extension to the X.25 standard, which may be included or omitted at the subscriber's choice. Under this procedure, the destination and precedence are signaled when an X.25 virtual circuit (VC) is set up from source to destination. All data sent on that VC is considered

to have the precedence indicated at the time the VC was set up. A single VC can be (and normally is) used for bidirectional (full-duplex) communication, since using two VCs unidirectionally consumes additional subscriber and network resources without providing any additional functionality. When a DDN X.25 subscriber wants to communicate with a destination at more than one precedence level, separate X.25 VCs are used for each precedence level, and it is the responsibility of the traffic source to use the correct VC to carry data of a particular precedence level.

The discussion above shows that a DDN X.25 subscriber host must use different X.25 VCs for its traffic based on dynamic requirements. Therefore, an X.25 host, when it has traffic to send, should check its existing X.25 VCs to identify one with the appropriate destination and precedence value. If no suitable VC exists already, then it needs to set up a new VC with the appropriate destination and precedence. After a suitable VC is found or created, the packet is sent on it.

7.1.4 Issue: Incoming Call Precedence Signalling

Proper use of precedence by DDN X.25 subscribers requires that they send all traffic on X.25 VCs of known correct precedence level. The means of doing so depends in part on whether the DDN PSNs report the precedence level of incoming calls. DDN X.25 service was first implemented in PSN software version 6, which is believed to have included an indication of incoming call precedence in Incoming Call packets sent from the PSN to the X.25 host. PSN 7 and PSN 8 software do not provide this indication. As a result, DDN X.25 hosts are currently unable to determine explicitly the precedence of an X.25 VC which was originated elsewhere.

This poses a problem for a host that needs to send traffic to another host for which it has an existing VC that was originated by the other end, because it does not know whether the precedence level of the existing VC is appropriate for the traffic that it intends to send. If the precedence levels of the VC and the outgoing data match, the existing VC should be used. If they do not match, another VC is needed. The host is not able to determine which of these actions is correct.

7.1.4.1 Alternative Strategies

Four alternative ways of addressing this issue are considered:

(1) The straightforward approach is to change the PSN software to provide a precedence indication in Incoming Call packets (to destination hosts which have elected to accept this facility). The indication delivered would be encoded identically to the precedence indication signaled by the originator of the VC.

(2) Subscriber equipment could be required to maintain an association between X.25 VCs and the applications which eventually receive data from incoming calls. Any return data from the same applications could then be routed to the same VC, on the assumption that the precedences must correspond. This would be a change to the normal procedure for assigning outgoing traffic (IP datagrams) to X.25 VCs, which was described in section 5.2. This alternative also implies that the subscriber equipment will not multiplex additional

applications on a VC that it did not initiate, since it is unsure of the precedence of such a VC.

(3) Subscriber equipment could be required to check the IP precedence (or in the future, the CLNP priority) of incoming data packets and label the X.25 VCs on which they arrive with that precedence level. Subsequent assignment of outgoing traffic to X.25 VCs would be done in the manner described in section 5.2 above. This alternative only works when each VC is either assigned to a single non-multiplexed use, or the multiplexing scheme contains an explicit precedence or priority indication visible to the destination.

(4) Subscriber equipment could be required to use X.25 VCs in a simplex (one-way) manner. All outgoing traffic would have to be sent on a VC originated by the sending host. The effect is that typical communications between two hosts (such as TCP or TP4 sessions) would always involve the use of two VCs.

7.1.4.2 Discussion of Alternatives

Although it is not possible to be certain, it is likely that Alternative 1 represents the original design intention. The DDN X.25 Host Interface Specification (DCA, 1983) does not state whether the DDN X.25 Precedence facility is included in Incoming Call packets, but the qualification test procedures for a DDN Data Circuit-terminating Equipment (DCE) described in DCA Circular (DCAC) 370-P195-5 (DCA, 1989), Supplement 1, Chapter 11, Module 23A, Test 3, call for the DCE to send a Precedence indication in an Incoming Call packet, and the corresponding DTE tests in Chapter 10, Module 23A, Tests 4 and 5 test the DTE's ability to receive it. It is suspected that the non-delivery of the precedence facility to the called DTE was an inadvertant oversight in the implementation of the new End-to-End protocols first deployed in the PSN 7.0 release. The BFE encryption device was designed to recognize a precedence facility in an incoming (Black side) call.

Alternative 4 would make it impossible for two-way traffic flows between a host pair to use piggyback acknowledgements in the X.25 Network Layer. The result would be a greater number of Receive Ready (RR) and Receive Not Ready (RNR) packets on the host access links, and some additional processing by the source and destination PSNs and hosts to handle the increased number of packets. The adverse impact on intermediate PSNs and trunk capacity would be considerably less since the BBN-private internodal End-to-End (EE) protocol enables acknowledgements to be piggybacked for multiple X.25 VCs.

Alternatives 2 and 3 both require introducing connection-like linkages between the Network and Transport layer functions of the subscriber equipment (or BFE). This is contrary to the connectionless nature of IP and CLNP protocols and could be expected to be a more difficult change for host implementors than alternatives 1 and 4.

Alternatives 1 and 3 require only a single VC, opened by either end, to support any number of multiplexed applications between a pair of hosts. Alternative 2 would require two VCs, one opened by each end, in the case of concurrent applications initiated at both hosts of a communicating pair.

Alternative 4 would require two VCs to support any application that requires bidirectional communication, even if there is only a single application involved.

Alternative 3 does not solve the problem for Basic X.25 uses which allow multiplexing different sessions on a VC but don't use a higher layer protocol with an explicit precedence or priority indication. It is not known whether any such applications exist on the DDN.

7.1.4.3 Evaluation

Alternative 1 as described above provides the most direct solution to this problem. It provides the required service with minimal resource consumption in the network and it does not require changes to subscriber equipment that passes the established qualification testing procedure.

In order to provide appropriate implementation guidance to DDN subscribers (i.e., the vendors who supply their equipment), a firm decision is needed on whether the PSNs will be modified to provide the X.25 precedence indication on incoming calls, as required by Alternative 1. An early resolution of this question is desirable to reassure the vendor community as to how their devices should be designed to work. The implementation guidelines presented in this document assume that a favorable decision will be made on this recommendation.

7.1.5 Issue: Logical Channel Preemption at Host Interfaces

Each X.25 host interface on a PSN is configured to support a certain maximum number of X.25 logical channels (which correspond to VCs). This value is usually chosen on the basis of host limitations (how many connections the host software can support simultaneously). Since one VC is needed for each combination of destination and precedence level, a system that communicates with a large number of other systems simultaneously will occupy many logical channels. The logical channel limit is most likely to become a factor for gateways and for large timesharing computers, because they are most likely to be communicating with a large number of other DDN systems simultaneously.

When all allowed logical channels to a host are in use, it is not possible to form a new connection to that host. The destination PSN will cause an error indication at the X.25 level to be returned to the calling system. This indication specifically identifies the cause as ‘‘no free logical channel available,’’ but in a DDN Standard X.25 (DOD protocol stack) implementation, full diagnostic information is usually not delivered to the end user attempting to send to the busy host. An error message indicating ‘‘host unreachable,’’ ‘‘connection failed,’’ or some equally generic report is more likely to be provided.

Thus, in the current PSN implementation, logical channels are allocated using the FCFS discipline without any attention to precedence. When all configured logical channels between a DDN host and the PSN are in use, it is not possible to establish a new call with that host regardless of the precedence specified by the calling host, since there is no provision for preempting VCs based on precedence. Thus a high precedence call could be blocked by low precedence calls.

7.1.5.1 Alternative Strategies

(1) A preemption capability could be implemented in the PSN to reclaim logical channels when needed for higher precedence incoming calls. This implies that the maximum number of logical channels may be used for low precedence calls when there are no higher precedence calls.

(2) The host could be responsible for keeping a free logical channel available, probably by preempting a VC of the lowest currently active precedence whenever all logical channels become busy. This would ensure a free channel for high precedence incoming traffic.

(3) The PSN could reserve one or more logical channels on each host interface for high precedence incoming calls. This would be similar to alternative 2 except that the work would be done in the PSN.

7.1.5.2 Discussion of Alternatives

Alternatives 1 and 3 require PSN software changes, but no subscriber changes. Alternative 2 requires subscriber changes, but no PSN changes.

Alternative 2 can only work if the subscriber system knows the precedence of incoming calls. This general problem was discussed in section 7.1.4 above. Of the alternatives offered in section 7.1.4, only the first fully meets this requirement. Alternative 3 of section 7.1.4 is inferior because the precedence level could remain undetermined for a period of time after connection establishment, and alternatives 2 and 4 in that section do not give the host any precedence information about incoming calls.

Alternative 2 has the limitation that the PSN will deliver any incoming call on the free channel, regardless of its precedence. If the incoming call is at the lowest precedence then in use, the host will need to clear this call (or another at the same precedence) immediately to maintain the free channel. Whichever connection is preempted, it is likely that an attempt will be made to reestablish the connection. Although repeated attempts are likely whenever calls are blocked, they consume more resources in alternative 2 because the attempts are passed through to the destination host system. Under alternatives 1 and 3, the attempt can be terminated in the destination PSN.

Under any of these alternatives, a host supporting multiple levels of precedence should implement a ‘‘mirror image’’ counterpart to Alternative 1 or 2 to provide logical channels for high precedence outgoing traffic, as discussed in section 5.2.

There is no solid information on which to base an estimate of the likelihood of call blocking due to logical channel exhaustion.

Alternatives 2 and 3 will not work for a system that supports only one logical channel. Such systems exist in the commercial market, but are probably rare or nonexistent on the DDN.

7.1.5.3 Evaluation

The PSN vendor should be tasked to investigate cost and technical issues relevant to implementing logical channel preemption or reservation along the lines of Alternatives 1 and 3.

Statistics on frequency and cause of logical channel exhaustion and connection block exhaustion should be collected during normal network operations and should be analyzed periodically on a routine basis, as well as when potentially related problems are reported.

7.2 DDN-PROVIDED APPLICATION SERVICES

Application services are those specific services furnished to subscribers of the DDN above and beyond the basic capability to exchange data between subscriber devices connected to the DDN. These services consist of terminal access equipment used by subscriber terminals to access remote host computers and a variety of network information services provided by DDN-controlled host computer systems, such as ‘‘WHOIS’’ (directory of network users) and domain name system (DNS) name servers.

7.2.1 Terminal Access Controller (TAC)

The TAC device allows up to 62 asynchronous terminal connections to be made to a co-located PSN for purposes of remote login to hosts on the DDN and interconnected networks. It does not provide any capability for the terminal user to select any precedence other than the default ROUTINE precedence. There are no precedence processing actions in the TAC.

Whether precedence features are needed in the TAC is a policy decision. The updated DDN Program Plan now under development calls for new terminal access requirements to be met by the Network Acces Component (NAC) rather than the TAC. If the NAC is fielded wherever requirements for non-default precedence exist, the TAC’s limitations are irrelevant.

7.2.2 Mini-TAC/Network Access Component (NAC)

The NAC device is used in the DDN in its Mini-TAC configuration. This allows up to 16 synchronous or asynchronous devices to establish up to 64 terminal connections to remote hosts on the DDN and interconnected networks. (The difference between 16 and 64 is accounted for by the possibility of attaching synchronous devices such as 3274 cluster control units, which may have multiple devices attached to the cluster control unit.) As described in the NAC documentation (Aydin, 1988), each of the 16 physical terminal ports on the NAC may be configured with a precedence level from 1 to 4 by command to the microterminal built into the NAC chassis. These values are arbitrary to the NAC and are not numerically equal to the precedence levels signalled in any protocol. The mapping to DDN X.25 Precedence values is shown in Table 4.

There is no feature in the NAC to allow precedence to be altered by the terminal user (that is, to select a precedence level lower than the maximum authorized for that port). This is an unfortunate design feature which limits the usefulness of the precedence support provided.

Table 4. Mapping Between NAC (Mini-TAC) Precedence and DDN X.25 Precedence

NAC Precedence Value	DDN X.25 Precedence Facility	Descriptive DOD Name
1	xxxx xx11	Flash
2	xxxx xx01	Immediate
3	xxxx xx01	Priority
4	xxxx xx00	Routine
x = Don't care when received, set to 0 when generated		

In the case of a hardwired connection to a single terminal, there is no special problem with allowing the terminal user to select different precedence. However, when many users share a terminal port on the NAC (e.g., when terminal lines are allocated to TAC ports by a data switching device, or are connected to dialup lines), it would seem that an additional function is also desirable to validate the specific user's authorization to claim a particular precedence level. This would probably be easiest to implement as an extension to the TAC Access Control System (TACACS) login host (which has not yet been deployed, even without such a feature). The login host database would presumably contain a maximum authorized precedence for each TAC User ID.

The requirements for such capabilities need to be reviewed by DISA and network users to determine whether this type of feature should be implemented.

7.2.3 Network Information Services

For this analysis, the range of services provided by the DDN Network Information Center and equivalent activities with DISA and supporting contractors can be grouped into two categories:

- Common application services—those using application protocols and software similar to subscriber hosts, but used in a network-wide customer support or service provisioning role. Examples include:
  - Messaging service—between users and network information services staff. Specific uses of this service would include resolution of user questions or problems, and registration of network users.
  - File transfer service—used as a distribution vehicle for certain network data such as host tables and DNS databases, and for files of wide utility such as network documentation.

- Remote terminal access (enabling network users to use the interactive information retrieval software to obtain information about the network).

- Directory services, such as the domain name servers, used to translate between host names and network addresses, or the ‘‘WHOIS’’ server, which supports a ‘‘white pages’’ lookup of address, telephone, and mailbox information for network users.

7.2.3.1 Common Application Services

The technical considerations involved in supporting precedence access to normal application services are the same for the systems that provide network-wide services as they are for ordinary subscriber systems. Since no action has been taken to implement precedence functions for these services at the DDN Network Information Center (NIC), it can be assumed that all of the actions described in section 5 need to be carried out for NIC systems if any requirement exists for a particular service at higher than ROUTINE precedence.

The next problem is to determine where a requirement for high precedence might arise. There is no rigorous way to answer this question, but it seems reasonable that such requirements would be most likely to occur in services that support ongoing network operations. The following cases may present such a requirement:

- Messaging at other than ROUTINE level may be used to request or report database updates, such as TAC user registrations or actions on network configuration changes.

- The same argument suggests a possible requirement for precedence in file transfer services. Since the duration of file transfers is sometimes long compared to messages, and since there is usually a limit on the number of file transfer users permitted at a time (to control the effect on system load), it may be appropriate to provide some special precedence handling in the file transfer server so that a high precedence request will not be refused due to low precedence demand. Detailed discussion of technical issues pertaining to precedence implementation in client/server applications such as file transfer appears in section 5.3.

7.2.3.2 Directory Services

Directory services are especially likely to require support for precedence features. The problem is not so much that demand will increase in a stress situation, but rather that the proper precedence needs to be selected to ensure that a user who is using the directory in support of a high precedence activity will receive priority service from the network when the answer to the directory query is being sent back.

The current implementations of the DNS software generally use UDP protocol to carry both queries and responses. Since UDP is a connectionless protocol, the DNS server implementation needs to specifically act to generate the appropriate precedence signal for responses. Failure to do so could result in the DNS being inaccessible in a congested situation. A large number of network

applications are highly dependent on a working and accessible DNS service to perform their application function. At present, no DNS server implementation is known to save the precedence of queries and generate appropriate responses. This problem is quite serious and must be corrected in order for subscriber access to precedence service to be useful.

Similar logic applies to other directory services, but no other service is as crucial to the network as a whole as is the DNS.

7.3 NETWORK OPERATION AND MANAGEMENT SERVICES

In addition to the data traffic between end users, the DDN network backbone also carries management traffic used to control the operation of network components. This traffic includes unsolicited status reports from the PSNs to the Monitoring Center (MC), configuration management traffic in both directions between the PSNs and MCs, statistics collection control and reporting (including throughput data used for billing), routing and topology information, line test traffic, Terminal Access Controller (TAC) and Network Access Component (NAC) access control messages, and traffic generated by various debugging and monitoring tools used by network operators and other technical support personnel to investigate problems in the network.

7.3.1 Relationship of Existing Precedence Mechanisms to Management Traffic

Some of the DDN backbone management traffic is handled outside of the normal procedures for processing user traffic. The Packet Core Protocol (PCP) provides the operating facility used for most of this traffic. PCP packets do not carry any precedence indication, but they are afforded special treatment by the PSN. However, many of the management functions rely partially or entirely on the normal communication mechanisms of the network to carry data between management entities. This traffic has a precedence level in the same sense as user traffic.

To the extent that management functions use the normal End-to-End protocol of the PSN subnetwork, the existing precedence implementation in the PSNs will use precedence-ordered queueing to the associated data packets just as is done with ordinary traffic (the same queues are used). Therefore, such management traffic will suffer or be protected from congestion-induced delay or loss according to the precedence level assigned. If an important management function uses a low precedence level, it may fail to operate when a major resource shortage occurs.

For management traffic source which operate under direct human control, a range of precedence levels may be allowable, so the responsible network operations person can select an appropriate precedence level when that function is used. When the traffic source is an automatic function, such as the reporting of error conditions by network components to the Monitoring Center, the maximum precedence level that might be necessary should be used in all cases.

7.3.2 Issue: Management Protocol Precedence Selection

An unknown number of DDN network management functions rely on the normal End-to-End protocol within the PSN backbone to convey network management (control and reporting) traffic between PSNs, TACs, NACs, and MCs. These functions may fail to operate in stressed situations if their traffic is not assigned to a high precedence level. Conversely, they may adversely affect user traffic if the precedence level is too high. The appropriate level depends on the criticality of the specific function to the provision of effective user service.

7.3.2.1 Discussion

This issue was raised by the PSN support contractor, BBN Communications Corporation, to the DISA engineering staff during pre-release testing of PSN 8.0 software. The BBN test report and concurrent discussions with DISA referred specifically to the global precedence cutoff feature, which is a capability in the MC to command the PSNs to reject all traffic below a specified precedence level. As a result of those discussions, the global cutoff feature was disabled pending definition of the long-term solution.

However, any network management problems that would result from a global cutoff would also be experienced locally in the event of connection block exhaustion at a PSN. This event is uncommon in general and could be made less likely for selected systems (the management-related components) by appropriate tuning of the connection block pool allocations. The effect of this tuning is to reserve connection blocks for certain systems by denying them to others. The effort needed to compute and implement this tuning may be comparable to the effort needed to set up the correct precedence levels so it will not be needed.

7.3.2.2 Evaluation

A list of all types of network management traffic should be compiled, covering all DDN-controlled assets.

Appropriate precedence levels should be determined for each type of traffic. In general, functions essential to continued operation or control of communication service to network users should use the highest precedence level. Less essential functions (if any) may be assigned to a lower precedence.

Software in all DDN network elements (PSN, TAC, NAC, Monitoring Center (MC)) should be modified to use the assigned precedence levels, or to allow the operator to select the precedence level, whichever applies. A similar requirement should be included in specifications for any newly developed components or software for network elements.

7.4 INTERNET SUPPORT SERVICES

Section 6 discussed precedence design and implementation issues for internetwork services in general. This section describes the specific issues that apply to the current DDN. At present, the DDN Mailbridges are the only vehicle for providing supporting facilities for internetwork

communication between the DDN and other networks (whether those other networks are ‘‘internal’’ DOD networks that use the DDN as a backbone long-haul network, or ‘‘external’’ networks such as the National Science Foundation Network (NSFNET) that service other subscriber communities). The Mailbridges provide two general types of service:

- Packet forwarding to and from external networks (especially NSFNET), currently for IP-based services, and potentially for CLNP-based services in the future
- Distribution of internetwork routing information within the set of DDN subscriber networks

7.4.1 Requirements

No formal requirements have been documented specifically for the support of precedence features in the Mailbridge. Since DOD policy calls for the use of local distribution systems (LDSs) to furnish DDN access to the DDN in many subscriber situations, and since distribution of internetwork routing information among DDN subscribers is necessary to enable the systems behind these LDSs to communicate with each other, the general requirement for precedence support in the DDN (as described in section 7.1.1) might be considered an implicit requirement for precedence support in the routing information traffic of the Mailbridge.

7.4.2 Design As Implemented

The functional specification for the Mailbridge (BBN, 1987) does not indicate any precedence-related activity occurring in the Mailbridge. It can therefore be characterized as employing FCFS processing of IP datagrams without respect to their IP precedence level.

7.4.3 Issue: Routing Protocol Precedence

Since the precedence of routing protocol exchanges between Mailbridges and other IP routers on the DDN is not set by the Mailbridge, the default Routine precedence is implicitly used. This poses a risk of loss of internetwork routing among subscriber or external networks interconnected by the DDN when extreme congestion or preemption of ROUTINE traffic occurs. For example, if the MILNET PSN connected to a Mailbridge becomes overloaded, routing information flowing to or from the Mailbridge might be greatly delayed or dropped. Repeated loss of routing packets might be interpreted by the Mailbridge or by the other router involved as an indication that the connection to the other network has failed completely. If a high precedence were selected for routing traffic and appropriately signalled to the PSN by the Mailbridge, the PSN would be more likely to maintain the flow of routing data.

7.4.4 Evaluation

The Mailbridge should set IP precedence level 6, Internetwork Control, in its routing-related traffic, as discussed in section 6.2.1.

The Mailbridge should map the IP precedence of forwarded datagrams to the DDN subnetwork precedence according to the standard DDN mapping, as illustrated in Table 1. This may already be occurring, but the functional specification does not state that it occurs.

The use of precedence-ordered queueing of IP datagrams in the Mailbridge at the entry points to the forwarding function and the output function would be beneficial.

It may be desirable to implement datagram filtering capabilities based on precedence levels.

SECTION 8

RECOMMENDATIONS

Using the technical information and analytical methodology described in this report, the following recommendations are made for actions to support the effective implementation and usage of precedence in DOD data networks. These recommendations are based on available information about the capabilities of existing components, projected evolution of the DDN and other military data communications systems employing similar technology, and applicable government policy decisions and technical standards. The non-technical factors were not studied in depth in this effort, nor are they necessarily stable over time. Therefore, reexamination or changes in these areas might justify different conclusions. These recommendations should be considered as a current snapshot to be updated as warranted by additional or revised information.

8.1 REQUIREMENTS ANALYSIS ACTIVITIES

The recommendations of this section concern administrative decisions about the use of precedence by specific users or devices.

8.1.1 Administrative Assignment of Precedence

Policy guidance already exists for the process of authorizing the use of higher than ROUTINE precedence. User requirements for such authorization need to be determined and the appropriate administrative actions taken to implement the decisions both in the network and in subscriber organizations.

8.1.2 Requirements for Enforcement of Precedence Validation

A decision is needed as to whether the DDN and DDN-connected DOD networks need to implement a precedence validation mechanism (i.e., dynamically check whether a given source is administratively allowed to use a precedence level that it is attempting to use). Although there is an implementation of this capability in the DDN PSNs now, it has some technical problems dealing with reestablishment of interrupted activities (section 5.1.3.2) and is rendered ineffective for many subscribers due to the need to allow routing, management and control traffic between routers to use high precedence.

If network enforcement of precedence is deemed necessary, the restriction should be implemented on a pairwise basis as described in section 5.1.3, in order to resolve the reconnection problem. However, it would be simpler to enforce precedence restrictions through external means such as usage sensitive billing or equivalent management reports which would identify the users of high precedence communication and the volume of their usage.

It is further recommended that the enforcement of precedence restrictions on users accessing the DDN through subscriber facilities (such as concentrators) be performed by subscriber organizations, and that enforcement of restrictions on external connections (such as the NSFNET interconnection) be performed by the DDN. DDN operationss and management organizations should review the extent of usage of different precedence levels by concentrators and similar traffic sources to detect possible excessive use of high precedence levels. Suspected problems should be investigated and resolved in conjunction with the subscriber organization.

8.1.3 Requirements for Precedence Selection on Dialup Terminal Access

DOD data communications users should determine whether there will be a requirement to allow dialup terminal access (through the TAC or NAC or similar component) at other than ROUTINE precedence. If such a requirement exists, the dynamic validation requirements need to be determined. If dynamic enforcement is needed, the appropriate mechanisms need to be implemented in the affected components (TAC or NAC, and TACACS) and procedures established for effecting authorizations to use this capability.

8.2 PRECEDENCE ENHANCEMENTS FOR EXISTING DDN COMPONENTS

The software enhancements deemed necessary for adequate support of precedence requirements are listed below for each major DDN component.

8.2.1 DDN PSN Precedence Implementation Enhancement

Signalling of connection precedence on incoming X.25 calls delivered to a DDN-connected device is essential to proper operation of the precedence service in the PSN backbone. This PSN software deficiency needs to be corrected to make the precedence feature useful.

Monitoring and control traffic originated by the PSN needs to be assigned precedence levels and these assignments need to be implemented in the PSN software. This is also a high priority requirement.

Implementation of logical channel preemption (perhaps using the logical channel reservation approach in the PSN) is also desirable but is less urgent.

8.2.2 DDN Mailbridge Precedence Implementation Enhancement

The Mailbridge should be modified to map IP precedence to subnetwork precedence signalling in accordance with the DDN standard mapping, if this has not already been done. It is recommended that the use or non-use of this mapping function be made a configurable option of the Mailbridge software.

Next most important is to administratively determine, and then implement in software, appropriate precedence level settings for management and control traffic, and for routing traffic, originated by the Mailbridge.

Precedence-ordered queue service is a desirable but less critical feature for the Mailbridge so long as it is used only for access to non-operational (in the military sense) activities in external networks. If the Mailbridge ever needs to perform packet switching functions between operational military users who employ high precedence, precedence-ordered queue service should definitely be provided. A specific design needs to be developed by the vendor, but based on the functional specification, it appears that the relevant queueing points in the Mailbridge are at the entry to the forwarding function and at the queues for outgoing links.

8.2.3 DDN TAC Precedence Implementation Enhancement

Precedence levels need to be administratively assigned and then implemented in the TAC software for the monitoring and control traffic originated by the NAC.

The requirement for precedence support for TAC users depends on policy decisions about whether the TAC is to be used to service users with such a requirement. If it is to be so used, a means of selecting precedence for TAC ports or users will have to be incorporated in the TAC software.

8.2.4 DDN NAC Precedence Implementation Enhancement

Precedence levels need to be administratively assigned and then implemented in the NAC software for the monitoring and control traffic originated by the NAC.

8.2.5 Network Information Service Component Precedence Implementation Enhancement

At a minimum, domain name servers servicing DOD domains need to be enhanced to set the precedence of responses or recursive queries to the precedence level of the incoming query that triggered them. The hosts supporting these servers must, at a minimum, provide the IP precedence of incoming requests to the DNS server, and map the precedence selected by the DNS server for outgoing datagrams or connections to the appropriate IP and subnetwork precedence signalling. Any internetwork components (routers) that must be traversed by DOD users to reach DNS servers also must implement the standard mapping of IP precedence to subnetwork precedence if connected to a subnetwork that provides precedence service. It is not currently known which specific components will require enhancement and which ones already provide the necessary features.

Since no existing DNS software is known to provide the precedence features described above, it is recommended that implementation of precedence in the most widely used DNS software be considered for DISA funding. The most logical candidate software for such enhancement is the BIND DNS software provided with the Berkeley networking software for UNIX.

8.3 PRECEDENCE REQUIREMENTS FOR FUTURE DDN COMPONENTS

This section notes the general requirements that would be appropriate to apply when new components are introduced into the DDN. As part of any such acquisition, the employment scenario and user requirements should be carefully reviewed to determine the exact capabilities needed. These proposed requirements represent a starting point for such a review.

8.3.1 Future DDN X.25 Subnetwork Component Requirements

Precedence-ordered queue service should be used for user control and data packets in the network. The subnetwork precedence service provided should be based on X.25 precedence signalling provided by the interface. Signalling options should be configurable per interface to use any one of the following signalling procedures: the DDN X.25 Precedence facility, the CCITT-Specified DTE Priority facility, or a future CCITT X.25 Priority Facility if one has been defined. Four levels of precedence are required for internal processing purposes. Conversion between four-level and 16-level priority signalling should use the CLNP to DDN value mapping.

All management, control, and routing functions should either be set to the highest precedence or should be configurable separately per function.

The network should either provide substantial assurance that all logical channels configured can be used simultaneously (at whatever level of throughput is attainable), or should provide for preemption of logical channels by the DCE when no additional channels can be supported.

8.3.2 Future DDN IP and CLNP Router Requirements

Subnetwork mappings of IP precedence and CLNP priority to all applicable subnetwork technologies should be provided. It should be possible to activate or deactivate the use of the subnetwork mappings under management control, separately for each subnetwork interface.

Precedence-ordered queue service should be provided at the entry to the forwarding function and at the queue for each outgoing transmission link.

8.3.2.1 Further Requirements for IP Routers

The IP precedence level of ICMP messages generated by the router in response to received datagrams should be set as follows:

- Error reports (Destination Unreachable, Time Exceeded, Redirect, and Parameter Problem) should be sent with precedence 7, Network Control.

- All other ICMP responses should be sent with the same IP precedence as the datagram that elicited the response.

8.3.2.2 Further Requirements for CLNP Routers

The Priority option should be included in any ER PDU generated by the router and the priority value should be set as specified by ISO 8473.

ISO 9542 ES-IS datagrams should include a priority option. The priority value should be configurable per interface or per subnetwork. The (recommended) nominal default priority for DOD usage is 14 decimal.

Other routing protocols should use the Priority option, if available. The priority value should be configurable per peer, and the nominal (recommended) default priority for DOD usage is 12 decimal. If no priority option is not defined for use with a particular protocol, the router should assume the maximum precedence for received data units associated with that protocol.

8.3.3 Future Network Management System Components

All operator-controlled management actions should provide the capability for the operator to specify the precedence, unless specific analysis has determined an appropriate fixed precedence for a particular management function and the management component always uses that fixed value for that function.

All automatic management reporting or control actions should use a preassigned precedence level, which should be configurable to any value allowable in the protocol employed.

8.4 PRECEDENCE REQUIREMENTS FOR SUBSCRIBER DEVICES

The following recommendations are provided for subscriber devices and the software used in those devices. Subscriber requirements obviously vary, but these recommendations form a basis for analyzing any special requirements. The full technical description of these features is provided in section 5.

8.4.1 Subscriber Host Computers

All systems acquired by DOD that support the TCP protocol should implement TCP precedence negotiation correctly as described in the TCP protocol specification.

All systems acquired by DOD that have an application programming interface should provide the precedence control features described in section 5.2.2.

All systems that are connected to a DDN PSN and use TCP/IP protocols should implement the standard DDN subnetwork mapping of IP precedence.

Subscriber systems providing domain name service should meet the same requirements as DDN-provided service. Specifically, the precedence level of each query received should be used in all derivative queries and responses to that query. User and application programming interfaces to a domain resolver should allow the user or calling program to select the precedence level of the query. Authorization checks may be made on the precedence value selected.

8.4.2 Subscriber Routers (Concentrators)

All systems that are connected to a DDN PSN and use TCP/IP protocols should implement the standard DDN subnetwork mapping of IP precedence.

The requirement for precedence-ordered queue service for the datagram routing function is a local issue. However, routers used to support sizable communities should have this feature since it is reasonable to expect that some of the users will have requirements for precedence service during the economic life of the device. Routers supporting very small communities may not have such a requirement. When provided, precedence-ordered queue service should be provided (at least) at the entry to the forwarding function and at the queue for each outgoing transmission link.

8.4.3 Subscriber LANs

Subscribers acquiring LANs should consider the bandwidth limitations of the proposed LAN technology, potential traffic surge requirements, and resulting degradation of service under surge conditions. If saturation of the network is a possibility and some users of the network require availability for high precedence activities, a LAN technology that supports a subnetwork priority mechanism should be strongly considered.

LIST OF REFERENCES

References Cited

Aydin Monitor and Computer Systems, October 1988, *Users Guide for the Terminal Access Controller CP-1757/U*, 298-0854-A.

BBN Communications Corporation, August 1987, *Computer Program Functional Specification—Butterfly Mailbridge (MB) Gateway*, Report No. 6513.

Braden, R. and J. Postel, June 1987, *Requirements for Internet Gateways*, RFC 1009, DDN Network Information Center, SRI International, Menlo Park, CA.

Braden, R., ed., 1989a, *Requirements for Internet Hosts—Communication Layers*, RFC 1122, DDN Network Information Center, SRI International, Menlo Park, CA.

Braden, R., ed., 1989b, *Requirements for Internet Hosts—Application and Support*, RFC 1123, DDN Network Information Center, SRI International, Menlo Park, CA,

Defense Communications Agency, May 1982, *Defense Data Network Program Plan*.

Defense Communications Agency, December 1983, *Defense Data Network X.25 Host Interface Specification*.

Defense Communications Agency, June 1989, *Defense Data Network Host Interface Qualification Testing—Link and Network Layers*, DCAC 370-P195-5.

Kleinrock, L., 1975, *Queueing Systems, Volume I: Theory*, New York: Wiley-Interscience.

Kleinrock, L., 1976, *Queueing Systems, Volume II: Computer Applications*, New York: Wiley-Interscience.

Peden, J. H. and A. C. Weaver, August 1988, ‘‘Are Priorities Useful in an 802.5 Token Ring?,’’ *IEEE Transactions on Industrial Electronics*, Vol. 35, No. 3, pp. 361-365.

Scott, D., 1990, *Elements of Subnetwork Access for DDN OSI Protocols*, MTR-90W00125

Other References

Case, J.B., M. Fedor, M.L. Schoffstall, and C. Davin, May 1990, *Simple Network Management Protocol (SNMP)*, RFC 1157, DDN Network Information Center, SRI International, Menlo Park, CA.

Hedrick, C., June 1988, *Routing Information Protocol*, RFC 1058, DDN Network Information Center, SRI International, Menlo Park, CA.

Lougheed, K., and Y. Rekhter, June 1990, *Border Gateway Protocol (BGP)*, RFC 1163, DDN Network Information Center, SRI International, Menlo Park, CA.

Mills, D., April 1984, *Exterior Gateway Protocol Formal Specification*, RFC 904, DDN Network Information Center, SRI International, Menlo Park, CA.

Mockapetris, P., November 1987, *Domain Names—Specification and Implementation*, RFC 1035, DDN Network Information Center, SRI International, Menlo Park, CA.

Postel, J., August 1980, *User Datagram Protocol*, RFC 768, DDN Network Information Center, SRI International, Menlo Park, CA.

U.S. Department of Defense, 12 August 1983, *Internet Protocol*, MIL-STD-1777.

U.S. Department of Defense, 12 August 1983, *Transmission Control Protocol*, MIL-STD-1778.

APPENDIX

OSI PROTOCOL SUPPORT FOR PRECEDENCE AND PRIORITY SELECTION

An approach to implementing precedence signalling using the OSI protocol family is presented here. Table A-1 provides a summary of this information, which is described in greater detail below.

A.1 PHYSICAL LAYER PROTOCOLS

The physical layer service does not provide any functions to which precedence can logically be applied. It simply transmits and receives whatever data stream is provided by the data link layer. Therefore, any precedence-related behavior needed at the ‘‘lowest layer’’ would logically be implemented as a function of the data link layer.

A.2 DATA LINK LAYER PROTOCOLS

A.2.1 Point-to-Point Link Technologies

The use of data link layer protocols for point-to-point connections varies in different networks. It is generally unnecessary to use prioritization in data link layer protocols since the same effects can be obtained by submitting traffic to the data link in prioritized order. However, in networks (such as many X.25 implementations, including the DDN implementation) where data link protocol data units are switched actively in the network, an explicit priority indication on the data link PDU is desired to allow precedence processing to work at intermediate nodes.

The HDLC protocol (ISO 7776) does not provide any priority signaling features. At this time, there is no generally accepted point-to-point link protocol with such features. X.25 networks often use unique protocols within their network backbones, and prioritization may be a feature of such a protocol, as is true in the DDN.

A.2.2 Multipoint Data Link Technologies

In multipoint technologies, it is possible for each transmitter to apply precedence ordering locally and still result in low priority traffic interfering with high priority traffic. This occurs because different transmitting stations may have different mixes of offered load; one transmitter may have only low priority users and another only high priority users. Some multipoint schemes provide a means of arbitrating access among different stations by propagating information about the highest priority data waiting at each station. Other schemes have no such feature. The specifics for several LAN technologies, including the popular protocols developed by Project 802 of the Institute of Electrical and Electronics Engineers (IEEE), are provided below.

Table A-1. Precedence Signalling and Processing Requirements for OSI Layers

Aspect of Precedence Signalling/Processing		Layer 1	2	3	4	5	6	7
Signalling								
Precedence parameter in layer service interface [1]	Req:			R	R	R	R	R
	CO:	N	N	Y	Y	Y	Y	D
	CL:	N	N	Y	Y	Y	Y	D
Signal precedence to peer entity in layer PDUs	Req:		O	R	R			R
	CO:		D	Y	Y	N	N	D
	CL:		D	Y	N	N	N	D
Map precedence parameter to lower layer service	Req:			R	R	R	R	R
	CO:	N	N	A	A	R	R	R
	CL:	N	N	A	A	R	R	R
Report precedence actions to peer entity in layer PDUs [2]	Req:			R	R			R
	CO:	N	N	Y	N	N	N	D
	CL:	N	D	Y	Y	N	N	D
Processing								
Precedence-ordered queueing for short-term resources [3]	Req:		R	R	R			R
	CO:	N	D	U	U	U	U	U
	CL:	N	D	U	U	U	U	U
Preempt connections based on precedence [4]	Req:			R	R			R
	CO:	U	U	U	U	U	U	U
	CL:	U	U	U	U	U	U	U
Block communication below precedence cutoff level	Req:			O	O			O
	CO:	U	U	U	U	U	U	U
	CL:	U	U	U	U	U	U	U
Validate precedence level specified by higher layer	Req:			O	O			O
	CO:	U	U	U	U	U	U	U
	CL:	U	U	U	U	U	U	U

Key:

Req: Required functionality
CO: Provided by connection-oriented protocol
CL: Provided by connectionless protocol
R: Required (in model)
O: Optional (in model)
D: Depends on specific protocol
N: No, incompatible with standard
U: Unspecified by standard
Y: Yes, specified in standard

Notes:

[1] Implies receiving selected precedence from higher layer entity and sending confirmation or error reports back to higher layer entity.
[2] Actions may include cutoff level change, connection preemption, etc.
[3] Includes controlling order of processing within a switch, order of transmission on link, etc.
[4] Or any long-term allocated resource.

To the extent that data link layer traffic is coupled to network layer flow control mechanisms, it is possible to obtain a degree of precedence support without any special data link layer features.

A.2.2.1 IEEE 802.3/Ethernet

These technologies, referred to as Carrier Sense Multiple Access with Collision Detection (CSMA/CD), are essentially random service schemes and do not provide any priority feature. The contention process is quite similar to that in a radio network.

A.2.2.2 IEEE 802.4 Token Bus and IEEE 802.5 Token Ring

These technologies define eight levels of signalled data link layer priority and allow for four ‘‘access classes,’’ which are handled in a queued fashion (the highest priority outgoing packet is sent first).

The 802.4 token bus prioritization scheme has been altered during the lifetime of this standard. The latest version uses token circulation timing to detect whether a priority level is able to be fully served. The token carries an indication of the current priority level being served. If the token takes longer than a (configured) time to complete a transit of the entire bus, the station will not transmit traffic of priority lower than the token. Clearly the effectiveness of the scheme depends on a suitable choice of timer values. A procedure for calculating these times is given in (Gorur and Weaver, 1988).

The 802.5 token ring uses a reservation scheme in which each station signals to the other stations the highest priority of data awaiting transmission from that station. Other stations with only lower priority data recognize this signal and refrain from transmitting for a period of time. Since some time is required for the reservation to be seen by all stations on the ring, there is some delay in gaining access for high priority traffic. One study (Peden and Weaver, 1988) suggests that priorities on 802.5 networks are not useful for delay control, but they do function as a sort of blocking scheme if the total offered load is greater than the network can carry.

A.2.2.3 IEEE 802.6 Distributed Queue/Dual Bus

This new standard is just approaching a stable state. It is expected to be used in so-called Metropolitan Area Networks (MANs). One proposed use is to connect large users to telephone service providers. This is possible with 802.6 because it provides for a mixture of isochronous and asynchronous traffic. The isochronous traffic has (in effect) unconditional precedence, and the asynchronous traffic receives whatever capacity is left. Eight levels of priority are specified for asynchronous traffic.

An unanswered question with this protocol is how to provide useful prioritized bandwidth allocation with a mixture of synchronous and isochronous users. The obvious case of interest would be an 802.6 link which supports both voice telephony using the isochronous service, and data transfer using the asynchronous service. There is no mechanism internal to the 802.6 standard for dynamically varying the portion of the bandwidth assigned to isochronous traffic, so if the preconfigured allocation of bandwidth is inappropriate, a case could arise in which a voice connection of high precedence could not be set up even though enough unused bandwidth was available in the asynchronous service, or

conversely, high precedence data could not take bandwidth away from low precedence voice. The allocations can be changed by management action, but further analysis will be needed when 802.6 implementations are planned to determine whether this characteristic of 802.6 presents a practical problem.

A.2.2.4 Switched Multi-megabit Data Service (SMDS) Interface Protocol

The SMDS Interface Protocol (SIP) is a subset of IEEE 802.6 protocol, with some refinement of the semantics. To date, there are no vendor announcements of intention to support priority processing based on the asynchronous priority field of the 802.6 header.

A.2.2.5 Fiber Distributed Data Interface (FDDI)

FDDI networks provide a priority service of eight signalled levels mapped to four access classes that are actually processed distinctly. The so-called FDDI II, FDDI with the Hybrid Ring Control (HRC) feature, provides mixed isochronous and asynchronous service as described above for the 802.6 network. The priority management process for asynchronous traffic resembles that of the 802.4 token bus. Early FDDI implementations may not allow the target rotation times to be set, which reduces the utility of the priority feature.

A.2.2.6 Frame Relay

None of the frame relay specifications (currently being finalized by various standards groups) include any notion of MLPP-like prioritization. If frame relay is used as a bearer service in ISDN, it would be possible to use the Q.931 signalling elements to indicate precedence (see section A.3.3 below).

A.2.2.7 Broadcast Radio Networks

Broadcast radio networks are typically multiple access networks with properties similar to 802.3 networks (that is, random service). To provide prioritization requires auxiliary protocols for cooperation among transmitters and generally takes the form of a reservation scheme. Since there is no widely used standard for this application and environment, prioritization features are generally one-of-a-kind implementations. The problem is avoided if different transmitters are assigned to different frequencies so that the links are effectively one-to-many, which for priority purposes is equivalent to the point-to-point case. Schemes for prioritized carrier sense multiple access (CSMA) radio have been developed for use in military environments, but the full details are not publicly disclosed.

A.2.3 ISDN

At the data link layer in ISDN, it is only necessary to consider the Link Access Procedure for the Data channel (LAPD). There is no provision for prioritization of access to the D channel. This suggests that ISDN applications that may need precedence should rely on the B channel for data transfer. If this approach is taken, the signalling traffic on the D channel should be small enough not to require any prioritization to provide adequate service.

A.3 NETWORK LAYER PROTOCOLS

A.3.1 Precedence Selection in X.25 Networks

There is no CCITT-specified signalling construct for an X.25 DTE to signal desired precedence or priority to the network. There is a ‘‘CCITT-specified DTE facility’’ in X.25-1988 for Priority indication, but it is oriented to DTE-to-DTE signalling and has not been defined as applicable to the X.25 network’s own behavior.

The implementor of an X.25 network has discretion in how the internal operations of the network actually work. These decisions may result in a design in which data link usage is closely tied to network layer connections, or one in which the use of data links is largely independent of network layer events. In the latter case, the internal data link protocols would need to carry some priority indication in order to allow precedence-based queueing in intermediate nodes of the network. The X.25 standard does not specify any particular protocol stack for internal data links between packet switches of a single X.25 network.

A.3.1.1 DDN Standard X.25

DDN Standard X.25 provides a four-level Precedence facility using a subnetwork-specific facility encoding in the X.25 call request packet.

A.3.1.2 CCITT X.25-1988 Priority Facility

Beginning with ISO 8208 and the 1988 revision of the CCITT recommendations, a CCITT-Specified DTE Facility for Priority is provided. This is intended to be used by the endpoints and not by an intervening subnetwork, at least in the case of public networks. Six 8-bit fields are provided, allowing the caller to propose target and minimum acceptable priority for connection setup, data transfer, and (protection against) disconnection. The standard restricts the values to the range 0 to 14, with value 0 meaning ‘‘not specified’’ and 1 to 14 representing increasing priority.

A.3.2 CLNP

An 8-bit priority value may optionally be included in a CLNP PDU. The ISO 8473 standard currently restricts the coding to values 0 through 14 decimal, with 14 being the highest priority.

A.3.3 ISDN

The potential diversity of services under ISDN makes a general description almost impossible. Superficially, though, ISDN can be characterized as having circuit-switched connections, packet-switched connections using X.25, and some datagram-like services. Standards for MLPP features for the circuit-switched connections are now being finalized by CCITT. The feature is described in a forthcoming Recommendation I.255.3, and it operates by including a three-bit precedence value (with five levels defined) in the Q.931 call setup packet. The MLPP feature is strictly optional and it may or may not be supported by a particular ISDN network.

The application of priority to X.25 and frame relay services in ISDN remains to be worked out. The signalling procedure of Recommendation I.255.3 could be used by data service subscribers if supported by the network being used. If an X.25 facility for priority or precedence were to be defined by CCITT, the relationship of X.25 and Q.931 signalling would need to be defined. This problem does not arise for frame relay because there is no concept analogous to X.25 optional facilities in the frame relay protocol.

A.4 TRANSPORT LAYER PROTOCOLS

A.4.1 Connection-Oriented Transport Protocols

The ISO 8073 connection-oriented transport protocols provide an optional 16-bit priority value in which value 0 is defined to have the highest priority. This ordering is different from the Network Layer ordering of values. This option is available in all classes of the protocol. The standard does not specify how the calling and called transport entities reconcile a disagreement about the desired priority level.

A.4.2 Connectionless-mode Transport Protocol

The ISO 8602 Connectionless Transport Protocol does not include any syntax for signalling priority. Whether this is a problem depends on user requirements. One possible implementation solution is to map the priority specified at the transport service interface to the network service interface unaltered (and perform the mapping in reverse for received data). This approach seems technically acceptable, but no standard requires any implementation to act in this manner.

A.5 SESSION LAYER PROTOCOLS

Priority is defined as a QOS feature provided at the session service interface, and the standard requires that it be mapped to the transport layer unaltered.

A.6 PRESENTATION LAYER PROTOCOLS

As with the session layer, priority is defined as a QOS feature provided at the presentation service interface, and the standard requires that it be mapped to the session layer unaltered.

A.7 APPLICATION LAYER PROTOCOLS

A.7.1 X.400

X.400 provides three levels of importance for messages. The mapping of these levels of importance to the presentation service is not identified. There is a question as to whether three levels are enough for military messaging applications. The Defense Message System (DMS) project is considering this

issue. The encoding provides enough room to accept larger values, and it would also be possible to define another option in the message PDU to contain a special military precedence indication.

The 1988 X.400 specification provides the mapping of message importance indications to Remote Operations Service Element (ROSE) priorities. This might need to be reworked if more message precedences are to be supported.

A.7.2 X.500

X.500 also allows for three importance levels, which correspond directly to those in X.400.

A.7.3 Other Applications

In other applications such as File Transfer, Access, and Management (FTAM) and Virtual Terminal (VT), the protocol specifications make no reference to priority. However, it would be possible for implementations to allow users to select priority levels and then map the selected level into a priority value to be presented to the presentation service interface. This is essentially the same process used by an application in the TCP/IP environment.

A.8 REFERENCES

Gorur, R. M. and A. C. Weaver, August 1988, ‘‘Setting Target Rotation Times in an IEEE Token Bus Network’’, *IEEE Transactions on Industrial Electronics*, Vol. 35, No. 3, pp. 366-371.

Peden, J. H. and A. C. Weaver, August 1988, ‘‘Are Priorities Useful in an 802.5 Token Ring?’’, *IEEE Transactions on Industrial Electronics*, Vol. 35, No. 3, pp. 361-365.

GLOSSARY

AHIP	ARPANET Host Interface Protocol
ARP	Address Resolution Protocol
ASC	AUTODIN Switching Center
AUTODIN	Automatic Digital Network
AUTOVON	Automatic Voice Network
BBN	Bolt, Beranek and Newman
BFE	BLACKER Front-End
BGP	Border Gateway Protocol
BIND	Domain name server software package developed at the University of California at Berkeley
BLACKER	NSA programs to produce end-to-end encryption equipment
CCITT	International Consultative Committee for Telegraphy and Telephony (Comite Consultatif Internationale de Telegraphique et Telephonique)
CLNP	Connectionless Network Protocol (ISO 8473)
CLNS	Connectionless Network Service
CONP	Connection-Oriented Network Protocol (ISO 8208)
COTP	Connection-Oriented Transport Protocol (ISO 8073)
COTS	Commercial Off-the-Shelf
CPU	Central Processing Unit
CSMA	Carrier Sense Multiple Access
CSMA/CD	Carrier Sense Multiple Access with Collision Detection
DCA	Defense Communications Agency
DCAC	Defense Communications Agency Circular
DCE	Data Circuit-terminating Equipment
DDN	Defense Data Network
DIS	Draft International Standard
DISA	Defense Information Systems Agency
DISNET	Defense Integrated Secure Network
DMS	Defense Message System
DNS	Domain Name System
DOD	Department of Defense
DSN	Defense Switched Network
DSNET	Defense Secure Network (either DSNET1, DSNET2, or DSNET3)
DT	Data Transfer (in OSI protocols)
DTE	Data Terminal Equipment

ECP/CRITIC	Emergency Command Procedure/Critical
EE	End-to-End
EGP	Exterior Gateway Protocol
ER	Error Report (in OSI protocols)
ES-IS	End System to Intermediate System (routing protocol)
FCFS	First Come, First Served (scheduling discipline)
FDDI	Fiber Distributed Data Interface
FTAM	File Transfer, Access and Management Protocol
FTP	File Transfer Protocol (MIL-STD 1780)
GOSIP	Government Open Systems Interconnection Profile
HDLC	High-level Data Link Control
HOL	Head of Line (scheduling discipline)
HRC	Hybrid Ring Control
ICMP	Internet Control Message Protocol
ID	Identification
IDRP	Inter-Domain Routing Protocol
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol (MIL-STD 1777)
IS-IS	Intermediate System to Intermediate System (routing protocols)
ISDN	Integrated Services Digital Network
ISO	International Organization for Standardization
ITDN	Integrated Tactical-Strategic Data Network
LAN	Local Area Network
LAPD	Link Access Procedure for the D channel
LCFS	Last Come, First Served
LDS	Local Distribution System
LLC	Logical Link Control
MAN	Metropolitan Area Network
MC	Monitoring Center
MILNET	DDN network for unclassified DOD traffic formed from military portion of ARPANET
MLPP	Multi-Level Precedence and Preemption
MTA	Message Transfer Agent
MTU	Maximum Transmission Unit

NAC	Network Access Component
NIC	Network Information Center
NSFNET	National Science Foundation Network
OSI	Open Systems Interconnection
OSPF	Open Shortest Path First (routing protocol)
PC	Personal Computer
PCP	Packet Core Protocol
PDU	Protocol Data Unit
PSN	Packet Switch Node
QOS	Quality of Service
RIP	Routing Information Protocol
RNR	Receive Not Ready (X.25 protocol)
ROSE	Remote Operations Service Element
RR	Receive Ready (X.25 protocol)
RS	Random Service
SIP	SMDS Interface Protocol
SJF	Shortest Job First
SMDS	Switched Multi-megabit Data Service
SNMP	Simple Network Management Protocol
TAC	Terminal Access Controller
TACACS	MILNET TAC Access Control System
TCP	Transmission Control Protocol (MIL-STD 1778)
TCP/IP	Transmission Control Protocol/Internet Protocol
TDMA	Time-Division Multiple Access
TP	Transport Protocol (ISO 8073)
TP4	Transport Protocol class 4 (ISO 8073)
UDP	User Datagram Protocol
UNIX	Operating system developed by Bell Laboratories
VC	Virtual call
VT	Virtual Terminal (protocol)
WHOIS	DDN user directory service

DISTRIBUTION LIST

INTERNAL

W010

R. P. Granato
J. S. Quilty
J. M. Ruddy

W030

J. Brand
J. N. Daigle
G. K. Holt
T. H. Nyman
F. J. Powers
J. Rubin
A. M. Schoka
L. L. Stine
D. C. Wood
D. P. Woodall

W031

W. W. Barns (20)
S. G. Bhanji
P. G. DeShazo
M. L. Fidler
D. A. Gomberg
R. V. Hansen
R. Khare
W. D. Lazear
E. B. McCoy
R. K. Miller
B. C. Mishra
G. A. Reichlen
R. Wilmer
Technical Staff

W033

J. M. Gravallese
J. E. Just
S. J. Turner

W034

R. C. Evans
R. C. Pesci

W035

T. E. Bachand
C. E. Bowen
M. C. Meltzer
A. I. Messeh
C. J. Pennington
J. B. Price
R. J. Ward
H. W. Williams
E. D. Zeisler

W036

D. J. Jurenko
W. C. Kinzinger

W037

H. S. Marsh
D. E. Zugby

G110

P. J. Brusil
A. B. Murphy

MITRE Washington Library

EXTERNAL

DISA/DNSO

COL S. Thacher, Code DIS
Mr. E. Schonborn, Code DISA
Mr. J. Mensch, Code DISB
Mr. E. Doughty, Code DISD
CDR A. Montemarano, Code DISDA
LtCol J. Bennett, Code DISDC
Mr. R. Torezan, Code DISDG
Ms. G. Wix, Code DISI
Mr. T. Clarke, Code DISM
LtCol L. Marier, Code DISP
Mr. J. Milton, Code DOD
Mr. W. Hawrylko, Code DR
Mr. E. Hanz, Code DRF
Mr. D. Ribar, Code DRF
LtCol M. Payne, Code DRF
LtCol D. Winchell, Code DRFD
Mr. G. Bradshaw, Code DRFD
Mr. E. Cain, Code DRFE
Mr. R. Cleary, Code DRFE
Mr. M. Gross, Code DRFE
Mr. C. Morgan, Code DRFE
Mr. D. Morris, Code DRFE
Mr. W. Showalter, Code DRFE
Mr. J. Tontonoz, Code DRFE
Mr. H. Perkins, Code DRFF
Mr. W. Arey, Code DRFFE
Mr. J. Nowakowski, Code DRFFE
Mr. R. Price, Code DRFFE
Mr. J. Thomas, Code DRG

ASD(C3I)

Ms. O. Elliott, ASD(C3I)/IS
Ms. D. Fountaine, ASD(C3I)/IS