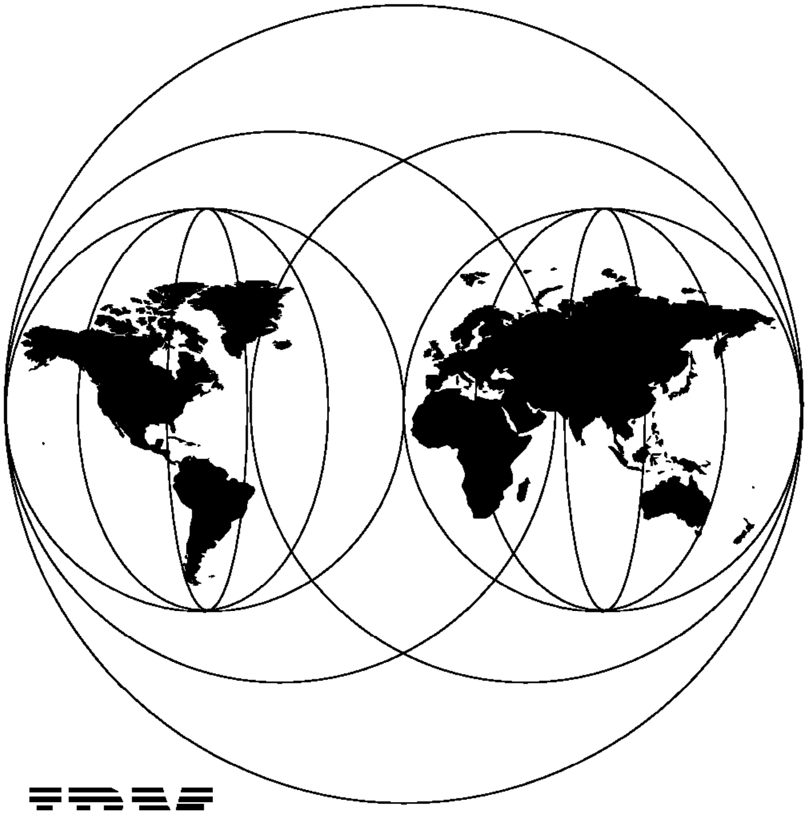
SG24-4753-00

Local Area Network Concepts and Products: LAN Architecture

May 1996



International Technical Support Organization
Raleigh Center

International Technical Support Organization

SG24-4753-00

Local Area Network Concepts and Products: LAN Architecture

May 1996

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix B, “Special Notices” on page 233.

First Edition (May 1996)

This edition applies to the most recent IBM LAN products and LAN architectures.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. HZ8 Building 678
P.O. Box 12195
Research Triangle Park, NC 27709-2195

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Preface

Local Area Network Concepts and Products is a set of four reference books for those looking for conceptual and product-specific information in the LAN environment. They provide a technical introduction to the various types of IBM local area network architectures and product capabilities. The four volumes are as follows:

- SG24-4753-00 - *LAN Architecture*
- SG24-4754-00 - *LAN Adapters, Hubs and ATM*
- SG24-4755-00 - *Routers and Gateways*
- SG24-4756-00 - *LAN Operating Systems and Management*

To obtain all four books, order the set SK2T-1306.

These redbooks complement the reference material available for the products discussed. Much of the information detailed in these books is available through current redbooks and IBM sales and reference manuals. It is therefore assumed that the reader will refer to these sources for more in-depth information if required.

These documents are intended for customers, IBM technical professionals, services specialists, marketing specialists, and marketing representatives working in networking and in particular the local area network environments. Details on installation and performance of particular products will not be included in these books, as this information is available from other sources.

Some knowledge of local area networks, as well as an awareness of the rapidly changing intelligent workstation environment, is assumed.

How This Redbook Is Organized

The redbook is organized as follows:

- Chapter 1, “Today′s LAN”

  LAN overview and introduction to today′s LAN environment.

- Chapter 2, “Physical LAN Attachment”

  This provides a conceptual overview of physical LAN attachment options, transmission techniques and network topologies.

- Chapter 3, “LAN Architectures and Standards”

  This chapter describes LAN architectures and standards. Information on the latest IEEE 802 standards as well as ATM and wireless forums is also provided.

- Chapter 4, “LAN Protocols”

  This chapter provides information about LAN protocols, which includes network, transport, management and router protocol summaries.

- Appendix A, “IEEE 802 Documents Summary”

  This appendix describes the currently available IEEE 802 documents.

The Team That Wrote This Redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization Raleigh Center.

The advisors for this project were:

Ricardo Haragutchi
International Technical Support Organization, Raleigh

John Parker
International Technical Support Organization, Raleigh

The authors of this document were:

Edmilson Barbosa
IBM Brazil

Ingvar Hyleborg
IBM Sweden

Jefferson da Silva
IBM/GSI Brazil

Klaus Wichmann
ITSO Raleigh

Marcello Belloni Gomes
IBM Brazil

Thanks to the following people for their invaluable advice and guidance provided in the production of this document:

Toshi Shimizu
International Technical Support Organization, Austin

Aroldo Yai
Barry Nusbaum
Donna Fox
Fergus Stewart
Jose Boo
Juan Rodriguez
Mark DeCain
Mohammad Shabani
Robert Macgregor
Stephen Breese
Volkert Kreuk
International Technical Support Organization, Raleigh

Alan Millard
Arthur Bond
Bert Wendle
Carol Carson
Dean Stockwell
Erik Dixon
H. Parrish

Paul Carter
IBM Research Triangle Park, Raleigh NC.

Comments Welcome

We want our redbooks to be as helpful as possible. Should you have any comments about this or other redbooks, please send us a note at the following address:

redbook@vnet.ibm.com

Your comments are important to us!

Chapter 1. Today′s LAN

Local area networks or LANs enable computer-based equipment to communicate and share resources. The products that make up a LAN consist of computers, adapters, media and software.

Most modern businesses today are using computers and intelligent workstations to remain competitive and take advantage of the information necessary to run a profitable business in a world of information technology that is evolving rapidly.

Industry trends suggest a dramatic increase in the use of intelligent workstations such as notebooks, portable and desktop computers. Vital data is being moved from the mainframe out to these machines to provide efficient and cost effective client/server functionality.

The resizing and reshaping of this information infrastructure will require faster and more reliable access to data which in turn will drive new applications based on multimedia incorporating voice, data and video. LANs are part of this emerging environment and will accommodate high-speed, multiuser connections.

New business applications are now being considered due to new services becoming available such as Asynchronous Transfer Mode (ATM) and frame relay. The environment we have today is represented by a vast array of protocols, topologies, transmission speeds, access interfaces, physical media and equipment and it is with these new transport technologies that businesses will be able to choose the most appropriate applications without being restricted by networking infrastructure.

Today′s LAN is evolving rapidly due to technological and business requirement changes. Voice, data and video are already a reality giving LANs the capability to be more than just a connection facility for workstations and printers. Competition in the business world will be a major driver in determining how fast the technology will evolve as we move into the gigabit and real-time data streams of voice, data and video.

1.1 LAN Capabilities

One of the main capabilities of a local area network is resource sharing, such as data and expensive peripherals. This ability to share resources can mean a decrease in the cost of an individual workstation, since not every workstation may need its own printer or hard disk.

A LAN may also provide better reliability and availability than a centrally controlled network. The failure of a workstation does not affect other users on the LAN unless the failure point is a server. A workstation on a LAN usually has more flexibility and function than a fixed function terminal connected to a host system. For example, the LAN user can frequently work as a stand-alone workstation, share resources on the LAN, or be in session with systems external to the LAN. Because the LAN attachments must have some degree of intelligence to support the LAN protocols, they may also provide management or problem determination support through those protocols.

1.2 Planning for a LAN

In planning for a local area network consider the following:

- The organization′s objectives
- End-user requirements
- Number and location of workstations
- Requirements for special functions or capabilities such as backup
- Security
- Network and systems management

1.2.1 Objectives for Providing the LAN

The LAN planner must know what problems are to be resolved and what expectations have been identified as reasons for installing the LAN. Will a LAN improve productivity and reduce costs? Most users will agree that beyond two or three workstations and a couple of printers, having all components connected is logical and cost effective.

1.2.2 End-User Requirements

The planner must determine which functions end users require from their workstations, and from LAN-provided servers. If an office system is required, what exactly does the end user want from the system? Is it local word processing, or is access to other office systems required via some sort of gateway or server? If the end users need to access a host system, how is it to be done and how many users require this function? It′s a good idea to get your local systems engineer in at this point to talk about the issues and work out what may be nice to have and what is really needed.

The person planning the LAN will need to look at the number and locations of users that require these types of function so that gateways and servers can be planned and sized accordingly. There is no point in having a very fast piece of cable between workstations and printers if the LAN operating system services or even the hardware used for the server is inadequate.

The number of workstations and their locations need to be considered so that the topology of the LAN can be worked out. It is important to do this to provide sufficient capacity for performance, and adequate availability and recovery capabilities.

The number of users wanting to share the same file(s) and printers needs to be considered. Some files may need to be kept on multiple file servers to distribute the load and ensure acceptable response times. For performance and security reasons, it may be beneficial to have users divided into workgroups or location groups, each on their own LAN with bridges or gateways between the LANs. If gateways to other systems are needed, the number of users and the type of usage should be looked at to determine the amount of traffic flow through the gateway, and hence the number and capacity of gateways required. LAN interconnection is described in more detail in Chapter 1 of *Local Area Network Concepts and Products:Routers and Gateways,* SG24-4755.

1.2.3 Special Functions

The need for the following types of functions should be identified:

- Access to public switched networks
- Access to external databases
- Application-to-application communication
- The ability to store files and use printers on a system external to the LAN
- Remote LAN Access - Dial-in and Dial-out LAN services

1.2.4 Security

The need for security can affect the topology and type of LAN used. One form of security is to use a number of small LANs (divided by workgroup) and to limit the connectivity between the LANs using gateways. Another consideration is the choice of protocol and media on which the LAN is based. For example, in a LAN using a carrier sense multiple access with collision detection (CSMA/CD) protocol like the IEEE 802.3 standard, signals are broadcast simultaneously to all stations on the LAN. This may present security problems if device implementations do not ensure uniqueness of addresses, and acceptance of only those frames addressed to the device. Intelligent hubs today do provide a lot more security and management features to help overcome this issue. Unless care is taken to ensure that duplicate addresses do not exist, it is possible for messages to be received by multiple stations. In an IEEE 802.5 token-ring LAN, the signal is passed serially from one station to another around the ring, but is received only by stations with matching destination addresses. The adapter initialization process prescribed by the IEEE standards ensures that duplicate addresses do not coexist on a local ring segment.

1.2.5 Network and Systems Management

As the size and complexity of the local area network grows, the need for network management and supporting tools becomes more important. Network management implies more than problem determination; it also includes system and change management.

While fast resolution of problems is important, the ability to fix potential problems before they occur is even more desirable. This involves monitoring the network and looking at the utilization of servers, routers, bridges, and gateways. Systems and change management to support planning of application and network configuration changes is important in all networks, and particularly, large networks.

Consideration should be given to the degree of network management required. LAN designs and supporting tools should be selected to achieve this.

1.3 Networking Model

The 802 Project of the Institute of Electrical and Electronics Engineers (IEEE) has produced a set of standards for LAN architectures, giving vendors guidance for producing LAN products and users a choice of standardized local area networks with a certain degree of interconnectivity. The IEEE LAN standards align with the bottom two layers of the International Organization for Standardization (ISO) model for Open Systems Interconnection, referred to as the OSI Reference Model.

In general, OSI defines standards by which computers can communicate together. OSI describes the architecture, protocols, and services that are needed to achieve this goal. There are multiple OSI standards. Some of these are complete, while others are still evolving.

The term *open system* in OSI defines a computer system that can communicate with another computer system using the OSI protocol. What this means is that computer systems, having different operating systems which process data differently, can still communicate and interpret information upon receipt, if the information that passes between their processors conforms to the ISO international standards.

The foundation of the OSI architecture is a layering concept, called the OSI Reference Model. Each layer in the OSI Reference Model has a name, a number, protocols that provide specific functions, and defined services. Because the various intended uses of OSI are very broad, spanning terminals, personal computers, and very large mainframes, different services and protocol options are available at each layer. This range of support can accommodate different connection requirements and environments.

The Reference Model defined by OSI is also an excellent model to understand how networks work. Although there are many different architectures, standards and models the OSI Refernce Model is mostly used to explain the different functions implemented in protocols from different layers and how these protocols work together. The concept part of this book is structured by the OSI Reference Model, starting with the lowest layer (the physical layer).

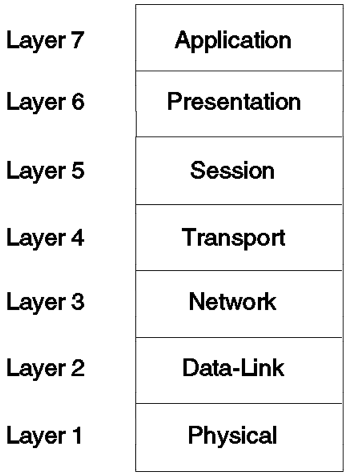


Figure 1. The Seven Layers of the OSI Reference Model

The ISO Reference Model, depicted in Figure 1 became an international standard in 1983 (IS 7498). Each layer addresses a well-defined section of the total architecture. The layers of the OSI Reference Model are, from top to bottom:

- **Application Layer**

  The application layer gives the user access to all the lower OSI functions. The purpose of this layer is to support semantic exchanges between applications existing in open systems.

- **Presentation Layer**
  The presentation layer is concerned with the representation of user or system data. This includes necessary conversions (for example, printer control characters) and code translation (for example, ASCII to or from EBCDIC).
- **Session Layer**
  The session layer provides mechanisms for organizing and structuring interaction between applications and/or devices.
- **Transport Layer**
  The transport layer provides transparent and reliable end-to-end data transfer, relying on lower layer functions for handling the peculiarities of the actual transfer medium.
- **Network Layer**
  The network layer provides the means to establish connections between networks. The standard also includes procedures for the operational control of inter-network communications and for the routing of information through multiple networks.
- **Data Link Layer**
  The data link layer provides the functions and protocols to transfer data between network entities and to detect (and possibly correct) errors that may occur in the physical layer.
- **Physical Layer**
  The physical layer is responsible for physically transmitting the data over the communications link. It provides the mechanical, electrical, functional and procedural standards to access the physical medium.

This layered approach was selected as a basis for the OSI Reference Model to provide flexibility and open-ended capability through defined interfaces. The interfaces permit some layers to be changed while leaving other layers unchanged. In principle, as long as standard interfaces to the adjacent layers are adhered to, an implementation can still work. For example, a system implementation could use either HDLC or local area network protocols as the data link layer. Similarly, a particular layer such as the presentation layer, can be implemented as a null layer for the time being. This means the layer is functionally empty, providing only the mandatory interfaces between the upper and lower layers (application and session layers respectively).

Chapter 2. Physical LAN Attachment

The purpose of physical LAN attachment is to provide an interface that allows the transmission medium (cable) and the LAN station′s access protocol to communicate. It provides mechanical, electrical, functional and procedural specifications for implementing a LAN.

The mechanical definition specifies the type of connectors to be used. The electrical definition, if applicable, states what voltages are to be used, while the functional definition defines the meaning of the voltages on the different pins. Finally, the procedural definition states the sequence of events to be followed to transmit and receive data, including the encoding scheme specifying how digital data is to be represented by electrical or optical pulses. Many different types of media can be used for the physical layer. For example, telephone twisted pair, coax cable, shielded copper cable and fiber optics are the main types used for LANs. Different transmission techniques generally categorized as baseband or broadband transmission may be applied to each of these media types.

2.1 Cable Types

This section will discuss the cable types that are used in constructing networks. Included are coaxial cable, unshielded twisted pair, shielded twisted pair, and fiber.

2.1.1 Coaxial Cable

Coaxial (coax) cable, although not as popular today, has low attenuation characteristics and can drive signals at high data rates for relatively long distances. Coax is an unbalanced cable (the two conductors have a different impedance to ground) with the shield being used as one of the conductors. This structure (a central core with a shield around it) does not generate as much Radio Frequency (RF) noise as Unshielded Twisted Pair (UTP) does when used with high data rates. Coax cable can be used for both baseband and broadband transmissions. 50 ohm coaxial cable is not recommended for new cabling installations and is expected to be removed from the next revision of the ANSI/TIA/EIA standard.

2.1.2 Unshielded Twisted Pair (UTP Category 3, 4, 5)

Telephone twisted pair (unshielded, voice-grade twisted pair) cable can be used for data transmission when data signal strength is filtered, and distances are restricted. It is used by many Private Automatic Branch Exchange (PABX) manufacturers to carry voice and data. Many advances have occurred in this area in the last five years with a lot of effort concentrated on improving performance of UTP cables.

Unshielded telephone twisted pair (UTP) cable tends to suffer from high attenuation (loss of signal strength due to inherent media characteristics) and is very susceptible to noise if located near strong electromagnetic fields (power cables, etc.). A result of attenuation is a reduction in the drive distance, number of attachments and bandwidth potential of the LAN using this medium. When used for high rates of data transmission (for example, 1 Mbps or higher), it will radiate RF emissions. Filters can be used to reduce this. However, filters increase loss in signal strength and add to the cost of the cabling and attachments. UTP wire also suffers from crosstalk between adjacent twisted pairs.

The type 3 specification for telephone twisted pair was the first to provide high frequency capabilities. It has no crosstalk specifications and no attenuation specifications above 1 MHz. It was basically designed for customers who wanted to use existing cabling for token-ring transmission.

The next advance came in 1991 with the development of EIA/TIA 568 (Commercial Building Telecommunications Wiring Standard). The specifications in this standard were developed with input from the IEEE Project 802 Local Area Network Standards committees. The UTP specifications are a lot more comprehensive than the type 3 specifications in that both crosstalk and attenuation characteristics through 16 MHz are specified. EIA/TIA 568′s category 3 UTP cable specification should be used as an update and a replacement for IBM′s type 3 specification.

Shortly after these specifications were adapted, cabling manufacturers developed higher performance UTP cables. The EIA/TIA issued the Technical System Bulletin TSB-36 which defined 5 categories of UTP cables. Categories 1 and 2 for voice and low-speed data on category 3 as described above, and categories 4 and 5 with specifications up to 20 and 100 MHz respectively. The categories 4 and 5 cables, in addition to being specified at higher frequencies, have lower attenuation and crosstalk.

IBM has announced the following enhancements to its 16 Mbps token-ring network offerings to allow for operation on UTP:

- A new family of token-ring 16/4 media filters and a new active hub technology that supports a re-clocked token-ring.
- Existing token-ring network 16/4 adapters are being supported for use with this new active hub technology by connection through new external media filters.
- Support for the token-ring active hubs with these new filters.
- Support for the token-ring passive hubs with new filters.
- New IBM adapters now offer on-board media filters.

Recent tests with data-grade UTP cabling from the desktop to the wiring closet have shown the potential to support reasonable transmission distances. The performance capabilities of a 16 Mbps token-ring using UTP depend on the specific electrical characteristics of the cable and its terminations. Older cable installations should be tested for their suitably to support 16 Mbps transmission.

2.1.2.1 Attenuation of Horizontal UTP Cable

Table 1 (Page 1 of 2). Maximum Attenuation dB per 305 m (1000 ft) at 20°C

Frequency MHz	Category 3 dB	Category 4 dB	Category 5 dB
0.064	2.8	2.3	2.2
0.256	4.0	3.4	3.2
0.512	5.6	4.6	4.5
0.772	6.8	5.7	5.5
1.0	7.8	6.5	6.3

Table 1 (Page 2 of 2). Maximum Attenuation dB per 305 m (1000 ft) at 20°C

Frequency MHz	Category 3 dB	Category 4 dB	Category 5 dB
4.0	17	13	13
8.0	26	19	18
10.0	30	22	20
16.0	40	27	25
20.0	--	31	28
25.0	--	--	32
31.25	--	--	36
62.5	--	--	52
100.0	--	--	67

2.1.2.2 NEXT Horizontal UTP Cable

Table 2. Minimum NEXT Loss Worst Pair at 305 m (1000 ft) at 20°C

Frequency MHz	Category 3 dB	Category 4 dB	Category 5 dB
0.150	54	68	74
0.772	43	58	64
1.0	41	56	62
4.0	32	47	53
8.0	28	42	48
10.0	26	41	47
16.0	23	38	44
20.0	--	36	42
25.0	--	--	41
31.25	--	--	40
62.5	--	--	35
100.0	--	--	32

All TIA standards, including TSBs, are available for a fee if you need additional detail through Global Engineering Documents, USA and Canada (1-800-854-7179) or International (1-714-261-1455).

2.1.2.3 Two IBM 8230 Solutions Meeting the Needs of UTP Categories

IBM is, in fact, supporting 16 Mbps token-ring speed using IBM 8228 Multistation Access Units (MAUs) with new external filters as well as the IBM 8230; however this section will focus only on the new 8230 solutions. For additional information on 8228s as well as detailed information on the 8230 cabling, please refer to the supplement to the wiring guidelines, *IBM Token-Ring Network Introduction and Planning Guide,* GA27-3677.

There are two models of the IBM 8230 Controlled Access Unit:

- The IBM 8230 Model 1 supports higher grade categories 4 and 5 UTP cables.

- The IBM 8230 Model 2 provides concentrators with retiming for each lobe to support categories 3, 4, and 5 cables at lobe lengths that will meet the needs of most installations.

2.1.2.4 Media Filters for 16 Mbps Token-Ring on UTP

Media filters must always be used for categories 3, 4 and 5 UTP cable. Transmission of high data rate signals on UTP requires filtering to stay within national emission limits for electrical energy. External filters are required on IBM′s 16/4 Token-Ring adapters and corresponding concentrator ports to permit attachment to UTP lobes operating at either 4 or 16 Mbps. In addition to the filters that attach a token-ring adapter, a filter is needed for the IBM 8230 Controlled Access Unit for 4 or 16 Mbps token-ring operation on UTP lobes.

2.1.2.5 Configurations without Retiming Concentrators

Current testing of the IBM 8230 Controlled Access Unit Model 1 with the new 16/4 media filters, indicates that the maximum allowable attachment limits and maximum lobe lengths as a function of UTP grade are as shown in Table 3.

Table 3. Maximum Lobe Length per UTP Grade

	At 4 Mbps:	**At 4 Mbps:**	**At 16 Mbps:**	**At 16 Mbps:**
UTP Category	MAX number of attach.	MAX lobe length (meters)	MAX number of attach.	MAX lobe length (meters)
3	72	100	n/a	n/a
4/5	132 *	100	132	100

Note: The maximum attachment limit is 72 stations if there are any 4 Mbps only adapter cards in the ring.

2.1.2.6 IBM 8230 Model 2 Support for Categories 3, 4, and 5 UTP at 16 Mbps

To achieve desired distances and station counts over category 3 UTP cable, an active concentrator such as the IBM 8230 Model 2 that provides signal regeneration with retiming is required. Table 4 gives the maximum supportable lobe lengths using retiming concentrators.

Table 4. Maximum Supported Lobe Lengths Using Retiming Concentrators

	At 4 Mbps:	**At 4 Mbps:**	**At 16 Mbps:**	**At 16 Mbps:**
UTP Category	MAX number of attach.	MAX lobe length (meters)	MAX number of attach.	MAX lobe length (meters)
3	132 *	100	132	90
4/5	132 *	100	132	100

Note: The maximum attachment limit is 72 stations if there are any 4 Mbps only adapter cards in the ring.

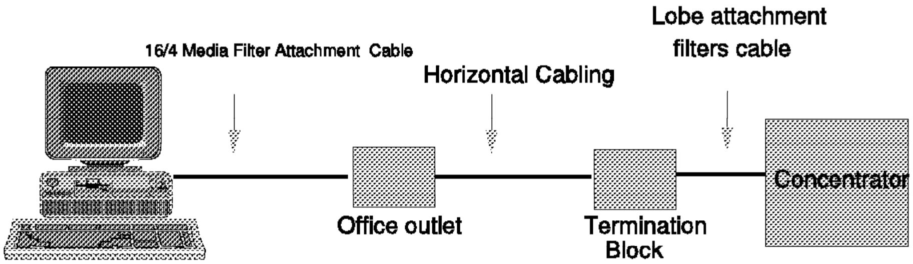
As shown in Table 3 and Table 4 , the supported cabling distances for all 4 Mbps operation and for all 16 Mbps operation on data-grade (Category 4 or 5) UTP, meet the requirements of the EIA/TIA Commercial Building Telecommunications Wiring Standard. In addition, the 8230 Model 2 with category 3 UTP will meet almost all the distance requirements for installations using category 3 cable that adhere to the wiring standard.

2.1.2.7 Multiple Wiring Closets

It is recommended that the main ring path of UTP 16 Mbps token-rings remain within a single wiring closet. However, if required, main ring paths can span multiple wiring closets with the use of repeater functions. STP or multimode optical fiber must be used for all main ring connections.

2.1.2.8 Termination of UTP Cabling

To achieve the distances described above, it is necessary to terminate the UTP cabling properly. Figure 2 shows a block diagram of a token-ring UTP lobe connection.



Lobe attachment filters cable
16/4 Media Filter Attachment Cable
Horizontal Cabling
Concentrator
Office outlet
Termination Block

Figure 2. A Block Diagram of a Token-Ring Network UTP Lobe

Both the device attachment cable and the concentrator attachment cable must be the specially designed high-quality cables specified by IBM and must be installed following IBM′s specified procedures. The horizontal cabling must be terminated at the telecommunications outlet in a high-quality modular connector. To minimize crosstalk, the cable twist should be maintained up to the termination points. The termination block should be a high-quality insulation displacement-style telephone block. Again, the fixed cabling should be untwisted no more than is necessary to make the termination. To assist in limiting crosstalk and attenuation, no intermediate blocks should be used. The horizontal cabling from the termination block to the telecommunications outlet must be dedicated to the token-ring signals for that single lobe connection. No other voice or data signals are permitted in the same cable sheath, to prevent crosstalk among the pairs.

2.1.2.9 Presently Installed UTP Cabling

If you are planning a network on previously installed UTP cabling, you need to know whether that cable was manufactured according to any of the applicable EIA/TIA UTP cabling category specifications. Fortunately, that assessment should be relatively straightforward. All cables should be clearly labeled by both the manufacturer and the installer. Markings on the cable sheath (accessible in the wiring closet) should clearly indicate the manufacturer of the cable and the manufacturer′s model number for that cable. The manufacturer can tell you if that model number meets any of the EIA/TIA cable category specifications.

The installer′s label should provide a way of determining the length of the permanently installed cable and its termination points. If any part of the manufacturer or installer′s labeling is missing, a hand-held cable tester, available from a variety of vendors, can help you determine cable length, attenuation, and crosstalk over a frequency range from 1 to 16 MHz. Some testers are designed to test the suitability of the cable for a specific application, such as 16 Mbps token-ring on UTP. Testers can alert you to incorrect installation procedures, as well as poor quality cable. A good general rule is to test any cable that has been installed for three or more years.

2.1.3 Data-Grade Media (DGM), Shielded Twisted Pair Cable (STP)

This type of cable has one or more twisted pairs within a shield. The shielding reduces its susceptibility to low levels of noise and its own generation of radio frequency interference, thus making it more suitable for data transmission. Shielded data-grade cable can be used with data rates in excess of 20 Mbps over most distances encountered within buildings. The cable can be constructed with two twisted pairs within a shield, while maintaining a low level of crosstalk due to the shielding and the way in which the pairs are twisted around each other. In addition to providing two data paths, the twisting and shielding of DGM cable provides greater immunity to external interference than coaxial cables that use the shield as one of the conductors. Data-grade twisted pair cable is a balanced medium better suited to the differential encoding schemes used by some LANs.

While this type of cable can be used for baseband and broadband transmission, it is primarily used for baseband. The IBM Token-Ring Network uses DGM STP (150 ohm) media for both 4 and 16 Mbps LAN segments.

2.1.3.1 New Improvements for Shielded Twisted Pair

Overall, 150 ohm STP cable, designed in the early 1980s, still stands out as a first class cabling choice for high-speed data transmission. Its high characteristic impedance is fundamental in achieving low transmission attenuation. The shield is needed to transmit the highest possible data rates while staying within country emission standards. 16 Mbps token-ring operation on shielded 150 ohm cable has been a standard since 1989. 100 Mbps Fiber Distributed Data Interface (FDDI), running on that cable for at least 100 meters from the workstation to the wiring closet, is available from several manufacturers. IBM also announced the F-Coupler in October 1991, which allows STP to be used for broadband transmissions in the 50 to 550 MHz band while simultaneously running 4 or 16 Mbps token-ring signals. For further information on the F-Coupler, see the *F-Coupler Planning Guide*, GA27-3949.

As imaging applications begin to play a rapidly increasing role in our computing environments, we may ultimately reach the absolute bounds of either the physical or regulatory laws that govern data transmission. When we reach those bounds, optical fiber is waiting for wider use in office environments. But for some time to come, both the IBM Cabling System shielded twisted-pair and data-grade UTP have an important place in data transmission systems. Each of the two twisted pairs are shielded from one another by a signal grounded polyester aluminum tape shield. The whole is further shielded in copper braid.

Figure 3 on page 13 shows a cross section of a token-ring STP lobe cable.

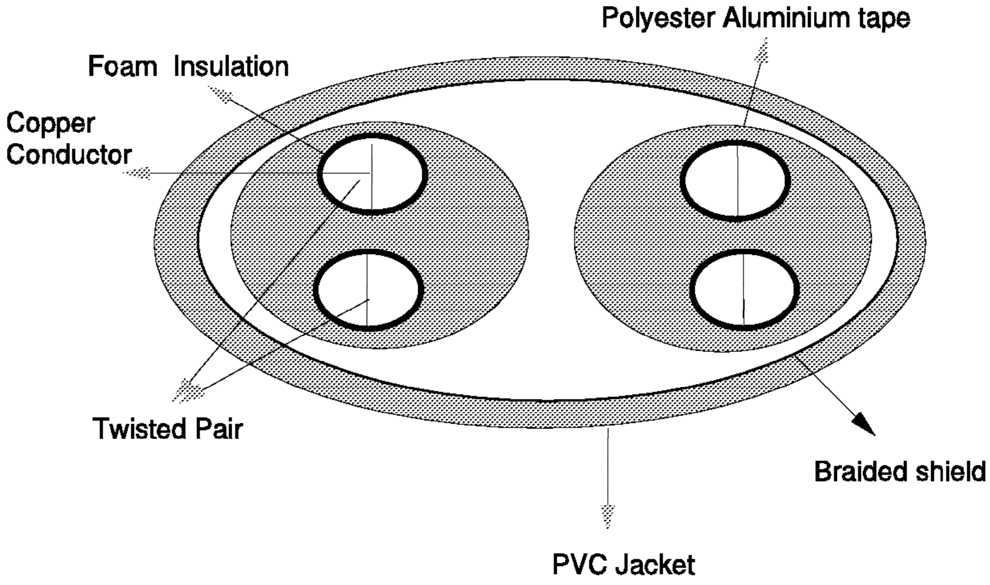


Figure 3. Type 1 Shielded Twisted Pair Cross Section

2.1.4 Fiber Optic Cable

Fiber optic cable presents an attractive solution for high-speed transmission rates used in backbone local area networks. The cable is relatively immune to the types of electrical noise and grounding trouble that can plague metallic conductors in some environments. Thus, it is also an ideal medium for outdoor connections or for factories or locations in which cabling has to run near higher voltage wiring. Fiber optic cable also has extremely high data transfer capability (hundreds of megabits per second, terabits per second) with very little signal attenuation (signal loss due to the medium). Because of the high data rates and the distances that fiber optic cables can carry a signal without regeneration, its use in telephone networks, channel extenders on mainframe computers, and backbone LANs is rapidly increasing (for example, with the use of Enterprise System Connection (ESCON) capable mainframes and communications devices). In comparison with transmission of electrical signals on copper media, it is difficult to tap an optical signal from a fiber optic cable without the inherent optical signal loss being detected. Therefore, fiber optic cable has potential for greater security than metallic conductors.

Fiber optic cable is normally used for baseband transmission. IBM includes fiber optic cable as part of the IBM Cabling System, referred to as ″Type 5 cable″. Optical fiber is produced in varying sizes. Typically these are 62.5/125 microns (10(-6) meters), 50/125 and 9/125. The dimensions refer to the diameter of the core and of the cladding. The core is the central tube within the fiber. The cladding is the fiber wall. Figure 4 on page 14 demonstrates the cable structure.

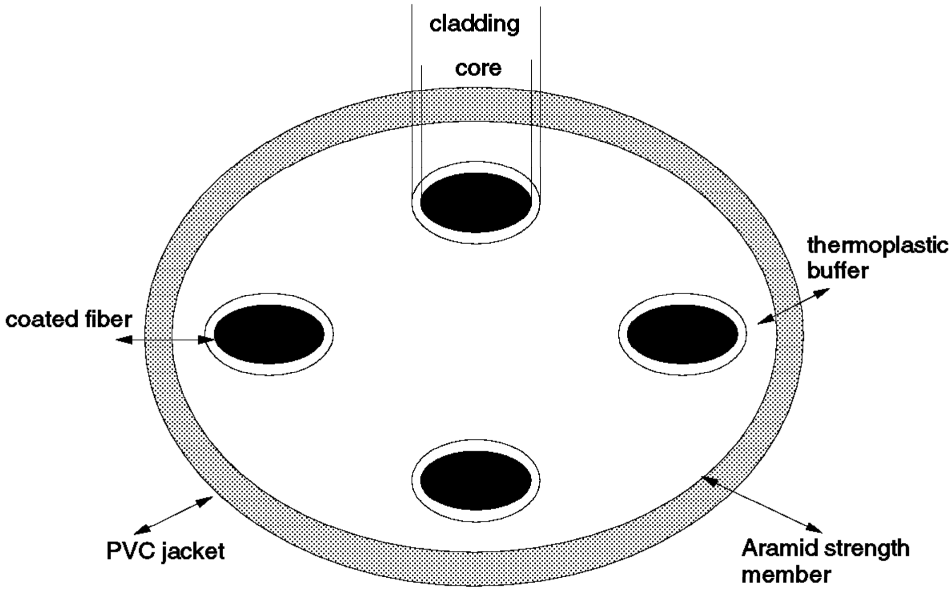


Figure 4. Fiber Optic Cable Structure

Individual fibers are normally bundled together into a sheathed cable, strengthened internally to protect the weak contents. There are basically two types of fiber. These are *multimode* and *single mode.* Single mode is often referred to as monomode. Multimode fibers are suitable for using Light Emitting Diodes (LEDs) as the light source and can handle data rates in the order of hundreds of megabits per second. They are suitable for general purpose LAN configurations, over distances of two kilometers. Because LED technology is relatively cheap, the cost of the transmitters is also kept at a minimum.

Monomode fiber requires the use of lasers as a light source. These devices are more expensive than LEDs, but drive distances can be longer. This technology is normally found within the long-haul networking environment. It is used, for example, by the public telephone companies for high-speed trunk links.

IBM Recommendation: IBM recommends the 62.5 micron optical fiber for most establishment cabling applications. The 100/140 optical fiber, which was part of the original IBM Cabling System offering, will continue to be supported for token-ring and future FDDI networks.

This 62.5/125-micron fiber specification is patterned after the fiber specification in the Commercial Building Wiring Standard (developed by the TIA 41.8.1, and by the ISO SC25/WG3 working groups) for meeting most intra-building and campus link requirements. It is expected to become the accepted multimode standard for government and commercial buildings and will meet future FDDI application requirements. The FDDI standard also provides information for attaching FDDI cable plants using 50/125, 100/140, and 85/125 micron multimode optical fibers as alternatives. IBM also recommends 62.5/125 multimode optical fiber for FDDI connection. However, IBM also supports 50/125 (preferred fiber in Japan and other countries), 85/125, and 100/140 micron multimode optical fibers, as defined in the ISO 9314/ANSI X3T9.5 standard for both token-ring network and FDDI application.

Each cable specification parameter must be met over the full range of operating temperatures. A suggested temperature range of 0 to 52 degrees Celsius is an appropriate choice for many installations. Maximum summer and minimum winter temperatures may differ from this range, particularly in installations where the fiber cable will be installed in uninsulated and unheated areas (typically building attics).

Customers should select a grade of fiber that will perform to specification in those instances where the temperature may exceed the suggested range. Please refer to the SP-2840 Commercial Building Telecommunications Cabling Standard (ANSI/TIA/EIA) for recommendations and considerations. Also please refer to the *IBM Cabling System Planning and Installation Guide,* GA27-3361 for more information. IBM also recommends that customers should consider installing single-mode fiber along with multimode for future high bandwidth applications.

Although fibers are normally installed using prepackaged bundles (cables), there is a new technology available known as blown fiber. With blown fiber, installation is achieved by ″blowing″ the optical fibers through tubes or microducts that have been previously installed between points in the building or on the campus site. Compressed air is used to blow the fibers down the ducts. This approach allows optical fibers to be installed or replaced as required, quickly and with minimum disruption to the customer′s facilities. Blown fiber is available from British Telecom (the inventor of the technology), Corning Incorporated and Sumitomo.

Optical fiber technology requires new skills to install, and subsequently test, the installation. It introduces new ranges of connectors, jumper cables and fiber splicing techniques, as well as a careful handling requirements, both of the fibers themselves and also at the patch panel. Its use is now widespread, and growing.

2.2 IBM Cabling System

When the IBM Cabling System was introduced in 1984, the electrical specifications for the shielded twisted pair cabling supported signals up to 20 MHz. These specifications have been adequate up to the present time for the range of frequencies found in office environments. However, with IBM′s announcement of support for 100 Mbps FDDI networks on STP cables between concentrators and workstations or between workstations in a ring of dual-attached workstations, the specifications for IBM Cabling System types 1, 2, 6, and 9 cable should be extended to maintain quality and to ensure suitable operation at higher frequencies than were contemplated when the specifications were first released.

2.2.1 Cable Specifications

When you are planning an installation on IBM Cabling System STP cable certified before 1993 and marked as type 1, 2, 6, or 9, IBM′s implementation of FDDI on this cable should operate on the lengths of cable described in the following section. However, since these cable types were certified for operation up to 20 MHz only, some anomalies may exist in cables longer than 60 m (198 ft) when signals at higher frequencies are used. In the overwhelming majority of cases, cable certified to 20 MHz will perform satisfactorily. If you are concerned about some of your longer lengths of cable, contact your IBM representative for further information about certifying such cables for operation at FDDI frequencies. Alternatively, the FDDI Diagnostic Tool can be used during installation to identify cables that cannot be used for transmission of FDDI signals.

When planning to install new IBM Cabling System STP cables, you should select cables that have been certified as meeting the specifications for types 1A, 2A, 6A, and 9A as set forth below. In the wiring closet, cables must be terminated in a distribution panel certified for 100 Mbps connections. Connections made at the existing metal distribution panels would be in violation of country electromagnetic interference regulations. IBM Cabling System shielded twisted pair (STP) installations that will be used for FDDI must use one of three plastic panels to meet country radiation requirements. Except in Germany, you may use either part number 33G2778, which provides 64 attachments in an 8-by-8 configuration, or part number 33G2763, which provides 16 attachments in a 2-by-8 configuration. Installations in Germany should use part number 33G2779, which provides 36 attachments in a 6-by-6 configuration. When using optical fiber cabling, follow the specifications and instructions for installation and testing as stated in the *IBM Cabling System Optical Fiber Planning and Installation Guide,* GA27-3943.

2.2.2 IBM FDDI Network Cabling Rules

These rules for cable lengths between devices must be followed to ensure a reliable network configuration.

2.2.2.1 New Specifications for STP Cables and Connectors

- For IBM Cabling System type 1, 1A or the shielded pairs of IBM Cabling System type 2 or 2A cable, the 8240 will support drive distances of 100 m (330 ft).

- For IBM Cabling System type 9 or 9A cable, the 8240 will support drive distances of 67 m (220 ft).

- For IBM Cabling System type 6 or 6A cable, the 8240 will support drive distances of 75 m (250 ft). This cable should be used only in office environments where no permanently installed cabling is available.

- For multimode optical fiber cable, the IBM 8240 will support drive distances of up to 2 km (6600 ft) with a link loss of no more than 11 dB, measured as described in the *IBM Cabling System Optical Fiber Planning and Installation Guide,* GA27-3943.

These distances are measured from the faceplate in the work area to the distribution panel in the wiring closet for both optical fiber and STP types 1, 1A, 2, 2A, 9, and 9A cables. For inter-wiring closet installations of optical fiber cable the distances are measured from distribution panel to distribution panel. Type 6 or 6A STP cabling is supported only for connecting rings of dual-attached workstations. Type 6 or 6A is generally not used for permanently installed cable.

Permanently installed IBM Cabling System STP cables may be used only between concentrators and single-attached stations. In addition to the permanently installed cabling, you will need to use an IBM FDDI Copper Adapter Cable and an IBM Concentrator Copper Port Adapter Cable between each station and concentrator. The lengths of these cables do not need to be added to the length of the permanently installed cable to determine the total allowable length.

IBM Cabling System type 6 or 6A cabling may be used between devices in a ring of dual-attached workstations. In addition to the type 6 or 6A cable, an IBM FDDI Copper Adapter Cable and an IBM FDDI Copper Adapter Reversing Cable will be used between each pair of stations. The lengths of these two cables do not need to be added to the length of the type 6 or 6A cable.

Optical fiber rules apply to all 8240-to-8240 connections in both ring and tree topologies. They also apply to the distances between single-attached stations and concentrators where optical fiber is employed as the transmission medium.

For optical fiber cables, patch cables of up to 9 m (30 ft) may be used at either end without adding their lengths to the length of the permanently installed cable to determine the total allowable length.

2.2.3 Transmission Characteristics of IBM Cabling System Cable

At IBM′s request, Engineering Testing Laboratories (ETL) and Underwriters Laboratories (UL) have begun certifying cable to electrical specifications. Since no redesign of current cable seems to be required to meet the new specifications, we expect that all cable will be certified to these specifications. Contact your IBM representative or nearest branch office for a list of cable manufacturers whose STP cables have been qualified using the new specifications. These specifications cover capacitance imbalance, resistance, resistance imbalance, balanced mode attenuation, and common mode attenuation. For further details please refer to SP-2840 from the ANSI/TIA/EIA standard.

2.2.3.1 Proposed New Connector Specifications

Like the IBM Cabling System STP cables, the IBM Data Connector was originally specified only up to 20 MHz. IBM′s testing has shown that currently available data connectors provide adequate performance at 100 Mbps as described in the *IBM FDDI Network Introduction and Planning Guide,* GA27-3892. Nevertheless, IBM is working with other vendors to provide a new specification for the IBM Data Connector that will allow the cabling subsystem performance to match that of the cable by itself.

2.2.3.2 STP Cable Information and Testing Service

A number of cable manufacturers′ cables have already been tested to the new type 1A, 2A, 6A, and 9A specifications. Customers can call IBM directly to obtain information on the options below:

- A list of the FDDI copper cabling part numbers and descriptions
- Wrap plugs
- Jumper cables
- Plastic patch panels
- New type ″A″ cable specification numbers

2.2.3.3 Selecting New Cables for Data Transmission

The *EIA/TIA-SP-2840 Commercial Building Wiring Standard* suggests three basic kinds of cable to use between work areas and wiring closets for data transmission. The three kinds of cable are: shielded twisted pair, unshielded twisted pair (divided into three acceptable categories), and 62.5/125 micron multimode optical fiber. The standard mandates that two copper paths must always be present: one for data, and the other for voice. Any optical fiber cable installed should be in addition to the copper cables.

The LAN standards groups use the EIA/TIA-568 standard to define transmission media for the LAN protocols. Table 5 compares these LAN Standards with the Commercial Building Wiring Standard.

Table 5. Cable Types and Relative Standards Conformance

Media	**4 Mbps (802.5)**	**10 Mbps (802.3)**	**16 Mbps (802.5)**	**100 Mbps (FDDI)**
Coax	C	S C	C	None
Fiber	C	S C	C	S C
STP	S C	C	S C	u C
UTP Cat 3	a C	S C	u C	u
UTP Cat 4	a C	S C	u C	u
UTP Cat 5	a C	S C	u C	u C

Note:

- S - Meets appropriate standard
- C - Commercial product currently available
- a - Described in appendix of a standard
- u - Under investigation by standards committee

2.2.3.4 Cable Choice and Standards

As demonstrated in Table 5 when rewiring or when wiring new installations, 150 ohm STP cable should be considered as the premier telecommunications cabling choice. It is often the most cost-effective solution when maintenance and future applications are considered. However, because each cabling decision is driven by its own set of unique cost considerations, installation of UTP is sometimes the appropriate choice. When this cable is chosen, only the high performance UTP cables (categories 4 and 5) should be used. Because their performance is so much greater than that of category 3 cables, the small additional cost should not be a consideration. Note that category 3 cable is still a very acceptable choice for telephone and low-speed (less than 1 Mbps) applications.

The terminating hardware and standard installation practices associated with 150 ohm STP cabling provide end-to-end cabling integrity and quality. With UTP cabling, both the terminating hardware and the installation practices can strongly affect the overall performance of the cabling system. Wall outlets and wiring closet terminating blocks that retain the overall transmission quality of category 4 cable are currently available. Category 5 cable pushes the limits of today′s technology; no generally recognized hardware ensures that the benefits of category 5 cable are preserved. In addition, careful workmanship is required to prevent significant degradation of the system′s electrical performance parameters from those for the category 5 cable itself. Still, category 5 cable is worth consideration in new installations, even if the connecting hardware may have to be replaced at a later date.

2.2.4 Commercial Building Telecommunications Cabling Standard

This standards document was put together by the American National Standards Institute, the Telecommunications Industry Association and the Electronic Industries Association. The current standard (SP-2840) has been prepared by the Working group TR-41.8.1 and replaces ANSI/EIA/TIA -568 standard of July 1991. This standard covers additional specifications for categories 3, 4 and 5 UTP cables and connecting hardware as well as additional specifications for 150 ohm STP cables and connectors. New specifications for 62.5/125 micron optical fiber and single-mode optical fiber cables, connectors and cabling practices have also been included.

The purpose of this standard is to specify the minimum requirements for telecommunications cabling within a commercial building which includes the telecommunications outlet and connections between buildings in a campus environment.

2.2.4.1 Horizontal Cabling

The current standard specifies this as the portion of cabling extending from the desk port to the wiring closet. Design and layout of this is very important as it′s more difficult to replace than a backbone cable layout. The maximum distance is 90 m (295 ft.) Please consult the standard for further information.

2.2.4.2 Cables

There are three types of cables recognized for the horizontal cabling system:

- Four pair 100 ohm UTP
- Two pair 150 ohm STP
- Two fiber, 62.5/125 micron optical fiber cable

Coaxial cable of 50 ohm is not recommended for new cabling installations. Performance characteristics for each of these cables can be referenced in the standard. Hybrid cables which consist of more than one type under a common sheath may also be used.

2.2.4.3 Backbone Cabling

This is basically the wiring that connects the wiring closets to each other within a building and between buildings using a star topology. Each horizontal cross connect is connected to a main cross connect. Please refer to the standard for details on supported media and distance limitations. For further details on each of the recommended cabling systems that cover the mechanical features of the cable such as attenuation, capacitance, termination, and cabling practices please consult the SP-2840 ANSI/TIA/EIA standard.

2.2.5 Other Cabling Specifications

The following examples are of other cables typically found in a commercial building installations.

2.2.5.1 100 Ohm Screened Twisted-Pair Cable

This may be used where improved crosstalk isolation or improved shielding is required.

2.2.5.2 Multipair UTP Cables

This is 25-pair inside cabling, usually used for voice services and is not generally recommended. Performance cannot be guaranteed. Intelligent hubs with repeater and signal regeneration functions may be required for full utilization. Costs for fiber today make putting fiber in the riser a more realistic solution even if the Multipair UTP cables exist.

2.2.5.3 Multiuser Telecommunications Outlet/Connector

This terminates multiple horizontal cables at a central location and can be of value in situations where office plans frequently change. 50 ohm coax, 100 ohm STP, other multimode optical fiber cables such as 200/230 micron step graded index optical fiber, 100/140 micron and 50/125 micron graded index optical fiber are all considered in the latest standards document.

At the time of writing a new 120 ohm standard was being considered by the European standards bodies.

2.2.5.4 Token-Ring and SDDI Cable Accessories

Some accessories are needed to have a complete installation. For the token-ring and the SDDI network these accessories are:

- Shielded twisted pair patch cable

  When a hub with RJ-45 connectors is used with the IBM cabling system to build token-ring and SDDI networks a twisted pair patch cable with an RJ-45 connector at one end and a universal data connector at the other end is needed. This cable connects the hub ports with their shielded RJ-45 connectors to the patch panel with the IBM Data Connector.

- STP cable for RI/RO ports

  The purpose of this cable to connect the RI/RO ports of token-ring modules when cable monitor mode is enabled. This allows the modules to sense a cable fault, and automatically wraps the ring to keep it up and running.

- Token-ring UTP media filter

  The token-ring UTP media filter links a network station to 4 or 16 Mbps token-ring networks which are using unshielded twisted pair (UTP) cabling.

The filter provides the following functions:

- It converts the connector on a token-ring adapter card from DB-9 to 8-pin modular connector (RJ-45).
- It matches the impedance from 150 ohms to 100 ohms.
- It reduces the radio frequency emissions for FCC class A compliance.

2.2.5.5 Asynchronous Transfer Mode (ATM) Campus Wiring

IBM′s key ATM campus strategy is to allow customers to use existing cabling, be it unshielded twisted pair, shielded twisted pair or fiber.

The 25 Mbps adapters make use of technology from IBM′s 16 Mbps token-ring system and allows the use of UTP category 3 cables.

For higher speed IBM is making use of a unique technology known as *Partial Response IV (PRIV)* which enables use of UTP category 5 and STP media.

ATM uses the same cable pairs as FDDI and token-ring.

IBM uses multimode fiber for distances up to 2 km and single-mode for distances up to 20 km. The SC connector has become ATM-standard.

2.3 Common LAN Topologies

Numerous topologies are used for local area networks. This section briefly discusses various LAN topologies. The term topology refers to the way in which network devices are interconnected. For example, if every station is directly connected to every other station, this is called mesh topology. Figure 5 on page 22 shows the basic network topologies.

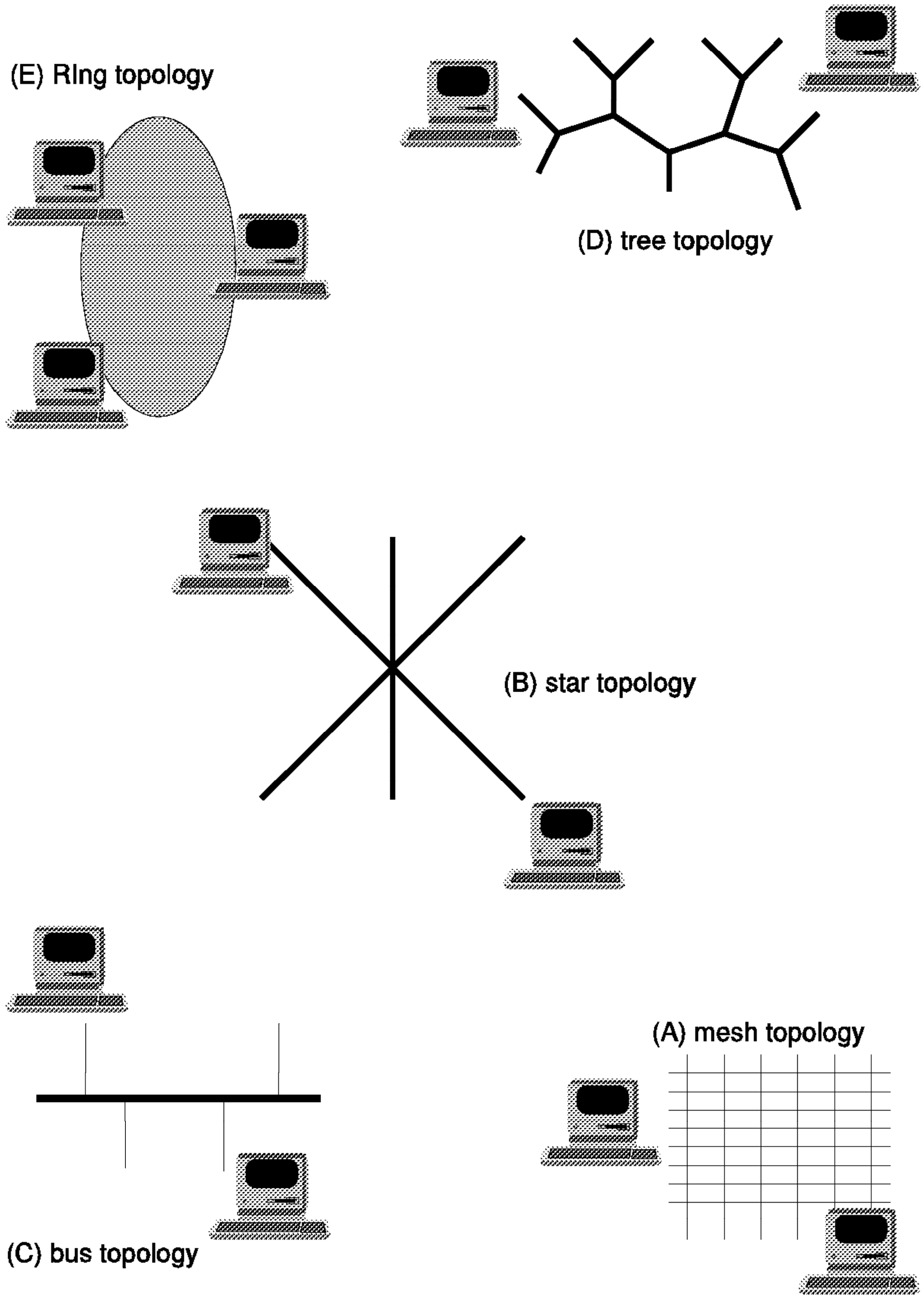


Figure 5. Common LAN Topologies

2.3.1 Mesh Topology

Part A of Figure 5 on page 22 shows a mesh topology network. This topology involves some wiring overhead since every network station is directly connected to all the other stations. It also means that each station has to have (N-1) I/O ports, where N is the number of stations in the network.

However, a mesh network topology has excellent fault tolerance, since, when a link fails, message traffic can be routed through an intermediate node.

2.3.2 Star Topology

In a network that has a star topology, each station is connected to a central controlling point (also called a switch) via point-to-point lines. Part B of Figure 5 on page 22 shows a star topology network. A PABX is a good example of this type of network.

For connection-oriented data transmission to another network station, the sending station must send a connection request to the central switch. The switch will then set up the connection. Once the connection is established, the two stations can transfer information. The structure of a star network is very simple, but to overcome the disadvantage of having a single point of failure, the central switch must use very reliable components and usually provides some form of redundancy.

2.3.3 Bus and Tree Topology

The bus topology is very common for local area networks. Part C of Figure 5 on page 22 depicts a bus topology network. Network stations are attached to a transmission medium, called a bus. When a station transmits a frame or segment of information on the bus, transmission occurs in both directions so that the frame is received by all other stations attached to the bus. Frames are said to be broadcast on the medium. A popular protocol used with this LAN topology is the Carrier Sense Multiple Access with Collision Detection (CSMA/CD) protocol, which will be discussed in 3.1, “Ethernet / 802.3” on page 28.

Unlike a multipoint line, in which a control point polls the other devices, there is no controlling station on a broadcast bus topology LAN. Control functions are distributed to all stations on the LAN. Each station must also be capable of detecting faults.

The tree topology is a variation of the bus topology. A CSMA/CD protocol can be applied to both topologies, and in both cases transmitted frames are broadcast to all stations active on the shared medium. As with the bus topology, there is no controlling station on the LAN. Part D of Figure 5 on page 22 illustrates a tree topology network.

2.3.4 Ring Topologies

In a network that has a ring topology, each station is attached to its adjacent station by point-to-point links, thus forming a physical ring. Each station′s adapter regenerates the signal as it retransmits a data packet that is circulating on the ring. A popular protocol used with ring topology is token passing, in which access to the medium is controlled by possession of a circulating token. Different token passing access protocols are defined for ring topology LANs, although token passing is also applicable to a bus local area network. Part E of Figure 5 on page 22 illustrates a ring topology network.

The major disadvantage of a physical ring topology is its sensitivity to single link failure. If one connection between two stations fails or a bypass for a particular inactive station is malfunctioning, the ring traffic is down.

A variation of a ring topology to avoid this problem is the star-wired ring topology, also referred to as radial hierarchical wiring.

Although the cable layout resembles a regular star topology, physically and logically the network stations form a ring. Each station is connected to a relay center with transmit and receive paths. The transmit path of one station is connected inside the relay center to the receive path of the next active station, thus bypassing inactive devices.

The relay center can be implemented as a passive wiring concentrator, in which case the relay mechanisms are powered by the ring stations.

The relay center can also be implemented as a ring wiring concentrator containing active, powered components that materially increase the ring′s error recovery ability, including automatic wrapping to a backup path and some degree of error message processing.

The previous discussion on different LAN topologies applies to a single LAN segment. A LAN segment consists of a single common medium and all the LAN devices connected to this medium. LAN segments can be combined to form larger local area networks or to interconnect different types of LAN segments into one network. LAN segment interconnection can be achieved through bridges, routers or gateways. LAN segments can only be combined in a single network using bridges or routers if they share a common communications layer (logical link control sublayer for bridges, network layer for routers). Gateways interconnect network segments which may use a totally different communications architecture. The different network interconnection techniques are discussed in greater detail in Chapter 1 of *Local Area Network Concepts and Products:Routers and Gateways,* SG24-4755. Specific LAN type topologies are covered later in this chapter.

Chapter 3. LAN Architectures and Standards

In Chapter 2, “Physical LAN Attachment” on page 7, various technological approaches were described for physical LAN attachment. Medium access methods and some alternative approaches to LAN interconnection and topology will be introduced in this chapter. Common LAN architectures, standardized by the Institute of Electrical and Electronics Engineer (IEEE), will be discussed.

In February 1980, the IEEE Computer Society established ″Project 802″ to draft standards for local area networks. In keeping with the OSI approach, IEEE Project 802 created a reference model with two layers (which correspond to the data link and physical layers of the OSI model). In the IEEE model the data link layer is further divided into two sublayers: the logical link control (LLC) sublayer, and the medium access control (MAC) sublayer.

Due to the variety of technically sound approaches proposed for local area networks, the IEEE Project 802 decided to draft standards for the most probable implementations:

- CSMA/CD
- Token-passing bus
- Token-passing ring

The commonly used names for these standards are derived from the project′s initial designation ″802″. Hence, we have:

- IEEE 802.0 - LAN and MAN (Metropolitan Area Network)
- IEEE 802.1 - Higher level interface standard
- IEEE 802.1k - Supplement to LAN and MAN management standard (6/93)
- IEEE 802.2 - Logical link control standard
- IEEE 802.3 - CSMA/CD standard
- IEEE 802.4 - Token-passing bus standard
- IEEE 802.5 - Token-passing ring standard

Over the years, the IEEE work has expanded, with the formation of new subcommittees. These include:

- IEEE 802.6 - Metropolitan Area Networks (MANs)
- IEEE 802.7 - Broadband Technical Advisory Group
- IEEE 802.8 - Fiber Technical Advisory Group
- IEEE 802.9 - Integrated Voice/Data on LAN
- IEEE 802.10 - Interoperable LAN Security
- IEEE 802.11 - Wireless LAN
- IEEE 802.12 - 100Base-VG
- IEEE 802.30 - 100Base-X

The IEEE 802.1 Higher Level Interface subcommittee is currently finalizing the draft IEEE 802.1 standard, also presented to the ISO as an ISO draft proposal. This draft includes the following subjects:

- IEEE 802.1 Part A which describes the relationship of the IEEE 802 work to the ISO Open Systems Interconnection Basic Reference Model.
- IEEE 802.1 Part B which specifies an architecture and protocol for the management of IEEE 802 LANs.
- IEEE 802.1 Part D which specifies an architecture and protocol for the interconnection of IEEE 802 LANs below the MAC service boundary.

- IEEE 802.1 Part E which deals with a Station Load Protocol, used for IPL of workstations from a server.
- IEEE 802.1 Part F which addresses guidelines for the development of layer management standards and definition of managed objects.
- IEEE 802.1 Part G which deals with remote bridges.
- IEEE 802.1 Part H which deals with Ethernet bridging.

The ISO has since adopted the IEEE 802 standards as part of the OSI Reference Model, and has given them the following ISO numbers:

- IEEE 802.2 Currently an OSI International Standard, IS 8802-2
- IEEE 802.3 Currently an OSI International Standard, IS 8802-3
- IEEE 802.4 Currently an OSI International Standard, IS 8802-4
- IEEE 802.5 Currently an OSI International Standard, IS 8802-5

A summary of the 802 standards is shown in Figure 6 on page 27.

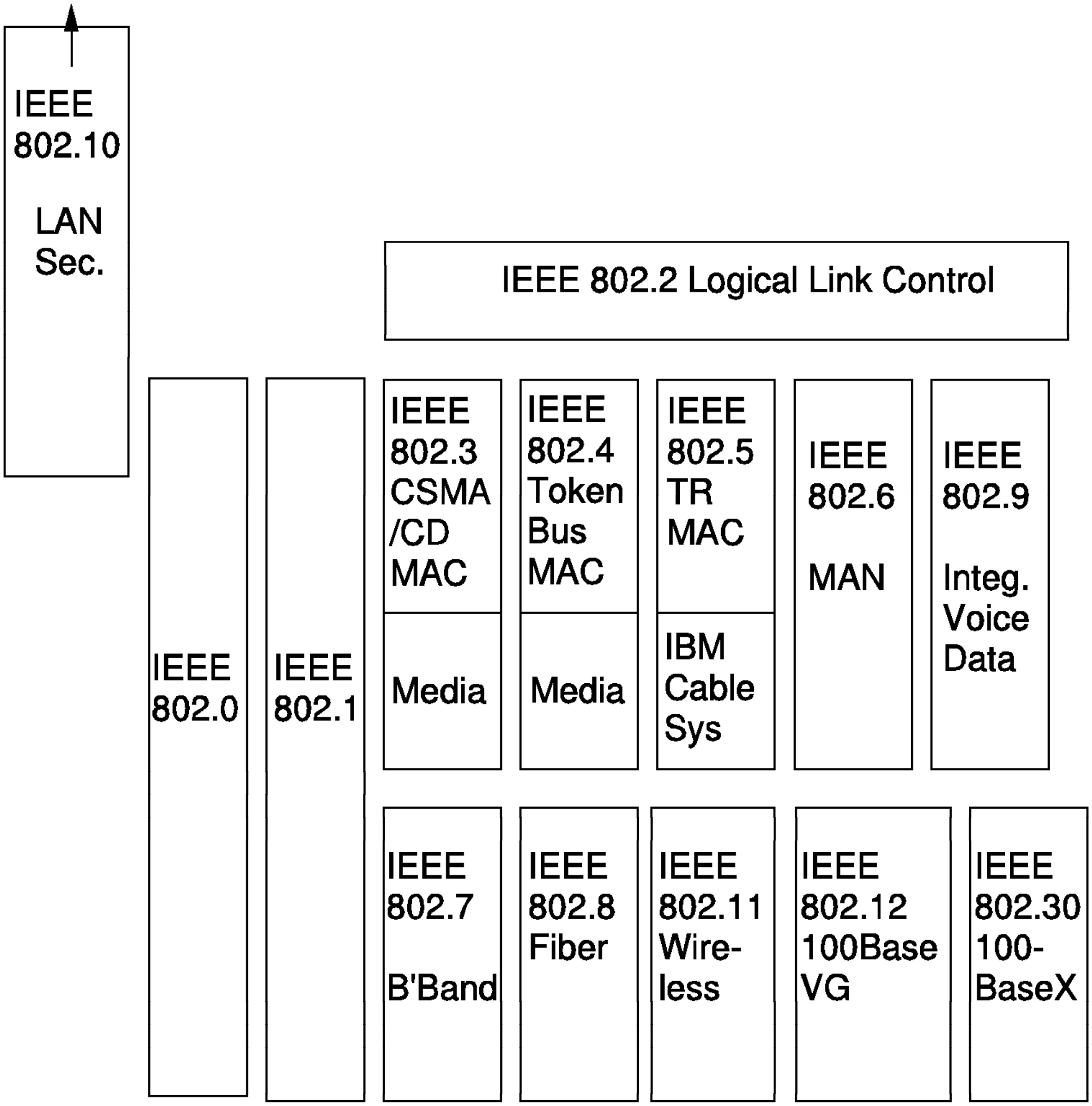


Figure 6. 802 Standards Summary

The data link layer provides the protocols used to physically transmit data on a communications link. These protocols must be able to detect when a transmission has been corrupted by errors and to retransmit the data. The IEEE 802 project divided the data link layer into two sublayers, the medium access control (MAC) layer and the logical link control (LLC) layer. These sublayers are discussed in the following sections.

Today, there is a growing requirement to further divide the physical layer into a physical and a physical medium dependent layer.

As previously mentioned, the IEEE 802.1 subcommittee work has not yet been completed, especially in the areas of the management and interconnection of IEEE 802 LANs.

Physical Layer and MAC Sublayer: The medium access control (MAC) sublayer contains mechanisms to control transmissions on the LAN so that two or more stations don′t try to transmit data at the same time, logic to control whether a station on the LAN is in transmit, repeat, or receive state, and addressing schemes to control the routing of data on the LAN. It also provides some or all of the following functions:

- **MAC addressing:** A MAC address is the physical address of the station (that is, device adapter) on the LAN, a predefined group address which is recognized by adapters of devices belonging to the group, a broadcast address which is recognized by all adapters on the LAN, or a null address for frames which should not be received by any station on the LAN. The MAC address is used to identify the physical destination and source of anything transmitted on the LAN.
- **Frame copying:** This is the ″receipt″, that is, copying of a frame into the buffers of an attached adapter which recognizes its own address in the destination address field of a frame.
- **Frame type recognition:** This is the identification of the type (for example, system or user) and format of a data frame.
- **Frame control:** This function ensures that frames can be processed accurately by providing frame check sequence numbers and starting or ending frame delimiters.
- **Priority management:** This function allows preferential access to the medium while maintaining fairness with respect to all participating stations.
- **LAN management:** A collection of protocols has been defined to support monitoring of a LAN and the ability to handle error conditions at the access control level.

Some access protocols do not provide priority management or frame control functions, but rely on higher layer services to provide these functions.

3.1 Ethernet / 802.3

Ethernet (802.3) is currently the most widely used LAN protocol in the world. Since its introduction to the marketplace in the 1970′s it has been established among a wide range of users.

Invented by Xerox in the early 1970′s and brought to the marketplace as Ethernet V.1, the protocol was then developed by a consortium of DEC, Intel and Xerox. This consortium brought out a new version of Ethernet in 1980 called Ethernet (DIX) V2. They also published the architecture and took it to the Institute of Electrical and Electronics Engineers (IEEE) to have it accepted as an international standard. The IEEE ratified the Ethernet DIX V2 standards with some slight modifications as IEEE 802.3. The 802.3 standard has since been approved by a number of other organizations including the American National Standards Institute (ANSI) and the International Organization for Standardization (ISO 8802-3).

Today both Ethernet and 802.3 LANs are widely implemented across all areas of the marketplace.

Although the protocol used by Ethernet/802.3 LANs has not changed, the physical topology over which they can be implemented has changed significantly. This has enabled users to have access to some of the benefits (such as manageability) offered by other topologies, such as token-ring, while still enjoying the perceived advantages of Ethernet/802.3, which include:

1. Wide choice of equipment
2. Low cost of equipment

Though Ethernet and 802.3 are not identical, the term *Ethernet* is widely used to describe LANs that use either protocol. As most of the information in this chapter applies equally to both Ethernet and 802.3 LANs, the term Ethernet (802.3) will be used throughout this section. However, where there are differences, they will be indicated by using the appropriate terminology.

Note

Both Ethernet V2 and 802.3 frame types can be used on the same physical Ethernet simultaneously. However, stations using one frame type cannot interoperate with stations using the other frame type unless they have both frame types configured. In practice, applications tend to use Ethernet V2 frame types while management programs tend to use 802.3 frame types. The differences between these two frame types will be explained later in this section.

Please note that this section will cover baseband Ethernet only.

3.1.1 Ethernet Concepts

Carrier Sense Multiple Access with Collision Detection (CSMA/CD) is the name of the protocol used on the Ethernet (802.3) bus to control the operation of the network. An example of CSMA/CD is shown in Figure 7 on page 30.

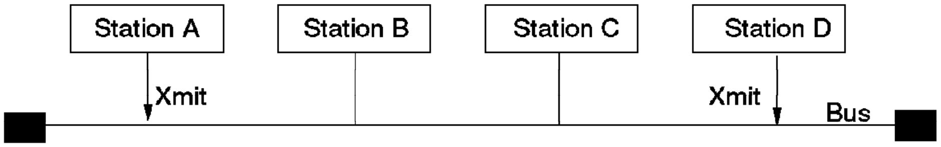


Figure 7. Ethernet CSMA/CD Bus

In a CSMA/CD bus, when a station wants to transmit data on the network bus, it first listens to see if the bus is free (that is, no other station is transmitting). If the bus is available, the station starts transmitting data immediately. If the bus is not available (that is, another station is transmitting), the station waits until the activity on the bus stops and a predetermined period of inactivity follows before it starts transmitting.

If there is a *collision* after transmission (that is another station starts to transmit at the same time), the stations will stop transmitting data immediately after the collision is detected, but they continue to transmit a jamming signal to inform all active stations about the collision.

In response to this signal, each transmitting station stops transmitting and uses a binary exponential backoff algorithm to wait before attempting to transmit again. This causes each station to wait for a random amount of time before starting the whole process again beginning with the process of carrier sensing. If a station′s subsequent attempt results in another collision, its wait time will be doubled.

This process may be repeated up to 16 times, after which the station, if still unsuccessful, reports a transmission error to the higher layer protocols.

The process of *collision detection* varies according to the type of media used in the LAN. This process is described in “Medium Attachment Unit (MAU)” on page 38.

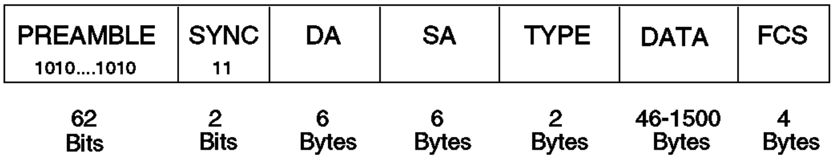
The probability of a collision occurring is proportional to the number of stations, the frequency of transmissions, size of frames and length of the LAN. Therefore, care must be exercised in designing LANs with an excessive number of stations which transmit large packets at frequent intervals. Also, you must ensure that the length of individual *segments* and total length of the LAN does not exceed a certain length as defined by the 802.3 standards. These limitations are discussed later in this topic.

According to Ethernet and the 802.3 standard, to be able to detect collisions, a transmitting station should monitor the network for a period of time called a *slot time*. Slot time is the time during which a collision may occur and is the maximum delay for a transmission to reach the far end of the network and for a collision to propagate back. Slot time is defined to be 51.2 microseconds (512 bit times in a 10 Mbps LAN). This time imposes a maximum length on the size of the network. It also imposes a minimum (64 bytes, excluding preamble and FCS) on the size of the frames transmitted by each station.

3.1.1.1 Ethernet and IEEE 802.3 Frame Formats

The frame formats for Ethernet and IEEE 802.3 are not the same. However, both protocols use the same medium and access method. This means that while LAN stations running these protocols could share a common bus, they could not communicate with each other.

Ethernet Frame Format: The layout of an Ethernet frame is as follows:



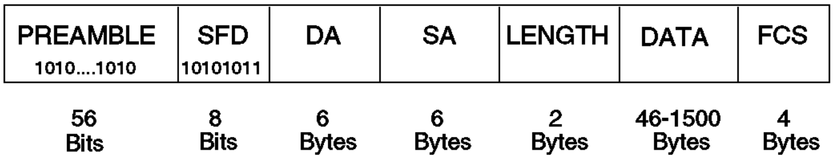
PREAMBLE 1010....1010	SYNC 11	DA	SA	TYPE	DATA	FCS
62 Bits	2 Bits	6 Bytes	6 Bytes	2 Bytes	46-1500 Bytes	4 Bytes

Figure 8. Ethernet Frame Format

- PREAMBLE - 62 bits, allows the Physical Layer Signalling (PLS) circuitry to synchronize with the receive frame timing circuitry.
- SYNC (Synchronize) - 2 bits, indicates that the data portion of the frame will follow.
- DA (Destination Address) and SA (Source Address) - 48 bits, Media Access Control (MAC) address. Three types of destination addressing are supported:
  - Individual: The DA contains the unique address of one node on the network.
  - Multicast: If the first bit of the DA is set, it denotes that a *group* address is being used. The *group* that is being addressed will be determined by a higher layer function.

  - Broadcast: When the DA field is set to all 1′s, it indicates a *broadcast*. A broadcast is a special form of multicast. All nodes on the network must be capable of receiving a broadcast.
- TYPE (Type Field) - 16 bits, this field identifies the higher layer protocol which is used. Vendors must register their protocols with the Ethernet standards body if they wish to use Ethernet Version 2.0 transport. Each registered protocol is given a unique 2-byte *type* identifier. As this field is used as the *length* field by the 802.3 frames, the value assigned to the *type* field in Ethernet is always higher than the maximum value in the *length* field for the 802.3. This is to ensure that both protocols can coexist on the same network.
- DATA (Data field) - This contains the actual data being transmitted and is 46-1500 bytes in length. Ethernet assumes that the upper layers will ensure that the minimum data field size (46 bytes) is met prior to passing the data to the MAC layer. The existence of any padding character is unknown to the MAC layer.
- FCS - 32 bits, the result of a cyclic redundancy check algorithm (specific polynomial executed against the contents of DA, SA, length, information and pad fields). This field is calculated by the transmitting station and is appended as the last four bytes of the frame. The same algorithm is used by the receiving station to perform the same calculation and the results are compared with the contents of the FCS field in the received frame to ensure that transmission was error free.

IEEE 802.3 Frame Format: The layout of the IEEE 802.3 frame format is as follows:



PREAMBLE 1010....1010	SFD 10101011	DA	SA	LENGTH	DATA	FCS
56 Bits	8 Bits	6 Bytes	6 Bytes	2 Bytes	46-1500 Bytes	4 Bytes

Figure 9. 802.3 Frame Format

- PREAMBLE - 56 bits, allows the Physical Layer Signalling (PLS) circuitry to synchronize with the receive frame timing circuitry.
- SFD (Start Frame Delimiter) - 8 bits, indicates that the data portion of the frame will follow.

- DA (Destination Address), SA (Source Address) - 48 bits, Media Access Control (MAC) address. Three types of destination addressing are supported:
  - Individual - The DA contains the unique address of a node on the network.
  - Multicast - If the first bit of the DA is set, it denotes that a group address is being used. The *group* that is being addressed will be determined by a higher layer function.
  - Broadcast - When the DA field is set to all 1′s, it indicates a *broadcast*. A broadcast is a special form of multicast. All nodes on the network must be capable of receiving a broadcast.
- LF (Length Field) - 16 bits, indicates the number of *data* bytes (excluding the PAD) that are in the data field.
- DATA and PAD field - IEEE 802.3 (and Ethernet) specify a minimum packet size (header plus data) of 64 bytes. However, 802.3 permits the *data field* to be less than the 46 bytes required to ensure that the whole packet meets this minimum. In order to ensure that the minimum packet size requirement is met, 802.3 requires the MAC layer to add *pad* characters to the LLC data field before sending the data over the network.
- FCS - 32 bits, the result of a cyclic redundancy check algorithm (specific polynomial executed against the contents of DA, SA, length, information and pad fields). This field is calculated by the transmitting station and is appended as the last four bytes of the frame. The same algorithm is used by the receiving station to perform the same calculation and the results are compared with the contents of the FCS field in the received frame to ensure that transmission was error free.

3.1.1.2 Ethernet (802.3) Network Model

Figure 10 shows the components of an 802.3 network and its relationship with the OSI reference model.

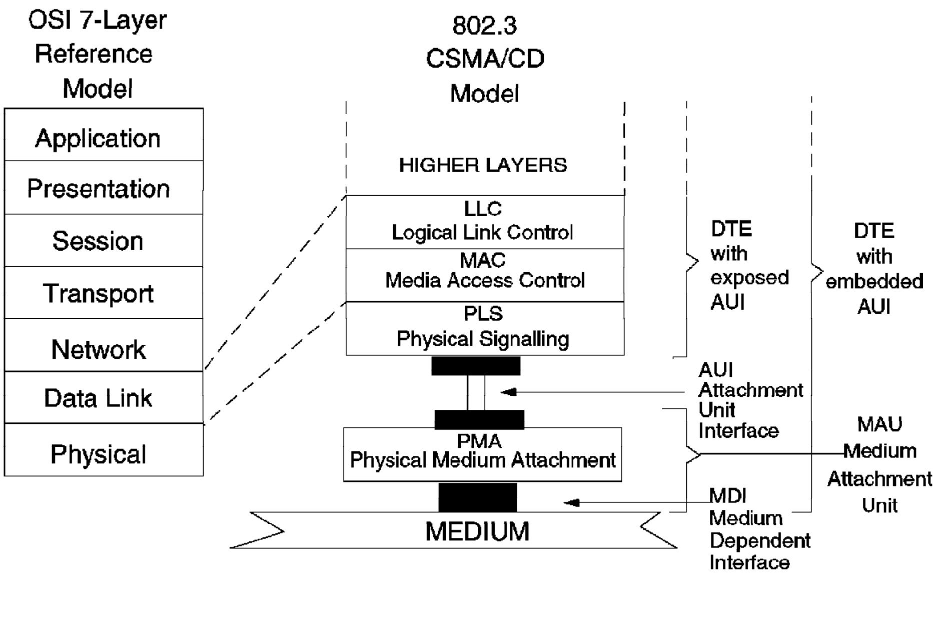


Figure 10. OSI Relationship to IEEE 802.3

In this model, *DTE* is the device that connects to the network and uses the network to exchange information with the other DTEs attached to the same network.

The following sections provide a brief description of the various components of this model.

Media Access Control (MAC) Sublayer: The MAC sublayer controls the routing of information between the *physical layer* and the *Logical Link Control* (LLC) sublayer by enforcing the CSMA/CD protocol. It provides:

- Frame transmission

  MAC sublayer is responsible for constructing a frame containing the data passed to it from the LLC sublayer. The transmitted frame will contain four bytes of FCS which is computed by the MAC sublayer based on the contents of DA, SA, Length field and Information field. The constructed frame will be transmitted on the physical medium.

  > **Note**
  >
  > In 802.3, the data field passed to the MAC sublayer can be less than 46 bytes. In that case the MAC sublayer will append pad characters to the data field to ensure that the minimum frame size is 64 bytes before sending the data over the physical medium.
  >
  > In Ethernet, a higher layer is responsible for ensuring that the data field is a minimum of 46 bytes.

- Collision detection and recovery

  To transmit the frame, the MAC sublayer must sense if the medium is currently active (in use). The status of the medium is sensed by the PLS and is passed to the MAC sublayer. If the medium is busy, the MAC sublayer will defer the transmission until the medium becomes *idle* and a period of time known as *Inter Packet Gap* expires. After this period, the MAC sublayer will start transmitting the frame. IGP is 9.6 microseconds and its purpose is to allow all the stations in the network to detect the *idle carrier*.

  If two or more stations attempt to transmit at the same time, a *collision* will occur. The collision will cause all the stations to *backoff* and restart transmission at some random time in the future as explained in 3.1.1, “Ethernet Concepts” on page 29.

- Frame recognition and copying

  When receiving a frame, the MAC sublayer identifies the Destination Address within the received frame and compares it with the address of the DTE (including Group and Broadcast addresses supported by that DTE). If a match is found, it will copy the frame, compute the FCS, and compare the result with the FCS contained in the received frame. If the frame is received error free, it will be passed to the higher layers; otherwise, a *CRC Error* will be reported.

Service Primitives: The basic MAC service primitives used in this and all IEEE MAC standards are:

- Medium access data request (MA_DATA.request) This primitive is generated whenever the LLC sublayer has data to be transmitted to another station(s) on the LAN. The MAC sublayer formats it in a MAC frame and transmits it.

- Medium access data confirm (MA_DATA.confirm) This primitive is generated by the MAC sublayer in response to MA_DATA.request from the local LLC sublayer. A status parameter is used to indicate the outcome of the associated MA_DATA.request.
- Medium access data indicate (MA_DATA.indicate) This primitive is sent to indicate that a valid frame arrived at the local MAC layer. The frame was transmitted without errors and was correctly addressed.
- Medium access data response (MA_DATA.response) This primitive is used as a response to MA_DATA.indicate.

The use of these primitives is shown in Figure 11 and shows two stations, A and B, with the LLC layer of station A requesting transmission of a frame to the MAC service interface. Upon receipt of the frame by station B, an indicate is generated, notifying the LLC of an incoming frame. LLC generates a response to indicate that it has the frame. The confirm indicates to station A that the frame was transmitted without error.

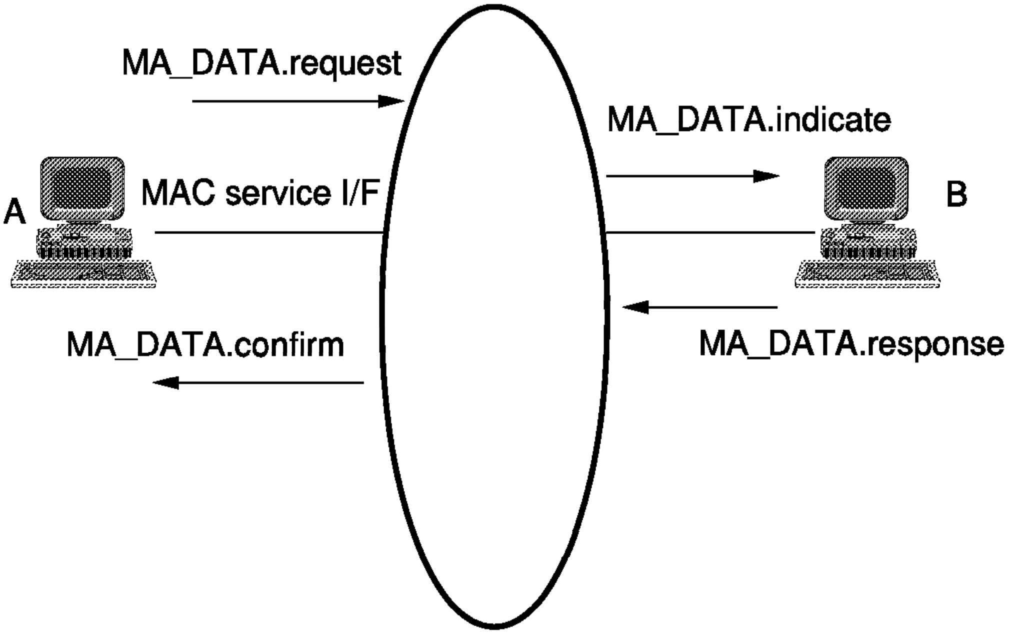


Figure 11. IEEE MAC Primitives

Physical Signalling (PLS): PLS is part of the physical layer which resides in the DTE and provides the interface between the MAC sublayer and the *Attachment Unit Interface* (AUI). The functions provided by the PLS are:

- Transmit data output

  The frame received from the MAC sublayer will be transmitted, over the AUI cable, to the *Medium Attachment Unit* (MAU) using the *Differential Manchester Encoding* technique. This technique enables a single bit stream to contain both the clock and the data by ensuring that there is a signal transition in the center of each bit. Also, at the bit boundary, there is a signal transition if the transmitted bit has a value of B′1′ while there is no transition at the bit boundary for B′0′.

- Receive data input

  PLS receives the *Manchester Encoded* data bit stream via AUI, decodes it to NRZ format, and provides it to the MAC sublayer.

- Perform carrier sense

  The PLS is responsible for passing status of the carrier to the MAC sublayer. This will enable the MAC layer to determine if there is network activity on the network. The PLS will inform the MAC that the carrier is active:

  - When it is receiving a frame
  - When there is a collision on the network
  - When the node is transmitting

  This is used by the MAC to determine that there is an AUI and/or MAU malfunction if there is no carrier sensed during the transmission from this station.

- Perform error detection

  After each frame is transmitted, the MAU is required to send a 10 MHz signal to the DTE to inform it that the MAU is connected and functioning properly. This signal is also used to check that the collision detection circuitry within the DTE is functioning properly.

  PLS is responsible for passing the presence of this signal to the MAC sublayer. The absence of this signal from PLS, will be interpreted by the MAC as a malfunction in the MAU.

Attachment Unit Interface (AUI): The connection between DTE (PLS function) and MAU (transceiver) is made by an Attachment Unit Interface (AUI) cable. This cable, which is also commonly known as the *transceiver cable*, provides the signal paths for:

- Data Out (DO) - DTE to MAU
- Data In (DI) - MAU to DTE
- Control In (CI) - Collision signal from MAU to DTE
- Power - DTE to MAU

AUI cables use individually screened AWG 22 wire for signal and power pairs. Since the 802.3 standard specifies that the DTE should have a female connector and MAU should have a male connector, the AUI cable requires opposite mating connectors to provide the connection between DTE and MAU. The connectors at the end of the AUI cable are 15-pin D-type connectors. The maximum allowed length for the AUI cable is 50 meters.

The AUI cables for Ethernet and 802.3 AUI attachment are not identical and have the following differences:

1. IEEE 802.3

   All shields of the signal pairs and the power pair are connected to pin 4. The overall AUI cable shield is connected to the AUI connector shell to provide a cable earth. Pin 1 is not used. See Figure 12 on page 37 for details.

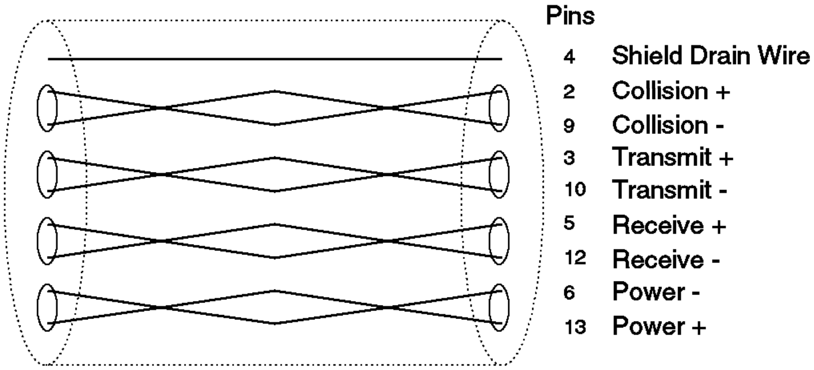


Figure 12. AUI Cable for IEEE 802.3

2. Ethernet Version 2.0

   All shields are connected to pin 1 and the AUI connector shell. Pin 4 is not used. See Figure 13 for more details.

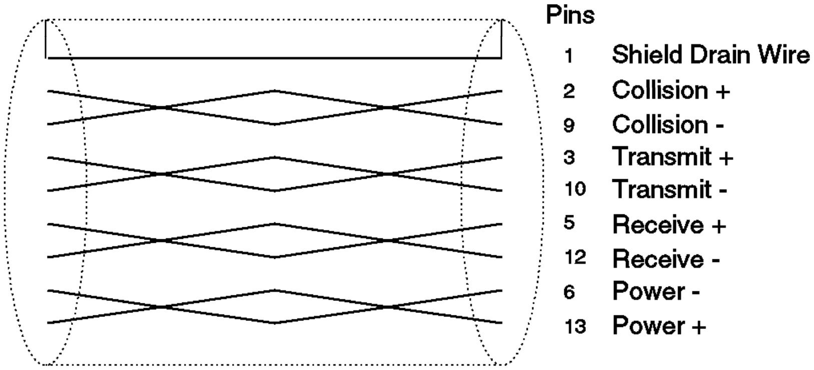


Figure 13. AUI Cable for Ethernet V2.0

3. Ethernet Version 1.0

   The point-to-point wiring of V1.0 and 2.0 is the same but the electrical requirement of the cable is different. Shielding of individual signal or power pairs is not required.

The overall AUI cable shield provides both shielding and signal ground. It is connected to pin 1 and the AUI connector shell.

The gauge of wire used is AWG 22 for signal pairs and AWG 20 for the power pair.

In fact most Ethernet equipment uses the Version 2 cable due to its superior construction.

Medium Attachment Unit (MAU): MAUs (also known as transceivers) provide the mechanical, electrical and functional interface between the DTE and the particular media used on the Ethernet (802.3) bus. Therefore, there is a different type of transceiver for each media type.

The use of a transceiver means that all the functions within a DTE (that is MAC, PLS and AUI) are identical regardless of the type of media used to provide the connectivity between the DTEs. The only component which requires changing, when the DTE is moved to another LAN which uses a different type of media, is the transceiver.

Transceivers perform the following functions:

- Transmit and receive data

  The transceiver will transmit data from the DTE onto the segment. It is also responsible for receiving data from the segment and passing it onto the DTE. It is important to note that the transceiver is not an intelligent device and will pass all the data to the DTE regardless of whether it is addressed to the DTE. The transceiver does not decode the data it sends or receives.

- Collision detection

  In accordance with 802.3 rules, it is the transceiver′s responsibility to detect collisions, and to inform the DTE of their occurrence. The transceiver does this by constantly monitoring the segment and reporting collisions. Collisions are reported via a 10 MHz signal which is sent on the *Control In (CI)* pair of the AUI cable.

  The collision detection mechanism used by the MAU varies according to the type of LAN medium used. In a coax network (thick or thin), since all the DTEs are connected to the center conductor of the cable, the transceiver can detect two or more devices simultaneously transmitting on the network by just monitoring the *voltage level* on the center conductor. If the voltage seen is more than the allowed threshold (-1.6V nominally), there is a collision on the network.

  In a 10Base-T network, there are two pairs of twisted copper between the DTE and the hub. One pair is for *transmit* and the other is for *receive*. During the normal transmission, the receive pair is idle. If the MAU detects activity on the receive pair while it is transmitting, it will report a collision to the DTE.

- Jabber protection

  Jabber occurs when a DTE sends more data than is allowed under the Ethernet (802.3) rules. This could be caused by hardware failure within the DTE or a running process attempting to send too large a frame. It is the transceiver′s responsibility to prevent DTE from monopolizing the network.

If a device transmits a legal size frame, it should take no more than a certain period of time to send that frame. 802.3 specification states that a transceiver must *cut off* the DTE after 20-150 microseconds by interrupting the transmission of the data on the network and indicating collision on the the *Control In (CI)* pair of the AUI cable. The transceiver should remain in this state until the DTE stops transmitting data on the *Data Out (DO)* pair of the AUI cable.

- SQE testing

  Also known as *heartbeat*, the SQE test is a 10 MHz burst that is sent to the DTE by the transceiver after each frame is transmitted. The purpose is to inform the DTE that the transceiver is working properly. The SQE test signal is sent on the *Control In (CI)* pair of the AUI cable.

  Note

  Repeaters must be attached to the network by transceivers that have SQE test disabled. Failing to disable SQE test will prevent the repeater from operating properly. This is because the repeater cannot discriminate between an SQE test and real collisions. If SQE is not disabled, the repeater will eventually partition that port and will prevent traffic from crossing the repeater.

- Link integrity

  During a normal transmission on a coax network (thick or thin), a transceiver will receive its own transmissions because of the fact that all the nodes on a coax network are connected to the same center conductor on the bus. The transceiver will return this to the DTE on the *Data In (DI)* pair of the AUI cable. This signal will be used by the DTE as an indication of transmit to receive integrity.

  In networks such as 10Base-T and 10Base-F, which have a separate transmit and receive path, the transceiver should provide an internal loopback path so that the transmitted data is received on the *receive path*. This is to ensure that the operation of various media types is transparent to the DTE and consistent network behavior is observed by the DTE across all the media types.

  However, it is still necessary in these networks, to ensure that a break in the transmit or receive path is detected. To do so, the MAU will start transmitting a *link test pulse* as soon as it has no data to transmit. If the MAU at the other end does not see either data packets or a link test pulse within a predefined time known as *link loss time*, (50 to 150 microseconds for 10Base-T), the transceiver will enter the *Link Test Fail* state. This will disable the transmit, receive, loopback, collision presence and SQE test functions. During the *Link Test Fail*, the transmission and reception of the link test pulses will continue.

  Receiving a minimum of two consecutive link *test pulses* or a single data packet, will cause the transceiver to exit the *Link Test Fail* state and re-establish the link.

3.1.2 Ethernet (802.3) Topologies

In an Ethernet (802.3) network, various types of cables can be used to provide the physical link between the DTEs. The media used can be thick or thin coax, twisted pair, or fiber optic cable.

Thick coax is also known as 10Base5 or Ethernet. Thin coax is also referred to as 10Base2 or Cheapernet. When using coax (thick or thin), this cable acts as the bus to which the DTEs are connected. In the case of thick coax, the transceiver is an external device, while in the case of thin coax, the transceiver can be an external device or mounted on board the adapter card (NIC).

Coax networks do not require structured wiring in the building, which makes them ideal for use in old buildings. However, they have the disadvantage of not providing management capability and fault isolation. For example, a break in the bus cable will render the whole network idle.

To enable the use of structured wiring in an Ethernet environment, a standard known as 10Base-T has been developed which provides a point-to-point link between the DTE and a central *hub* over twisted pair wiring. The hub contains a Multistation Access Unit (MAU) function on each of its ports. It also contains a repeater function which allows these point-to-point segments to communicate with each other. The hubs can also be connected to extend the size of the network and the number stations that can be attached to them.

Because, of the existence of hub(s), a 10Base-T network provides a much better management and fault isolation capability than the coax-based networks.

Fiber optic cables are used to provide point-to-point links, typically as a *backbone* between concentrators, to interconnect buildings or cross long distances within a building. However, it is also possible to use fiber optic cables as a means of providing connections to workstations. There are various standards covering the use of fiber optic cables in an Ethernet (802.3) environment. These standards are described briefly in the following sections.

The physical size of a network and the number of stations attached to it varies according to the type of medium used to construct the network. However, users can build a network consisting of mixed topologies by using repeaters and bridges. Also, such mixed topologies are made possible by intelligent hubs such as the IBM 8250 which provide various repeater, bridge, media and management functions via a number of modules which can be installed on the hub as required. The following sections provide a brief description of the various standards used in Ethernet (802.3) networks.

3.1.2.1 10Base5 (Thicknet)

The names given to the IEEE 802.3 standards provide some information as to the capabilities and requirements of the implementation. In the case of 10Base5 they have the following meaning:

- 10 indicates the data rate (10 Mbps)
- Base indicates the transmission type (Baseband)
- 5 indicates the maximum cable length (500 meters)

10Base5 (thicknet) uses a very high quality coaxial cable for the bus. This cable is very thick (10 mm in diameter) which makes it difficult to manipulate particularly if it is being run into work areas and needs to go in and out of ducting. The cable is generally marked every 2.5 meters to indicate where transceivers can be attached.

Attachment of DTEs to the coaxial cable is done by attaching a transceiver to the cable and attaching the DTE to the transceiver via an 10 Mbps AUI cable. This is shown in Figure 14.

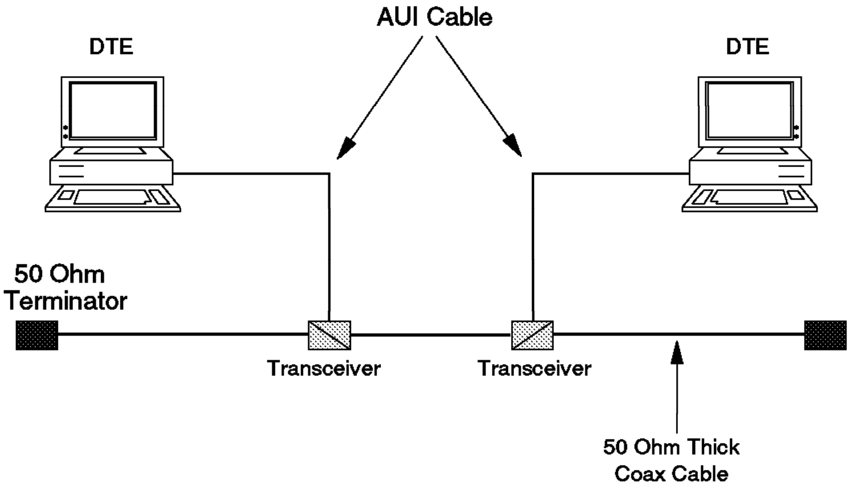


Figure 14. 10Base5 Segment

Note that terminators are used at both ends of the segment to prevent the signal from being reflected back when it reaches the end of the segment.

The transceivers used with this type of installation come in two main types:

1. Piercing Tap Connectors or Vampire Taps

   These are the most common types of transceivers used on 10Base5 networks. They are known as *vampire taps* because the center connection is made by drilling or piercing through the outer shield and dielectric of the cable and inserting a tap screw.

   Making this type of connection is not a trivial task. This makes adding/removing transceivers a job for a skilled person. Figure 15 on page 42 shows a cross section of a tapped thicknet cable.

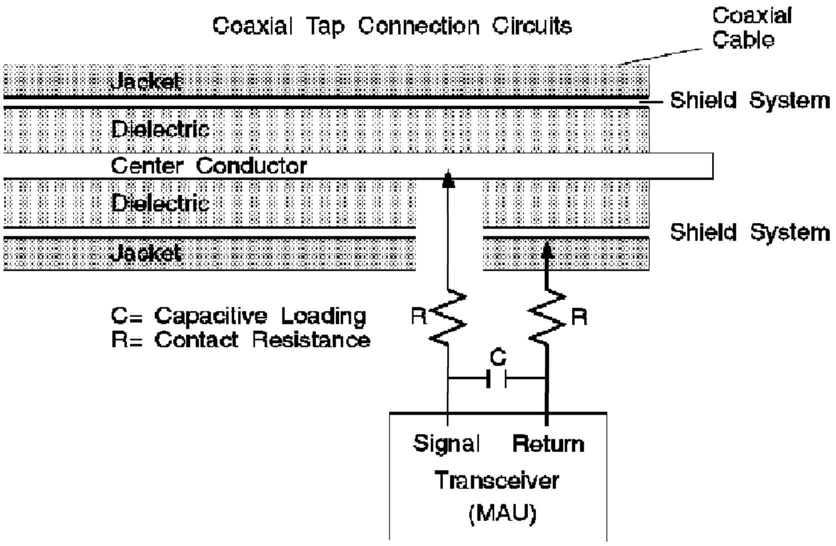


Figure 15. Cross Section of Tapped Thicknet Cable

2. N Type Connector

   This type of connection requires the cable to be segmented. As cutting the cable, while the bus is in operation, renders the network unusable, these transceivers are not as common as the *vampire tap*. However, as manufacturing techniques have improved, various manufacturers do offer 10Base5 segmented cables terminating in N connectors and transceivers capable of being attached via this method.

In modern environments 10Base5 topology is not very practical. The difficulties of manipulating the bus cable, rerouting AUI cables, attaching transceivers etc., means that installations of this nature are inherently inflexible and unable to accommodate the rate of change that is expected on most local area networks today.

Despite the drawbacks associated with this type of installation, 10Base5 has been widely installed. The use of multiport transceivers with a thinner and more flexible five meter transceiver cable has made it somewhat easier to add/remove DTEs and enable most connections to be made without having to manipulate the thick coaxial cable. Also, despite the fact that 10Base5 has become less popular for providing access to the LAN directly it is still widely used, particularly in situations where relatively few attachments are required and change is limited.

Table 6 provides a summary of the 10Base5 specification.

Table 6 (Page 1 of 2). 10Base5 Specification

Item	Specification
Cable type	Ethernet 50 ohm PVC or teflon FEP coaxial
Connectors	N-series
Termination	Segment ends not attached to repeaters must be terminated with 50 ohm terminators

Table 6 (Page 2 of 2). 10Base5 Specification

Item	Specification
Transceiver cable	Four-strand, twisted-pair conductors with an overall shield and insulating jacket
Data rate	10 Megabits/sec
Max segment length	500 meters
Distances between transceivers	2.5 meter multiples
Max no. of transceivers	100 transceivers
Max no. of stations per network	1024 adapters
Max transceiver cable length	50 meters
Impedance	50 ohms (+/- 2)
Attenuation	8.5 dB for 500 meters at 10 MHz
Max propagation delay/segment	2165 nanoseconds
DC resistance	5 ohms per segment

3.1.2.2 10Base2 (Thinnet/Cheapernet)

As a means of addressing the problems associated with 10Base5, the 10Base2 standard was defined.

The name 10Base2 was chosen because of the characteristics of this type of network as shown below:

- 10 indicates the data rate (10 Mbps)
- Base indicates the transmission type (Baseband)
- 2 indicates the maximum cable length (200 meters)

Note: The actual length permitted on a 10Base2 segment is 185 meters.

10Base2 uses a much lower grade of coaxial cable than 10Base5. The cable is also a lot thinner and more flexible which makes it easier to manipulate and capable of being brought right up to the DTE. This, in conjunction with the fact that the 10Base2 transceiver function is generally integrated into most of the Ethernet adapters, provides the user with the option to connect the DTE to the bus directly and avoid the use of AUI cable. However, because of the lower quality of the cable, there is a reduction in both the segment length available and number of transceivers supported when compared to 10Base5.

A 10Base2 segment consists of a number of thin coax cables connected to each other via a number of T-connectors. In addition to connecting the two cables together, a T-connector provides a BNC connection for attaching the DTE. The use of BNC type connectors makes adding and removing transceivers a straightforward task in a 10Base2 network. Figure 16 on page 44 shows an example of a typical 10Base2 segment.

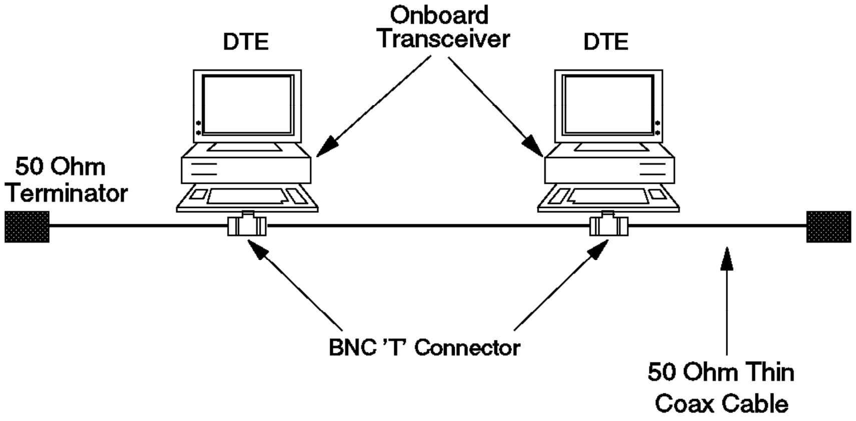


Figure 16. 10Base2 Segment

Note that terminators are used at both ends of a segment to prevent the signal from being reflected back when it reaches the end of the segment.

Table 7 provides a summary of the 10Base2 specification.

Table 7. 10Base2 Specification

Item	Specification
Cable type	RG-58A/U, 50 ohm coaxial cable
Connectors	BNC type
Termination	Segment ends not attached to repeaters must be terminated with 50 ohm terminators
Transceiver cable type	Four-strand, shielded twisted-pair conductors with overall shield and insulating jacket.
Data rate	10 Megabits/sec
Max segment length	185 meters
Min distance between transceivers (or T-Connectors)	0.5 meter
Max no. of transceivers/segment	30 transceivers
Max transceiver cable length	50 meters
Max no. of stations per network	1024 adapters
Impedance	50 ohms (+/- 2)
Attenuation	8.5 dB for 185 meters at 10 MHz
Max propagation delay/segment	950 nanoseconds
DC resistance	10 ohms per segment

Because of the relative simplicity of running and attaching stations to it, 10Base2 is often used to extend the services offered by an existing 10Base5 network.

The advantage of 10Base5 in terms of the segment length available can be utilized for parts of the LAN where change will be minimal such as through ducts and risers to provide a backbone bus.

The advantage of 10Base2 in terms of the cable itself being easier to manipulate plus the relative ease with which transceivers can be added and removed can be utilized in areas of the LAN where changes will be made more frequently to the configuration of the network.

Repeaters and/or bridges must be used to connect the segments.

3.1.2.3 10Base-T

During the 1980s, due to the decrease in the cost of hardware, the concept of a workstation on every desk became a reality and many businesses became dependent on the need to attach these workstations to local area networks to enable them to exchange information with the other workstations and provide them with access to the available services within the organization. This stretched implementations of 10Base5/10Base2 networks to their limits. Manufacturers of the attachment devices came up with numerous ingenious ways of making these systems more flexible but could not hide the fact that the basic requirement of connecting the workstations in series was incapable of providing the flexibility needed.

It was also becoming clear that a more structured approach to the whole subject of providing services into the business environment was required. Many organizations were discovering that a majority of buildings were not able to make adequate provisions for business services such as telephone, telex, fax, data and video.

A number of companies were marketing structured cabling systems designed to provide a single cabling infrastructure over which most services could be provided. These were mainly star-wired systems, in which the cabling radiated from a wiring closet to service a defined area within the building. The wiring closets were also linked together to allow services to be connected between any two points in the building.

The EIA/TIA (Electronics Industries Association/Telecommunications Industries Association) brought out a standard for the physical cabling of buildings to provide data and voice services. This provides the standards for:

- Topology - structured wiring using a star topology between the work area and the wiring closet.
- Maximum cabling distance (point-to-point) between the wiring closet and the work area.
- Recommended media and connectors:
  - 100 ohm Unshielded Twisted Pair (UTP) consisting of four 24 AWG wire pairs.
  - 8-pin modular jack and plug such as RJ-45 pairs.
  - 150 ohm Shielded Twisted Pair (STP) consisting of two individually shielded pairs and a common shield around both pairs.
  - Media connector as specified in the IEEE 802.5 standard.

The standard also provides recommendations for the use and application of fiber and coaxial cables within the building.

The 10Base-T standard was defined by IEEE to address the requirement of running Ethernet/802.3 over the structured cabling systems using twisted pair copper wires. Although, actually completed prior to the EIA/TIA 568 standard, a 10Base-T Ethernet (802.3) LAN requirement would be met by a cabling system that conformed to EIA/TIA 568.

The term 10Base-T was chosen for this standard because:

10	indicates the data rate	(10 Mbps)
BASE	indicates the transmission type	(Baseband)
T	indicates the medium	(Twisted Pair)

10Base-T is a star topology in which the DTEs are attached to a central *hub*. The hub acts as a multiport repeater between a number of segments in which each segment is a point-to-point connection between a DTE and a port on the hub.

A segment can also be a point-to-point connection between two hub ports. This would allow you to set up a network consisting of multiple hubs. Also, by taking advantage of bridges and repeaters (which normally are offered as modules that can be installed on these hubs), networks consisting of mixed topologies of 10Base-T 10Base5, 10Base2, etc. can be constructed.

Table 8 provides a summary of the 10BaseT specification.

Table 8. 10Base-T Specification

Item	Specification
Cable type	2 unshielded twisted-pairs (UTP)
	0.4 mm AWG 26
	0.5 mm AWG 24 (most widely used)
	0.6 mm AWG 22
Connectors	RJ-45
Termination	No external terminators are required
Data rate	10 Mbps
Single segment length	100 meters (point-to-point)
Max no. of repeaters/segment	2 multiport repeaters
Impedance	85 - 111 ohms (nominal 100)
Attenuation	8.5 - 10 dB for 100 m at 10 MHz
Max. propagation delay/segment	1000 nanoseconds

3.1.2.4 FOIRL and 10Base-FL

Fiber Optic Inter Repeater Link (FOIRL) was the first standard to be defined for the use of fiber optic cables in an Ethernet LAN. Although it was originally intended as a repeater-to-repeater link only, providing a long-distance connection of up to 1 km between two repeaters, it has also been used to allow fiber connectivity to the desktop.

The use of FOIRL for desktop connectivity was originally excluded from the standard, but the 10Base-FL standard, which is specified to supersede FOIRL, permits such connections.

FOIRL is similar to the 10Base-T standard. It requires the use of a separate transmit and receive path. It also requires the use of repeaters in a central hub acting as the concentration point for a group of nodes.

Similar to 10Base-T, an FOIRL MAU is required to perform link integrity. FOIRL link integrity is performed by each MAU transmitting a 1 MHz signal when no data transmission is taking place. If the MAU at the other end fails to detect this signal it enters *link fail* state and prevents the DTE from transmitting onto the network.

10Base-FL extends the allowable distance between two MAUs to 2 km.

3.1.2.5 10Base-FB

10Base-FB is designed to provide a superior technology using synchronous signalling over the fiber cables. A 2.5 MHz active idle signalling is used to indicate that the transmit path is idle. In addition, the transmit data from the repeater is synchronized to this idle signal, enabling the receiving MAU to remain locked to the active/idle packet data transition.

3.1.3 Ethernet Design Rules

The following is a summary of the considerations that should be taken into account when designing an Ethernet network. But, please note that it is not a complete review of all the design considerations for Ethernet networks as there are many more factors that should be considered when designing local area networks in general and Ethernet LANs in particular.

These considerations for designing Ethernet networks are used to ensure that data transmitted by the source will be received by the destination error free and any collisions that occur can be reliably detected.

Maximum segment lengths for each medium type must exceed the stated limits for that medium:

- 10Base5: 500 m
- 10Base2: 185 m
- 10Base-T: 100 m
- 10Base-FL: 2000 m
- 10Base-FB: 2000 m

Note that certain products will allow you to go beyond these limits. For example, the 8250 10Base-T module will allow segments to be longer than 100 m. The maximum number of stations allowed on a segment varies according to the type of medium used:

- 10Base5: 100 stations
- 10Base2: 30 stations
- 10Base-T: 2 stations
- 10Base-FL: 2 stations
- 10Base-FB: 2 stations

The maximum number of stations in a *collision domain* is 1024.

- Repeaters can be attached at any position on the *coax segments* but should be at the ends of a *link segment*.
- Each repeater takes *one* attachment position on the segment and should be counted towards the maximum number of stations allowed on that medium.
- You can have many segments and repeaters within a single *collision domain* as long as no two DTEs in the same collision domain are separated by more than *four repeaters*.
- No two DTEs in the same collision domain can be separated by more than three *coax segments*. The other two segments in a maximum configuration must be *link segments*.
- *Link segments* can be 10Base-T, FOIRL, 10Base-FL and 10Base-FB.
- 10Base2 and 10Base5 segments cannot be used as *link segments*.
- 10Base5 and 10Base2, 10Base-T and fiber segments can be mixed in a single collision domain allowing you to take advantage of the facilities offered by the most appropriate medium for different parts of your network.

3.1.4 Ethernet Summary

The IEEE 802.3 standard is a contention-based protocol. For low-traffic volumes and relatively short messages, an IEEE 802.3 LAN can provide the best response time. However, because of the instability introduced by collision recovery during periods of high utilization, response times and performance cannot be predicted reliably. At the MAC level, the IEEE 802.3 protocol cannot guarantee access to the medium. These performance considerations make IEEE 802.3 protocol LANs less desirable than other protocols for backbone LANs.

All stations have equal rights to transmit on an idle bus. In low-traffic situations, this minimizes the need to handle priority requests. However, with heavier traffic, the lack of priority support within the protocol may be a problem for some applications. Special considerations may be required for applications or data with security requirements, because IEEE 802.3 stations broadcast their frames simultaneously to all other LAN stations on the network which therefore can receive and copy any transmitted data.

Network management is not defined within the IEEE 802.3 standard. It is under consideration as part of the IEEE 802.1 sub-project.

3.2 Token-Ring/802.5

The IEEE 802.5 standard describes the token-ring medium access protocol and its physical attachments.

In a token-ring network the stations on the LAN are physically connected to a wiring concentrator usually in a star-wired ring topology. Logically, stations are connected in a pure ring topology. Each station has driver/transmitter as well as receiver circuitry; see Figure 17 on page 49.

Differential Manchester code is used to convert binary data into signal elements, which are transmitted at 1, 4, or 16 Mbps (IEEE standard speeds). The standard does not prescribe the type of cabling to be used. In IBM′s token-ring network implementation, shielded twisted pair cabling is recommended although UTP may now be used.

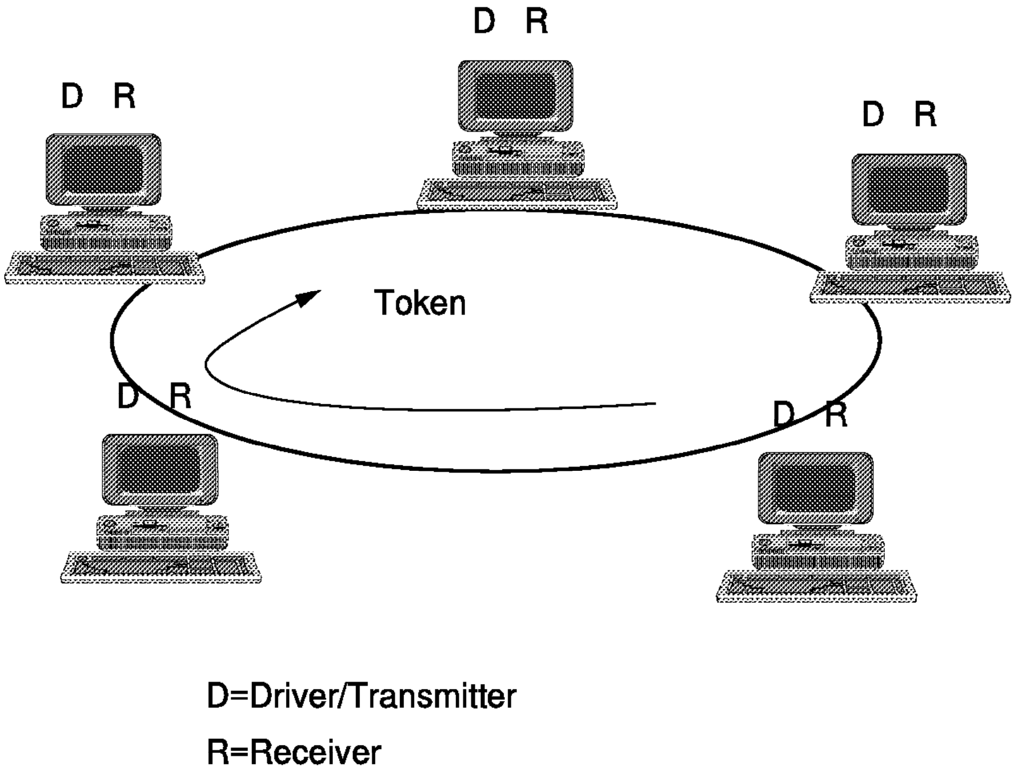


Figure 17. Sample Ring Configuration

Access to the ring is controlled by a circulating token. A station with data to transmit waits for a free token to arrive. When a token arrives, the station changes the token into a frame, appends data to it and transmits the frame. If the destination station is active, it will copy the frame and set the frame copied and address recognized bits, providing MAC level acknowledgment to the transmitting station. The sending station must strip the frame from the ring and release a new token onto the ring.

An option in the architecture allows the sending station to release a token immediately after transmitting the frame trailer, whether or not the frame header information has already returned. This is called early token release and tends to reduce the amount of idle time in 16 Mbps token-passing rings.

The token-passing protocol provides an extensive set of inherent fault isolation and error recovery functions, for implementation in every attaching device. The adapter network management functions include:

- Power-on and ring insertion diagnostics
- Lobe-insertion testing and online lobe fault detection
- Signal loss detection, beacon support for automatic test and removal
- Active and standby monitor functions
- Ring transmission errors detection and reporting
- Failing components isolation for automatic or manual recovery

The token-passing ring medium access protocol will be described in the following sections.

In summary, the token-passing ring protocol is based on the following cornerstones:

- Active monitor
  - Ensures proper ring delay
  - Triggers neighbor notification
  - Monitors token and frame transmission
  - Detects lost tokens and frames
  - Purges circulating tokens or frames from the ring
  - Performs auto-removal in case of multiple active monitors
- Standby monitor (any other ring station)

  Detects failures in the active monitor and disruptions on the ring.

- Token claiming process

  A new active monitor is elected when the current active monitor fails. This process may be initiated by the current active monitor or by a standby monitor.

Figure 18 shows the format of an 802.5 standard MAC frame as well as the token format and the format of the abort delimiter.

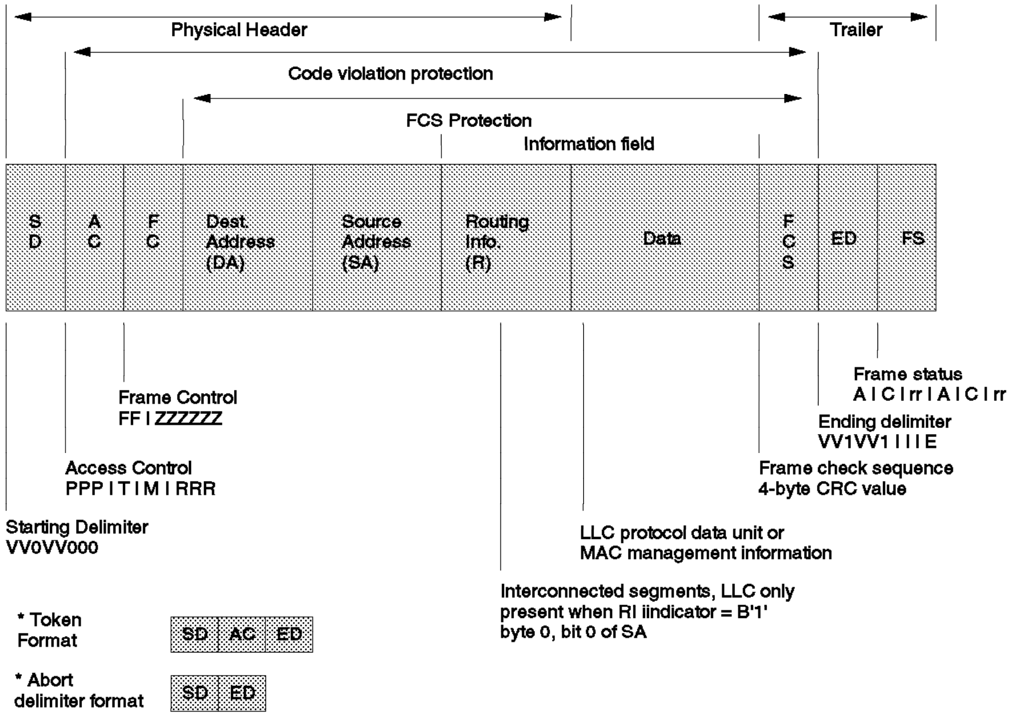


Figure 18. 802.5 Standard MAC Frame

The architecture describes 28 different MAC control frames, each identified by a unique major vector identifier (MVID). The main ones will be described in the following and are referred to as:

- Active Monitor Present MAC frame

- Ring Purge MAC frame
- Standby Monitor Present MAC frame
- Claim Token MAC frame
- Lobe Media Test MAC frame
- Duplicate Address Test MAC frame
- Request Initialization MAC frame
- Beacon MAC frame
- Soft Error Report MAC frame

3.2.1 Token-Ring Concepts

A token-ring network consists of the attaching medium and ring stations (devices able to attach to the ring and to use the link access protocols). A token-ring network uses one of several twisted pair media specifications, each having its own price/performance ratio, and all suitable to carry most other data communications signals. A token-ring network may also use optical fiber media.

A token-ring LAN installation is illustrated in Figure 19.

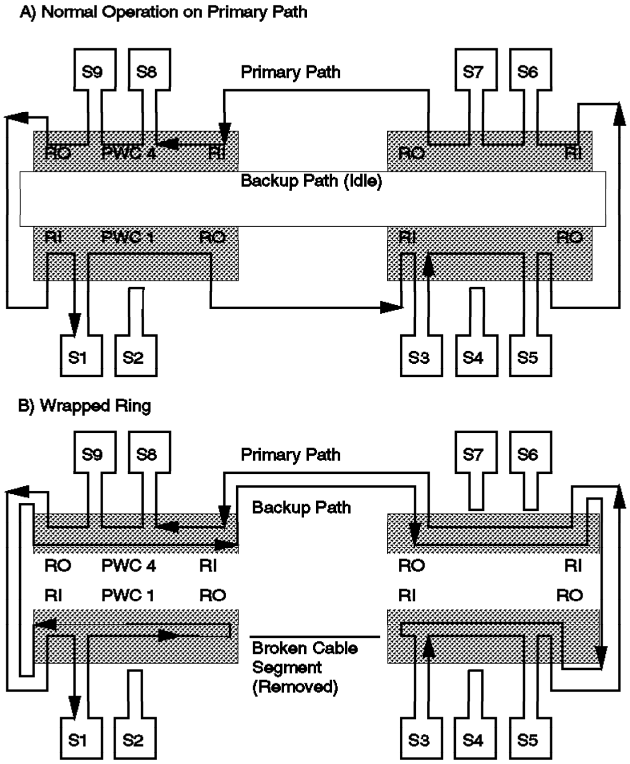


Figure 19. Normal and Wrapped Token-Rings

This figure shows four passive wiring concentrators (PWCs) and nine physically attached nodes. The power from the attached nodes when transmitted to the concentrator activates relays in the concentrator to allow the station to send signals across the LAN to other stations. In Part A of Figure 19, adapters (nodes S2 and S4) have not powered their respective PWC relays and, therefore, their lobe wires are internally bypassed. In part B of Figure 19, four adapters (nodes S2, S4, S6 and S7) are not actively inserted into the ring (their lobe wires are internally bypassed) and the primary path is wrapped to the backup path in PWCs 1 and 2.

When a cable segment between PWCs fails, manual removal from the appropriate ring-in and ring-out connectors causes automatic wrapping of the primary path to the backup path. How such a permanent wire fault is reported for LAN management is explained in the discussion of beaconing later on in this section. Recovery from the same error is automatic when using the IBM 8230 Controlled Access Unit, which is an active, or powered, wiring concentrator.

A ring station transfers data to the ring, in a data transmission unit called a frame. Frames are sent sequentially from one station to the next station physically active on the ring. This station is called the downstream neighbor. Each ring station repeats the frame. While doing so it performs error checking on the bit stream and it will copy the data if its own address, either its MAC or any of its functional addresses, are identified as a destination station in the frame. Upon return of the frame to the originating station, the latter will remove the data from the ring. In a token-passing protocol, a ring station can only transfer data to the ring while it is holding a token. The token is a specific bit sequence (24 bits) circulating around the ring at a rated speed of (4 Mbps or 16 Mbps in current implementations) 100 Mbps according to the Fiber Distributed Data Interface specification; see Chapter 2, “Physical LAN Attachment” on page 7. Because of the high transmission speed with respect to the total ring length, a short ring might contain only a few bits at any point in time. Only one token may exist on a ring segment at any given point in time. Therefore, a delay equivalent to the time it takes for a token to circulate the ring is required to ensure that no overrun occurs which would result in a station receiving a token that it is transmitting and thinking that a second token exists on the ring. For a 24-bit token this means a minimum 24-bit delay. In addition to this delay an additional elastic buffer is introduced to support the token protocols and speed.

In order to establish communication between any two ring stations, addressing mechanisms are needed. At the same time the integrity of the transmitted frames between ring stations must be preserved. Therefore data checking capabilities are required at the medium access control level of a ring station.

3.2.1.1 MAC Addressing

All ring stations are identified by a unique individual address. This address can be universally administered, assigned by the IEEE organization (see Chapter 3, “LAN Architectures and Standards” on page 25). Because it is set in read-only memory (ROM) on a token-ring adapter card, the universally administered address is also called a burned-in address.

Some manufacturers have been assigned universal addresses that contain an organizationally unique identifier. For instance, IBM has an identifier of x′10005A′. All IBM token-ring cards that use IBM token-ring chip sets, have the first 6 digits of their address begin with those characters. Other identifiers are x′000143′ for IEEE 802, and x′1000D4′ for DEC. IEEE universal addresses, whether for token-ring or 802.3 stations are all allocated out of the same common pool, but uniqueness is guaranteed.

A ring station′s individual address can also be locally administered, that is set at adapter-open time and typically defined by a network administrator. A number of destination ring stations can be identified by a group MAC address. Some standard group addresses have been defined. These are listed in Table 9 on page 53

Table 9 (Page 1 of 2). Standardized Group Addresses

Bridge	X′8002 4300 0000′
Bridge management	X′8001 4300 0008′
Load server	X′8001 4300 0088′
Loadable device	X′8001 4300 0048′
ISO 10589 level 1 1 intermediate stations	X′8001 4300 0028′
ISO 10589 level 2 1 intermediate stations	X′8001 4300 00A8′
FDDI RMT directed beacon	X′8001 4300 8000′
FDDI status report frame	X′8001 4300 8008′
OSI network layer end stations	X′9000 D400 00A0′
OSI NL intermediate stations	X′9000 D400 0020′
Reserved for transparent bridging	X′8001 4300 000x′
All LANs bridge mgt group address (802.1D)	X′8001 4300 0008′
All cons end systems (ISO 10030)	X′8001 4300 0068′
All cons snares (ISO 10030)	X′8001 4300 00E8′
FDDI all root concentrator MACs (ANSI X3T9.5)	X′8001 4300 1004′
Reserved for FDDI	X′8001 4300 10X0′
Loopback assistance	X′F300 0000 0000′
AppleTalk support	X′9000 E000 0000′
AppleTalk highest address within range except broadcast	X′9000 E000 003F to X′9000 E0FF FFFF′
Novell IPX	X′9000 7200 0040′
Hewlett Packard probe	X′9000 9000 0080′
HP DTC	X′9000 9000 0020′
Apollo domain	X′9000 7800 0000′
Vitalink diagnostics	X′9000 3C40 00A0′
Vitalink gateway	X′9000 3CA0 0080′
LANtastic	X′FFFF 0006 0020′
LANtastic	X′FFFF 0002 0080′
LANtastic	X′FFFF 8007 0020′
Concord DTQNA	X′0000 9640 XXXX′
DEC DNA Dump/load assistance (MOP)	X′D500 0080 0000′
DEC DNA remote console (MOP)	X′D500 0040 0000′
DNA level 1 routing layer	X′D500 00C0 0000′
DNA routing layer end nodes	X′D500 0020 0000′
Customer use	X′D500 2000 XXXX′
System Communication Architecture	X′D500 2080 XXXX′

Table 9 (Page 2 of 2). Standardized Group Addresses

VAXELN	X′D500 D400 0040′
LAN traffic monitor	X′D500 D400 00C0′
CSMA/CD encryption	X′9000 D400 0060′
NetBIOS emulator (PSCG)	X′9000 D400 00E0′
Local area transport (LAT)	X′9000 D400 00F0′
All bridges	X′9000 D480 0000′
All local bridges	X′9000 D480 0080′
DNA level 2 routing layer routers	X′9000 D440 0000′
DNA naming service advertisement	X′9000 D440 8000′
DNA naming service solicitation	X′9000 D440 8080′
LAT directory service solicit (to slave)	X′9000 D440 8020′
FDDI ring purger advertisement	X′9000 D440 80A0′
LAT directory service solicit - X service class	X′9000 D440 80D0′
Local area system transport (LAST)	X′9000 D420 XXXX′
UNA prototype	X′5500 C000 XXXX′
Prom 23-365A1-00	X′5500 C080 XXXX′
Misc.	X′5500 C040 XXXX′
H400 - TA Ethernet transceiver tester	X′5500 C040 0000′
NI20 products	X′5500 C0C0 XXXX′
DECnet phase IV station addresses	X′5500 2000 XXXX′
Prom 23-365A1-00	X′1000 D40X XXXX′
Prom 23-365A1-00	X′1000 D48X XXXX′
Bridge mgt.	X′1000 D444 0000′
Prom 23-365A1-00	X′1000 D4C4 XXXX through X′1000 D4CX XXXX′
Shadow for prom 23-365A1-00	X′1000 D42X XXXX′
Shadow for prom 23-365A1-00	X′1000 D4AX XXXX′
Shadow for prom 23-365A1-00	X′1000 D4B6 XXXX′ through X′1000 D4EX XXXX′
VAXft 3000 fault tolerant LAN addresses	X′1000 D407 XXXX′
VAXft 3000 fault tolerant LAN addresses	X′1000 D40F XXXX′

A token-ring LAN also provides a special case of a locally administered group address called functional addresses. Each (bit-significant) functional address represents a well-identified server function within the access protocol. Of 31 possible functional addresses, 22 have been defined while the remaining ones are reserved for future use or may be user-defined. They are listed in Table 10.

Table 10 (Page 1 of 2). New and Current IEEE and IBM Functional Addresses

Active monitor	X′C000 0000 0001′
Ring parameter server	X′C000 0000 0002′
Network server heartbeat	X′C000 0000 0004′

Table 10 (Page 2 of 2). New and Current IEEE and IBM Functional Addresses

Ring error monitor	X′C000 0000 0008′
Configuration report server	X′C000 0000 0010′
Synchronous bandwidth mgr	X′C000 0000 0020′
Locate - directory server	X′C000 0000 0040′
NetBIOS	X′C000 0000 0080′
Bridge	X′C000 0000 0100′
IMPL server	X′C000 0000 0200′
Ring authorization server	X′C000 0000 0400′
LAN gateway	X′C000 0000 0800′
Ring wiring concentrator	X′C000 0000 1000′
LAN manager	X′C000 0000 2000′
User-defined	X′C000 0000 8000′ through X′C000 4000 0000′
ISO OSI ALL ES	X′C000 0000 4000′
ISO OSI ALL IS	X′C000 0000 8000′
IBM discovery non-server	X′C000 0001 0000′
IBM resource manager	X′C000 0002 0000′
TCP/IP	X′C000 0004 0000′
6611-DECnet	X′C000 2000 0000′
LAN Network Mgr & 6611	X′C000 40000 0000′

The most relevant protocol server functions will be described in greater detail in the Bridge and LAN management sections of *Local Area Network Concepts and Products:Routers and Gateways,* SG24-4755.

In addition two special destination address values have been defined. The all-stations broadcast group address X′FFFFFFFFFFFF′ identifies all ring stations as destination stations. A frame carrying the individual null address X′000000000000′ as its destination MAC address is not addressed to any ring station; therefore, it can be sent but not received.

IEEE allows vendors to implement either 16-bit or 48-bit MAC addresses. The actual address field formats are shown in Figure 20 on page 56.

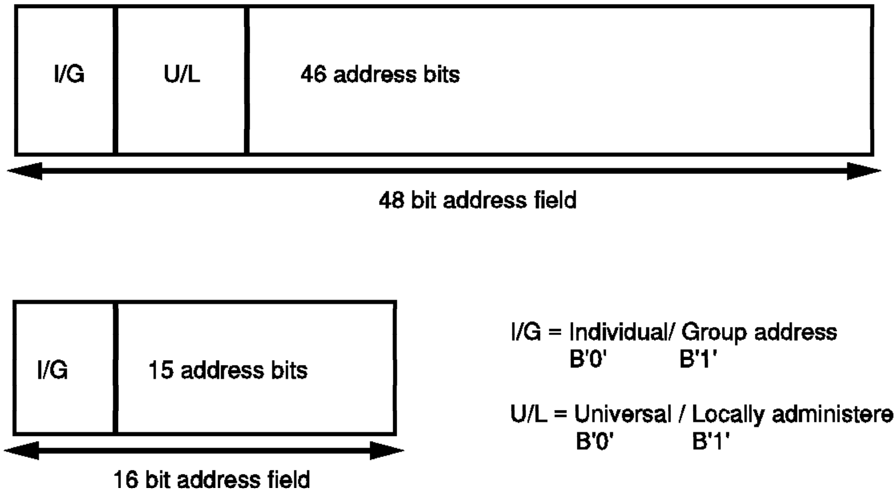


Figure 20. IEEE LANs - MAC Address Format

For the IBM LAN implementations, 48-bit addressing has been selected. The implementation format is shown in Figure 21.

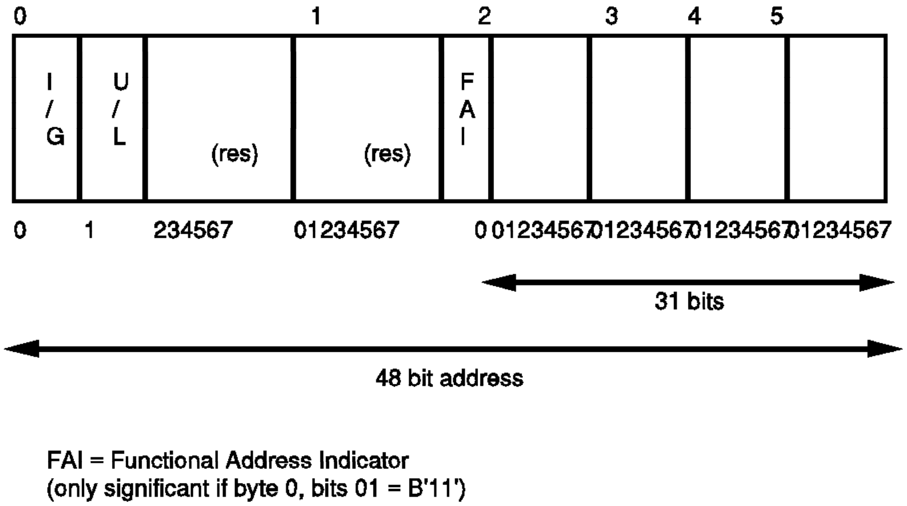
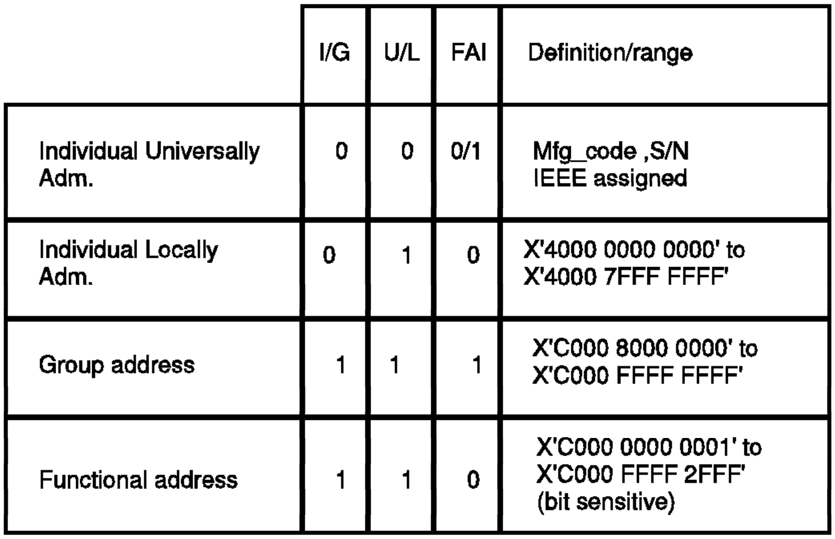


Figure 21. IBM Token-Ring Network - MAC Address Format

- The reserved bits are set to B′0′ for locally administered addresses.
- Functional address indicator = B′0′ indicates a functional address if I/G = B′1′ (indicating a group address).

- For individual locally administered addresses, FAI must be B′0′ by convention. This is an addressing anomaly.

These rules yield valid address ranges as described in Figure 22 for any IBM Token-Ring Network adapter.



	I/G	U/L	FAI	Definition/range
Individual Universally Adm.	0	0	0/1	Mfg_code ,S/N IEEE assigned
Individual Locally Adm.	0	1	0	X'4000 0000 0000' to X'4000 7FFF FFFF'
Group address	1	1	1	X'C000 8000 0000' to X'C000 FFFF FFFF'
Functional address	1	1	0	X'C000 0000 0001' to X'C000 FFFF 2FFF' (bit sensitive)

Figure 22. Valid Address Ranges

3.2.1.2 Data Transmission

The transmission technique used in token-passing rings is baseband transmission. In a token-ring LAN, high-order bytes/bits are transmitted first; that is, byte 0 is transmitted before byte 1 and high-order bit 0 within a byte (of 8 bits) is transmitted first. This transmission order can be different for other types of LAN segments using different access protocols, for example, CSMA/CD. Opposite transmission order may be a diagnostic consideration when evaluating trace information from LAN segments of a different nature because of the possible need to reorder the bits. The ability to reorder the bits without significant performance degradation may also be a functional requirement of the bridge products being considered for a LAN segment interconnection. Figure 23 on page 58 shows the format of the information to be transmitted on a token-passing ring.

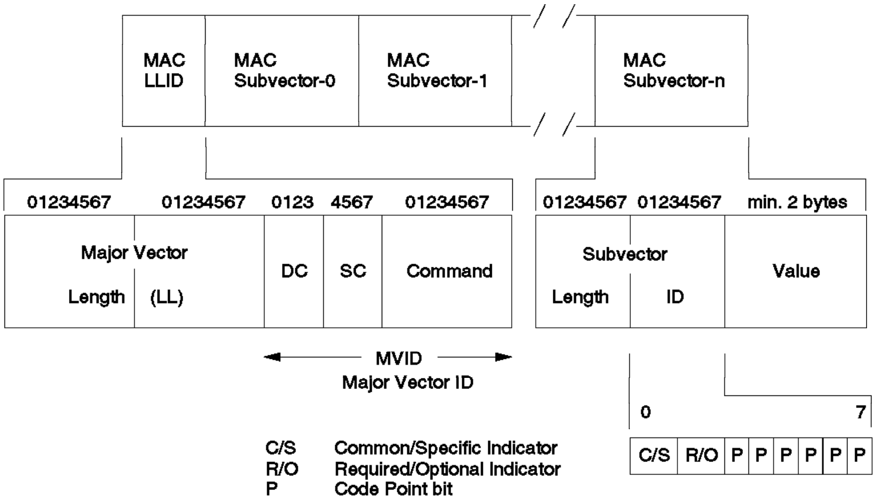


Figure 23. Token-Ring MAC Frame - Data Field Format

Examples of MVID code points are X′05′ to indicate an Active Monitor Present MAC frame, X′02′ for a Beacon MAC frame, etc. Unique MVID values for 28 different MAC frames have been defined. A complete description can be found in *IBM Token-Ring Network Architecture Reference,* SC30-3374. Sub-vectors provide additional information depending on the specific major vector identifier. MAC frames are processed according to destination and source function classes, including:

- **Ring Station**

  The functions necessary for connecting to the LAN and for operating with the token-ring protocols. A ring station is an instance of a MAC sublayer in a node attached to a ring.

- **DLC_LAN_MGR**

  The manager function of the data link control component activates and deactivates ring stations and link stations on command from the physical device. It also manages information transfer between data link control and the physical device.

- **Configuration Report Server (CRS)**

  The CRS function can reside on each ring in a multisegment environment in which configuration is being monitored. This function receives notifications about station insertion and removal, and notifications about active monitor failures.

- **Ring Parameter Server (RPS)**

  The RPS function can reside on every segment in a multisegment ring environment on which operational parameters are centrally managed. It may provide operational values to attaching stations upon request. For example, a ring station will request such parameters as ring number upon insertion into the ring.

- **Ring Error Monitor (REM)**

  The REM server function is present on segments for which errors are to be monitored or analyzed. It collects error information from LAN stations attached to the local ring, analyzes soft error reports and possibly forwards error reports to a LAN Manager.

- **RPL Server**

  The RPL server function and its RPL functional address are involved during the power-on process of a LAN station equipped with the remote program load feature. Such a station will insert into the ring to find a control program server on the ring from which to download its control program and complete its initialization processes.

- **The Token-Passing Ring Protocol**

  To transmit data on the LAN medium, a ring station captures a token and sets the token bit in the access control field to identify that the data being transmitted is a frame. To this header, the transmitting station appends destination and source MAC addresses, data, a newly calculated frame check sequence field, and the ending delimiter and frame status fields.

  Any subsequent station will receive and retransmit the frame while performing a CRC check. Such a station is said to be in normal repeat mode.

  In general, a ring station in normal repeat mode checks the data in the tokens and frames it receives and sets the error-detected, address recognized or frame copied bits in a frame (bits E, A, or C) as appropriate while repeating the signal. A destination station will copy the data (frame copied) and pass the frame on. While processing the frame trailer, the destination station marks the A and C bits. Upon return to the originating ring station, the frame is removed from the ring and the A and C bits in the frame trailer′s FS field are checked to see if the frame was recognized and read by the destination station or a bridge (this occurs for MAC frames only). When the frame header is received by the originating station the originating ring station must release a new token, possibly at a different priority level, for another ring station to capture and proceed with data transmission. The priority reservation bits in the access control field of the returned frame together with stored priority levels in the originating station determine the priority of the new token. See the description of access priority later in this section for details.

  This protocol is called a single token protocol, since only one token can circulate on the ring at any time.

3.2.1.3 Early Token Release

If an originating station releases a new token only when the frame header has circulated around the ring back to the source and the frame transmission time is shorter than the ring transmit time, then the originating station must generate idles until a header is received.

Token-passing ring protocols define the length of a token to be 2 bits and the shortest possible MAC frame to be 200 bits long. On a 4 Mbps token-ring LAN where the length of 1 bit is roughly 50 meters, a complete token is 1,200 meters long while the shortest frame length would be 10,000 meters. Therefore at 4 Mbps, the percentage of potential bandwidth which remains idle can be extremely small (that is high bandwidth utilization can be maintained at higher traffic levels).

If, however, we consider a 16 Mbps token-ring LAN, 1 bit being 12.5 meters long, along with a complete token and the shortest possible MAC frame both become four times smaller (300 and 2,500 meters respectively). Now we may wish to optimize the utilization of the medium by reducing the idle time required by waiting for a header. Obviously, when moving to even higher transmission speeds, (for example, a 100 Mbps FDDI LAN) the token-passing protocol must be adjusted to achieve better utilization of the potential bandwidth.

The architecture provides an option called *early token release.* With this option a transmitting station will release the token after completing the transmission of the data frame before the receipt of the header of the transmitted frame; they thereby eliminate the idle time while waiting for the header to reappear. When such early release has occurred, an adapter indicator is set to prevent the adapter from releasing another token upon return of the frame header. This allows multiple frames but still only one token on the LAN.

The early token release option is enabled by default on the 16 Mbps IBM Token-Ring Network. It is an option for each station, and it is not required that all stations implement the option but is recommended.

3.2.1.4 Token Monitoring

Token-passing protocols provide relatively greater control and management at the medium access control (MAC) level than that provided by CSMA/CD protocols. The token-passing ring protocol concepts, described in the following sections, are implemented in the adapters themselves. They contribute to the availability, performance and manageability of a token-ring LAN. At any point in time, one and only one station per segment performs an active monitor function. Any ring station can act as the active monitor. Only one will have this function enabled. Active monitor tasks support the monitoring of the token and other ring management functions such as the following:

- Detection and recovery of a lost token or frame, including initiation of a token when a ring is started
- Detection and recovery of a circulating priority token or frame
- Detection and recovery of multiple tokens on the ring
- Detection and recovery of multiple active monitors
- Timing control to ensure accurate transmission regardless of the ring length

All other ring stations are said to be *standby monitors*, prepared to take over the active monitor function should it fail. The following description summarizes how the active monitor performs its ring management tasks:

- In every transmitted token or frame, the monitor bit (M) in the access control field is initially set to B′0′. As the active monitor repeats a frame or non-zero priority token, the M-bit is set to B′1′. If the M-bit had already been set to B′1′, the active monitor assumes that the frame or token has already circled the ring once and that the originating station has not been able to remove the frame or priority token. The active monitor will purge the ring and generate a new token.
- To ensure that a complete (24-bit) token can be transmitted before the token returns to the originating ring station, the active monitor introduces a 24-bit ring delay.
- The active monitor periodically broadcasts (every seven seconds) an Active Monitor Present MAC frame. This forces each station on its ring to acquire the address of its nearest active upstream neighbor (NAUN) and to initiate a

number of control timers within each station. This information is used when isolating errors on a segment.

- Loss of a token or frame is detected by expiration of an any-token timer whose time-out value exceeds the time required for the longest possible frame to circle the ring. The active monitor restarts this timer each time it transmits a starting delimiter. Upon expiration of this timer, the active monitor assumes a lost token or frame, purges the ring and originates a new token. The any-token timer value is defined as the sum of the physical trailer transmission delay plus the delay to transmit the longest frame. The IEEE 802.5 name for this timer is Valid_Transmission_Timer.

3.2.1.5 Ring Purge

To purge the ring, the active monitor broadcasts a Ring Purge MAC frame (indicated by X′04′ in the frame control field) before originating a new token. Return of the Ring Purge MAC frame indicates proper signal propagation around the ring. The Ring Purge frame resets the ring stations to normal repeat mode, canceling or restarting all the appropriate timers. The active monitor starts a Ring-Purge timer when sending the purge frame. This timer will expire if the frame can′t circulate and the monitor will enter a recovery process called a Claim Token.

3.2.1.6 Neighbor Notification

The neighbor notification process begins when the active monitor transmits an Active Monitor Present MAC frame to all stations on the ring (single ring broadcast). The first ring station that receives the Active Monitor Present MAC frame copies it (if possible) and sets the address-recognized (A) and frame-copied (C) bits to B′1′. It then saves the source address from the copied frame as its NAUN address (the address of the active monitor) and starts a timer called the Notification-Response timer.

All other active stations on the ring repeat, but do not otherwise process the Active Monitor Present MAC frame because the frame′s A and C bits have already been set.

When the Notification-Response timer of the first station downstream from the active monitor expires, it broadcasts a Standby Monitor Present MAC frame.

The next ring station downstream copies its NAUN address from the source address field of the Standby Monitor Present frame, sets the A and C bits to B′1′, and starts its own Notification-Response timer. When this timer expires, this station in turn transmits its Standby Monitor Present MAC frame.

In this way, neighbor notification proceeds around the ring, with each ring station receiving and transmitting Standby Monitor Present MAC frames until the active monitor copies its NAUN address from a Standby Monitor Present MAC frame. The active monitor then sets the Neighbor - Notification Complete flag to B′1′, indicating that the process has been successfully completed. Neighbor notification thus enables a ring station to learn its NAUN address, and to provide its address to its downstream neighbor.

3.2.1.7 Standby Monitor

Any ring station that is not performing the active monitor function acts as a standby monitor. Its purpose is to detect a failing active monitor and disruptions on the ring.

Each time a token or frame is repeated, a standby monitor restarts its good-token timer to verify the presence of an active monitor.

A second timer, the Receive-Notification timer, is restarted by a standby monitor each time it copies an Active Monitor Present MAC frame. If any of these two timers expires, the standby monitor station will initiate the token-claiming process.

3.2.1.8 The Token-Claiming Process

This process, also called the monitor-contention process, is the procedure by which ring stations elect a new active monitor. This process is started upon any of the following conditions by:

- The active monitor detecting the following:
  - Loss of signal.
  - The Active Monitor Present MAC frame doesn′t return (Receive-Notification timer expires).
  - Failure of Ring-Purge MAC frames to return completely (Ring Purge timer expires).
- A standby monitor detecting the following:
  - Loss of signal
  - Absence of active monitor′s token management functions (good_token timer expires).
  - Missing Neighbor_Notification process (Receive-Notification timer expires).
- A ring station attaches to the ring and does not detect an active monitor (for example, when it is the first station on the ring).

The ring station detecting one of these conditions enters Claim-Token-Transmit mode by broadcasting a Claim Token MAC frame and repeating it at a defined interval. Each participating ring station compares the address in the Claim Token MAC frame′s source address field to its own.

- If the source address is greater than the ring station′s address, the station enters Claim Token Repeat operating mode.
- If the source address is less than the ring station′s address, the station transmits its own Claim Token MAC frames.
- If the source address is the same as the ring station′s address, it continues broadcasting until it has received three of its own Claim Token MAC frames. This indicates that the ring is viable and the station has won token-claiming.

  The station then adds the token delay to the ring, purges the ring, starts its active monitor timers, and releases a new token. It is now the new active monitor.

3.2.1.9 Ring Station Insertion

This process is executed by any ring station when entering the ring. It is also known as the five phase insertion process.

- Phase 0: Lobe testing

  A series of Lobe Media Test MAC frames are sent on the lobe wire to the multistation access unit. The signal is wrapped at the entry into the multistation access unit causing the frames to return to the station. Then the receive logic is tested. If the tests are successful, a 5 volt DC current (also called phantom current) is sent to open the relay and attach to the ring.

- Phase 1: Monitor check

  The attaching station starts its Insert timer, and watches for an Active_Monitor_Present, Standby_Monitor_Present or Ring Purge MAC frame before this timer expires. If the timer expires, token-claiming is initiated. When it is on the first station on ring, the attaching station will become the active monitor.

- Phase 2: Duplicate address check

  The station sends a Duplicate Address Test MAC frame (destination address = source address = station′s unique address). If a duplicate address is found (A-bit = B′1′), the station detaches from the ring.

- Phase 3: Participation in neighbor notification

  The station learns its nearest active upstream neighbor (NAUN) and reports its own address to its nearest active downstream neighbor.

- Phase 4: Request initialization

  A Request Initialization MAC frame is sent to the ring parameter server, if present (if not, default values will be used). The ring parameter server responds with an Initialize_Ring_Station MAC frame. Parameters which can be set are physical location, soft error report timer value, ring number and ring authorization level. In this way the last three parameters may be set to the same values for all stations on the ring.

3.2.1.10 Hard-Error Detection and Reporting

A hard error is a permanent fault that stops normal traffic on the ring. It is usually detected first at the receive side of the ring station downstream from the fault. A change in ring configuration is required to bypass such a fault and to restore normal operation. Reconfiguration may be the result of automatic recovery or, if this process fails to bypass the error, it may require manual intervention.

When a ring station detects a hard error, it starts transmitting Beacon MAC frames at a specified time interval until its input signal is restored or until it removes itself from the ring. The detecting station also starts a Beacon timer. All other stations enter Beacon Repeat mode when they receive a Beacon MAC frame.

A beacon frame identifies the address of the nearest active upstream neighbor of the beaconing station as well as error type information. When the beaconing station′s NAUN has copied a number of these beacon frames, the NAUN will go offline and perform microcode and lobe tests. If the tests are successful, the station reattaches to the ring immediately. If the tests fail, the station stays offline.

When the beacon timer expires in the detecting (beaconing) station, and normal traffic has not been restored, the station assumes that its NAUN went offline, found no errors and came back online. It will now go through the same process as its NAUN. If the tests fail, the beaconing station remains detached. If successful, the station reattaches immediately. In the latter case, normal traffic may not have been restored during automatic recovery. Network management will be informed and manual intervention will be required. While reporting a permanent hard error, a set of adapter addresses is provided to identify the faulty part of the ring as a small fault domain.

3.2.1.11 Soft-Error Detection and Reporting

Intermittent faults that temporarily disrupt normal operation of the ring are called soft errors. They are usually tolerated by error recovery procedures but they may impair normal ring operation if excessive or non-random. The most critical soft errors are monitored in each ring station by a set of counters. Every two seconds the values of the soft error counters are sent as a Soft Error Report MAC frame to the Ring Error Monitor functional address (typically residing in a bridge or LAN Manager station), where the values for each counter are accumulated. If a soft-error counter exceeds a predefined threshold, a LAN Manager will be informed through its link with the LAN reporting mechanism. The LAN Manager may reconfigure the ring to bypass a faulty node, if the fault can be located.

Soft errors are said to be *isolating* if a fault domain can be specified. If not, they are called *non-isolating* soft errors.

3.2.1.12 Access Priority

The following discussion on access priority applies both to 4 Mbps and 16 Mbps token-ring LANs and is an integral part of the token-passing ring protocol. This access priority architecture is not applicable to the FDDI protocol where access priority is based upon timers rather than the contents of an access control field.

As stated earlier, access priority in a token or frame is indicated by the first three bits (PPP) of the access control field (AC). Any reservation of a priority level is indicated in the last three bits (RRR) of the AC field by a station requiring higher transmission priority.

A ring station wishing to transmit a frame at a given priority can use any available token with a priority level equal to or less than the priority of the frame to be transmitted. If such a matching token is not available, the ring station may reserve a token of the required priority in a passing token or frame according to the following rules:

- If the passing token or frame already contains a priority reservation higher than the desired one, the ring station must leave the RRR bits unchanged.
- If the RRR bits have not yet been set (RRR = B′000′), or indicate a lower priority than the desired one, the ring station will set the reservation bits to its required priority.

Upon removal of a frame by its originating station, the reservation bits in the header are checked. If they show a non-zero value, the station must release a non-zero priority token. The actual priority of the new token is based on the priority used by the ring station for the recently transmitted frame, the reservation received upon return of the frame and any stored priority.

A ring station originating a token of higher priority enters priority-hold state, (also called a stacking station in the IEEE 802.5 token-passing ring standards).

Table 11 lists the priority definitions as provided by the IBM Token-Ring Network architecture.

This protocol option however, impacts the priority handling mechanism, since a new token may be transmitted by the originating station before it is able to verify the access control field in its returned frame.

If the frame header was already received, the token will be issued according to the priority and reservation requested in the AC field of the frame and the resulting priority levels stored in the station.

If the frame header has not yet been completely received by the originating station, the token will be released with the same priority and reserved priority as the transmitted frame.

Table 11. Token-Passing Ring Protocol - Priority Allocation Table

B′000′	Normal user priority-MAC frames that need no token Response type MAC frames
B′001′	Normal user priority
B′010′	Normal user priority
B′011′	Normal user priority - MAC frames that need token
B′100′	Bridge
B′101′	Reserved for IBM
B′110′	Reserved for IBM
B′111′	Specialized station management

To prevent a high-priority station from monopolizing the LAN medium and to make sure the priority eventually can come down again, the protocol provides fairness within each priority level.

3.2.1.13 Additional Token-Ring Considerations

Using an average frame size of 1000 bits to simulate the performance of a 4 Mbps token-passing ring with 100 active LAN devices results in a maximum throughput of about 3.6 Mbps. The token-passing protocol appears to be particularly stable and most efficient even under high load conditions.

The impact of increased transmission speeds, increased numbers of attached stations, or increased transmission distances on a token-passing LAN is significantly less than similar changes on a CSMA/CD LAN. Because each station regenerates the signal, increased distances are easier to support, while transmission speed is primarily limited by the choice of media. The use of bridges to provide additional device capacity and/or distance is an attractive growth option because the absence of collisions simplifies the processing requirements of bridges and maintains the deterministic characteristics of the protocols.

In a token-passing ring, fairness in the access protocol and high priority utilization by the bridge helps avoid frame loss. Even when a frame is rejected due to bridge congestion, successful recovery is simplified by the protocol.

3.2.2 Token-Ring Summary

The token-passing protocol provides for efficient use of the media under both light and heavy traffic loads. It guarantees fair access to all participating stations. This fairness is enhanced by an eight-level priority mechanism, based on priority reservations made in a passing token or frame. A key benefit of the token-passing ring protocol is its ability to handle increased traffic loads or peaks, making it an ideal protocol for larger and/or more heavily used LANs (including backbone rings). This also makes it a good base LAN for connection to even higher bandwidth LANs such as FDDI.

3.3 Token Bus/802.4

The IEEE 802.4 standard describes the token-passing bus access protocol and its physical layer specifications. In this type of LAN, stations on the network are physically connected to a bus, but access to the bus is controlled as if it were a logical ring. Each station keeps track of the addresses of its logical neighbors (that is, those that immediately precede and follow it in a logical sequence). The physical connection sequence on the bus is independent from the order of the logical ring.

A token is used to control access to the bus. Once a station has control of a token, it has complete control of the bus for a defined period of time (the token holding time). In addition to transmitting one or more frames, the controlling station can poll other stations and receive responses or acknowledgments during this time period. When it wishes to relinquish control of the bus or when the token holding timer expires, the station passes the token on to the successor station.

The IEEE 802.4 standard defines broadband transmission on shielded coax cable at data rates of 1, 5 or 10 Mbps. It provides a choice of different modulation techniques and cable types. There are three standard topologies, each involving different modulation techniques and cable types used for the trunk and drops. The three topologies have the following characteristics:

- Omni-directional bus at 1 Mbps, using Manchester encoding
- Omni-directional bus at 5 or 10 Mbps
- Directional bus with active head-end repeater at 1, 5 or 10 Mbps

In 1984, IBM issued a statement of direction stating IBM′s intention to implement the Manufacturing Automation Protocol (MAP) standard based upon approval of the MAP Version 3 specifications. IBM also refers to MAP as the industrial LAN, which is part of the IBM Computer-Integrated Manufacturing (CIM) product offerings. The ultimate goal of CIM is integration of all processes involved in a manufacturing environment. Integration is based on controlling the information flow for which MAP, as a full seven-layer communications architecture, will provide networking and applications support. The ISO 8802-4 standard provides the two lowest layers of this seven-layer architecture.

3.3.1.1 Manufacturing Automation Protocol

Work on a MAP standard started in 1980 as a task force within the General Motors Corporation. Gradually other manufacturers, including IBM, began to participate in the definition of the protocols and the implementation of products.

The following milestones have occurred in the development of the MAP standard:

- MAP 1.0 was released in April 1984 as the result of the General Motors task force.
- Based upon the participation of additional companies in the standardization process, MAP 2.1 was issued in March 1985 as a set of interim specifications to allow early product development and experience.
- In December 1985, MAP 2.1 was updated to correct some mistakes in the earlier release.
- MAP 2.2 was released in March 1987, with the addition of carrierband network specifications.
- The ultimate standard, referred to as MAP 3.0, was released as a draft in April 1987. Final release occurred in 1988.

MAP 2.1 specifications have been subject to the following two main criticisms:

1. When multichannel capability was not required, the full broadband implementation was excessively expensive and not required. Therefore some potential users and vendors preferred a carrierband LAN implementation.

2. The processing of a full seven-layer implementation of the architecture may require too much processing time for smaller systems in time-critical applications. Initially, this critique was addressed by two different approaches, MAP/Enhanced Performance Architecture (MAP/EPA) and MiniMAP .

MAP Version 3 enhances MAP 2.1 by:

- Imbedding the MAP/EPA and MiniMAP approaches into the standard to meet the time-dependent requirements of the process industry
- Replacing MMFS by the Manufacturing Messaging Specifications (MMS)
- Including OSI presentation layer 6 support
- Extending layer 4 class 4 service dealing with error detection and recovery

With respect to network interconnection, the MAP approach is to provide a full, seven-layer MAP 3.0 backbone configuration at the plant floor interconnected with other carrierband MAP LANs and other standard LANs such as as IEEE 802.3 and IEEE 802.5. This approach addresses the objective to integrate technical design, office and business applications, computer-aided production management and manufacturing with the communications requirements of the individual production cells.

An important concern which appeared during the evolution of the Manufacturing Automation Protocol is associated with migration from MAP 2.1 local area networks to a MAP 3.0 environment. This may be difficult because of the nature of the changes.

3.3.2 Token Bus Concepts

The token-passing bus medium access protocol is differentiated from the token-passing ring protocol mainly by topology requirements and by a timer-based token management approach. In a bus topology there is no physical sequence between stations. To be able to pass the token from one station to its successor station in a logical ring fashion, a specific medium access function must be provided to initiate and maintain the logical ring sequence allowing non-disruptive station insertion or removal.

Instead of relying on a token monitor function as in the token-passing ring access protocol, token management is distributed among all stations involved in token-passing based on specific timer values and a continuous measurement of the traffic load on the bus.

Some stations which do not require the ability to transmit frames are said to be non-token holding stations. They can only receive frames and are bypassed by the access protocol at token-passing time. These stations are only allowed to respond to requests for acknowledgment.

As in any token-passing protocol, a station is only allowed to transmit data while it holds the token. Possession of a token can therefore be described as the right to transmit.

In a token-passing bus, the appearance of a logical sequence between stations is provided by a concept of predecessor and successor stations. Each station on the medium, capable of participating in token holding, establishes the identity of its logical predecessor and logical successor station (independent of physical arrangement on the medium) on the basis of the descending numerical order of MAC addresses of each station. Thus the logical sequence of the stations in Figure 24 would be A - C - D - B.

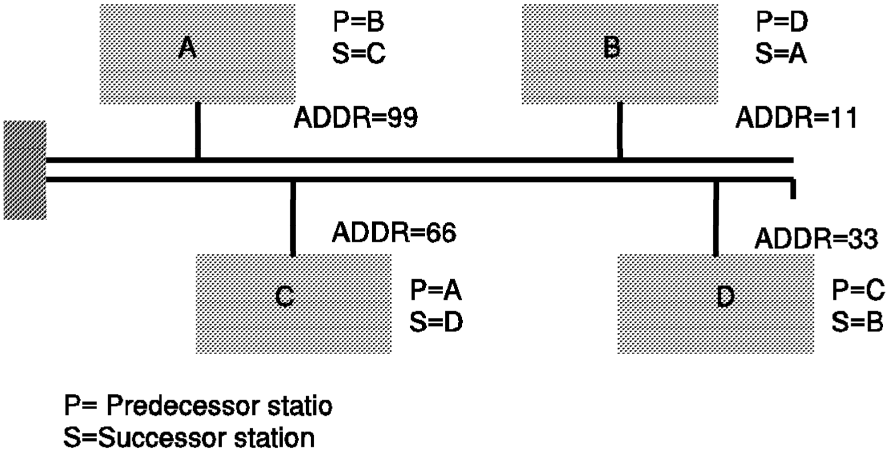


Figure 24. Token Bus Logical Sequence of Stations

The token is normally passed from one station to its successor station using a short token pass frame. If a station fails to pick up the token, the sending station uses a sequence of increasingly comprehensive recovery procedures to find a successor station.

3.3.2.1 The Token-Passing Bus Protocol

The major token-passing bus MAC functions cover the following:

- **Ring initialization**, performed when the network is first powered up and after a catastrophic error.
- **Station addition**, optionally performed when a station holding the token accepts the insertion of a new successor station, (that is, a new station with an address that is between that of the station holding the token and its current successor station.
- **Station removal**, achieved by sending a new successor identification to its predecessor, or by just disconnecting from the LAN. n the latter case, recovery mechanisms will establish the proper new logical ring configuration.
- **Recovery and Management,** including recovery from the following types of failures:
  - Bus idle (lack of activity on the bus)
  - Token-passing failure (lack of valid frame transmission)
  - Duplicate token (detected by the token-holding station)
  - Detection of a station with a faulty receiver
  - Duplicate MAC addresses

When the network is first powered up, or after a catastrophic error, the logical ring needs to be initialized. This process is triggered by the bus idle timer expiring in a LAN station. The detecting station sends a Claim Token MAC control frame.

As described above, each participating station knows the address of its predecessor (the station that transmitted the token to it and its successor) and station to which the token should be sent next.

After a station has completed transmitting data frames and has completed other logical ring maintenance functions, the station passes the token to its successor by sending it a token MAC control frame. Any failure in reaching a successor station triggers a staged recovery mechanism, using other MAC control frames (set solicitor 1 and set solicitor 2). If the token holder does not receive a valid token after sending the token the first time, it repeats the token pass operation. If the successor doesn′t transmit after a second token frame, the sender assumes that its successor has failed and sends a Who Follows MAC control frame containing its successor′s address in the data field. The station detecting a match between its predecessor and the address in the Who Follows frame data field will respond by sending its address in a Set Successor MAC control frame. In this way, the token holding station establishes a new successor excluding the failing station from the logical ring.

Stations are added to the logical ring through a controlled contention process using response windows, a specific interval of time common to all stations, and based on numerical address comparisons. The actual procedure is referred to as the Solicit Successor procedure.

This procedure raises a concern with respect to excessive delay experienced by a station before gaining access to the LAN when many stations attempt to insert into the logical ring and perform the solicit successor procedure. The time a station has to wait between successive passes of the token is called the token rotation time (TRT). To avoid an excessive TRT, every station measures the rotation time of the token. If the time exceeds a predefined value set by station management, the station will defer initiation of the solicit successor procedure and verify, at the next appearance of the token, whether it is now rotating fast enough to perform the procedure.

A station can remove itself from the logical ring simply by not responding anymore to the token passed to it. Ring station sequence recovery procedures will adjust the successor and predecessor information in the predecessor and successor stations respectively. A more efficient way of leaving the logical ring is to have the exiting station send a Set Successor MAC control frame to its predecessor, containing the address of its successor.

Access Priority: The token-passing bus protocol provides an optional 8-level priority mechanism by which higher layer data frames, LLC sublayer or higher level protocols, are assigned to eight different service classes according to their desired transmission priority. Service classes range from 0 (low) to 7 (high).

At the access method level, there are four access classes, corresponding to four request queues to store pending priority frames. Access classes are 0 (lowest), 2, 4 and 6 (highest priority).

Token bus MAC maps the service class requested by the LLC sublayer into a MAC access class (service classes 0 and 1 into access class 0, service classes 2 and 3 into access class 2, etc.).

The purpose of this priority mechanism is to allocate bandwidth to the higher priority frames and to send lower priority frames only when there is sufficient bandwidth left. To prevent any station from monopolizing the medium with access class 6 frames, a station can only send class six frames for a maximum time defined by station management. When this time expires, a station must release the token even if it has additional class six frames ready to transmit.

Similarly, each access class is assigned a target token rotation time (TTRT). For each access class a station measures the time it takes the token to circulate the logical ring. If the token returns to a station in less than the target token rotation time, the station can transmit more frames of that access class until the expiration of the TTRT. If the token takes more than the TTRT for a specific access class, no frames of this priority class can be sent at this pass of the token. The actual algorithm consists of loading the residual value of a token rotation timer into a token holding timer. If an access class′s queue is empty or if the token holding timer expires, the station begins to serve the next lower access class. When the lowest priority class is serviced, the station performs any required logical ring maintenance operation and releases the token for its successor station.

3.3.3 Token Bus Summary

The token-passing bus access method is predominately used for industrial LANs. It uses a logical ring on a physical bus, and has a number of transmission techniques and data rates to choose from.

The protocol optionally provides eight priority service classes to higher level data frames, handled at the MAC level by four access classes based upon the current traffic load on the bus.

With respect to a full seven-layer architecture for the manufacturing environment, application implementation is very limited because of the immaturity of the standards, but will likely increase as the standards are finalized.

3.4 FDDI

The American National Standards Institute (ANSI) standards working group X3T9.5 has formulated a draft proposal for an international standard called Fiber Distributed Data Interface (FDDI). The ANSI X3T9.5 documents are listed below:

- FDDI Single-Mode Fiber Physical Layer Medium Dependent (SMF-PMD)
- FDDI Token-Ring Media Access Control (MAC) X3.139-1987
- FDDI Station Management (SMT) X3T9.7/88 (Rev 7.2)
- FDDI Hybrid Ring Control X3T9.5/89-43 (Rev 6.1)
- FDDI Media Access Control (MAC-2) X3T9.5/88-139 (Rev 4.0)

A detailed description of the FDDI LAN will be given in the following. In terms of current standards, FDDI is described by ISO 9314, which is subdivided into 6 parts. These are:

- 9314-1 - the physical layer.
- 9314-2 - the MAC layer.
- 9314-3 - the physical medium dependent layer.
- 9314-4 - use of FDDI on single-mode fiber.
- 9314-5 - Hybrid Ring Control - FDDI II, which divides the FDDI frame into slots, such that frames may contain both synchronous and asynchronous data. This implementation is more suitable for voice traffic.
- Station Management has as yet not been allocated a 9314 number.

In 1982 the American National Standards Institute (ANSI) created the X3T9.5 committee, which began studies on high-speed communications. Originally envisioned as a standard for high-speed host channels, FDDI rapidly became viewed as a new generation of LANs which use optical fiber to provide a high-speed communication network.

Today, it is possible to implement standardized FDDI LANs based on the physical backgrounds and the logical links which are defined in the ISO 9314 and the ANSI X3T9.5 standards. The ANSI X3T9.5 and ISO 9314 committees describe FDDI as a dual counter-rotating ring which operates at a rate of 100 Mbps. In many ways, FDDI is similar to the IEEE 802.5 token-ring, although there are some differences.

3.4.1 FDDI Concepts

FDDI uses a token-passing protocol in which each station has the chance to transmit data when a token passes. A station can decide how many frames it will transmit using an algorithm which permits *bandwidth* allocation. FDDI also allows a station to transmit many frames without releasing the token.

An FDDI network consists of a set of stations/devices connected to each other as a serial string of stations/devices by a transmission medium to form a physically closed loop. Information is transmitted sequentially, as a stream of suitably encoded symbols, from one active station/device to the next active one. Each station/device regenerates and repeats each symbol. The method of actual physical attachment to the FDDI network may vary and is dependent on specific application requirements.

FDDI uses two rings:

- The *primary ring*, which is similar to the main ring path in token-ring terminology
- The *secondary ring*, which is similar to the backup ring path of a token-ring

Note each ring consists of a single fiber path which is equivalent to a pair of copper conductors.

FDDI permits many attachment units (stations, concentrators, and bridges) to be attached in various ways.

From a wiring point of view, FDDI is similar to a fiber optic token-ring network; however, there are the following differences between the token-ring and FDDI techniques:

- A device can be attached directly to the ring without requiring a concentrator such as the Multistation Access Unit (MAU) on a token-ring.
- A device can be attached to either or both of the primary/secondary rings.

To differentiate between devices that attach to one ring or both rings, FDDI defines two classes of devices:

- A *Class A* device attaches to both of the rings directly. It may be a station and it is called a Class A station or a Dual Attachment Station (DAS). It can be a concentrator and it is called a Dual Attachment Concentrator (DAC).
- A *Class B* device attaches to only one of the rings directly or through a concentrator. It may be a station and it is called a Class B station or a Single Attachment Station (SAS). It can be a concentrator and it is called a Single Attachment Concentrator (SAC).

Concentrators are active devices that act as *wiring hubs* and are similar to an active token-ring access unit (such as an IBM 8230 Controlled Access Unit).

During normal ring operation, the primary ring is active while the secondary ring is idle.

In the wake of a failure on the primary ring, the secondary ring will become active when a class A station or a Dual Attachment Concentrator wraps the primary ring to the secondary ring, establishing a single ring. This functionality is mandatory to maintain the reliability of the LAN.

3.4.1.1 FDDI over Copper (SDDI and CDDI)

An alternative to FDDI is a Shielded twisted pair Distributed Data Interface (SDDI). This proposal is for transmitting FDDI directly on copper wires without converting the electrical pulse stream to optical signals. The data stream remains at the rate of 100 Mbps.

This type of solution is envisioned to provide a copper-based FDDI solution which could be implemented on the existing cabling system and will cost approximately less than 50% of the equivalent fiber solution.

The FDDI signal is put on the copper wire as a baseband signal exactly as it would have been transmitted on a fiber. There is no special coding, modulation or modification of the signal. A one bit is signaled as a voltage of between .35 and .7 volts and a zero bit is the absence of voltage. Transmission distance between workstation and hub (or between two directly connected workstations) is limited to 100 meters. The IBM 8240 FDDI Wiring Concentrator, the IBM 8250 Intelligent Hub and the PS/2 FDDI Workstation Adapter are examples of available products built to this SDDI specification.

There is also a proposal to the ANSI FDDI TP-PMD workgroup for a copper solution running FDDI over UTP category 5 (CDDI). In 1993 this was still a proposed option. The proposal involved changing the data link encoding scheme and thus required more extensive modification to the chip sets than for the STP proposal. CDDI, as it′s commonly called, is very important because of the large amount of installed UTP cable.

3.4.1.2 Port Types

The port at the end of a physical connection determines the characteristics of that connection. These characteristics include whether the connection will be accepted or rejected.

The standard specifies four port types for FDDI ports:

- A-Type

  For Dual Attachment Stations: Primary Ring-In, Secondary Ring-Out

- B-Type

  For Dual Attachment Stations: Secondary Ring-In, Primary Ring-Out

- M-Type

  On a Concentrator, to attach a Single Attachment Station

- S-Type

  On a Single Attachment Station, to attach to a Concentrator

The connection rules for the different port types are the following:

- A-to-B and B-to-A are peer-to-peer trunk connections.
- M-to-S is a master-to-slave connection.
- M-to-A and B provide dual homing.
- S-to-S is a point-to-point connection.

In a dual ring configuration, one end of the link is configured as an A-Type and the other end as a B-Type port. When configured as a tree, one end of the link is an M (Master) port and the other end is an S (Slave) port.

Only Dual Attachment Stations (DAS) reside on both rings. Single Attachment Stations (SAS) reside on the FDDI tree and connect to the network via concentrators. In this case, the concentrator can be a DAC or a SAC.

Table 12 on page 74 shows the connection rules for Single and Dual Attachment Stations.

Table 12. Connection Rules for SAS and DAS

Port Type	Port Type	Rule
A	A	Undesirable peer connection that creates twisted primary and secondary rings.
A	B	Normal trunk ring peer connection.
A	M	Tree connection with possible redundancy. Port B shall have precedence for connecting to port M in a single MAC node.
A	S	Undesirable peer connection that creates a wrapped ring.
B	B	Undesirable peer connection that creates twisted primary and secondary rings.
B	M	Tree connection with possible redundancy. Port B shall have precedence for connecting to port M in a single MAC node.
B	S	Undesirable peer connection that creates a wrapped ring.
M	M	Invalid configuration.
M	S	Normal tree connection.
S	S	Connection that creates a single ring of two slave stations.

3.4.2 FDDI Ring Topologies

The FDDI network topology may be viewed at two distinct levels:

- The physical level
- The logical level

Physical topology describes the arrangement and interconnection of nodes with physical connections. The logical topology describes the paths through the network between MAC entities. An FDDI network forms one of the two following physical topologies:

- A dual ring of trees
- A subset of dual ring of trees

The implication on the logical topology is that at the most, two logical sets of MAC entities (for example, two independent token/data paths) exist in a single fault free network. A set of MAC attachments can be called a logical ring. The two types of topologies may not necessarily be similar.

3.4.2.1 Dual Attachment Station Ring

Figure 25 on page 75 shows an FDDI dual ring configuration consisting of Dual Attachment Stations (DAS). Each station will have both ports (A and B) attached to the rings. The cabling between the stations has to be all fiber or shielded twisted pair (STP).

The FDDI network consists of the primary ring on which the data flows from port B on one station to port A on next station, and the secondary ring in which the data flows in the opposite direction of the primary ring. The secondary ring provides the backup path for failure conditions. In normal conditions, there is no data flow on the secondary ring.

If any station fails, the ports on the adjacent stations will wrap the primary and secondary rings and the network will continue to operate as a single ring. For example, in Figure 25 on page 75, if station 1 fails, station 4 will wrap its B port and station 2 will wrap its A port, resulting in a single ring which connects the remaining stations (2, 3 and 4).

If there is a cabling failure, the stations at either end of the broken cable will wrap their corresponding ports to restore the operation of the ring. For example, in Figure 25 if the main ring between station 1 and station 4 is broken, the A port on station 1 and the B port on station 4 will wrap and all four stations will continue to operate on the same FDDI ring. However, if there was a second break on the ring (for example between stations 2 and 3) then the ring would be fragmented, forming two rings, one with stations 1 and 2, and another with stations 3 and 4.

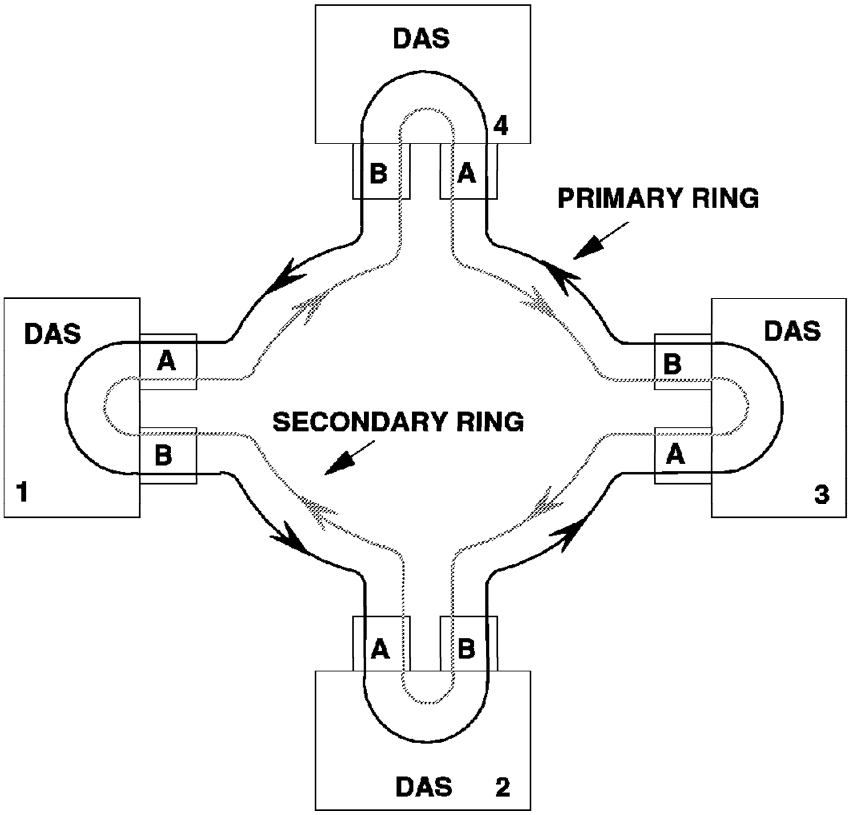


Figure 25. FDDI Dual Attachment Station Ring

3.4.2.2 Dual Attachment Concentrator Ring

Figure 26 on page 76 shows an FDDI Dual Attachment Concentrator ring.

In this configuration, the signal will enter the concentrator at the A port and after flowing through the M ports, it will exit the concentrator at the B port.

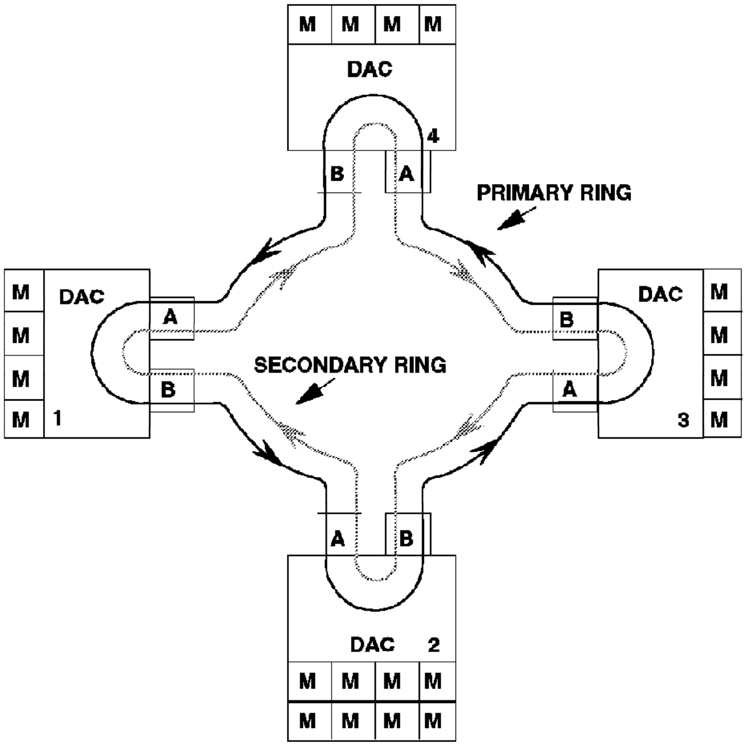


Figure 26. FDDI Dual Attachment Concentrator Ring

3.4.2.3 Dual Attachment Concentrators and Workstations

Figure 27 on page 77 shows the topology of workstations attached to a Dual Attachment Concentrator (DAC). Each workstation is attached as a Single Attachment Station (SAS) to the master port (M) of a concentrator (DAC). The M port attaches each SAS to the primary ring.

Data enters each DAC from the A port and, after passing through each M port (and its attached workstation), exits through the B port.

The benefit of a concentrator attachment is that it allows an SAS to enter and leave the ring without the risk of disrupting the ring.

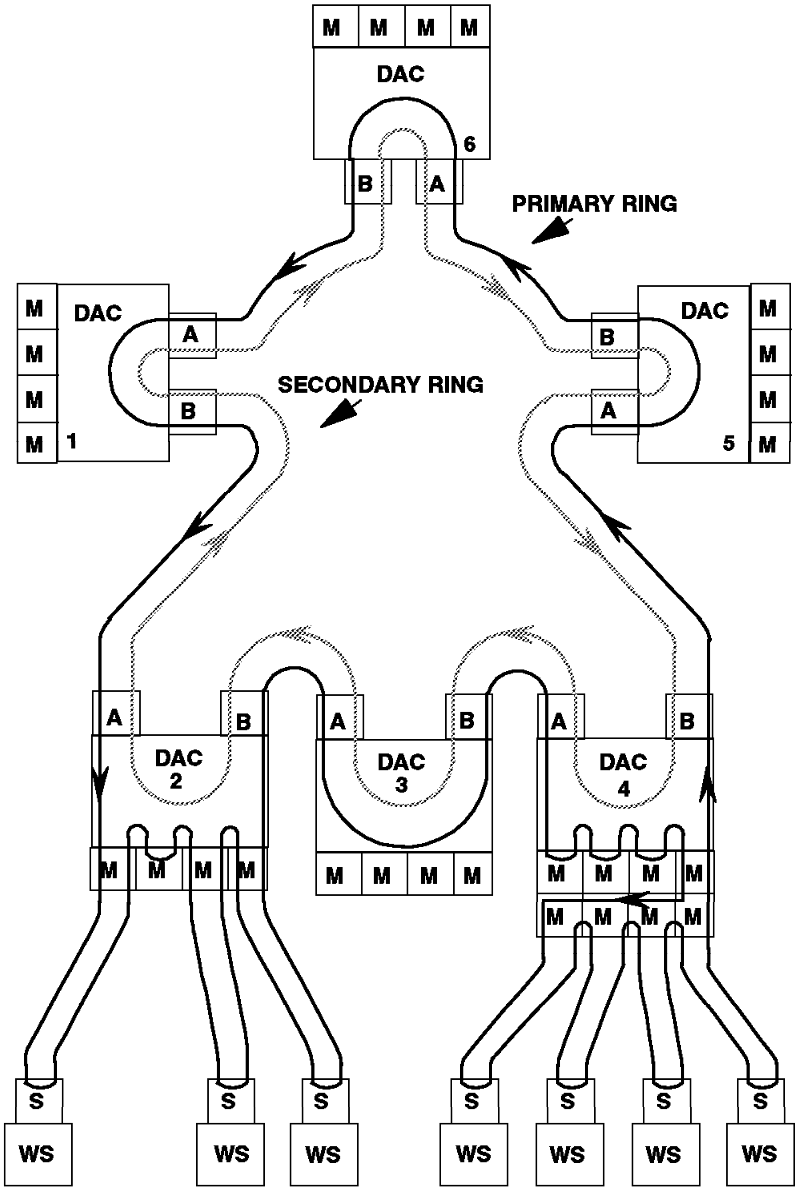


Figure 27. Dual Access Concentrator and Workstations

3.4.2.4 Tree Topology

Figure 28 on page 78 shows an FDDI tree topology.

Note: The purpose of the diagram pictured here is to show the signal flow through a complex tree topology. It is not intended to suggest a typical installation.

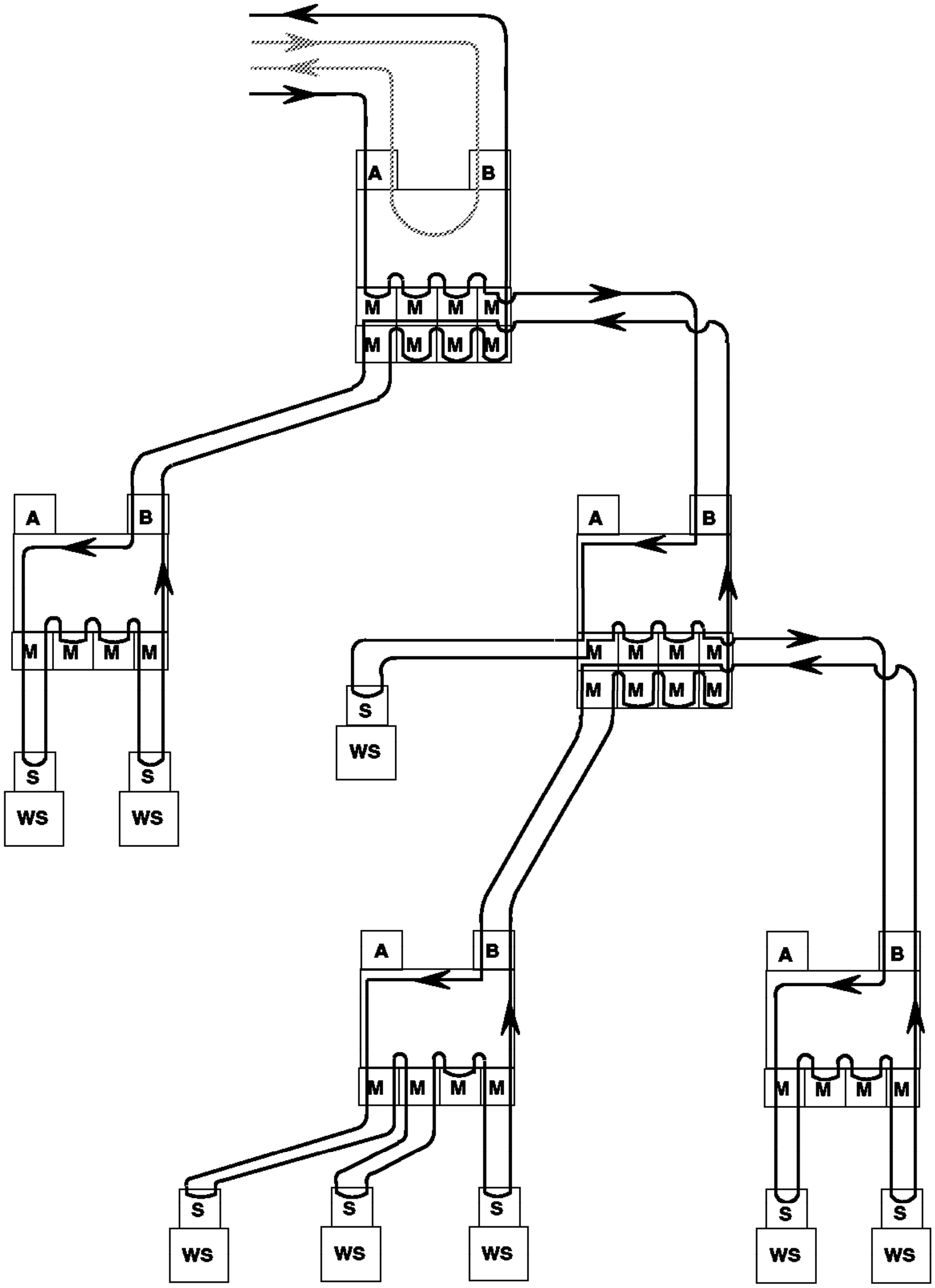


Figure 28. FDDI Tree Topology

3.4.2.5 Dual Homing

Figure 29 shows an FDDI *dual homing* topology. A concentrator, that is not part of the main ring, may be dual attached via one or two other concentrators to provide greater availability. When connected in this manner, a concentrator is described as a Dual Homing Concentrator (DHC).

Similarly a Dual Attachment Station can be connected to one or two concentrators using both A and B ports to provide high availability. The station connected in this manner is considered a Dual Homing Station (DHS).

In both of these cases, only port B is active, and the connection to port A remains in standby mode. Should the connection to port B fail, port A would become active without any impact on the users of the Dual Homed Station or concentrator.

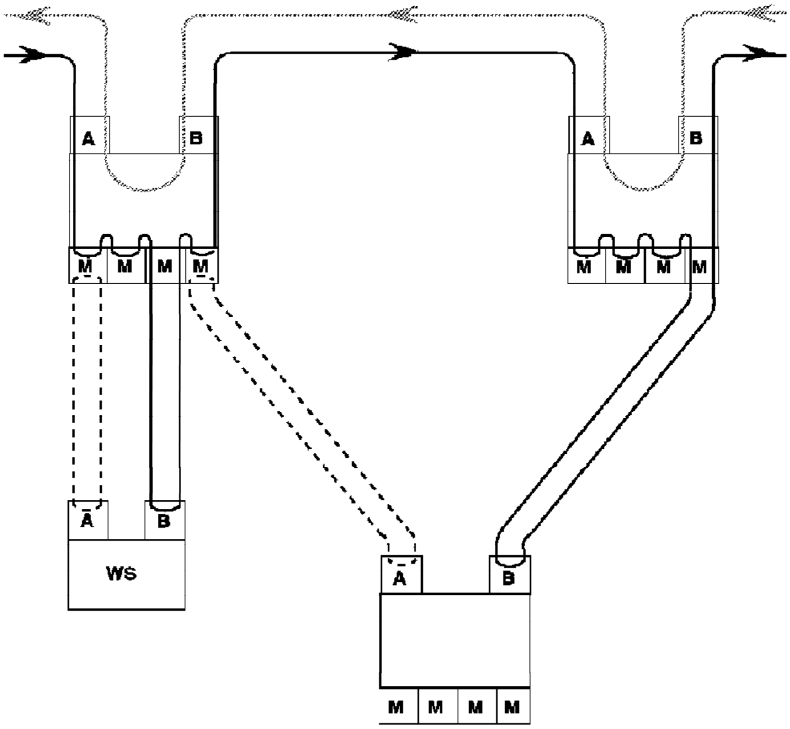


Figure 29. FDDI Dual Homing Topology

3.5 ATM

The telecommunication industry began in about 1988 to develop a concept known as *B-ISDN* or Broadband Integrated Services Digital Network. The aim was to get an integrated way to transfer information (voice and data) at higher speed. This idea was used by the industry and the technology is called *Asynchronous Transfer Mode (ATM)*.

IBM submitted three proposals aimed at speeding the acceptance of ATM network technology on the desktop by addressing customer concerns about interoperability, cost and network management issues.

The ATM Forum is the organization responsible for driving the development of standards-based ATM specifications for both local and wide area networks. In its proposals, IBM emphasized the need for pragmatism, both in cost and interoperability with existing types of networks.

IBM′s proposals have addressed:

- Specifications on how an ATM network can support local area network migration and interconnection services
- Development of a fast, cost-effective, mid-range user interface (MUNI) that provides a migration path for currently installed LANs, using existing voice-grade, unshielded twisted pair, category 3 cable
- Creation of a network management subgroup of the ATM Forum technical committee to address the requirements and technology for network and systems management

Local Area Network Emulation and Interconnection Services: One way ATM is delivered to the desktop is through emulation, so that a PC user can gain access to the powerful ATM network while continuing to participate in existing Ethernet and token-ring LANs. A major benefit of emulation is that it will require no change to existing LAN applications.

In order to provide a simple and easy means for porting or transferring existing local area network applications to the ATM environment, IBM proposed an ATM service that emulates services of existing LANs across an ATM network. The emulation enables workstations and servers on a LAN to connect to an ATM network, while software applications interact as if they were attached to a traditional LAN. Emulation also allows interconnection of ATM networks with traditional LANs using today′s bridging methods.

The LAN emulation and interconnection services, as proposed by IBM, is supported via a software layer in end systems (workstations or servers).

To emulate a LAN-like service, different types of emulation can be defined, ranging from the MAC service (such as IEEE 802.x or FDDI LANs) up to the services of network and transport layers (such as TCP/IP, OSI and APPN). IBM proposed concentrating first on emulation of the MAC service, which provides support for the majority of existing LAN applications.

Cost-effective ATM Interface with Flexible Cabling Choices: IBM and Chipcom Corporation jointly proposed that the ATM Forum placed high priority on standard, low-cost interfaces that provide a fast migration path for ATM to the desktop. This included the adapters, the hub ports and the wiring structure. The proposal resulted from research showing that corporations place more importance on cost, interoperability and network management as criteria in evaluating ATM, than on sheer data rates and increased bandwidth.

IBM and Chipcom proposed the ATM technical committee to undertake a project to produce specifications for a mid-range user interface (MUNI) at a data rate of 25 megabits per second. This rate is consistent with proven, available, cost effective technology for operation over a variety of network environments including fiber optic, shielded twisted pair and unshielded twisted pair, category 3 (UTP-3).

With the same wiring users get the flexibility to mix and match ATM and non-ATM networking products so that they can evolve to ATM as applications and networking requirements dictate.

Trends: In local area network evolution you can see five major trends:

- Client/server environment pushes bandwidth requirement for server access beyond the limit of current LAN technologies.
- Applications tend to require more and more bandwidth. The access of the LAN-shared media induces delays, and a way to overcome this is to reduce the number of users that contend for the bandwidth. This leads to fewer workstations per LAN segment.
- Having fewer workstations per LAN segment subsequently increases the number of LAN segments that belong to a specific subnet.
- You then have to isolate high-capacity servers or high-performance workstations on dedicated segments to avoid contention for bandwidth access from regular end users.
- You will possibly end up with an increasing complexity of inter-networking between independent segments and subnets. It will give you intricate topologies with federating or collapsed backbone networks, and directly affect in cost of ownership and network reliability.

A consequence of these trends is the appearance of various LAN-switched products that help address the clustering of bandwidth demand.

A number of various LAN switch products have appeared to help you solve the demand of bandwidth. The industry has accepted that only ATM can give you an unlimited segmentation, based on logical addressing (virtual circuit) rather than physical attachment. With ATM, each end point can be the termination of multiple segments, each associated to a ″virtual circuit″. You get the ultimate segmentation technique that resolves the increased complexity of the total campus interconnection.

ATM helps your campus interconnections be both manageable and affordable. ATM is the economical answer to higher capacities and functionality by minimizing the number of boxes.

The new multimedia applications, which offer both image manipulation and voice, add a new dimension to the networking architecture. This is not addressed by current LAN architectures. Multimedia applications using complex objects will not only require a high bandwidth, but the transportation of time-dependent information (audio and video signals) must also be supported. Multimedia applications have introduced a new requirement: isochronous communications. ATM has the characteristics (guaranteed delay, bounded jitter rate) you need to enable multimedia applications over your network. With ATM to each station you get a scalable bandwidth adapted to your application requirement and processor′s performance.

The structure of the ATM cell, which is constant whatever the network (LAN or WAN), simplifies your total connectivity with a seamless architecture.

You may have been told that ATM should be first used to build high-speed backbones or attach high-capacity devices (servers or scientific workstation).

But limiting ATM to backbone implementation ignores your need for high-speed, dedicated and isochronous bandwidth to the desktop to meet your multimedia application requirements.

It is IBM′s view that networking must enable a new world of applications to the desktop to increase your personal productivity or open the door to new opportunities. In order to achieve that goal, several criteria must be met:

- ATM technology must be affordable:
  - The new family of workstation adapters should match the price of current LAN adapters.
  - The installed infrastructure (wiring and wiring-closet equipment) should be usable without changes to let you attach your current LANs or ATM devices.
  - The price per port of the connectivity equipment should be competitive with current LAN implementations and include the cost of interconnecting devices.
- ATM should be *standard*.
- ATM should be *compatible*, which means remain compatible with the current installed base of applications and permit total interoperability:
  - You must be able to access data on workstations on current LAN servers attached to ATM, or to current LANs over an ATM backbone.
  - You must be able to run your workstations with new ATM adapters and current software (developed for current LAN support), *unchanged*.
- You must be able to manage ATM using the current resources for administration and operations.

3.5.1 ATM Concepts

Asynchronous Transfer Mode is the new evolving standard that addresses the demand for high-speed, high-bandwidth networks. With new demands emerging for image, client/server and multimedia applications, a new set of standards was required for sending large amounts of varied information (data, voice and video) over high-speed networks. ATM is seen as the new technology to fulfill these new business demands and is already touted as the definitive communications technology for the next decade.

The ATM Forum which currently has over 300 members consists of customers, carriers, service providers and network equipment providers. Although multimedia applications exist today, full maturity will depend on networks capable of delivering effectively a mix of text, voice, images and video, with interactive facilities. Some examples of emerging applications that will take advantage of this technology will be in the areas of collaborative work group sessions where business groups can share voice, data and video with each other via their personal computers. Engineers and scientists can work with complex drawings and satellite images and medical specialists can review X-Ray and CAT scan information from distant locations in real time. Other examples include kiosks for retail merchandising, audio-video distribution (media on demand, news, movies etc), and manufacturing for process control. As multimedia involves the use of several communication media together, it will add stress to existing networks.

ATM is a multimedia sub-networking technology that will have the ability to provide bandwidth on demand and will be protocol independent. In the past every connection in the network required a dedicated circuit. In the 80′s, *Time Division Multiplexing* (TDM) dedicated a fixed amount of bandwidth to each active station, regardless of the traffic generated by that station. TDM-units are completely protocol independent which is achieved by associating a specific outgoing port to a predetermined destination port on the remote TDM.

Frame-Relay and X.25 offer statistical multiplexing in that through a single network interface, multiple links can be established with multiple destinations carrying many protocols. Frame-Relay uses packet transfer to allocate bandwidth more efficiently to bursty data traffic; however, it is not suited to isochronous traffic because it supports variable frame sizes. It provides improved throughput over X.25 by doing less processing on each data frame and is transparent to higher layer protocols, which allows it to transport X.25, SNA, DECnet etc.

3.5.1.1 Cell Relay and ATM

ATM, as a cell relay technology, draws on some of the benefits of fast packet switching and provides basic multiplexing and circuit switching with its main design aimed at supporting isochronous (voice, video) information communications. Cell relay is a technology designed to relay data, voice and video in fixed size cells. This technology will be used by Broadband ISDN and metropolitan networks. Cell relay therefore combines the high throughput and bandwidth optimization of frame relay with the predictability of time division multiplexing making it suitable for both bursty data traffic and isochronous voice/video traffic.

ATM is the international standard for cell relay and defines a fixed length 53-byte packet or cell which consists of a 5-byte header and 48-byte payload. The ITU-T (formerly CCITT) has selected ATM as the basis for the future broadband network in view of its flexibility and suitability for both transmission and switching.

3.5.1.2 ATM Structure

As stated previously, ATM has a cell size of 53 bytes consisting of a 5-byte header and 48 bytes of user information or payload. Various types of traffic can be mixed without concern for delay due to this small cell size. The ATM structure consists of a service layer, adaptation layer, ATM layer and physical layer. Figure 30 on page 84 shows the different layers implemented for ATM services.

ATM Adaptation Layer: The *AAL* or ATM Adaptation Layer provides an ATM service to applications of the same type of which there can be five:

1. Circuit Emulation (CBR)
2. Video/Audio (VBR)
3. Connection-oriented data
4. Connectionless data
5. Simple data

It also segments and reassembles data into 48-byte payloads and hands these across to the ATM layer.

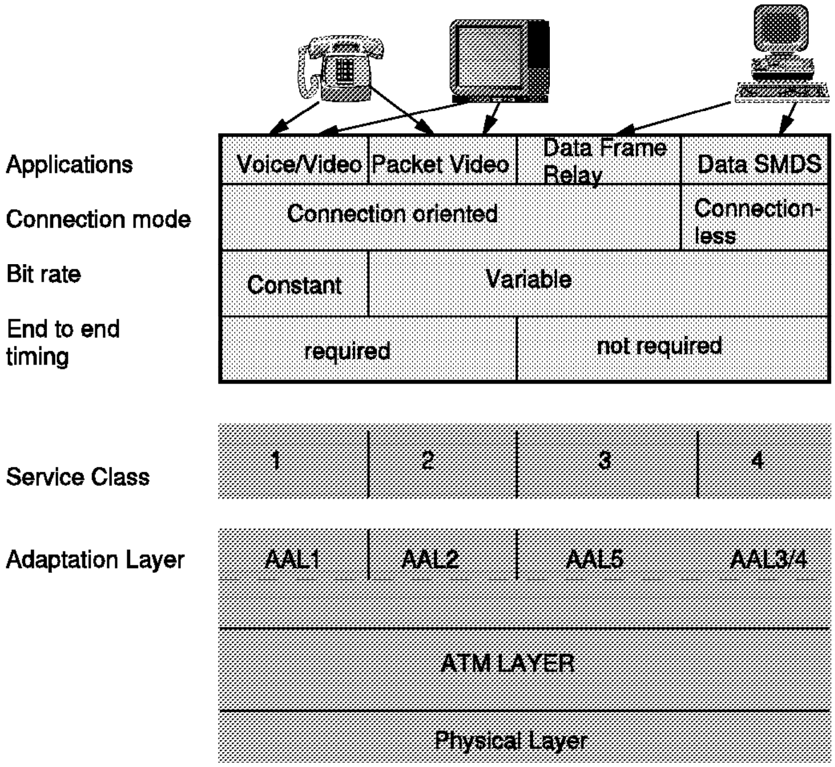


Figure 30. ATM Services

ATM Layer: The ATM layer then adds/removes the header to the payload which contains a connection identifier and multiplexes the payloads into virtual connections. The ATM header contains the virtual connection ID (VCI) and/or the virtual path ID (VPI). carries user data or network data information and if the cell has a high or low priority. Error calculation is also held on the header and not in the payload. Multiplexing is accomplished via the virtual circuits. Voice traffic can be carried on one circuit, video on another and data on another. Cells for each traffic type can be intermixed hence the term *asynchronous* in asynchronous transfer mode.

A virtual path is a group of virtual circuits and ATM intermediate nodes ignore the least significant bits of the VCI so a range of virtual circuits is carried between end systems. The end systems still deal with the full VCI. The virtual circuits are a vital element of ATM and a method is required to set up and clear these circuits. For B-ISDN standards, it is likely that a protocol based on the ISDN signalling standard, Q.931 will be defined, which includes accessing services from public carriers that are based on B-ISDN. A user′s equipment that interfaces to the public carrier network would set up and clear ATM VCs by sending appropriate messages over a virtual circuit with an agreed uponVCI value. Signalling for private ATM-based LAN networks is still being defined.

ATM Physical Layer: One principle of ATM is that it is not tied to any physical layer (that is, it is media independent) and therefore there are no new layers being proposed. One new function of this layer is to delineate the 53-byte cells within the bits of the payload. The receiver scans the incoming transmission payload bit-stream for a 5-byte block that has the correct result in the 5th byte. When found, the receiver skips 48 bytes to the next 5-byte block and checks the calculation to determine the cell boundaries.

The physical layer provides the necessary bit-time timing and physical medium support. This layer consists of a transmission convergence sub-layer and physical media dependent sub-layer for support of coax, copper, single-mode and multimode fiber. As ATM is connection oriented, two users communicate using a virtual circuit and send packets over this path. The table below details the user to network interface of the physical layer:

Table 13. User to Network Interface

Physical Layer	**UNI**	**Media**	**RATE**
Sonet-STS3C	Public/Private	Fiber	155 Mbps
DS3 (T3)	Public	Copper/Fiber	45 Mbps
100 Mbps Multimode Interface	Private	Multimode Fiber	100 Mbps
155 Mbps	Private	Multimode Fiber / Copper STP	155 Mbps

3.5.1.3 Broadband Network Services (BNS)

IBM has gained a lot of experience in high-speed multimedia networking with projects such as PARIS which commenced in 1986. This formed the basis of what today is our Broadband Network Services. BNS supports both variable packet (TCP/IP, APPN, FR) and fixed cell (ATM) networking. BNS, in simple terms, is an architecture for controlling multimedia networks at gigabit speeds. ATM is a transportation mode supported by BNS.

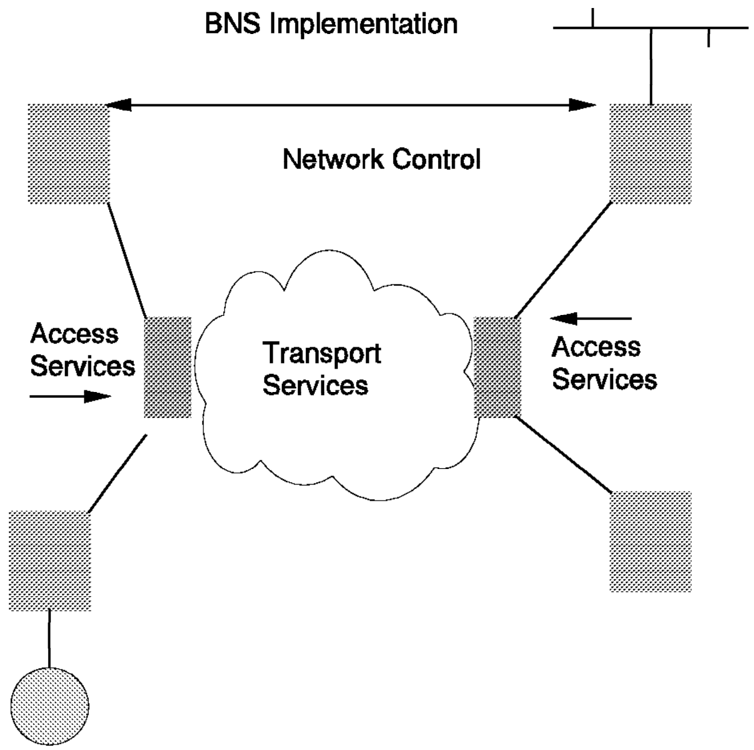


Figure 31. Broadband Network Services Implementation

BNS major functions are:

- To provide routing and network control functions
- To provide access services
- To support existing applications
- Provide efficient multicast support
- Guarantee Quality of Service requirements
- Economic bandwidth management
- Non-disruptive path switching
- Ability to support more connections across a finite bandwidth

Broadband Network Services as shown in Figure 31 is a comprehensive approach for meeting the challenges of high-speed, multimedia networking technologies. This architecture describes a set of Control Point Services, Access Services, and Transport Services. Control Point Services include bandwidth management, route computation, non-disruptive path switching, and multicast group management. Access Services provide the framework for the high-speed network to support a variety of standard open interfaces including ATM, frame relay, circuit emulation, and higher layer interfaces such as TCP/IP and APPN. Transport Services provide the reliable, bandwidth-managed pipes for user traffic.

Although standards bodies and the ATM Forum have been addressing ATM for some time, today′s standards pertain primarily to the network user′s interfaces and lower layer (physical and ATM cell) functions, but not to its internal architecture or implementation. This approach enables the community of ATM vendors to design products that interoperate yet maintain product distinction.

In order to take full advantage of ATM-based networks, today′s network architectures must evolve to tomorrow′s high-speed network architecture. Existing communications architectures were simply not designed to operate at gigabit speeds. The advent of ATM and a new set of networking requirements has led IBM to develop an evolutionary state-of-the-art high-speed networking architecture. This architecture and the products it yields will guide IBM in developing high-speed multimedia networks and applications. ATM will be prominent in our product plans and our ATM products will conform to existing ATM standards as well as providing functions not as yet addressed by the public standards process. IBM will continue to work with the standards bodies to ensure that our architecture and products comply with new standards as they emerge.

IBM is committed to licensing Broadband Network Services to interested vendors, who are invited to contact IBM for more information. A technical overview document is available that describes the principal functions and algorithms in the architecture and that outlines expected customer benefits.

IBM will continue to be active in ANSI, ETSI, ITU-T, and the ATM Forum, relative to ATM standards and implementer′s agreements. The timing and degree to which the Broadband Network Services is submitted to these organizations will depend on the interests of all participants and on the topics given priority by working groups and subcommittees.

In short, the Broadband Network Services will be open. Detailed publications are available beginning with a technical overview. IBM will license the architecture to interested vendors and will continue its active participation in standards organizations and in the ATM Forum.

As mentioned previously, IBM′s Broadband Network Services provides a set of Control Point Services, Access Services, and Transport Services.

The Control Point Services address:

- Address Resolution - An efficient distributed algorithm to translate external user names into internal network addresses thereby obviating the need for a centralized name server.
- Reservation and Management of Bandwidth - A connection requiring Quality of Service guarantees can reserve bandwidth in an amount which depends on the characteristics of the traffic. This amount is less than the peak rate of the connection but larger than its average rate. The architecture tracks the available bandwidth and decides whether new connections can be accommodated.
- Route Computation with Load Balancing - The architecture ensures that the traffic load on the network is reasonably distributed amongst the nodes of the network to guarantee quality of service to users. The most efficient routes through the network are computed to meet quality of service and overall traffic demands.
- Non-Disruptive Rerouting - Network control messages must be broadcast efficiently and quickly throughout the network. In the case of a line or node

failure, messages are broadcast to controlling network elements to route traffic around failures without disrupting established connections.

- Rapid Distribution of Control Information - The rapid distribution of control information is also critical for early notification of potential areas of congestion so they may be avoided in setting up new connections.
- Multicast Group Management - Via the architecture′s Set Management capabilities, resources or users may be grouped to receive messages simultaneously. These specialized routing techniques enable native multicast implementations.

Broadband Network Services describes a variety of Access Services for circuit emulation, frame relay, ATM, and other interfaces enabling these services to take advantage of the architecture′s control services. In addition, a general structure for the incorporation of additional access services is provided.

Broadband Network Services describes Transport Services: Standard fixed-length ATM cells are supported as well as an ATM complementary Packet Transfer Mode (PTM) of operation. PTM takes a variable length packet approach that is optimized for data traffic, while the ATM fixed length cells are optimized for mixed data and isochronous traffic. Either of these Transport Services can make use of the Dynamic Routing for Multicast and the other Broadband Network Services Control Point Services.

3.5.1.4 ATM and Existing LAN Architectures

A number of solutions to handle congested networks have emerged. ATM is seen as the common solution because it can both handle a number of existing solutions and offer new ones.

Advantages of ATM over FDDI, Token-Ring and Ethernet: FDDI, Ethernet and token-ring technologies are based on a shared medium concept. In other words, there is contention for bandwidth and each adapter must run at the same speed (100, 16 or 10 Mbps) even though the average usable bandwidth per user is less. With ATM, the adapters run at their designated speed because the connections are switched and the bandwidth is dedicated. Key advantages of ATM are its scalability, high network throughput, low latency, and multimedia support and speed adaptation.

Advantage of ATM over Fiber Channel as a LAN: Fiber Channel uses a switch fabric and can be used as a high-speed LAN, but Fiber Channel is better suited for processor-to-control unit or processor-to-processor communications, while ATM is better suited to networking.

While there is some potential for overlap between ATM and Fiber Channel in the LAN backbone and the high-end computing environment, it is likely that both will coexist and interconnect to provide complete solutions.

Advantage of ATM over Frame Relay: Frame relay is a high-speed packet-switching solution for wide area networking which is primarily data-oriented. The key differences from a user perspective are:

- ATM is capable of reaching Gbps speeds, while frame relay is limited by current specifications to 2 Mbps.
- ATM is designed to support isochronous (delay sensitive) traffic, while frame relay is not.

ATM is an emerging technology. You should not view ATM investments as ″future proofing″. The inherent characteristics of ATM are:

- Scalable for present and future network bandwidth requirements
- Bandwidth on demand
- Simplified network configuration, management and control
- Aggregation of voice, data, and video traffic into one transmission scheme

These reasons justify ATM on the basis of long term cost savings rather than cost expenditure. The available ATM adapters and LAN emulation provide a sensible migration path and coexistence with current applications, with better reliability and availability.

ATM, LAN to WAN: The trend is towards consolidation and simplification of enterprise networks. Since the same technology is used for the LAN, WAN, and MAN networks, end-to-end connectivity and management is simpler. This also simplifies your training of LAN and WAN network personnel. Handling of the ATM cell remains the same across the local, wide and metropolitan area networks.

3.5.1.5 ATM LAN Emulation Service

In order to use the vast base of LAN software, it is necessary to define an ATM service that emulates services of existing LANs across an ATM network. This service is called *LAN emulation* service and supports the interconnection of ATM networks with traditional LANs by means of today′s bridging methods. End systems (workstations, servers) connect to the ATM network while the software applications interact as if they were attached to a traditional LAN.

To emulate a LAN-like service, different types of emulation can be defined, ranging from emulating the MAC service (802.X LANs) up to emulating the services of network and transport layers like TCP/IP, OSI and APPN. In the LAN environment the main objective of the LAN emulation service is to offer the same MAC driver service primitives which will keep the upper layer protocol layers unchanged. This allows existing applications to access an ATM network via protocol stacks like TCP/IP, NetBIOS, etc. as if they were running over traditional LANs.

The LAN emulation service must support the use of multicast MAC addresses (broadcast, group, or functional MAC addresses). In case of multiple emulated LANs over a single ATM network, broadcast/multicast messages will be distributed only to the members of an emulated LAN, and not to all users of the LAN emulation service. The LAN emulation service provides not only the connectivity between ATM-attached end systems, but also allows connectivity with LAN-attached stations. This includes connectivity both from ATM stations to LAN stations as well as LAN stations across ATM. The bridging methods used for interconnection includes transparent and source-route bridging.

Note

LAN emulation service includes any-to-any bridging through the ATM network. This includes token-ring-to-token-ring, Ethernet-to-Ethernet, and token-ring-to-Ethernet bridging. These functions are provided by ATM bridge products.

Basic Configurations of the LAN Emulation Service: Figure 32 on page 90 illustrates the three basic configurations for LAN emulation service.

- ATM-ATM
- ATM-LAN
- LAN-LAN

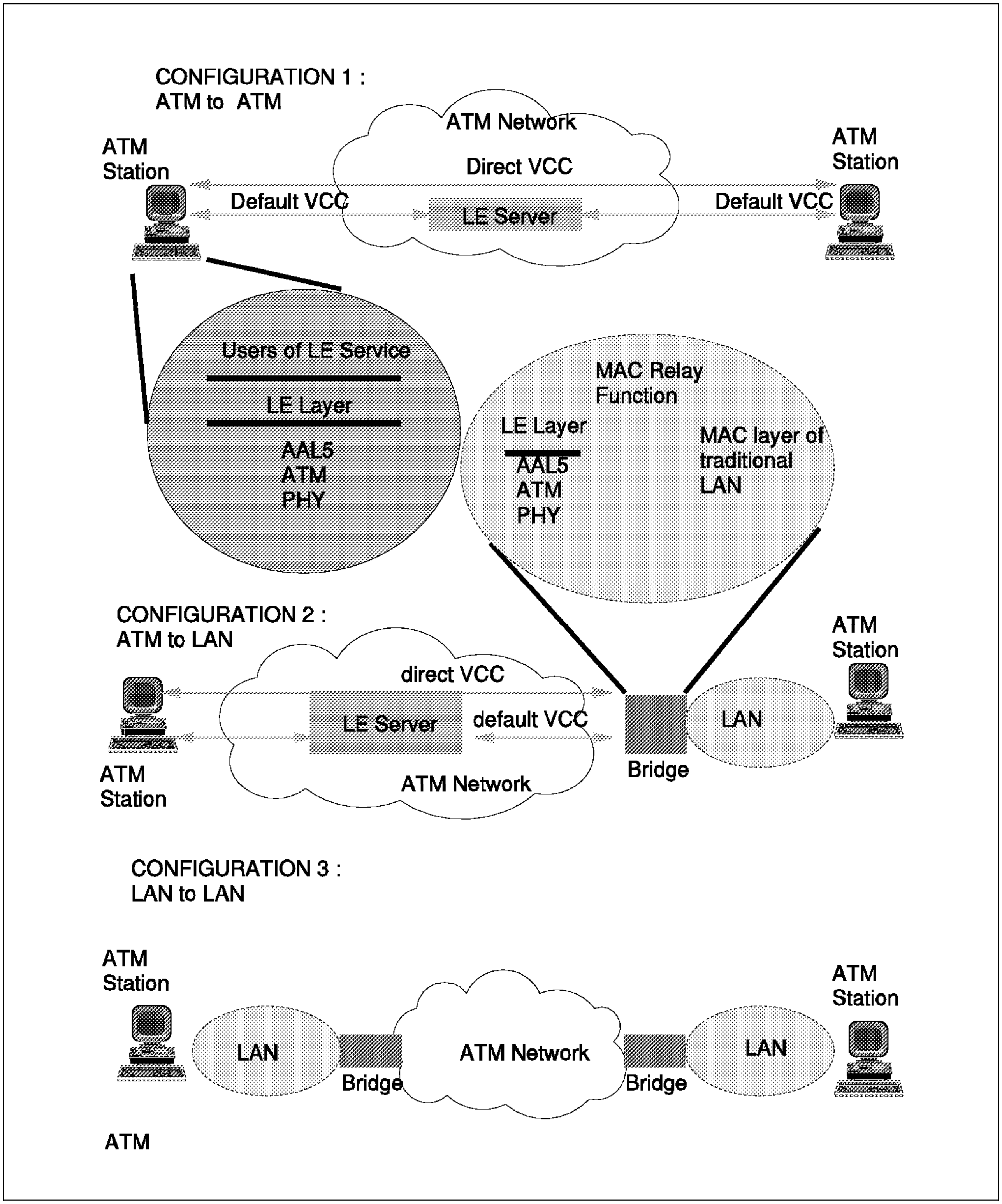


Figure 32. Proposed Configurations for LAN Emulation Service

The first configuration shows the interconnection of two ATM stations using the LAN emulation service. Within the ATM stations, the *LE* service is provided by the *LAN emulation layer* (LE layer). The IEEE 802.2 LLC, NetBIOS, and TCP/IP protocols use this layer, which shields them from the ATM network and gives the illusion of being attached to a traditional LAN.

In the network, the LE emulation service is supported by a LAN emulation server which may be external or integrated in the ATM network. All LE layers are connected to the LE server via a default VCC. Default VCCs are point-to-point VCCs connecting a LAN emulation layer to the LAN emulation server. These VCCs may be permanent or switched. Direct VCCs are point-to-point VCCs connecting two LAN emulation layers directly. Direct VCCs are always switched.

The LE layer inside an ATM station transfers LAN frames to their destination. These are carried over two types of VCCs:

- Default VCCs: Station sends LAN frames to the LE server who then forwards them to their destination based on address information in the frame header.
- Direct VCCs: LAN frames are sent to the destination LE layer. The direct VCCs are established and maintained by the LE layer for as long as they are needed.

The decision on which type of VCC to use depends on whether or not the sending LE layer knows the ATM address of the receiving LE layer.

The second configuration of Figure 32 on page 90 shows the interconnection of an ATM station with a station attached to a traditional LAN. The ATM bridge which is used also contains an LE layer on its ATM side. The bridge layer gives to its user, the MAC relay function of the bridge. Therefore today′s bridging methods can be used without modification.

The standard LAN-to-LAN interconnection over ATM is illustrated in the last configuration in Figure 32 on page 90 and there is no difference between the bridge component shown here and the one in the second configuration. In order to emulate both an IEEE 802.5 and an IEEE 802.3 LAN, two instances of the LAN emulation server must be implemented. In order to obtain LAN emulation services, an ATM endstation must implement the functions of the LAN emulation layer. Bridges must also implement a separate instance of the LE layer for each LAN that is emulated.

Multicast Traffic: Since MAC services are emulated, the LE service must also support broadcast, group or functional MAC addresses. There are several ways the LE service can support multicast traffic:

- No multipoint connection supported by the ATM network:

  The LE service implements the multicast function itself.

- Point-to-multipoint connection supported by the ATM network:

  This function uses a point-to-multipoint connection to distribute the frame instead of copying them.

- Multipoint connection supported by the ATM network:

  In this case no multicast function needs to be implemented by the LAN emulation service. The service uses a multipoint connection of the ATM network to distribute the multicast frames.

The LE multicast function can reside either within the LE server or within an additional multicast server. If it resides in the server, then the default VCCs can be used by the end systems for transferring their multicast frames to the server. In the additional multicast server case, point-to-point VCCs have to be established between the end systems and the server for this purpose.

Figure 33 on page 93 illustrates the three basic support configurations for multicast traffic.

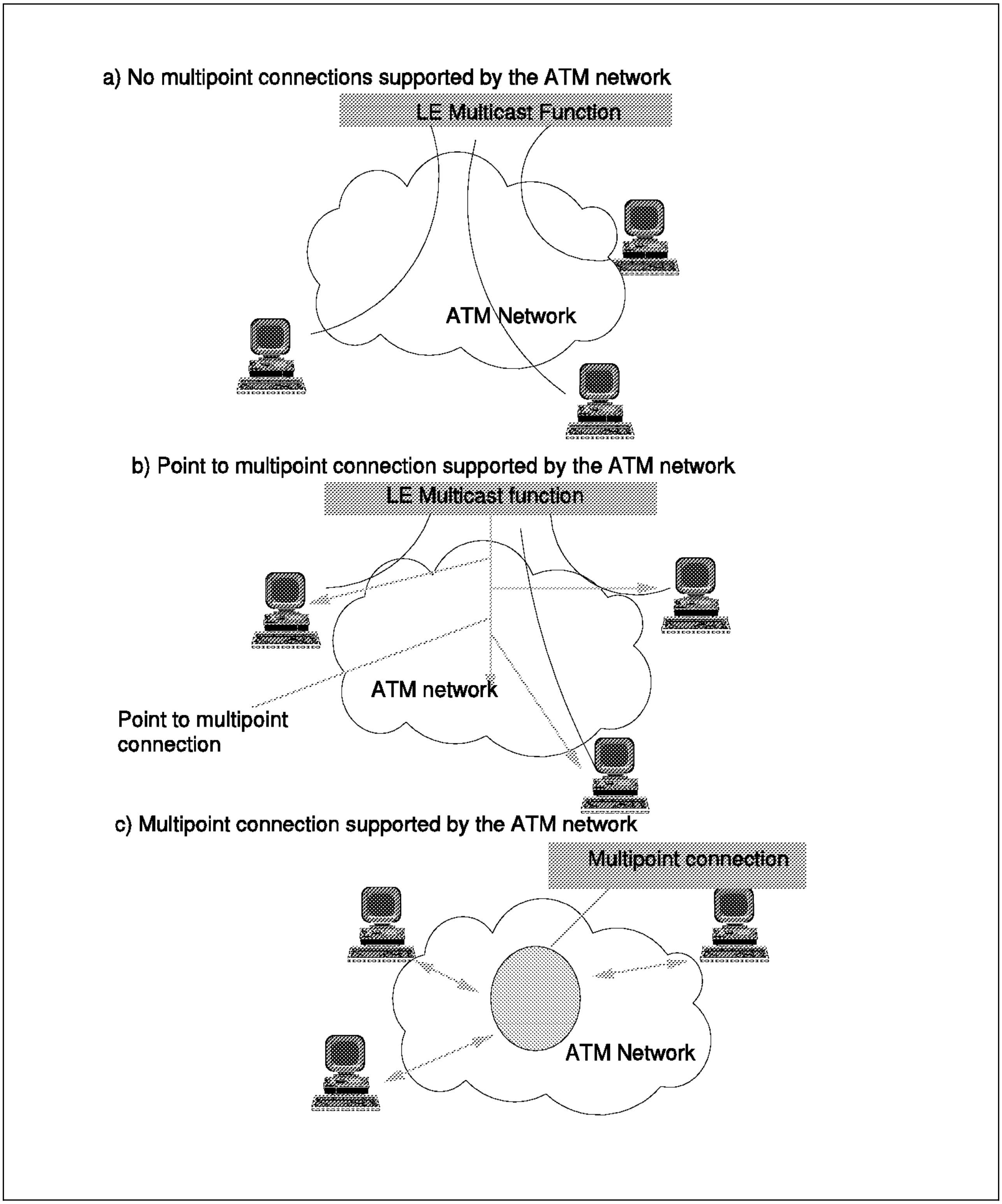

a) No multipoint connections supported by the ATM network
LE Multicast Function
ATM Network
b) Point to multipoint connection supported by the ATM network
LE Multicast function
ATM network
Point to multipoint connection
c) Multipoint connection supported by the ATM network
Multipoint connection
ATM Network

Figure 33. Proposed Support of Multicast Traffic

3.5.1.6 Functions of the LAN Emulation Service

Initialization: The initialization function provides the LE layer (residing in an ATM station or in a bridge) with the capability to set up and access the default VCC. In contrast, the initialization function is used by the LE server to inform the ATM switch (to which it is attached) about its functionality.

Address Registration: This provides the LE layer with the MAC addresses (individual or group) the LE layer has to use to filter the LAN frame it is receiving. It allows the LE server to know about the ATM station and bridges that participate in the LE service.

Address Management and Resolution: This function is only needed by the LE layer if it wants to use the direct VCC to send data. It provides the mapping of the destination MAC address or of a bridge number and segment identifier to the ATM address of the LE layer, which gives connectivity to that destination MAC address or bridge. Address resolution is also part of this function. The same function is provided to the LE server.

LAN Frame Transmit: For an ATM station or bridge, the LAN frame transmit function of the LE layer identifies the VCC (default or direct) over which the LAN frame has to be sent. It also manages the direct VCCs and provides setup, monitoring and encapsulation services. The same applies to the LE server transmit function.

LAN Frame Receive: For an ATM station, the LE layer determines whether the frame is destined to the endstation by comparing the destination MAC address of the LAN frame with the MAC addresses it has obtained at MAC address registration. It also informs the address management and resolution function about the ATM address of the LE layer that sent the LAN frame. For a bridge, the same is true, but with no filtering based on MAC address. The same applies for the LE server.

Server Interconnection: In the case where there is more than one LE server, each server having, within its own area, its own address registration, management and resolution functions, the server interconnection function provides connectivity between stations which belong to different areas. An ATM station may belong to more than one server area at any one time. In this case it is attached to each of its servers via a default VCC. This station does not need the server interconnection function when it communicates with another station which is also attached to one of its servers.

3.5.1.7 ATM in Campus

The early realization of ATM technology takes place in the campus LAN environment. LANs are interconnected by ATM switches to form high-speed backbones. These backbones have the capability of carrying voice, video and data and it is here that the initial growth of high-speed, high-volume multimedia applications will take place. ATM servers and workstations are attached to the ATM switch using the adapters best suited to required speed and wiring infrastructure. The LAN emulation function allows existing LAN devices to communicate with the new ATM devices. LAN emulation also provides a LAN appearance to LAN software allowing it to run unchanged over an ATM network.

The hub products must provide ATM support for the attachment of main frames, servers, desktop workstations, routers and other hubs. The switching fabric that is used must enable the hub to meet the high-speed switching demands between the hub′s ports as well as between hubs.

For IBM products, see Chapter 5 of *Local Area Network Concepts and Products: LAN Adapters, Hubs, and ATM,* SG24-4754.

3.5.1.8 ATM in Wide Area Network

ATM technology also moves into the wide area network allowing high-speed interconnection of existing ATM campus environments. The ATM wide area switches request and manage the connections that support the integration of voice, video and data traffic, and also manage requests for scalable bandwidth. As ATM technology is deployed in the public network arena, several technical, regulatory, and tariff issues must be resolved before widespread deployment is a reality. The WAN (wide area network) switches must support existing standards and connect ATM devices and campus networks. The WAN switches must use a high-speed switching fabric to produce a self-routing, non-blocking, scalable switch with multigigabit throughput, and growth to support gigabit port speeds. The transport network node has to:

- Provide bandwidth on demand to multiple users
- Manage the distribution of traffic across the elements of the ATM network
- Support multicast which is the simultaneous distribution of information to multiple users
- Avoid congestion
- Support high-speed switching of variable length packets

For IBM products, see Chapter 5 of *Local Area Network Concepts and Products: Lan Adapters, Hubs, and ATM,* SG24-4754.

3.5.2 ATM Summary

IBM delivers the ATM technology within the framework of a systems strategy synchronized with the anticipated customer demand. ATM is viewed by many customers as the solution for their ever growing demand for bandwidth. This demand is being generated by increasingly more sophisticated applications that integrate various combinations of voice, data, video and image. An important characteristic of ATM is that it is both a local and wide area transport technology. This suggests that ultimately ATM will provide seamless networking and all the flexibility and economies that LAN = WAN = MAN implies. It has been generally acknowledged that the initial implementation of ATM will address the local environment where its use of switched, rather than shared, media will make it the high-speed LAN of choice. The IBM hub strategy is to deliver these benefits of ATM switching to the desktop via a series of advanced capabilities that will be delivered in the Nways product families. These will address a design point in conformance with the ATM Forum specs and augmented by IBM contributions such as:

- Native ATM adapters ranging from 25 Mbps to 155 Mbps provides ATM connectivity for workstations, servers, and hub-to-hub interconnection. The low-speed 25 Mbps entry is an IBM/Chipcom contribution to the ATM Forum to provide a cost-attractive adapter solution that does not need rewiring since it can use both shielded and unshielded twisted pair cabling in use today for token-ring and Ethernet.

- IBM expects that dedicated 25 Mbps to the workstation will accommodate the vast majority of the applications. For the very few applications requiring more than 25.6 Mbps per workstation, it is possible to use faster adapters (hub-to-hub, server-hub or workstation-hub).

- Any ATM solution must provide support for existing devices and networks in order to preserve customer investment. Devices now connected via token-ring and Ethernet must not only continue to operate as they do today, but also be able to provide access to ATM backbones and communicate with devices that are on the ATM LAN. To address this, IBM has made another contribution to the ATM Forum, the LAN emulation, which permits existing applications (developed for token-ring or Ethernet) to use new ATM adapters and the powerful ATM backbone without changes.
- IBM has recommended common goals for a proposed network management subgroup, including addressing known management requirements for ATM technology, providing a common base for management, and producing a cost-effective response for heterogeneous networks.

IBM believes that ATM offers the promise of being the true information superhighway. IBM has been actively investing in ATM technology for a number of years. One of the fundamental driving forces behind the networking blueprint has been the recognition of the broad value that ATM can bring to our customers, and the need for IBM to bring that value to them in a way that protects and leverages their existing investments.

IBM is an active participant in the ATM Forum with several key submissions in the area of LAN support. IBM has also invested in several key prototype efforts to test the viability of the technology and the networking enhancements that will have to be made to help ATM deliver on that promise. Figure 34 on page 97 positions ATM with current and emerging networking technologies.

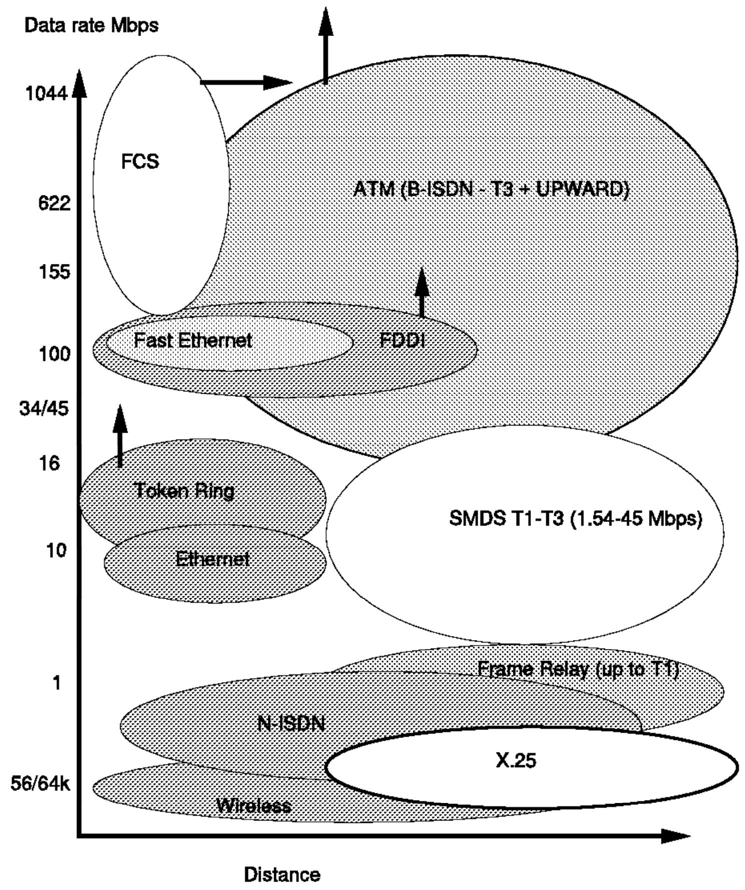


Figure 34. Network Technology Trends

3.6 Wireless LAN/802.11

Currently the IEEE 802.11 has specified wireless LAN standards for the U.S.A. There are other bodies designated for other geographical areas as defined below.

Europe

- **Spectrum allocation**

  Below are the CEPT recommendations:

  – The band 2400 - 2500 MHz shall be used on a non-interference and unprotected basis for wide-band data transmission systems using spread spectrum technology with a minimum aggregate bit rate of 250 Kbps.
  – The total power in this frequency range shall not exceed 100 mW.
  – The instantaneous peak shall not exceed 100 mW measured in a 100 kHz bandwidth for systems using frequency hopping spread spectrum technology.

- **Workgroups in ETSI for radio equipment and systems**
  - RES2: Responsible for tests and conformance testing for the wide-band wireless data system in 2.4-2.5 GHz ISM band. Draft published 1/93 for public inquiry.
  - RES3: Responsible for DECT
  - RES8: Low power devices
  - RES10: Radio LANs hiperlan (speeds above 10 Mbps)

U.S.A

- **FCC**
  - The FCC standards are already published for the ISM band of 2400 to 2500 MHz in the code of Federal Regulation 47 parts 0-19 (October 1991).
- **IEEE 802.11**
  - 93 members
  - IBM is proposing the MAC and PHY protocol
  - Focus is on the 2.4 GHz band design

3.6.1 Wireless LAN Concepts

The prediction today is that the 1990s will become the decade for wireless communications. IBM is taking an active role in ensuring that products both in the LAN and WAN environments are available to customers who wish to choose either infrared or radio frequency technologies.

With the miniaturization of personal computers such as laptops and notebooks and the growth of wireless networks like ARDIS, Circuit Cellular and CDPD (Cellular Digital Packet Data), more commercial applications are being developed to take advantage of the benefits wireless communications has to offer.

3.6.1.1 The Electromagnetic Spectrum

The electromagnetic spectrum has often been divided up, somewhat arbitrarily, by wavelength; the divisions have been assigned familiar names like VHF (very See Figure 35 on page 99.

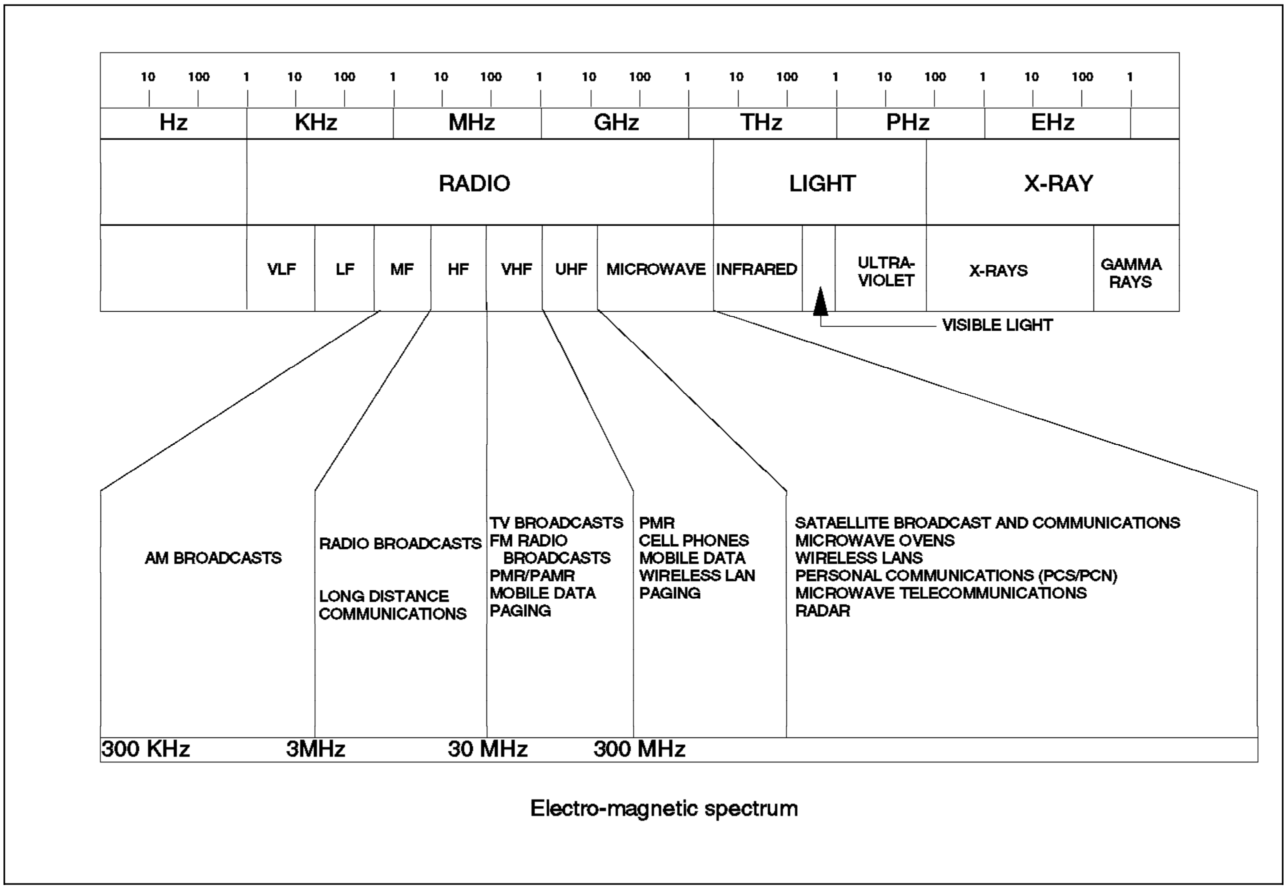


Figure 35. The Electromagnetic Spectrum. The divisions indicated here are not absolute. Other representations of the spectrum may show the divisions at slightly different wavelengths.

Superimposed upon some of these divisions are legacy names from the early days of radio (for example, short wave and long wave which refer to the particular region of the spectrum to which the radio transceiver is tuned).

Over the years uses have been found for every portion of the spectrum to the extent that every country has had to impose some form of licensing or regulation on users of transmitting devices. The uncontrolled use of spectrum leads to chaotic communications which in turn leads to potential life-threatening situations, where emergency messages cannot be received due to interference from other users on the same frequency. Less dramatic, but very damaging, would be the impact on business and human existence if radio communication could not be relied upon.

In some industrialized countries, the spectrum has been over-subscribed to such an extent that there is now an acute shortage of frequencies available for further exploitation. In the US, for example, there has recently been an auction of some under-utilized frequencies. A successful bidder for any of these frequencies has the exclusive right to build devices or services using them.

Internationally, there is a move to standardize the use of specific frequency ranges for particular applications. However, due to the ad hoc nature of the way spectrum has been used in the past, this is no simple task. One successful allocation has been in the ISM (industrial, scientific and medical) bands. There are in fact three ISM bands, but only the middle band from 2.4 to 2.48 GHz is available worldwide for exploitation by wireless systems such as LANs.

These bands require no licensing, and the middle band in particular is widely used by different devices. Many consumer devices are found in these bands, including such things as baby monitors and garage door openers. The most familiar usage of this band is for microwave ovens. It is also proving very popular with wireless LAN developers. In the latter case, a special transmission technique must be used to minimize the interference with similar devices in the vicinity. The technique described later in this chapter is known as a *spread spectrum*.

Frequency and Wavelength: An energy wave of a particular frequency will have certain physical characteristics. For instance, a radio wave below about 900 MHz will be able to propagate well through walls and floors and be generally able to penetrate buildings. Broadcast radio between about 1600 kHz (AM) and 108 MHz (FM) are good examples of this. Frequency(f) and wavelength(λ) are related to each other and to the speed of light by the following expression:

$$f \times \lambda = 3 \times 10\bullet$$

where: f is measured in cycles per second or Hertz (Hz)
λ is measured in meters
the speed of light is measured in meters per second

It can be seen from this expression that since the speed of light is constant, if the frequency increases, the wavelength must decrease proportionally. Frequency and wavelength are said to be in inverse proportion to one another. These characteristics are important when considering radio or wireless technologies of any kind. For example, the shorter the wavelength, the more the signal is attenuated by the air and the particles in it. This is why FM broadcasts tend to be localized, while AM broadcasts can be received over much greater distances at the same transmit power.

As the wavelength becomes shorter, the wave takes on more of the properties or characteristics of light. It is more easily absorbed by most materials, and so it attenuates quickly. The shorter the wavelength, the more pronounced the effect. However, light-like properties can be extremely useful to a wireless developer. Generally, the shorter the wavelength, the easier it is for the wave to be reflected, focused and generally controlled. A signal with a frequency greater than about 300 MHz can be focused using a parabolic reflector which is the familiar *dish* antenna seen at TV stations, on building roofs and even in domestic gardens. Focusing a signal in this manner and directing it towards a destination means the transmitting device needs far less power than if it were transmitting in all directions. Focusing an incoming signal with the same antenna means that a far weaker signal can be detected than with a conventional antenna.

3.6.1.2 Wide Area Cellular Networks

Cellular networks are now commonplace throughout the world alongside conventional wired networks. Some countries are electing to install cellular networks, in preference to wireline networks, in regions that have previously not had a telephone service.

Many people are familiar with the cellular phone; it is for voice transmission that these networks were developed.

Analog Voice Services: The basic design element of an analog service network is that of a cell, which consists of a base station and a number of mobile devices which it controls. To cover a larger geographic area without raising output power levels means distributing more base stations and hence creating more cells. This has the additional benefit of increasing channel reuse and therefore overall capacity. With the proper controls, users can roam between these cells without suffering a break in communication.

There is no worldwide standard for such cellular systems, but certain implementations do dominate. In the US, Canada and Japan, AMPS (Advanced Mobile Telephone System) is prevalent. In many European countries, TACS (Total Access Control System) and NMT (Nordic Mobile Telephone) are widely distributed, although some countries, France, Germany and Italy for example, use a system specific to each country.

The AMPS, TACS, and NMT systems share some common features (apart from all being analog services). They all utilize either the 450, 800 or 900 MHz band depending on which country they are operating in and which band has been allocated. They all use FM (Frequency Modulation) and FDMA (Frequency Division Multiple Access) as their means of voice transmission.

It is possible to use these voice networks to send data. However, this is not generally desirable from an end user point of view. The main reasons for this are as follows:

- **Cost**

  These are connection-oriented services which means that the user is paying for the entire connection period. Cellular phone charges are generally much higher than conventional PSTN (Public Switched Telephone Network) charges. The data rate can be up to 19.2 Kbps, but even at this speed a large file transfer would require a long connect time and high cost.

- **Reliability**

  These services were designed specifically for voice transmission, and absolute accuracy was not a consideration. The human ear is very tolerant when it comes to conversation. It can miss complete words and even sentences and yet still be able to infer what the missing pieces were with remarkable accuracy. Humans can maintain conversations over extremely bad communication lines.

  A data transmission is far less forgiving. Breaks in communication mean data has to be retransmitted which adds to the cost of the connect time. The receiving application may not be aware that it has missing data, so the end result may be incomplete. Noise on the channel may be interpreted as data and accepted, thereby corrupting the actual data.

  These are very real effects in an environment where a user may be moving into and out of areas of poor radio reception.

Digital Voice Services: A digital service network shares a similar design philosophy with an analog one in that there are a number of base stations controlling mobile users within the cells. As with analog services, there is no worldwide standard for a digital service. However, a widely accepted standard today is GSM (Global System for Mobile communication).

GSM operating in the 900 MHz range is now firmly entrenched throughout Europe and is the digital cellular standard there. A product that came out of the same development effort is DCS 1800 (Digital Cellular System 1800) which is also implemented in Europe but not to the same extent as GSM. It operates at 1800 MHz, hence the name, but otherwise it is essentially the same technology as

GSM. Users on either system are able to share the infrastructure of both networks because of this similarity.

The US rival to GSM is DAMPS (Digital Advanced Mobile Phone Service) which, due to the extreme lack of frequency available there, must share the same band of 800 MHz as the AMPS analog service. In Japan, JDC (Japan Digital Cellular) is the available phone service. There is, at present, no shortage of spectrum in Japan, so JDC has the luxury of operating in three pairs of bands with distinctly separate uplink (mobile to base) and downlink (base to mobile) sub-bands as follows:

- 810 - 826 MHz downlink
- 940 - 956 MHz uplink
- 1429 - 1441 MHz downlink
- 1447 - 1489 MHz uplink
- 1453 - 1465 MHz downlink
- 1501 - 1513 MHz uplink

Digital systems are designed to eliminate some of the shortcomings of the analog networks as follows:

- **Efficiency**

  An analog conversation ties up the entire frequency channel for the duration of the conversation. Much of the time is *dead time* such as the spaces between words, pausing for breath and thinking time. A digital system can use this dead time for other conversations at the same frequency using the time-slicing technique of time division multiple access (TDMA).

  Digital technology allows speech to be compressed prior to transmission thereby permitting even more users per channel. GSM, for instance, can support eight conversations in the same bandwidth where analog technology will only support one.

- **Quality**

  A digital data stream can have error correction built in so that short breaks in communication and spurious noise on the channel can be largely overcome.

- **Security**

  Using TDMA as a transmission protocol makes the conversation very difficult to monitor and follow. Specialized equipment is required to do this.

On a reliability basis alone, digital cellular services hold more promise as a means of transmitting data than analog cellular services. It is still a connection-oriented service and there is still a connect time charge to be paid. However, the channel can be shared by up to eight users concurrently, and tariffing should be more reasonable.

Note: IBM′s ARTour wireless product supports the GSM interface.

Packet Data Services: These networks were designed specifically for carrying data with no provision being made for voice traffic. As the name suggests they are packet data networks for the radio environment. Of the three services mentioned here, Mobitex and RD-LAP were designed to support mobile data services and in this capacity can be used by devices mounted in vehicles.

There is no worldwide standard for this technology and no common frequency on which to operate. However, Mobitex and RD-LAP are the two main services around which the world seems to be polarizing. A third and upcoming contender is CDPD.

- **Mobitex**

  This was developed by Ericsson Mobile Communications AB of Sweden and began operating commercially in 1986. Since then it has experienced a number of enhancements and is now implemented widely across Europe and the US. In the US it operates in the ranges of:

  - 935 - 940 MHz downlink
  - 896 - 901 MHz uplink

  In Europe the ranges vary by country, but all are within the range:

  - 425 - 459 MHz

  Mobitex has also been implemented in Australia and a number of South East Asian countries. The maximum data rate on a Mobitex network is 8.4 Kbps, but the actual throughput is closer to 8 Kbps.

- **RD-LAP**

  RD-LAP is the name given to the radio protocol which runs on a network developed by Motorola (with some input from IBM in the earlier stages). In the US, this network is known as ARDIS (Advanced Radio Data Information System) where it operates in the ranges of:

  - 851 - 869 MHz downlink
  - 806 - 824 MHz uplink

  Like Mobitex, it has been implemented in many parts of the world where it is known by various names. In Germany, for example, it is Modacom; in Australia it is Datatac.

  The data rate is 19.2 Kbps which gives an actual throughput of about 14.4 Kbps. An earlier protocol called MDC4800 (Mobile Data Communication 4800) is still used by some services on ARDIS, and as its name suggests it has a data rate of 4.8 Kbps. However, both protocols use exactly the same infrastructure.

The previous two radio networks have been developed independently of X.25, but they do resemble an X.25 network somewhat in their packet handling.

They have the same cell structure as the voice networks in that they have multiple base stations controlling mobile devices within their cells. The base stations then connect via landline (such as an X.25 network to a user′s home network) where the user then has access to whatever applications that normally reside there.

Note: IBM′s ARTour product supports both the Mobitex and RD-LAP interfaces.

- **Cellular Digital Packet Data (CDPD)**

  The CDPD service (originally Celluplan II) was initially developed in the US at IBM Boca Raton, with the following objectives:

  - To provide a digital packet (connectionless) network, able to transmit data reliably and cheaply
  - To use the same frequency allocation as the existing AMPS network without impacting the voice carrying capacity or conversations in any way
  - To utilize the dead time on channels in the AMPS network to transmit data packets
  - To use the existing AMPS network infrastructure

  The way CDPD works is to constantly monitor the existing AMPS channels for activity. When it has data to send, it selects the first available free channel and begins to transmit. It will continue to use this channel as long as it is not required to provide a voice call by the analog network.

  If the channel is required to provide a voice call, CDPD will cease to transmit on that channel immediately and will begin to scan the channels again until it finds another one free. At this point, it will switch to the free channel and continue its transmission. These breaks will not affect the integrity of the data being sent, and providing the break is not of a duration such that the receiving application times-out while waiting for some response, the data will arrive intact.

  Considering that the amount of dead time, which has been shown to be as high as 50% on even the busiest of analog networks, such time-outs are likely to be rare.

  So far, CDPD has only been implemented in the US, where it appears to be seeing some success. It is currently being offered by at least six carriers. Whether other countries choose to follow suit remains to be seen.

3.6.1.3 Wireless LAN Types

Wireless LANs and wired LANs operate in much the same way at the physical level. That is, they both use electrical energy to transmit data. One encodes its data upon an electrical impulse in a wire, and the other encodes it upon a radio or light wave.

With a waveform, the amount of data that can be transmitted is theoretically, directly proportional to the frequency of the wave. For example, if a particular piece of switching equipment in a transmitter is capable of encoding (modulating) one bit of data per cycle, then a wavelength with a frequency of 10 kHz (ten thousand cycles per second) is going to carry ten times as much data as a wavelength having a frequency of 1 kHz in the same period of time. This holds true for as long as the switching equipment can match the frequency.

Some modulating methods make it possible to encode up to 4 bits on a single cycle. Considering this and the high frequencies available, wireless LANs have the potential to transmit data at very high speeds.

As already mentioned, the characteristics of a wave change with respect to the rate of frequency change. The higher the frequency, the more the wave takes on the properties of light and the more it is absorbed or reflected. This has a number of implications; the main one being the fact that as the frequency increases, the range of the LAN decreases due to the greater attenuation of the signal. Also, the equipment used to transmit and receive radio waves begins to strike some physical limitations as the frequency increases, so there is an upper limit to its practical application.

There is a wide frequency range which could potentially be used for wireless LAN communication. The range from about 200 kHz, through the microwave range, all the way up to the top of the infrared range at around 200 THz, could be utilized. When selecting a range upon which to build equipment, consideration must be given to the required throughput versus transmission range. As one increases, the other decreases as mentioned previously. Greater range can be achieved at any frequency by increasing the transmission power. However, most governments impose strict controls on power output through their licensing and regulatory authorities. A more efficient antenna or a directional one can increase the range of any transmission, but this is nowhere near as significant a gain as simply increasing the power.

The choice of frequency range has largely been limited by regulation. The same frequency shortages that occur, at these higher regions also occur lower down in the spectrum. Most frequencies need to be licensed, which adds to the cost of the technology. In fact, practically all wireless LAN development has gone into supporting a limited number of specific bands in the microwave and infrared regions. These bands have been identified (in some countries at least) for specific purposes and have no licensing requirement. However, there are tight controls on permitted power output which effectively limits the range to a maximum of 400 meters.

Apart from the health consideration, this limitation attempts to reduce the amount of interference that would otherwise be experienced by different organizations employing the same technology in close proximity to each other.

Microwave LANs: In the US three bands have been reserved for ISM (Industrial, Medical and Scientific) use and require no licensing, but the restrictions, as mentioned above, do apply. A further restriction is that radio devices using these bands must employ a spread spectrum technique to further reduce the possibility of interference with other users. The first is in the UHF (ultra high frequency) band and the other two are in the microwave region. The frequencies of these bands are:

- 902 - 928 MHz
- 2.4 - 2.48 GHz
- 5.7 - 5.85 GHz

Other countries have recognized some or all of these designations, but only the 2.4 GHz band has been recognized internationally.

Today, wireless LAN vendors must decide whether to build and support more than one range of equipment to cover multiple bands or to stick to the one band internationally available. An international traveler should be aware that a device that supports the 902 MHz or 5.8 GHz specification may not be legal in all countries. With the 2.4 GHz specification there is less cause for concern.

One of the drawbacks with using the ISM band is that other devices besides the radio have been built to the same specification. The most widely known of these would be the microwave oven. These can have an undesirable effect on nearby radio transmissions. However, commonplace machinery like generators, welding equipment and transformers can all produce frequencies in this range. Operating a wireless LAN in these areas poses unique design considerations to overcome strong radiators.

Note: The IBM Wireless LAN Entry and IBM Wireless LAN products both operate in the 2.4 GHz band.

Infrared LANs: The infrared portion of the spectrum is used successfully in a number of familiar domestic devices. Infrared heating lamps use the lower end of the infrared band. The rest of the band is utilized for things like television remote controls, motion detectors and door openers.

From a wireless point of view, light has the potential for very high transmission rates; the present technology already approaches that of wired LANs such as a token-ring at 16 Mbps. As with the ISM bands there are no licensing requirements for infrared use, but there are controls nonetheless. Transmitters are only allowed to operate at fractions of a watt, and the maximum range that can be achieved is about 20 meters. A great deal of infrared energy arrives at the surface of the earth from the sun. This precludes an infrared LAN from being used outdoors. This is in contrast to microwave LANs which are unaffected by sunlight and are frequently employed in outdoor applications. Infrared signals tend to reflect off walls and ceilings, a characteristic that can be taken advantage of in a *diffused* design.

Infrared transmitters can be constructed with either a laser diode (LD) or a light emitting diode (LED). Both have their advantages and disadvantages and tend to find specific uses.

LDs are used in a line-of-sight placement, where a device needs to communicate with one other. The devices need to be precisely aligned for communication to take place. This is called *focused* infrared.

LEDs produce a spread of light rather than a beam and so lend themselves to transmission by diffusion. This is where the beam is transmitted over a wide angle. It may be reflected off multiple surfaces before being picked up by the receiver. When this happens, the same signal can arrive at the receiver from different directions and at slightly different times which is known as the multipath effect. It is this same effect which produces the phenomenon of *ghosting* on a television set.

However, one advantage of being *surrounded by signals* is that workstations have the ability to move around the local area without losing contact with the base. In addition, where the LAN topology supports it, roaming can occur.

The multipath effect can be overcome somewhat by a technique known as quasi-diffuse transmission. This is where the transmitters and receivers are all directed towards a particular spot on the ceiling known as a *satellite*. Instead of a diffuse spread, a focused beam is directed at the satellite from the transmitter. The reflected signal is then detected only once by all the receivers.

Note: The IBM Infrared Wireless LAN product is IBM′s offering in the infrared LAN arena.

3.6.1.4 LAN Access Methods

When a radio transmission is limited to one particular frequency, for example a radio broadcast, this is known as a narrow band transmission. When a band with a large frequency spread has been allocated, the transmission using it is known as a broad band transmission.

Broadband transmission can be either FDMA (frequency division multiple access) or spread spectrum. Spread spectrum techniques can be either CDMA (code division multiple access), direct sequence (DS) or frequency hopping (FH). Frequency hopping in its turn can use CSMA (carrier sense multiple access) or TDMA (time division multiple access) access methods. These last two techniques are equally applicable to narrow band transmissions if multiple stations wish to share the same frequency. The following list summarizes these techniques:

- Narrow band
  - TDMA
  - CSMA
- Broad band
  - FDMA
  - Spread spectrum
    - CDMA
    - Frequency hopping
      - TDMA
      - CSMA
    - Direct sequence

TDMA: Time division multiple access is a deterministic method of sharing a channel resource between multiple users on a *time-slice* basis. It follows the same philosophy as a token-ring in the wired LAN world. With TDMA, each station is allocated a time slot of a particular duration in which it can transmit data. If it has no data to send, the slot is allocated to another station.

Because a station only needs to listen at assigned times, there is a power saving potential for the device in question. TDMA tends to use bandwidth very efficiently and can handle isochronous data, such as video, in a very predictable manner. It also has the ability to prioritize data from particular stations.

TDMA is a technique equally applicable to wired and wireless LANs and also in wide area wireless where it is used in the GSM and CDPD technologies.

CSMA: Carrier sense multiple access is a contention-based system and, as a design philosophy, shares a common ancestry with the Ethernet LAN. Indeed, the chip sets in common usage in wireless LANs are frequently the same as those used in wired LANs.

CSMA depends upon each station listening to an allocated channel when it has data to send, and if the channel happens to be free, it then transmits. This is also known as *listen before talk* or *LBT*. If another station has done the same thing, then a collision occurs and both stations must retransmit but at different times. A *random backoff algorithm* is invoked on each station to ensure that they do not attempt a second transmission at the same time.

A wired LAN uses CD (collision detection) to determine whether a packet was sent successfully or not. In a wireless LAN, a station has no way of detecting a collision, so CA (collision avoidance) is implemented instead.

FDMA: FDM, also known as frequency division multiplexing, is a broad band method of transmitting multiple signals concurrently over very closely spaced frequencies within a given band. It is the technique used by cable television to transmit multiple programs over the same piece of wire. It works equally well using microwaves as the carrier, and it could be utilized over a large range of the spectrum.

Its success depends upon the receiving station being able to detect and lockon to the desired frequency.

Spread Spectrum: Spread spectrum techniques spread a transmission over multiple available frequencies within a given band. Simply stated, the goal is to have a low watts-to-hertz ratio. Usage of the ISM bands for radio transmission of any kind demands that one of the two spread spectrum techniques be employed. The technology has been around for some 40 years and was originally developed in the US for military use.

In this environment, spread spectrum provides good protection against channel jamming and intentional eavesdropping by unfriendly elements. In the business environment it offers protection from radio interference and casual eavesdropping.

CDMA: This spread spectrum technique uses direct sequencing (DS) to artificially spread the signal over a greater range of frequencies than it would actually need if it were transmitting user data alone. Multiple user data streams can be transmitted concurrently over exactly the same spread of frequencies.

The factor that distinguishes one data stream from another is the additional *pseudo-random bit stream* with which it is combined. Every random bit stream is unique, which in turn makes every user data stream unique. Only the sending and receiving stations of a particular bit stream know what the random pattern is. The receiving station strips off the random bits and the user data remains.

This technique makes a transmission relatively insensitive to the effects of background noise even when that noise is at quite high levels.

Frequency Hopping: Frequency hopping is another spread spectrum technique that divides up the available bandwidth into a number of discrete channels. A transmission is synchronized between the sending and receiving stations such that they switch channels or *hop* in a *pseudo-random* pattern. Transmission will recommence on every channel that is hopped to.

Note: Either CSMA or TDMA can be used for channel access.

The technology distinguishes between FH (fast hopping) and SH (slow hopping) with SH being by far the most widely implemented. The stipulation with SH is that hopping must occur at least every 400 ms and must cover all available channels.

3.7 Fast Ethernet / 802.30

This is the name of a technology that marries a modified version of the IEEE 802.3 CSMA/CD protocol with the 100 Mbps ANSI TP-PMD twisted pair signalling approach used in CDDI. While seeking to preserve much of Ethernet′s CSMA/CD protocol, 100 Mbps makes significant trade-off in the allowable topologies, cabling, and network design rules with which 10Base-T users are familiar. In addition, 100 Mbps offers no migration path for token-ring users. The IEEE 802 standards organization has chartered a 802.30 committee to consider 100 Mbps along with a number of competing 100 Mbps CSMA/CD proposals.

3.8 100VG-AnyLAN / 802.12

This is a new IEEE 802.12 technology for transmitting Ethernet and token-ring frame information at 100 Mbps. 100VG-AnyLAN combines increased transmission speeds with a simple yet efficient media access control that operates over Category 3, 4 or 5 unshielded twisted pair (UTP), shielded twisted pair (STP), and optical fiber. By supporting all of the network design rules and topologies of 10Base-T as well as token-ring, 100VG-AnyLAN allows organizations to leverage their existing network and cable infrastructure. In addition, 100VG-AnyLAN can provide guaranteed bandwidth for emerging time sensitive applications such as multimedia. Low costs and an easy migration path for existing 10Base-T and token-ring networks promise to make 100VG-AnyLAN the best alternative for upgrading 10Base-T as well as token-ring users to 100 Mbps speeds.

100VG-AnyLAN will operate at 100 Mbps, with the ability to bridge to either 10 Mbps to standard Ethernet environments or 16 Mbps to standard token-ring environments. This speed will be available on all wiring types.

3.8.1.1 Demand Priority

Demand priority provides a transfer with very low latency across the LAN hub. The low latency is achieved by means of an ″on the fly″ packet transfer. The demand priority protocol supports guaranteed bandwidth through a priority arrangement. A request to transmit data is made by a workstation. This is sent to a switch that handles it immediately if there is no other request active. This is done in a FIFO basis. When the packet is transmitted to the hub, the hub determines in real time, the correct output port and switches the packet to that port. Packet traffic can be flagged as priority ensuring guaranteed bandwidth.

The fact that neither Ethernet nor token-ring media access control (MAC) is used to send a frame allows either of them to go through a single switch. The use of Ethernet or token-ring frame formatting will allow the traffic to be transparent to applications running on today′s Ethernet or token-ring LANs.

3.8.1.2 Quartet Coding

Quartet coding is used in the transmission of 100VG-AnyLAN by segmenting transmissions into ″quartets″, and then sending them concurrently across four pairs of copper wires. A data transfer rate of 100 Mbps can be achieved over standard unshielded (UTP) copper wires without giving up topology distance.

100VG-AnyLAN is a dedicated media; shared bandwidth network and all devices attached to the network share the 100 Mbps bandwidth.

3.8.1.3 Cabling Support

100VG-AnyLAN will support four pair Voice Grade copper lines (Category 3) as well as Category 4 and 5. Two pair STP copper lines and fiber optic lines will also be supported. The topology for 100VG-AnyLAN is a 100 meter radius for Category 3 cabling, a 200 meter radius for Category 5 cabling and STP and a 2 km radius for fiber optic cabling.

A new workstation adapter card and new concentrator would be required to implement this technology. At this time IBM has not announced any specific product functions.

3.9 Switched and Full-Duplex LANs

Switching is a general term that describes numerous networking technologies. These include packet switching, port switching and some less obvious techniques such as high-speed bridging and routing.

Early LANs were typically small, both in network span and in the number of stations. There were few concerns over bandwidth requirements. These early LANs were typified by Ethernet 10Base5 and 10Base2 and can be categorized as having shared media and shared bandwidth, with bandwidth being shared on a common medium. The LAN adapters were half-duplex in operation.

In the last decade, LANs have grown in scale, speed and importance. Cabling practices have evolved from shared bus to star-wired shielded and unshielded twisted pair cable. Examples of these cabling schemes are the IBM Cabling System and the EIA/TIA-568 Commercial Building Telecommunications Cabling Standard. Ethernet 10Base-T and token-ring are examples of LANs that often employ a wiring closet for concentrating each star-wired LAN segment. The star wiring simplifies problem determination on a LAN. A wiring closet provides one place where all station attachments can be accessed. Thus, a defective station or cable, once identified, can easily be removed and repaired without the need to literally walk the wire as with Ethernet bus configurations.

Initially, wiring closets contained little “intelligence.” Token-ring had non-powered concentrators containing insertion relays. On these shared-media and shared-bandwidth LANs, any cable or station failure was liable to affect all stations on the same LAN segment.

With the continued growth and business importance of LANs and the decreasing cost of electronics, wiring closets started to become “intelligent.” For example, an intelligent token-ring concentrator can prevent a 4 Mbps station from inserting into a 16 Mbps LAN segment. 10Base-T concentrators provide a number of mechanisms including an auto-partition function to disable faulty devices and cabling faults.

LANs built with intelligent wiring closets can be classified as dedicated media and shared bandwidth. Bandwidth is still shared among all users, but each station′s media is “dedicated.” Because of the star-wired cabling topology, station failures are most often isolated to a single station and its media. Figure 36 on page 111 shows a five-port token-ring intelligent concentrator.

The active logic shown in the figure automatically removes failing stations from the main ring, so, for example, a station attempting to insert at a data rate different from the main ring data rate will not be permitted to join the network.

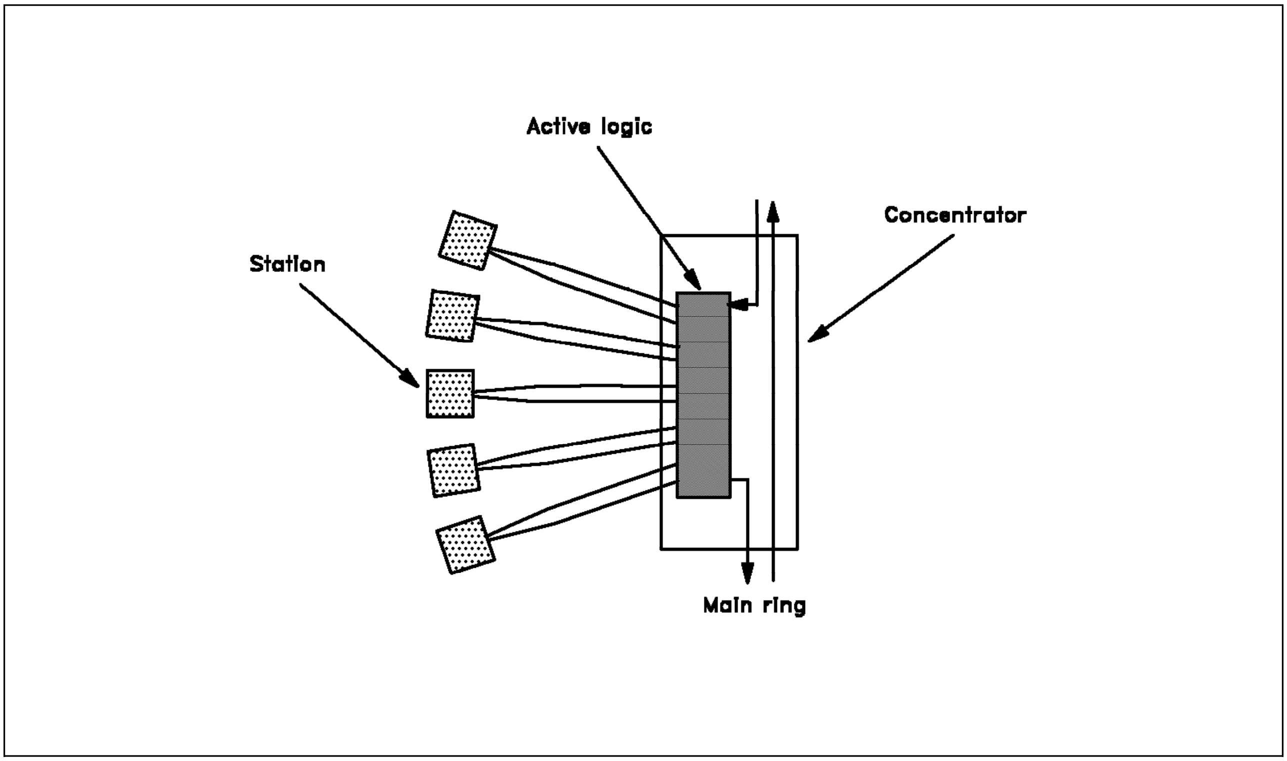


Figure 36. A Five-Port Token-Ring Intelligent Concentrator

Shared bandwidth with dedicated media increases the manageability and reliability of LAN segments, but does not increase the amount of bandwidth available per station. In a shared-bandwidth LAN all transmitted frames pass by all stations, even those frames not destined for a particular station.

By using parallel switching techniques in an intelligent concentrator, the amount of bandwidth per station can be increased. For example, a cross-bar switch can allow dedicated frame-by-frame connections between switch ports.

Note: When a frame is received at a port on a switch, it is automatically switched "on the fly" to the appropriate port for output based on the destination address contained in the frame.

A table at each switch port determines which port a given destination address should be switched to. Thus, each LAN segment attached to a switch port receives only the frames with destination addresses on that LAN segment. A switch port can also contain frame buffering to temporarily "hold" frames that cannot be immediately forwarded. For example, in an Ethernet LAN if the destination Ethernet segment is currently busy, a switched frame would be temporarily stored in the switch port buffers. Figure 37 on page 112 shows a four-port cross-bar switch.

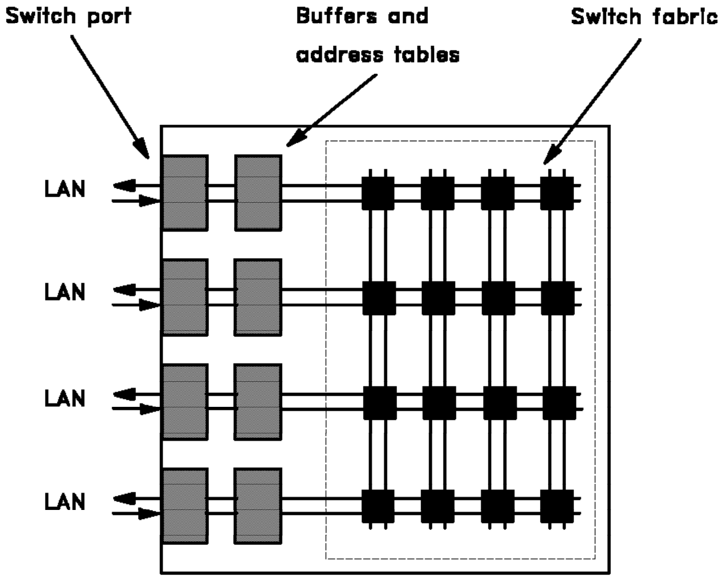


Figure 37. A Four-Port Cross-Bar Switch

Each switch port contains a media access control (MAC) unit to support attachment to a LAN segment. Attached to each switch port can be a dedicated-media single station or shared-media LAN segment with multiple stations.

With an Ethernet hub of the design shown in Figure 37 the presence of the buffers creates separate collision domains. In the example shown, there are four separate collision domains, where the span of each is governed by the normal four-repeater rule and 64-byte slot time. As far as Ethernet design rules are concerned, the device shown is a bridge.

All LAN adapters support a MAC protocol. The purpose of the protocol is to regulate access to a common media (for example, via tokens as in token-ring and CSMA/CD for Ethernet). In a switched LAN with a single link per station, the requirements for a MAC protocol are greatly minimized. No longer must a station “contend” with other stations for permission to transmit onto the media. Thus, in a switched LAN, stations can transmit whenever a frame is queued in the adapter. A station can also receive at any time. Full-duplex operation is now possible; a station can transmit and receive independently of other stations on the switched LAN. Full-duplex operation has been proposed for Ethernet in IEEE 802.3 and for token-ring in IEEE 802.5.

Full-duplex offers the following benefits over half-duplex operation:

- Increases the bandwidth of switch-attached single-station LAN segments
- Provides migration to higher bandwidths as existing and emerging applications require additional bandwidth
- Capitalizes on emerging LAN switching technology
- Positions the desktop personal computer or workstation for the future high-speed enterprise ATM backbone

- Preserves the investment in existing LAN adapters and infrastructure (for example, cabling, software applications, and network management platforms and protocols)

Port Switching: The term *port switching* is used in the context of IBM 8250 and 8260 hubs. These products support many LAN segments of the same or differing type within the same chassis. Port switching describes the process whereby a user device is physically moved from one of the backplane segments to another, without re-cabling at the hub.

An example of this switching technique is in the 8260 Token-Ring Active Port Switching card, where any of the 18 ports on the card can be independently assigned to one of the ten ShuntBus (backplane) rings.

In this context, port switching is a LAN segmentation tool rather than a segment interconnect tool.

3.10 Logical Link Control/802.2

The IEEE 802.2 standard describes the top sublayer of the data link layer. It is common to all the MAC sub-layers defined by the IEEE. This means higher layer protocols are shielded from the peculiarities of the physical medium as well as from the specific medium access protocol being used. Obviously, limitations such as throughput characteristics and number of attachments, inherent in each of the lower level standard protocols, apply equally to higher layer protocols. Local area networks (as well as wide area networks) need the function of a DLC layer to optimize the capacity, accuracy, and availability of the physical medium.

The function of a network layer, to route data through the network and set up a connection between two endpoints, is also needed in a LAN. Basic LAN routing, error control, and flow control is defined as part of the LLC sub-layer while the network layer may be a null layer.

In summary, logical link control provides a consistent view of a LAN to the upper layers regardless of the media and protocols being used and may support all required end-to-end connectivity services.

3.10.1 LLC Concepts

The interface to the upper layers is provided through LLC service access points (LSAP), just as service access points (SAPs) are the designed interfaces between any two adjacent layers in the OSI Reference Model as shown in Figure 1 on page 4. A station can have more than one SAP associated with it for a specific layer, just as a station may have more than one session active via one SAP. Any link level connection of one station to another is known as a link, while at each end of a link, associated control support is referred to as a link station. In the remainder of this section, an LLC service access point will be referred to as a SAP.

For example, a LAN station may have an 802.2 LLC-level session through SAP X′04′ with SAP X′08′ in a token-ring attached 3174-01L, and concurrently be in session with a file server in another LAN station through SAP X′F0′. Logical link control provides the capability to manage these independent sessions.

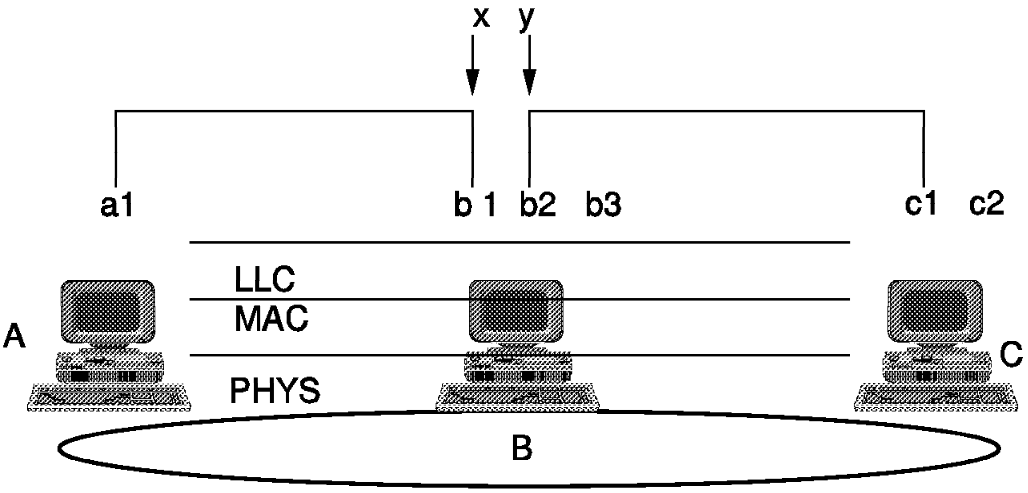


Figure 38. SAPs and LLC Connections

In Figure 38 application x is using SAP b1 to communicate with a resource in station A. Assume that this resource is a printer. A connection is set up between SAP b1 and a1. Once this connection is established, output can be sent to the printer. At the same time, application y, using SAP b2, can have a connection with a resource in station C using SAP c1. The use of service access points allows the applications and upper layer protocols to concurrently access the adapter(s) and media of the LAN.

Another way to describe an LLC SAP is to define it as a designed code point between DLC and the upper layers identifying an application to the LLC sub-layer. The IEEE standard defines a number of SAP addresses, while the IBM Token-Ring network architecture uses a number of user-reserved SAP addresses to interface with IBM proprietary protocols. Before detailing the architected SAP values in use on IEEE LANs, it is necessary to describe bit ordering techniques. These have equal applicability to MAC addresses, in fact, one of the easiest places to see the various techniques is to trace TCP/IP address resolution protocols.

3.10.1.1 Canonical Bit Ordering

In the documentation produced by the standards bodies, a different method of bit ordering is used to that commonly accepted in the IBM world. The standard way of describing the order of bits in a byte is known as canonical bit ordering and is the reverse of the way IBM usually writes binary values. Depending on whose documentation is being studied, and when trying to relate the sets of documentation to each other, confusion may result. Also, when tracing a token-ring, a field containing a station address field may be present within the data. This field may be in canonical bit order, and has to be translated into the

IBM order if sense is to be made of it. Hence an explanation of this technique here. For example:

The decimal number 71 has a hexadecimal value of 47. IBM would write this value, in binary, as:

```
0100 0111
```

In canonical form, this same binary value would be read right to left, so as to look like:

```
1110 0010
```

If we take a much longer bit stream, such as a station MAC address, then:

```
   1    0    0    0    5    A    1    7    5    8    6    9

0001 0000 0000 0000 0101 1010 0001 0111 0101 1000 0110 1001


Canonically

0000 1000 0000 0000 0101 1010 1110 1000 0001 1010 1001 0110

0    8    0    0    5    A    E    8    1    A    9    6
```

The order of the bytes remains the same, but each byte has its binary digits read from right to left when converting from the IBM form to the canonical form.

Canonical bit ordering has particular relevance when looking at SAP values, and the reasons why SAP values are numbered the way they are.

3.10.1.2 SAP Values

The structure of the format of the source SAP and destination SAP fields within the LLC protocol data unit is shown in Figure 39 on page 116.

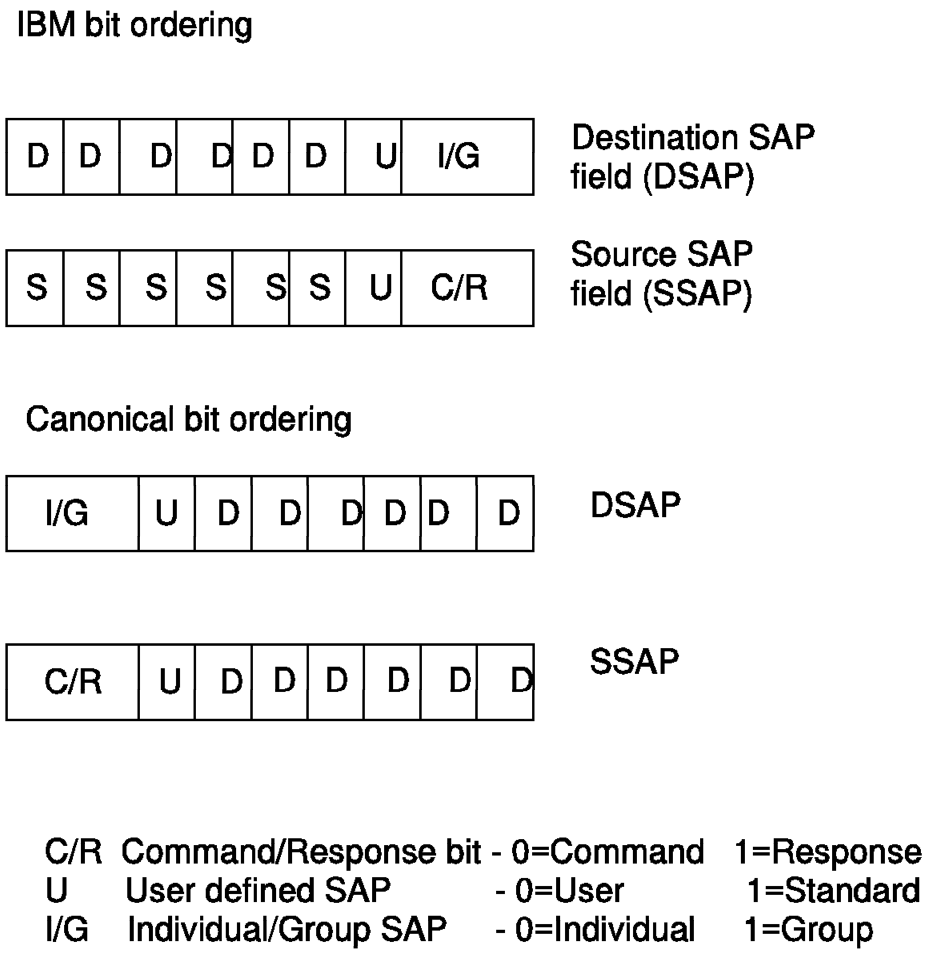

IBM bit ordering
D D D D D D U I/G
Destination SAP field (DSAP)
S S S S S S U C/R
Source SAP field (SSAP)
Canonical bit ordering
I/G U D D D D D D
DSAP
C/R U D D D D D D
SSAP
C/R Command/Response bit - 0=Command 1=Response
U User defined SAP - 0=User 1=Standard
I/G Individual/Group SAP - 0=Individual 1=Group

Figure 39. SAP Fields - Bit Ordering

Since the C/R bit of the SSAP is not used as a bit within a SAP address, there are in fact only 128 possible SSAP values. A group SAP can only be used as a destination SAP, and is a special case of one of the 128 possible SAP values. These 128 SAPs are then divided into two groups, each with 64 values. If the bits of the SAP are shown in canonical order, binary values ′100 0000′ to ′111 1111′ are reserved for IEEE use. SAPs ′000 0000′ to ′011 1111′ are for users. The pattern of SAP values can now be more easily seen, than if they were represented in IBM bit order. Whether the group SAP bit is on or not, makes no difference to the allocation.

3.10.1.3 LSAP Addresses

The LSAP is a 1-byte field in the LLC header of an 802 LAN frame that is used to multiplex data to the higher layer applications or protocol stacks within the same box. The 8th bit of the LSAP is the I/G (Individual vs group) indicator, leaving 7 bits for a unique SAP. The SAPs shown in the following tables are reserved by the International Standards bodies.

Table 14. LSAP Reserved Addresses

Bits	**HEX**	**Description**
0000 000X	00,01	Null SAP
bbbb 001x	X2,X3	Net Management (Standards)
nnnn 011x	x6,x7	General Standards
pppp 101x	xA,xB	General Standards
qqqq 111x	xE,xF	General Standards
1111 111x	FE,FF	All Stations Address
0000 0110	02	IEEE 802.1B LAN Management
0000 0110	06	Department of Defense Internet Protocol (IP)
0000 1110	0E	PROWAY Net Management and Initialization
0100 1110	4E	Layer messaging service for factory automation
0100 0010	42	MAC Bridges BPDUs
0111 1110	7E	ISO 8208 X.25 packet layer protocol
1000 1110	8E	PROWAY active station list maintenance
1010 1010	AA	Sub-Networking Access protocol (SNAP)
1111 1110	FE	OSI Network Layer Protocol

Table 15 (Page 1 of 3). LSAP User-Defined Individual Addresses

Bits	**HEX**	**Description**
0001 000x	10	Novell NetWare
0010 000x	20	
0011 000x	30	
0100 000x	40	
0101 000x	50	
0110 000x	60	
0111 000x	70	
1000 000x	80	User-defined - XNS
1001 000x	90	user defined
1010 000x	A0	OEM res.
1011 000x	B0	OEM res.
1100 000x	C0	IBM res.
1101 000x	D0	IBM res.
1110 000x	E0	Novell NetWare
1111 000x	F0	PC Network NetBIOS
0000 010x	04	SNA Path Control
0001 010x	14	
0010 010x	24	
0011 010x	34	
0100 010x	44	
0101 010x	54	

Table 15 (Page 2 of 3). LSAP User-Defined Individual Addresses		
Bits	**HEX**	**Description**
0110 010x	64	
0111 010x	74	
1000 010x	84	User-defined
1001 010x	94	User-defined
1010 010x	A4	OEM res.
1011 010x	B4	OEM res.
1100 010x	C4	IBM res.
1101 010x	D4	Resource Manager (8230 Discovery)
1110 010x	E4	PC Network Net Management
1111 010x	F4	LAN Network Management
0001 000x	08	SNA
0001 100x	18	
0010 100x	28	
0011 100x	38	
0100 100x	48	
0101 100x	58	
0110 100x	68	
0111 100x	78	
1000 100x	88	User-defined
1001 100x	98	User-defined
1010 100x	A8	OEM res.
1011 100x	B8	OEM res.
1100 100x	C8	IBM res.
1101 100x	D8	IBM res.
1110 100x	E8	Terminal Controller Comms for Retail POS
1111 100x	F8	Remote Program Load, Initial microprogram load (IMPL)
000 110x	0C	SNA
0001 110x	1C	
0010 110x	2C	
0011 110x	3c	
0110 110x	4C	
0101 110x	5C	
0110 110x	6C	
0111 110x	7C	
1000 110x	8C	User-defined
1001 110x	9C	User-defined
1010 110x	AC	OEM res.
1011 110x	BC	OEM res. VINES IP
1100 110x	CC	IBM res.

Table 15 (Page 3 of 3). LSAP User-Defined Individual Addresses

Bits	HEX	Description
1101 110x	DC	Address Discovery Protocol
1110 110x	EC	IBM res.
1111 110x	FC	GDS RPL Discovery
Note: It is important to remember that all of the addresses above are user defined in the industry.		

In the future, large multiprotocol environments will converge to standard protocol stacks and will therefore use industry standard LSAPs.

3.10.1.4 Types of Operation

To satisfy the broad range of potential applications, IEEE 802.2 logical link control defines different types of operation. These types of operations, each representing a number of link-level services, are combined in several service classes.

- **Type 1: Connectionless Operation**

  The connectionless services are also referred to as user datagram mode of operation. In this mode, there is no data link connection establishment between the SAPs of the end stations before actual information frame transmission starts. At the LLC level there is no guaranteed delivery of frames to the destination station. There is no correlation between frames in a particular data transmission, and the sender does not expect an LLC-level acknowledgment. In this mode of operation, higher layer services must provide recovery and frame sequencing.

- **Type 2: Connection-Oriented Operation**

  Prior to data exchange, a data link connection must be established, for which an LLC Type 2 control block has been defined. These control blocks together with the associated delivery and error recovery services are called link stations.

  The line protocol that is maintained on an established link between end stations is HDLC asynchronous balanced mode extended. In this way reliable end-to-end service for high traffic rates is provided.

  Essentially, the LLC Type 2 services may be summarized as:
  - Sequence numbering of data frames at the data link layer
  - Error detection and basic recovery
  - Flow control

  This also includes an LLC-level acknowledgment, for which a window size of up to 127 outstanding frames may be sent before expecting acknowledgment from the destination station.

- **Type 3: Acknowledged Connectionless Operation**

  This third mode of operation has been proposed as an enhancement to the international standard. In this mode, there is no connection establishment prior to data exchange, as in datagram operation. But link-level acknowledgment, consistent with the window size in use, is expected by the sending station.

  Type 3 operation seems particularly appropriate for LANs that experience traffic patterns with unpredictable bursts in message rates, such as a LAN with file servers or backbone traffic.

Currently two classes of LLC operation are defined within the IEEE 802.2 standard based upon the above Type 1, Type 2, and Type 3 services:

- Class I - Has Type 1 service only.
- Class II - Has Type 1 and Type 2 service.

However, as more types of LLC service are defined, there is a move to drop the concept of classes, due to the greater number of combinations of service, and hence classes, that would be created.

3.10.1.5 LLC Protocol Data Unit

Figure 40 shows the format of a logical link control protocol data unit (LPDU) and the code bits for identifying each of the LLC services for both connectionless and connection-oriented modes of operation.

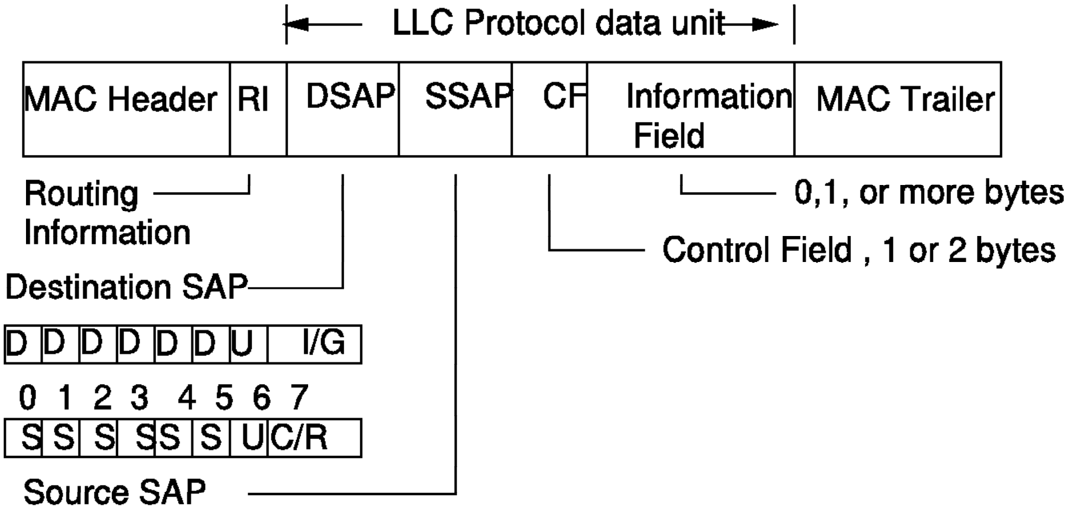


Figure 40. IEEE 802.2 LLC Protocol Data Unit (LPDU) Format

Each LPDU consists of a destination and source SAP address field, a control field and an information field of zero or more bytes. The following is a list of the abbreviations used in this figure:

- D = Destination SAP bits
- S = Source SAP bits
- U = User-defined address bit (B′1′ if IEEE defined)
- I/G = Individual (B′0′) or group (B′1′) SAP address
- C/R = Command (B′0′) or response (B′1′) LPDU
- F = Supervisory function bit

DSAP Field:

The destination service access point (DSAP) address field identifies one or more service access points for which the information is intended. This eight-bit field contains one bit to identify whether the address is an individual or a group address, and seven address bits. Two group addresses have specific IEEE standard definitions:

- Global address: If the DSAP field contains B′11111111′, it is a global address. The frame is addressed to all SAPs.
- Null address: If the DSAP field contains B′00000000′, it is a null SAP address, applicable only to connectionless operation. It supports response to a TEST or XID command in those situations where no SAP has yet been activated in the remote node.

SSAP Field:

The source service access point (SSAP) address field identifies the service access point that sent the frame. This eight-bit field contains a bit to identify whether the frame is a command or a response, and seven address bits. The command/response bit is not considered as part of the source SAP address in a received frame. The low-order bit in the seven address bits of both DSAP and SSAP indicate an IEEE defined SAP address when set to B′1′.

Control Field:

The control field is a one- or two-byte field used to designate command and response functions. It contains sequence numbers when required. The control field is very similar to the HDLC control field.

Information Field:

The information field contains zero or more bytes of information, depending on the particular LLC service contained in the control field.

An LLC PDU is invalid under the following conditions:

- It is identified by the physical layer or MAC sublayer as being invalid.
- The frame is not an integral number of bytes in length.
- It does not contain properly formatted address fields.
- The frame contains mandatory fields that are out of order.
- The frame length is less than three bits for a one-byte control field, or less than four bits for a two-byte control field.

Invalid protocol data units are ignored.

3.11 Summary

IBM participates in most of the major LAN standard-making organizations , including IEEE, ISO and ANSI, at both the international and local levels. IBM has provided many submissions in the area of LAN standardization, and has taken steps to ensure that IBM LAN products support or closely follow these standards. IBM′s active role in standards activity is particularly demonstrated by the SRT Bridge suggestions to IEEE 802.1D and the wireless and ATM Forum. IBM participates in the Manufacturing Automation Protocol (MAP) project, and has been active in the development of the FDDI standard, particularly within the area of the complex protocols of station management.

IBM Ethernet products conform to the IEEE 802.3 standard as well. The IBM Token-Ring network fully conforms to (and enhances) the IEEE 802.5 and 802.2 standards. IBM has participated in and supported the following enhancements to the IEEE 802.5 specifications:

- 16 Mbps as a standardized transmission speed
- Early token release as an option
- Source routing bridge protocols

These proposals have been included in the the international standard. IBM devices attach to IEEE 802.3, IEEE 802.4 and IEEE 802.5 networks, and implement IEEE 802.2 protocols.

IBM is supporting ISO management on its LANs, with the IBM LAN Station Manager product, and using the OSI CMIP protocol to convey management information. IBM is supporting international standards, and Heterogeneous LAN Management (HLM), a joint project between IBM and 3Com, which means that a managed LAN environment, no matter on what topology or technology it is based, could become a practical possibility. IBM is also very active in the Internet Engineering Task Force (IETF) which supports standards like frame relay, media and bridge management and TCP/IP.

Chapter 4. LAN Protocols

The purpose of this chapter is to discuss the major protocols that make up the LAN environment. These include transport, multiprotocol transport, router and management protocols.

Transport protocols allow data to be sent and received from end station to end station (for example APPN, TCP/IP, NetBIOS). Information that is wrapped around this data allows the intermediate systems to determine where the data is going. From the end user′s point of view, the transport protocols supported are the most important criteria as the end user made an investment in the applications best suited to meet their current business needs without the limitation of requiring all applications on one transport protocol base. Routers must be concerned with transport protocols due to the different implementations of the routing functions.

Multiprotocol Transport Networking (MPTN) enables applications using a certain transport protocol to communicate via a different transport protocol. For example, an application using TCP/IP is able to communicate over NetBIOS. MPTN makes it possible to choose applications regardless of the transport protocol used by the network.

Router-to-router protocols exchange routing information between intermediate systems in a domain as well as reachability information with other routing domains. The important criteria for the end user′s decision is the cost of these protocols in memory, CPU, and line usage when maintaining the route tables.

Management protocols allow error or failure information to be passed from router to router. Management protocols for the transport protocols and physical transport are also available and must be supported by the router.

4.1 Major Transport Network Protocols

Transport protocols allow data to be sent and received between end stations. They can be categorized as proprietary (vendor specific) or standard based. Transport protocols are layered by function and can be divided into two categories at the network layer:

- Connection-oriented
- Connectionless

Transport protocols support a wide array of functions:

- Reliable transport
- Alternate routing
- Broadcast
- Flow control
- Reachability
- Security

Proprietary versus Standards - Each has its advantages and its disadvantages. Proprietary may lock you into one vendor but also gives the customer some leverage as it is only one entity to persuade when a function is needed by the end user. There is also the controlled environment for assuring interoperability and management. Standards give the end user the theoretical ability to buy from multiple vendors as there is a common protocol to be implemented. However, given the current state of standards, there are many variables which may inhibit true interoperability and robust management.

Connection versus Connectionless protocols - These also have a set of advantages and disadvantages. Connection-oriented protocols have the overhead of keep alive messages and no route dynamics but give the system administrator the ability to understand the WAN resource requirements and maintain the system as connections remain stable. Connectionless protocols provide dynamics during routing problems but usually require a higher bandwidth to support their bursty nature. Fault determination is also more complex given route changes. Flow control is much harder given no consistent path.

Some of today′s major networking protocols, which will be reviewed in this section in this same order, are:

- OSI
- IBM SNA and APPN
- TCP/IP
- Novell IPX/SPX
- NetBIOS
- Apple Computer AppleTalk
- Digital DECnet
- Xerox XNS
- Banyan Vines

4.1.1 Open Systems Interconnection (OSI)

Although we discussed the OSI Reference Model in the first chapter, it is discussed again in this chapter, because of the importance it plays in both network interconnections and internetworking of applications. What should be highlighted is that:

- Each layer performs a unique, generic, and well-defined function.
- Layer boundaries are designed so that the amount of information flowing between any two adjacent layers is minimized. This is accomplished by having each layer within an open system use the services provided by the layer below. Conversely, each layer provides a sufficient number of services to the layer immediately above it.

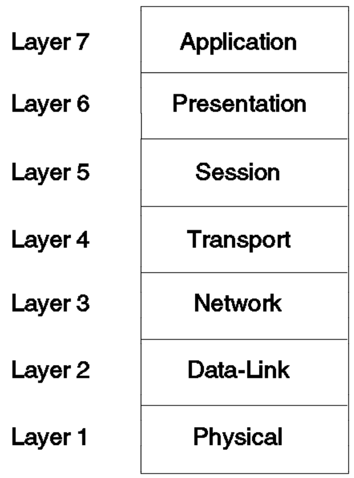


Figure 41. The OSI Reference Model

Below is a brief description of each of the layers:

- **Physical Layer** (layer 1)

  The physical layer describes the electromechanical characteristics for attachment of the open system to the physical medium. These include, for example, the definition of plugs and sockets, definitions of voltage levels and signaling rates, and the mechanism or technique for encoding of data.

- **Data Link Layer** (layer 2)

  The data link layer is responsible for data transfer between adjacent systems. It detects and possibly corrects errors that may occur in the physical layer. It may also provide flow control function.

- **Network Layer** (layer 3)

  The network layer provides the addressing, routing and relay information required to control the data-flow between source and destination end systems, and/or among multiple intermediate systems. Controlling the data flow involves establishing, maintaining, and terminating connections between these systems. Services that keep the transport layer independent of the data transmission technology are included. Multiplexing, segmenting, and blocking may also be involved. The network layer also provides internetworking for concatenated subnetworks, such as LAN′s, X.25 subnetworks, and ISDN subnetworks.

- **Transport Layer** (layer 4)

  The transport layer provides the end-to-end control of data exchanged between two open systems, which includes:

  - Establishing and releasing connections between two end systems
  - Receiving quality-of-service requirements from the session and negotiating them with the network-layer
  - Deciding whether to multiplex several connections on a single transport connection
  - Providing end-to-end flow control

  - Arranging the use of expedited data units
  - Establishing the optimum data-unit size
  - Segmenting single data units into multiple data-units and the reverse
  - Optionally concatenating several data-units into a single data-unit
  - Providing end-to-end sequencing, error detection and error recovery

  Depending on user needs and the lower layers′ capabilities, it provides the required level of reliability of the underlying transmission subsystem. It is this layer that controls the size, sequence, and flow of transport packets and the mapping of transport and network addresses. In addition, it enciphers data if security is needed.

  While the session layer (layer 5) is mainly under the control of the application, the transport layer (layer 4) exchanges primarily flows, signals, and operations with the transport facilities beneath.

- **Session Layer** (layer 5)

  The session layer coordinates the dialog between applications on two open systems by providing functions for negotiating, establishing, controlling, and releasing sessions between these applications. It provides mappings of session-to-transport and session-to-application connections between two cooperating applications.

- **Presentation Layer** (layer 6)

  The presentation layer provides a common representation or format to be used between application entities. This relieves application entities of any concern with the problem of ″common″ representation of information, that is, it provides them with syntax independence. It performs encoding and decoding between abstract syntax, which is written in ASN.1 (Abstract Syntax Notation One), and transfer syntax, which is encoded by means of the basic encoding rule for ASN.1. This is necessary because different vendors′ systems use various methods for the internal representation of data.

- **Application Layer** (layer 7)

  The application layer provides services to the actual applications to accomplish information transfer. This involves supporting and managing the communication between the end users on connected open systems, which includes not only the exchanges of data but also security checks and requesting other specific application services.

In reality, many implementations of OSI do not include a full seven-layer stack. Instead, a “short stack” of the lower three layers is used to avoid the cost of implementing and using the upper layers. For example, routers only use the lower three layers.

4.1.1.1 Peer-to-Peer Communications

The OSI Reference Model describes a completely peer-to-peer communications environment. A layer entity of one open system communicates through a set of protocols with its peer entity in the corresponding open system. A layer entity represents one of the seven layers within a given open system.

A given layer entity of Open System 1 communicates with its corresponding entity in the same layer of Open System 2. Both entities use the services of the layers below. This is illustrated in Figure 42 on page 127.

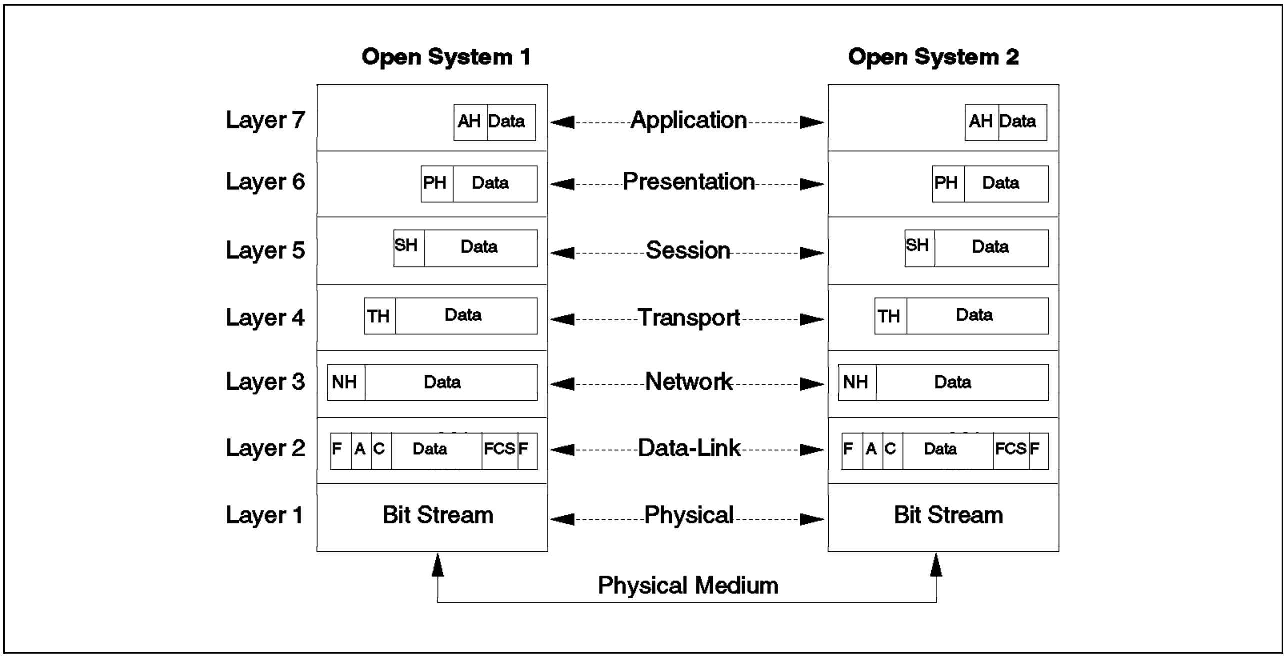


Figure 42. The OSI Peer-to-Peer Communication

To accomplish this logical peer-to-peer communication, each layer within an open system uses the services provided by the layer below, and then in turn, it provides services to the layer above. The layer providing the services to the layer above is called the *service provider*, and the layer using the services of the layer below it is called the *service user*.

Each layer accomplishes its function:

- By adding a header to the received data from the upper layer, or
- By providing data received from another open system to the above layer after handling it according to the protocol information in the header

As illustrated in Figure 42, each lower layer adds or removes the header information depending on the direction of data flow. The header contains the layer-dependent protocol to communicate with another open system using the same protocol.

For example, let′s assume that an application entity in Open System 1 wants to communicate with its peer entity, an application entity in Open System 2. To do this, it uses the services of its presentation layer entity (its service provider) for negotiating and defining the kind of presentation. The presentation layer entity, in turn, uses the services provided by its session layer entity. This process, referred to as adjacent layer communications, continues down through all layers in the open system.

4.1.1.2 OSI Connection-Oriented/Connectionless Modes

All protocols defined in the transport layer (layer 4) operate only between OSI end systems. This layer can be connection-oriented or connectionless, but usually it performs in the connection-oriented mode.

In the connection-oriented mode, the transport service provides the means to establish, maintain, and release transport connections, and is, therefore, a reliable and sequenced message flow between a pair of applications. This mode is appropriate in applications that call for relatively long-lived interactions between entities in stable configurations.

In the connectionless mode, datagrams (individual data transfers) are sent with no relationship one to another. There is no requirement to maintain the sequence of message flow, no guarantee of reliable delivery, no negotiation of options that govern the transport, and clearly no distinguishable connection. This mode might be associated with particular forms of data transmission, such as LANs, or digital radio, and particular types of applications, such as remote sensing.

4.1.1.3 OSI Adjacent Layer Communication

Figure 43 shows the series of interactions between adjacent layers. The service user in Open System 1 requests a service 1 from a specific service provider by using a service access point (SAP). A particular layer can provide more than one service to different service users. The service access point, which is a unique interface between two adjacent layers, allows the correct service user to be selected. The service user, in turn, is able to find the corresponding service provider. The service provider entities communicate and exchange data causing the service provider in Open System 2 to indicate 2 to the desired service user in Open System 2 that a service has been initiated.

Depending on the type of service, the service user in Open System 2 may send a response 3 indicating that an action has been taken on the information in the indication. The service providers again communicate and the service provider in Open System 1 will confirm 4 to its service user the action taken on the initial request.

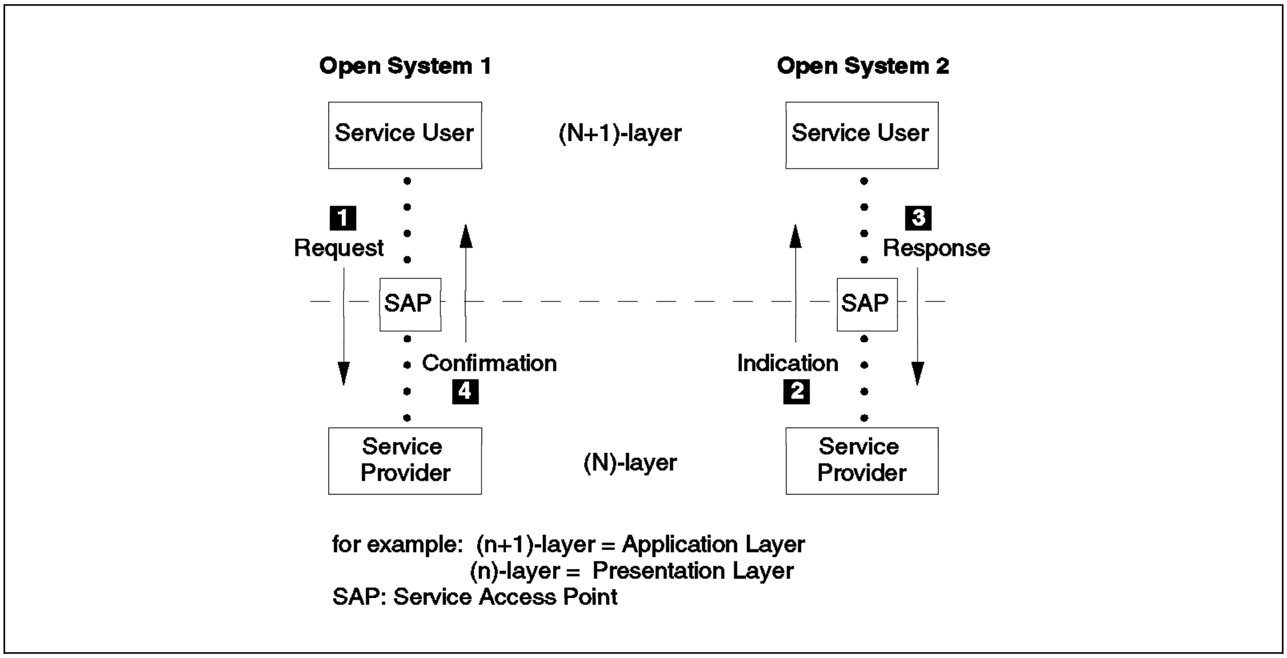


Figure 43. The OSI Adjacent Layer Communication

Layer N uses the services of layer (N-1) to communicate with its peer. This continues on down through the layers. At the lowest level, the physical layer entity in Open System 1 sends the data to its peer, the physical layer entity in Open System 2. When the data is received by the physical layer entity in Open

System 2, it passes it to the data link layer entity. This continues up through the layers in Open System 2 until the application layer entity receives the data.

As previously mentioned, adjacent layer communications use the service access point (SAP) to exchange data and to find the correct service user. A SAP is an internal address that identifies where a service is made available to a user. Each SAP has a *service access point identifier (SAP-ID)*. Every connection references a unique SAP, which allows multiple communication paths between adjacent layers to exist in parallel. Multiple simultaneous connections may also be established on the same SAP as well. These different connections are identified by their connection IDs.

SAPs are prefaced with the first letter of the service provider, for example, NSAP for the network layer, and TSAP for the transport layer. The PSAP (presentation service access point) is the point at which the services of the presentation entity are provided to the application layer using the presentation services.

4.1.1.4 OSI Profiles

Each layer in the OSI Reference Model has a name, a number, and a set of protocols that provides specific functions for defined services. Since the intended range and scope of OSI is very broad, each layer contains a multiplicity of protocols, and each protocol has a multiplicity of options. Therefore, it is possible, and *likely*, that two vendors could implement OSI protocols yet not be able to communicate with each other.

To ensure and maximize interoperability, multiple standards-oriented groups have identified specific OSI standards, protocols, and options for their environment. The documents provided by these groups are called OSI *profiles*. The profiles specify which standard, protocols and options within the set of standards must be implemented. An example of a profile is GOSIP (Government OSI Profile), which standardizes addressing formats and routing procedures in the United States. Currently, the most common instances of user organizations producing profiles are those produced by government IT functions such as NIST (National Institute of Standards and Technology) which is part of the U.S. Department of Commerce and GCTA (Government Centre for Information Systems) in the UK.

4.1.1.5 OSI Summary

OSI has been somewhat slow in gaining acceptance. Standards have taken years to reach full maturity and many areas still require extensive work. Many customers have made large investments in other technologies and continue to generate a high demand for compatible and familiar technologies.

One of the major reasons why customers are not moving to OSI is that OSI requires new and expensive development work. This means that there are not many available products. Not many vendors are willing to undertake an expensive development project unless they are convinced that there will be a huge demand and that the pricing of the product will provide adequate returns.

Another reason, that was cited earlier, was that many of the standards are complicated and very difficult to implement. A related reason is that many existing non-OSI systems have substantial investments in hardware that cannot adequately support the newer more powerful OSI implementations. Still others question whether it is really necessary to go through all seven layers and that this could result in poor performance.

Whatever the reason, customers are generally not moving to OSI in one seven layer cutover. Instead, they express interest in:

- Running OSI higher layer services, such as X.400, FTAM, or RDA over other networking protocols. For example, X/Open has seen a requirement to run RDA over TCP.
- Replacing other networking protocols with OSI layers 1 through 4, and continuing to run existing applications as well as new OSI applications.

4.1.2 IBM SNA and APPN

During the 20 years since its announcement in 1974, IBM′s Systems Network Architecture (SNA) has undergone continual improvements. Originally, SNA was hierarchical, but gradually, this model was modified, culminating in 1980 when IBM SNA introduced LU6.2 (peer-to-peer) communications. In 1986, SNA was enriched by the Advanced Peer-to-Peer Networking (APPN) function. This new function was added and first made available on the S/36. In 1992, APPN was extended and made seamless with the traditional Subarea SNA. Figure 44 illustrates this evolution. Today, SNA protocols are found in many applications, including those that execute on LANs.

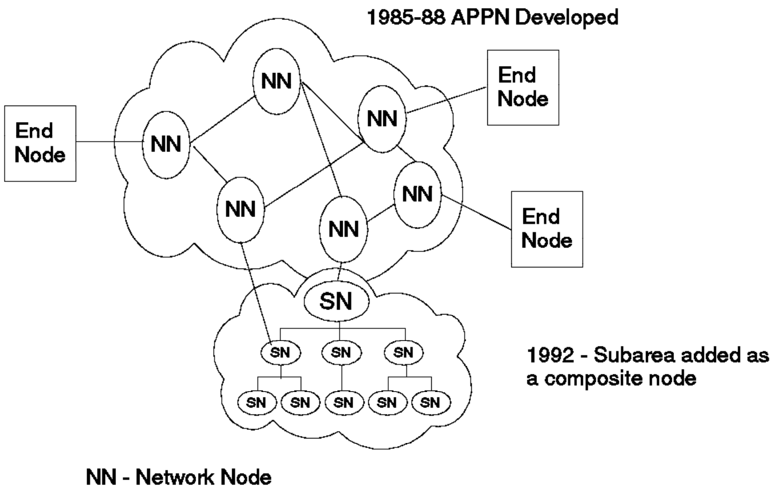


Figure 44. SNA and APPN Evolution

4.1.2.1 SNA Functional Layers

As Figure 45 on page 131 illustrates, SNA is also a seven-layered architecture, with each layer making use of the services of the layer below, and providing services to the layer above.

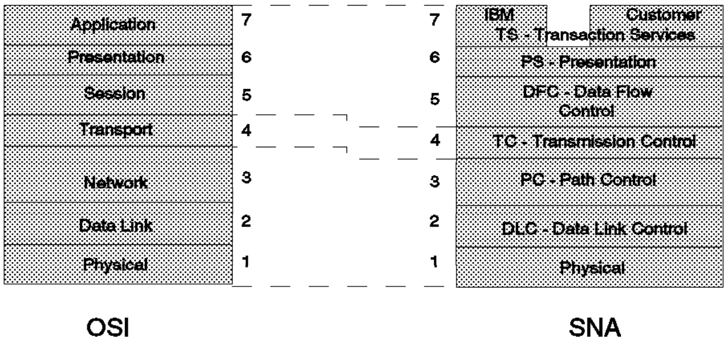


Figure 45. SNA Functional Layers

Although SNA architecture is continuously evolving to meet changing needs, the layer structures provide an important reference model for coordinated efforts. Basic functions of each of the seven layers are very similar to the OSI specifications, even though the details of the functions differ and the distribution of functions among the layers is different. SNA key functions are:

- **Physical (layer 1)**

  The Physical layer involves the physical plugs and associated electrical signals to provide a transparent transmission of any bit stream.

- **Data Link Control (layer 2)**

  This layer consists of the functions in link stations that schedule and initialize data transfer over a link between two nodes and performs error control, detection, and correction for the link. It also performs link disconnections. Examples of connections include point-to-point connections, multipoint, party-line connections, and local area networks.

- **Path Control (layer 3)**

  This layer manages the sharing of link resources of the SNA network and routes message units through it. This layer is responsible to select the network routing, provide class of service, segment and block messages between logical units (LUs) in the network, and resolve any path information conversions between them. (For a description of LUs and other SNA terms, please refer to 4.1.2.2, “SNA Basic Components” on page 133.)

- **Transmission Control (layer 4)**

  This layer synchronizes and controls the speed of session-level data traffic, checks sequence numbers of requests, and enciphers and deciphers end-user data, providing security if needed. It performs pacing of data exchanges and correlates buffer and processing capabilities of different node types.

- **Data Flow Control (layer 5)**

  This layer synchronizes the data exchange in an orderly interchange, and regulates the user′s send and receive flows. It generates and assigns sequence numbers and correlates requests and responses. This layer is primarily under the control of the application.

- **Presentation Services (layer 6)**

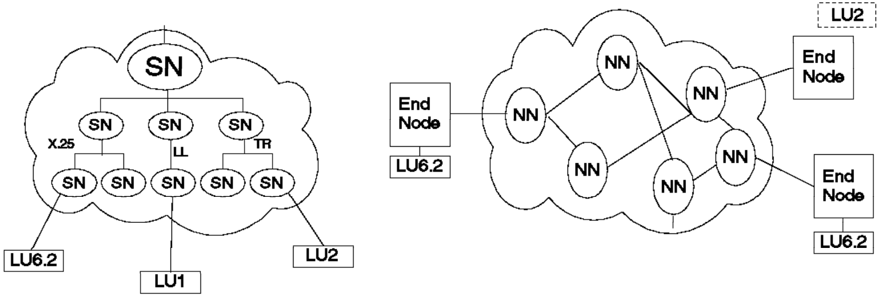
This layer provides services for transaction programs, such as controlling conversation-level communication between them, which may include converting data between different syntax and formats of the communicating end users, which could be a program or device. Therefore, it formats data for presentation (image or hard-copy, 3270, SNA character string, general data stream) and is then responsible for data representation between end users. It resolves network addresses and names, selects session profiles, performs sync-point processing, and offers network services, such as configuration and session management, maintenance, and measurement. It is also concerned with the interpretation of verbs received from APIs.

- **Transaction Services (layer 7)**

  This layer, also known as Application Services, controls such functions as the establishment of facilities for program-to-program communications, configuration activation services, directory services, distributed data services, messaging services, and document interchange. These are designed services and include common application services such as file transfer, store-and-forward messaging services, such as electronic mail, and remote database/file access.

All of the layers need to be part of a composite routing process. For these reasons, many of the layers have one or more of the following functions:

- Logical connection establishment with its peer, to prepare for message exchange and arrange for needed resources at that layer
- Address handling, which may involve interpretation or translation of the address for use within that layer
- Message size manipulation, possibly involving blocking, deblocking, segmenting, and re-assembly of messages
- Error detection and recovery, when the performance of the lower layers requires surveillance and correction
- Flow control



Classical SNA allowed mixing of Data Links and LU's (layers 4, 5 & 6) Path Control (layer 3) was a constant.

Basic APPN only allows LU6.2 across the network. Thus, layers 3, 4, 5 & 6 are constants. Recent extensions allow other LU types to be supported.

Figure 46. SNA and APPN Mixing Layers

Figure 46 shows how SNA and APPN differ structurally and in their processing of data across the network.

4.1.2.2 SNA Basic Components

SNA specifies data communication environments in terms of a collection of uniquely addressable, functional entities, called *Network Accessible Units* (NAUs). NAUs implement the upper four layers of a node, which could be defined as a set of hardware and associated software components that implement the functions of the seven architectural layers. Although all seven layers are implemented within a given node, nodes can differ based on their architectural components and the sets of functional capabilities they implement. Nodes with different architectural components represent different *node types*. Four types of nodes exist: type 5 (T5), type 4 (T4), type 2.0 (T2.0), and type 2.1 (T2.1).

Nodes that perform different network functions are said to act in different *network roles*. It is possible for a given node type to act in multiple network roles. A T4 node, for example, can perform an interconnection role between nodes at different levels of the subarea network hierarchy, or between nodes in different subarea networks. The functions performed in these two roles are referred to as *boundary function* and *gateway function*, respectively. T2.1 and T5 nodes can also act in several different network roles. Node roles fall into two broad categories:

- Hierarchical roles
- Peer-oriented roles

Hierarchical roles are those in which certain nodes have a controlling or mediating function with respect to the actions of other nodes. Within such networks, nodes are categorized as either *subarea nodes* (SNs) or *peripheral nodes* (PNs). Subarea nodes provide services for and control over peripheral nodes and also provide intermediate routing functions. Networks consisting of subarea and peripheral nodes are also referred to as *subarea networks*.

- **Subarea nodes** (SNs)

  Type 5 (T5) and type 4 (T4) nodes can act as subarea nodes. T5 subarea nodes provide the SNA functions that control network resources, support transaction programs, support network operators, and provide end-user services. T5 subarea nodes can also route and control the flow of data. Because these functions are provided by host processors, T5 nodes are also referred to as *host nodes*.

  T4 subarea nodes also provide the SNA functions that route and control the flow of data in a subarea network. Because these functions are provided by communication controllers, T4 nodes are also referred to as *communication controller nodes*.

- **Peripheral nodes** (PNs)

  Type 2.0 (T2.0) and type 2.1 (T2.1) nodes can act as peripheral nodes attached to either T4 or T5 subarea nodes. Peripheral nodes are typically devices such as distributed processors, cluster controllers, or workstations. A T2.1 node differs from a T2.0 node by the T2.1 node′s ability to support peer-oriented protocols as well as the hierarchical protocols of a simple T2.0 node. A T2.0 node requires the mediation of a T5 node in order to communicate with any other node. (There is no direct communication between peer T2.0 nodes.) Subarea nodes to which peripheral nodes are attached perform a *boundary function* and act as subarea *boundary nodes*.

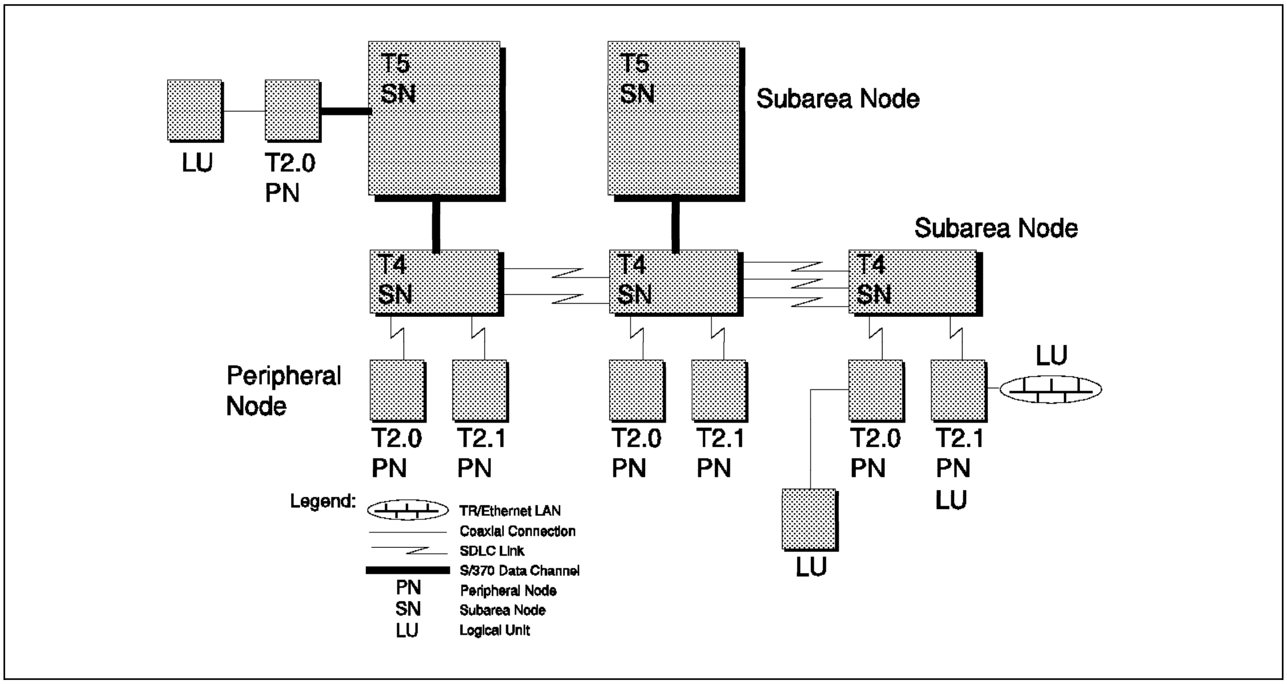


Figure 47. SNA Subarea Network

Figure 47 illustrates a subarea network containing the four node types acting as subarea and peripheral nodes. The network contains two type 5 subarea nodes, three type 4 subarea nodes, seven peripheral nodes, and two logical units (LUs).

Note: There is no architectural association between node type, or node role, and the kind of hardware that implements it, as a result Figure 47 uses symbols to represent node types and roles.

Peer-oriented roles enable nodes to communicate without requiring mediation by a T5 node, giving them increased connection flexibility. The Advanced Peer-to-Peer Networking (APPN) extensions to the T2.1 node allow greater distribution of network control by enhancing the dynamic capabilities of the node. T2.1 nodes with these extensions are referred to as *APPN nodes*, and a network of APPN nodes makes up an APPN network. A low-entry networking (LEN) node (which is just a type 2.1 node) can also attach to an APPN network. An APPN node can dynamically find the location of a partner node, place the location information in directories, compute potential routes to the partner, and select the best route from among those computed. These dynamic capabilities relieve network personnel from having to predefine those locations, directory entries, and routes. APPN nodes can include processors of varying sizes such as the AS/400, the PS/2 running under OS/2, and VTAM running under MVS/ESA. More will be discussed in 4.1.2.4, “Overview of Advanced Peer-to-Peer Networking” on page 136.

Certain components within the upper four architectural layers of a node use transport network services to establish temporary, logical connections with one another called *sessions*. Sessions can be established between two components residing in different nodes, or between components within the same node. Components that can establish sessions are referred to as *network addressable* *units* (NAUs) as was mentioned earlier in this section. NAUs in session with one another are referred to as *session partners*.

NAUs establish and control sessions in order to deliver control data and end-user data across the transport network. There are three types of NAUs:

- **Physical Units (PUs)**, which perform local node functions such as activating and deactivating links to adjacent nodes. PUs exist only in nodes within subarea networks. (In APPN networks, these functions are performed by control points.) To perform its functions, a PU must exchange control data with its controlling system services control point (SSCP) over an SSCP-PU session initiated by the SSCP.

- **Logical Units (LUs)**, which provide network access for end users by helping the end users send and receive data over the network. Nodes in both subarea and APPN networks contain LUs. LUs send and receive control data and end user data over the LU-LU sessions established between them. In a subarea network, an LU residing in a peripheral node is classified as either SSCP-dependent or SSCP-independent, depending on the protocols it uses for LU-LU session initiation. An SSCP-dependent LU sends a session-initiation request to its controlling SSCP over an SSCP-LU session, while an SSCP-independent LU sends a session-activation request directly to the partner LU. SSCP-dependent LUs (also known as dependent LUs) may reside in both T2.0 and T2.1 nodes, but SSCP-independent LUs (also known as independent LUs) may reside only in a T2.1 node.

  Over the years, IBM has defined several different LU session types: LU type 0, LU type 1, LU type 2, LU type 3, LU type 4, LU type 6.1, and LU type 6.2. The fundamental distinction between these types is the type of end-user pair served by each session type. The latest, which is LU type 6.2, is for transaction programs communicating in a distributed data processing environment. This type supports multiple concurrent sessions and can be used for communication between two type 5 nodes, a type 5 node and a type 2.1 node, or two type 2.1 nodes. Examples of the use of LU type 6.2 are:

  - An application program running under CICS communicating with another application program running under CICS

  - An application program in an AS/400 communicating with a PS/2

- **Control Points (CPs)**, which provide network control functions that include managing the resources in their domains and monitoring and reporting on the status of those resources. The functions of node CPs in subarea networks differ greatly from those in APPN networks. In subarea networks, the control point is referred to as a system services control point (SSCP). SSCPs control the PUs and dependent LUs in their domains by exchanging control data with them over SSCP-PU and SSCP-LU sessions respectively. They may also initiate SSCP-SSCP sessions for the control of cross-domain sessions (sessions that cross domain boundaries). In APPN networks, a control point in a type 2.1 node or a type 5 node is simply called a control point (CP). Like other control points, it can activate locally attached links, interact with a local operator, and manage local resources. It also initiates sessions with adjacent CPs, called CP-CP sessions, in order to exchange control data for its routing and directory services.

For an extensive explanation of SNA, including its architectural objectives, network components, link and node components, and the major functions they perform, please refer to *Systems Network Architecture, Technical Overview*, GC30-3073.

4.1.2.3 SNA Addressing

SNA addressing has evolved with the announcement of Type 2.1 and APPN. Originally, all SNA resources in a particular network were assigned in one of two ways:

- NCP generations (gens) that fixed a network address to a particular resource
- VTAM startup that allocated network addresses in order of activation

A *network accessible unit* in an SNA network could be a logical unit (LU), physical unit (PU), or system services control point (SSCP). Each has a unique network address, which consists of two parts:

- Subarea address, which is the same for all network accessible units in the same subarea node (T4 or T5)
- Element address, which is unique to each network accessible unit within that subarea node

Network addresses are used only within the SNA network. End users refer to network accessible units by their network names. Each network accessible unit within an SNA network must have a unique name. A network directory service is used to map the network names to their corresponding network addresses.

With the introduction of Type 2.1 node architecture, SNA addressing has evolved. Type 2.1 nodes do not assign network, or even local, addresses to LUs. Instead these nodes use an LU name-based session establishment scheme. Sessions are identified by a dynamically assigned session index value, which is placed in the Transmission Header (TH) field which once contained the destination and origin address.

With the announcement of APPN, addressing is even further simplified. With APPN, the LUs resident in an end node do not have to be defined to the network node within the APPN backbone network to which the end node is to be attached, as was the case with traditional SNA. The network node automatically queries the end node for the LUs resident in that node and then dynamically registers the LUs in its directory when a connection is first made between the end node and the network node. With APPN, a user can literally plug in a workstation or mini-computer into a port in an operational APPN network without the need for any predefinitions being done in the existing network. (The attaching workstation end node must only define its network node server.)

4.1.2.4 Overview of Advanced Peer-to-Peer Networking

Since APPN is the latest advancement in SNA, this short section highlights why APPN was developed and how it works.

As previously highlighted, APPN may be described as a network design in which ease of use and simplicity were achieved by:

- Decentralizing control so that each node maintains its own responsibility for membership in the network.
- Automating the processes of directory lookup, route finding, session bind, and congestion control in such a way as to make them as invisible as possible to users of the network.

- Using network control and data transport algorithms that minimize control message overhead and maximize the robustness of the connection between end users.

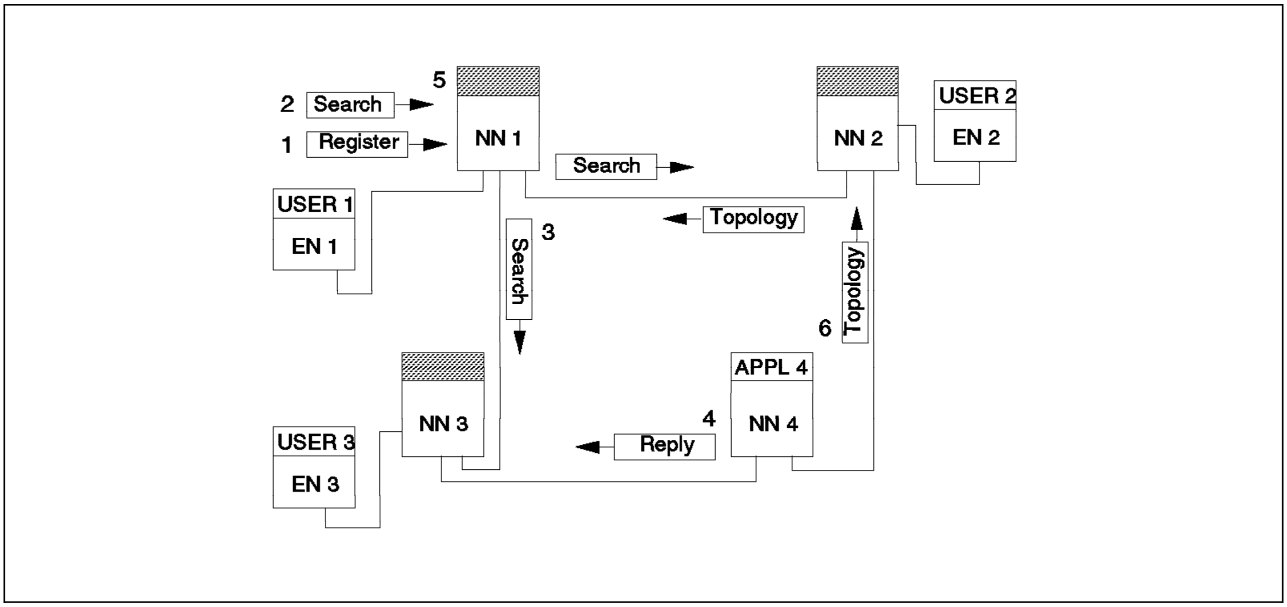


Figure 48. Peer-to-Peer Functions

Figure 48 illustrates an APPN network consisting of two parts:

- A set of network nodes (NN 1 to NN 4), each capable of intermediate session routing that collectively make up the backbone of the APPN network. Network nodes:

  - Locate network resources
  - Calculate routes
  - Maintain directory database of network resources
  - Maintain network topology database

  The topology database exists at each network node and contains a list of all of the network nodes and the links between adjacent nodes. The topology database at each network node is maintained to be exactly the same. In fact, the one network topology database is fully replicated in every network node. In this picture, the topology database at each network node will contain knowledge of the four network nodes (N1, N2, N3, N4) and the lines that interconnect them.

  A directory database also exists at each network node; it contains only the local resources that exist on that network node or any end node that uses that network node as its server. An end node may register its resources (LUs) with its network node server and the knowledge of the resource is kept in the directory database. In Figure 48, the directory database on network node 1 contains an entry for USER 1, a resource that resides on end node 1. This directory would contain any resources or LUs that actually reside on network node 1 as well.

- End nodes (EN 1 to EN 3) cannot perform the intermediate session routing function, but must have this done for them by a network node. An end node:

  - Uses directory and route calculation services of a network node

- May register local resources with its network node server
- Manages local links

Example network node and end node platforms include OS/2, AS/400, AIX/6000, 6611 Router, 3174, DPPX/370, ACF/VTAM, and third-party vendors.

Below is a step-by-step description of the flow outlined in Figure 48 on page 137:

1. End node EN 1 registers local resources, USER 1, with network node NN 1.
2. USER 1 requests a session with APPL 4. EN 1 sends the search request to its network node server NN 1.
3. NN 1 will locate the resource. Search logic may include:
   - Checking local directory
   - Sending search request to a central directory server
   - Sending broadcast search to adjacent network nodes, NNs (only if no central directory server exists)
4. NN 4 returns a positive reply back to NN 1.
5. By knowing the network topology, NN 1 can compute the best route from EN 1 to NN 4 and return it to EN 1.
6. After NN 1 passes this information to EN 1, EN 1 sends a BIND to start the session.

In addition to these basic functions, these are other actions to illustrate the capabilities of APPN:

1. NN 1 enters the location of APPL 4 into its directory database, since USER 1 may want to start a session with APPL 4 again later on.
2. NN 4 enters the location of USER 1 into its directory database. One of the concepts of APPN and APPC is that these applications are peers and any one of them can start a session, so even though USER 1 started the session this time, APPL 4 might start it next time so NN 4 should keep track of the location of USER 1.
3. Next time USER 1 wants to start a session with APPL 4, he requests the location from NN 1 (his or her network node), and this time when the network node looks into his or her directory database, he or she finds the location. USER 1 then sends a verify to make sure APPL 4 is still at that location; then, if the verify is successful, USER 1 sends the BIND.
4. Later, the link between NN 3 and NN 4 goes down. NN 4 notifies its adjacent network nodes, and they in turn notify their respective adjacent network nodes, so that each network node has up-to-date network topology knowledge.
5. Sometime later, APPL 4 is moved to a different end node. USER 1 wants to initiate a session with APPL 4 and asks NN 1 for APPL 4′s location. NN 1 responds that APPL 4 is at location NN 4. When USER 1 fails to find APPL 4 since it has moved, NN 1 verifies that APPL 4 is no longer where its directory database indicated, so it sends a new broadcast search to find APPL 4′s new location. The appropriate directory databases are then updated with the new information.

These are some of the other types of nodes which participate in an APPN network:

- **Low-entry networking (LEN) Node**, which can connect to the APPN network nodes to obtain its routing services, but do not automatically register all of their LU resources with the directory database.
- **APPN central directory server**, which maintains a central directory database, handles queries from network nodes about location of network resources, and is a focal point for broadcast searches.
- **VTAM composite network node**, which is a VTAM node with all its domain NCP nodes, and which appears as a single network node.

APPN protocol has the following important features:

- Dynamically locates resources in a network. When a new node is added or an existing node is moved in an APPN network, APPN recognizes that this has happened. It can discover the new location and determine a new route to get there.
- Determines best route through the network to get to the resource.
- Dynamic resource registration and directory.
- Provides dynamic topology changes.
- Provides complete APPN to Subarea interoperation.
- Provides intermediate session routing.
- Class of Service (COS) route selection.
- Adaptive pacing, which is the mechanism that allows for a smooth flow control mechanism to prevent the flooding of a resource.
- Multitail end node that allows connection to more than one network.

APPN will deliver the required link characteristics (performance, security) at the minimum cost. The customer does not have to define the path, and change all the path definitions as new resources (such as lines, controllers) become available. The optimal path will be selected each time a session is started, and if new resources are available, they will be utilized. In addition, if a session is active over one path, and the path fails, as the session is restarted it will pick the best path available at that time. New resources will be utilized at the next session startup.

4.1.2.5 SNA and APPN Summary

Although SNA has been considered proprietary, it is still the most popular and heavily used networking scheme. Today, its emphasis is on program-to-program communications and peer-to-peer oriented control and interaction. The critical step in accomplishing this is to dissolve the link between every SNA node and a systems service control point (SSCP) in a central host. This was introduced in 1982 when low-entry networking (LEN) architecture (called Type 2.1 node architecture at that time) was announced. This was the first time that an SNA node could function without any SSCP intervention. There has been much success with this approach. Type 2.1 nodes can be found on most, if not all, IBM and non-IBM mini-computers and workstations. There is, however, a significant drawback with the current Type 2.1 architecture, in that it does not support intermediate node routing and can only support a single link between any two adjacent nodes.

To overcome this, IBM introduced APPN. APPN is easy to implement, and a dynamically configurable, totally peer-oriented networking architecture. APPN does not depend upon on a central controlling host. It can automatically:

- Register end node resources
- Dynamically locate undefined remote resources via network-wide broadcast searches
- Continuously monitor and update network topology and path configurations

However, APPN does not yet support dynamic alternate routing. Thus, link failure will cause all the sessions routed over that link to fail, just as in SNA subarea. The sessions will be notified of a link failure; any LU can attempt to activate a new session, since the network node will attempt to find a new route to satisfy the request, based on the latest topological information.

APPN/HPR is a further enhancement to APPN. APPN/HPR is a promising new technology for network nodes and end nodes that transparently extends APPN, replacing the intermediate session routing function in selected nodes with the faster high-performance routing. APPN/HPR includes:

- A new connection-oriented transport layer protocol known as Rapid Transport Protocol (RTP), which is one of a new class of transport protocols that can also serve as a logical link (with priority) over multiple hops
- A new type of connectionless source routing called automatic network routing (ANR)
- A new type of flow control and congestion prevention mechanism called adaptive rate-based (ARB) flow control

APPN/HPR provides nondisruptive rerouting based on class of service, fast packet switching, minimal intermediate node storage, and drop-in migration to existing APPN networks based on an APPN/HPR boundary function, for seamless interoperation with current APPN products and protocols.

4.1.3 Transmission Control Protocol/Internet Protocol (TCP/IP)

TCP/IP is a networking protocol with a large community of users and over 400 supporting vendors. It is used in very large networks, including the Internet. It is built into UNIX, and is available for most other operating systems.

TCP/IP is not one protocol, but is a suite of many protocols. The protocols define applications, transport controls, networking, routing, and network management. It is today′s most widely used multivendor interoperability protocol. (The other major multivendor interoperability protocol, OSI, is not yet completely defined and not widely used.) TCP/IP was originally developed by the United States Department of Defense, U.S. Defense Advanced Research Project Agency (DARPA) and became a U.S. Department of Defense standard. Today, TCP/IP has become a de facto multivendor standard.

4.1.3.1 Internet Protocols′ Functional Layers

Focused is probably the best term to describe the Internet suite of protocols. When there was a need to connect diverse network technologies, the Internet researchers worked on it. As a result, there are several application protocols available for production use in the Internet suite. Before discussing these, it is important to briefly explain the Internet protocols′ four functional "layers," as illustrated in Figure 49 on page 141 .

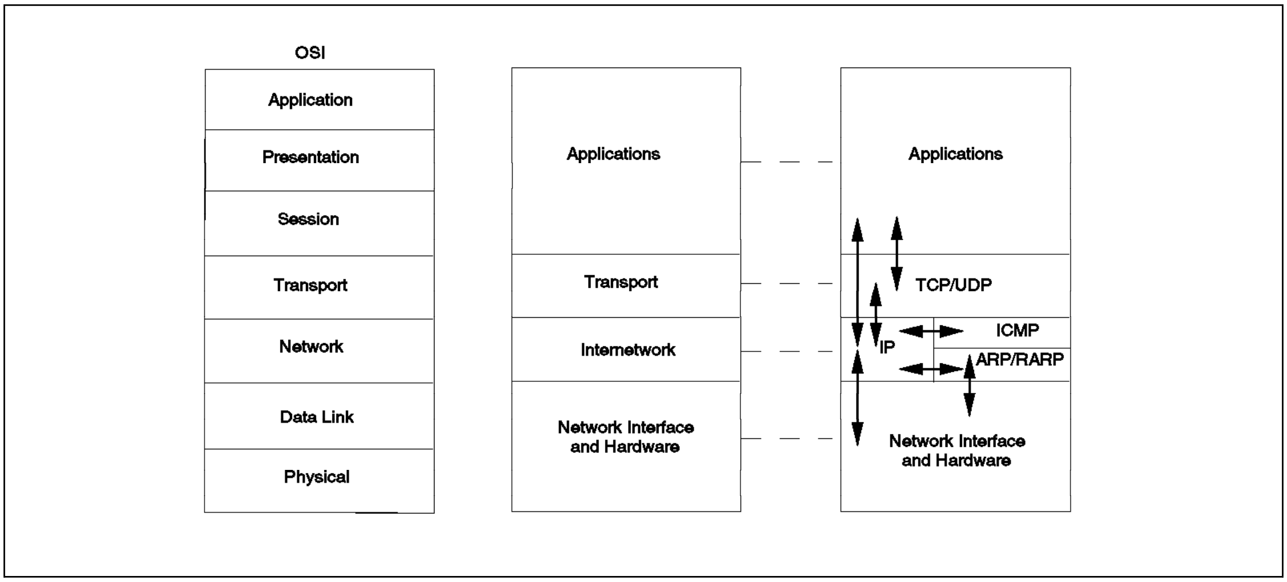


Figure 49. Internet Protocols' Functional Layers

Application

Application is a user process cooperating with another process on the same or a different host. Examples are Telnet (protocol for remote terminal connections), FTP (File Transfer Protocol) and SMTP (Simple Mail Transfer Protocol). Other examples are listed in 4.1.3.2, "Internet Suite of Protocols" on page 143.

Transport

Transport provides the end-to-end data transfer. The two protocols at this layer are:

- **Transport Control Protocol (TCP)**, which is the end-to-end reliable *connection-oriented* protocol providing flow-controlled, error-free logical connections between pairs of processes.
- **User Datagram Protocol (UDP)**, which is basically an application interface to IP. It is *connectionless*, and adds no reliability, flow-control or error recovery to IP. It simply serves as a "multiplexer/demultiplexer" for sending/receiving IP datagrams. An IP datagram (Internet datagram) is the base transfer packet in the Internet protocol suite. It has a header containing information for IP, and data that is only relevant to the higher level protocols.

Internetwork

Internetwork provides the "virtual network" image of the Internet. (That is, this layer shields the higher levels from the network architecture below it.) The protocols at this layer are:

- **Internet Protocol (IP)** is the most important protocol in this group. It is the "layer" that hides the underlying physical network by creating a virtual network view. It is an unreliable, best-effort *connectionless* packet delivery protocol that moves data between the intermediate nodes, referred to as routers. It adds no reliability, flow control or error recovery, and also does not assume reliability from the lower layers.
- **Internet Control Message Protocol (ICMP)** is the protocol used to report errors and control messages at the IP layer. Although ICMP uses IP as if it were a higher level protocol, it is actually part of the IP protocol.

- **Address Resolution Protocol (ARP)** is the end station to router protocol used to dynamically map Internet addresses to physical (hardware) addresses on local area networks. This protocol is limited to networks that support hardware broadcast.
- **Reverse Address Resolution Protocol (RARP)** is the protocol that a diskless host (an end node) uses to find its Internet address at startup. RARP maps a physical (hardware) address to an Internet address.
- **Open Shortest Path First (OSPF)**, which is a proposed standard protocol, is an IP routing protocol that uses an algorithm to build a routing table that is used to find the shortest path to every gateway and network the gateway can reach. This protocol uses a link-state algorithm to compute routes, which is more efficient than the vector-distance protocol used by RIP.

  Note: APPN also uses a link-state routing protocol.
- **Routing Information Protocol (RIP)**, which is a vector-distance protocol, is today′s most widely used routing protocol for IP networks. It does not, however, have the capability to handle larger internetwork configurations because:
  - The algorithm is run every 30 seconds.
  - The whole routing table is transmitted, which causes network overhead.
  - Routing decisions are based on a single criterion: the number of hops between stations, which may lead to an inefficient use of the network.

  As a result, there are two protocols that are vying to replace RIP: OSPF, which we have already defined, and Integrated IS-IS (intermediate system to intermediate system), which is described in RFC 1195 (Dec. 1990). The goal of Integrated IS-IS is to provide a single and efficient routing protocol for TCP/IP and for OSI. Its design extends the OSI IS-IS routing protocol (the ISO-specific protocol for routing within a single routing domain) to encompass TCP/IP. Both OSPF and Integrated IS-IS are link-state protocols, which route information differently than vector-distance protocols, the protocol used by RIP.

Network Interface

Network Interface is the interface to the actual network hardware. This interface may or may not provide reliable delivery, and may be packet or stream-oriented. In fact, TCP/IP does not specify any protocol here, but can use almost any network interface available, which illustrates the flexibility of the IP layer. Examples are IEEE 802.2 (for local area networks, such as the IBM Token-Ring network, or contention-type local area networks, such as Ethernet and CSMA/CD networks), X.25 (which is reliable in itself), and Packet Radio Networks (such as the AlohaNet, which is the granddaddy of broadcast networks.)

The actual interactions between the layers are shown by the arrows in Figure 49 on page 141.

4.1.3.2 Internet Suite of Protocols

The Internet Protocol suite includes a wide array of application support. Figure 50 illustrates the relationship between these applications available for production use and the end-to-end services. Also shown is how they map to the functional layers, discussed above.

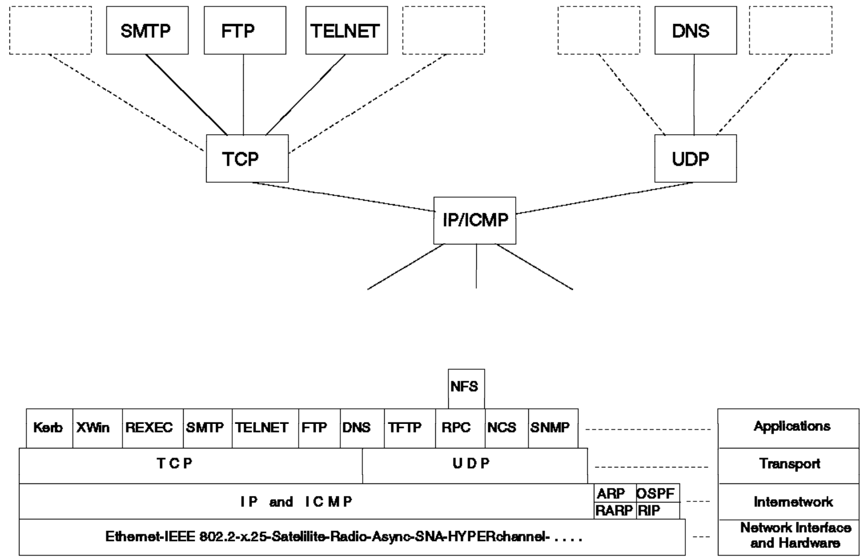


Figure 50. Overview of Internet Protocols

- **Kerberos** is a widely-used, encryption-based, third-party authentication mechanism for network security, which was developed by the Massachusetts Institute of Technology (MIT). It provides mutual security checking between clients and servers in a network environment.
- **X Window System** is a popular windowing system developed by Project Athena at the Massachusetts Institute of Technology and implemented on a number of workstations. It provides simultaneous views of local or remote programs/processes on a bit-mapped display and allows the application to run independently of terminal technology. For example, X-Windows may run on OS/2, DOS, and X-Windows terminals.
- **Remote Execution Protocol (REXEC)** is a protocol that allows the user to issue remote commands to a destination host implementing the REXEC server. The server performs an automatic login on a local machine to run the command. It is important to note that the command issued cannot be an interactive one; it can only be a batch process with a string output.

  Note: For remote login to interactive facilities, Telnet should be used.
- **Simple Mail Transfer Protocol (SMTP)** is the Internet electronic mail protocol, which provides store-and-forward service for textual electronic mail messages. Mail is sent from the normal local mail application to a background SMTP application that stores the mail and directly contacts the destination host′s SMTP application. Thus, mail is kept local until successfully transmitted. RFC822 defines the format of those messages.
- **Telnet** is the Internet virtual terminal protocol, which allows users of one host to log into a remote host and interact as normal terminal users of that host.

- **File Transfer Protocol (FTP)** is the Internet protocol and program, which is used to transfer files between hosts. It uses TCP. The user must be identified to the server with a user ID and a password before any data transfer may proceed.

- **Domain Name System (DNS)** primarily provides mappings between host names and network addresses, and is used extensively in the Internet.

- **Trivial File Transfer Protocol (TFTP)** is a file transfer application implemented over the Internet UDP layer. It is a disk-to-disk data transfer, as opposed to, for example, the VM SENDFILE command, a function that is considered in the TCP/IP world as a mailing function, where you send out the data to a mailbox (or reader in the case of VM). TFTP can only read/write a file to/from a server, and therefore, is primarily used to transfer files among personal computers. TFTP allows files to be sent and received, but does not provide any password protection (or user authentication) or directory capability. TFTP was designed to be small enough to reside in ROM, and thus can be used with BOOTP to boot a diskless workstation.

- **Remote Procedure Call (RPC)** is an API for developing distributed applications, allowing them to call subroutines that are executed at a remote host. It is, therefore, an easy and popular paradigm for implementing the client/server model of distributed computing. A request is sent to a remote system (RPC server) to execute a designated procedure, using arguments supplied, and the result returned to the caller (RPC client). There are many variations and subtleties, resulting in a variety of different RPC protocols.

- **Network File System (NFS)** uses a remote procedure call (RPC) to implement a distributed file system. It was developed by SUN Microsystems, and allows authorized users to access files located on remote systems as if they were local.

- **Network Computing System (NCS)** is a distributed computing environment developed by Hewlett Packard/Apollo Computer. It provides tools for designing, implementing, and supporting applications requiring distributed data and distributed computing. NCS also formed the base for the Open Software Foundation (OSF) Distributed Computing Environment (DCE) definition.

- **Simple Network Management Protocol (SNMP)** is the network management protocol of choice for TCP/IP-based Internets. It is used to communicate management information between the network management applications in the network management station, and the network management agents in the network elements. The network management applications may alter (set) or request (get) information from the Management Information Base (MIB) defined on each network element. In the other direction, the network element can send alerts (traps) to inform the management application as to the occurrence of asynchronous events.

4.1.3.3 Addresses

To be able to identify a host on the Internet, each host is assigned an address, the *IP address*. The Internet address consists of a pair of addresses:

```
IP address = <network address><host address>
```

which can take on several forms. The *network address* part of the IP address is assigned by a central authority, known as the Network Information Center (NIC), and is unique throughout the Internet.

IP addresses are 32-bit addresses usually represented in a dotted decimal form. For example, 128.2.7.9 where 128.2 is the network part and 7.9 is the host part of the IP address.

The binary format of the decimal address 128.2.7.9 is:

```
10000000 00000010 00000111 00001001
```

Internet addresses are used by the IP protocol to uniquely identify a host on the Internet. IP datagrams (the basic data packets exchanged between hosts) are to be transmitted by some physical network attached to the host. To send a datagram to a certain IP destination, the target IP address must be translated or mapped to a real physical address. Sometimes, this can be done by simply applying some algorithm to that IP address (as for X.25 networks), but sometimes it involves actual physical transmissions on the network to find out the destination′s physical network address. This is known as the Address Resolution Protocol.

Note: The Domain Name System (DNS) of TCP/IP provides Internet names associated with Internet addresses, which are easier to recall. Tables translate these names to their respective Internet addresses. For example, Internet name "RALYDPD.IINUS1.IBM.COM" could be translated into Internet address "9.98.1.2." The Internet name is used by many TCP/IP applications that execute on the Internet.

4.1.3.4 Internetworking with TCP/IP

Transmission Control Protocol (TCP) is a standards-based protocol. It is the most widely used higher-level protocol and provides reliability, flow control and some error recovery.

As previously discussed, IP is the best-effort connectionless packet (datagram) delivery system that forms the basis of the TCP/IP protocol suite. IP is also a standards-based protocol. The IP is the "layer" that hides the underlying physical network by creating a virtual network view for the higher layers. IP adds no reliability, flow control or error recovery to the underlying network interface protocol. There is no need for a connection establishment sequence. To send data, an end user has to put the destination address in front of the data and send it. Since the data can be sent over different paths, an implementation of flow and congestion control is much more difficult than in a connection-oriented network. Packets (datagrams) sent by IP may be lost, out of order, or even duplicated, but IP will not handle these situations. It is up to the higher "host-host" layer (TCP) to deal with such situations. IP also assumes little from the underlying network mechanisms, only that the datagrams will probably (best-effort) be transported to the addressed host.

IP allows each packet to select the least used routes (dynamic load balancing). Because the typical implementation of IP does not use priorities for different data types, like file transfer and interactive data, the interactive traffic cannot pass the batch traffic. This can produce erratic response times on heavily loaded links.

Most of the user application protocols, such as Telnet and FTP, use TCP as the underlying protocol. TCP is a connection-oriented, end-to-end reliable protocol providing logical connections between pairs of processes. Within TCP, a connection is uniquely defined by a pair of sockets on the same or a different system, that are exchanging information. The socket interface is one of several application programming interfaces (APIs) to the communication protocols. The socket interface is differentiated by the different services that are provided.

- **Stream socket interface** defines a reliable connection- oriented service (over TCP for example). Data is sent without errors or duplication and is received in the same order as it is sent. Flow control is built-in to avoid data overruns. No boundaries are imposed on the exchanged data, which is considered to be a stream of bytes. An example of an application that uses stream sockets is the File Transfer Program (FTP).
- **Datagram socket interface** defines a connectionless service (over UDP for example). Datagrams are sent as independent packets. The service provides non guarantees; data can be lost or duplicated, and datagrams can arrive out of order. No disassembly and reassembly of packets is performed. An example of an application that uses datagram sockets is the network file system (NFS).
- **Raw socket interface** allows direct access to lower-layer protocols such as IP and ICMP. This interface is often used for testing new protocol implementations. An example of an application that uses raw sockets is the PING command.

TCP can be characterized by the following facilities it provides for the applications using it:

- **Stream Data Transfer:** From the application′s viewpoint, TCP transfers a contiguous stream of bytes through the Internet without record boundary delineation. The application does not have to bother with chopping the data into basic blocks or datagrams. TCP does this by grouping the bytes in TCP segments, which are passed to IP for transmission to the destination. Also, TCP itself decides how to segment the data and it may forward the data at its own convenience.

  Sometimes, an application needs to be sure that all the data passed to TCP has actually been transmitted to the destination. For that reason, a push function is defined. It will push all remaining TCP segments still in storage to the destination host. The normal close connection function also pushes the data to the destination.

- **Reliability:** TCP assigns a sequence number to each segment transmitted, and expects a positive acknowledgment (ACK) from the receiving TCP. If the ACK is not received within a timeout interval, the data is retransmitted. As the data is transmitted in blocks (TCP segments) only the sequence number of the first data byte in the segment is sent to the destination host.

  The receiving TCP uses the sequence numbers to rearrange the segments when they arrive out of order, and to eliminate duplicate segments.

- **Flow Control:** The receiving TCP, when sending an ACK back to the sender, also indicates to the sender the number of bytes it can receive beyond the last received TCP segment, without causing overrun and overflow in its internal buffers. This is sent in the ACK in the form of the highest sequence number it can receive without problems. This mechanism is referred to as a *windowing* mechanism.

  A simple transport protocol might use the following principle: send a packet and then wait for an acknowledgement from the receiver before sending the next packet. If the ACK is not received within a certain amount of time, retransmit the packet. While this mechanism ensures reliability, it only uses part of the available network bandwidth. To better the use of the network bandwidth, ensure reliable transmission and flow-control, the window mechanism can be enhanced by not only having a sender group its packets to be transmitted, but also use the following rules:

  - The sender may send all packets within the window without receiving an ACK, but must start a timeout timer for each of them.
  - The receiver must acknowledge each packet received, indicating the sequence number of the last well-received packet.
  - The sender slides the window on each ACK received.

  The above window principle is used in TCP, but with the following differences:

  - The window principle is used at the byte level; that is, the segments sent and ACKs received will carry byte-sequence numbers and the window size is expressed as a number of bytes, rather than a number of packets.
  - The window size is determined by the receiver, when the connection is established, and is variable during data transfer. Each ACK message will include the window-size that the receiver is ready to deal with at that particular time.

- **Multiplexing:** Multiplexing is achieved through the use of ports. A port is a term used by Internet transport protocols to distinguish among multiple simultaneous connections to a single destination host. Each process that wants to communicate with another process identifies itself to the TCP/IP protocol suite by one or more ports. A *port* is a 16-bit number, used by the host-to-host protocol to identify to which higher-level protocol or application program (process) it must deliver incoming messages. Comparatively speaking, a TCP/IP port is equivalent to an OSI selector, which is used by an OSI entity to distinguish among multiple SAPs (service access points) to the layer above.
- **Logical Connections:** The reliability and flow control mechanism, which is outlined below, requires that TCPs initialize and maintain certain status information for each data stream. The combination of this status, including sockets, sequence numbers and window-sizes, is called a logical connection (or *virtual circuit*). Each connection is uniquely identified by the pair of sockets used by the sending and receiving processes.
- **Full Duplex:** TCP provides for concurrent data streams in both directions.

As noted above, the primary purpose of TCP is to provide reliable logical circuit or connection service between pairs of processes. It does *not* assume reliability from the lower-level protocols (such as IP) so TCP must guarantee this itself. Below is a step-by-step summary of TCP′s basic operation:

1. TCP establishes a connection.
    - Originator (Internet Address, Port Number)
    - Destination (Internet Address, Port Number)
2. Application data arrives as a stream of octets.
3. TCP segments the stream, appends a header to each segment and sends the results to IP.
    - Segment size is negotiated when connecting.
4. IP delivers received segments to TCP, which:
    - Checks (and corrects) errors
    - Properly sequences segments

- Removes headers
- Delivers the stream to the correct application
- Acknowledges data delivery

5. TCP terminates the connection.

Below is a step-by-step summary of IP′s basic operation:

1. Data segments arrive from TCP.
2. IP adds its own header to each segment to form an IP datagram and routes the datagrams toward the specified address.
   - Routes the datagrams to on-net or off-net hosts
3. As necessary, IP fragments datagrams to meet network limitations.
4. At the destination host, IP:
   - Reassembles fragments
   - Discards bad datagrams
   - Removes IP headers
   - Delivers in order received to TCP or UDP

Forming an internetwork by interconnecting multiple networks is done by gateways. This terminology differs from the generally used terminology where:

Bridge Interconnects LAN segments at the data link layer level and forwards frames between them. A bridge is independent of any higher layer protocol. It provides MAC layer protocol conversion, if required.

Router Interconnects networks at the network layer level and routes packets between them. The router must understand the addressing structure associated with the networking protocols it supports and take decisions on whether, or how, to forward packets. Routers are able to select the best transmission paths and optimal packet sizes. Example: any IBM host or workstation running TCP/IP may be used as a router.

Gateway Interconnects networks at higher levels than bridges or routers, ranging from network layer to application layer. A gateway usually supports address mapping from one network to another, and may also provide transformation of the data between the environments to support end-to-end application connectivity. Example: A VM host running TCP/IP may be used as an SMTP/RSCS mail gateway.

In the TCP/IP terminology, the terms gateway and Internet gateway are used to qualify what is defined above as being a router. It is also called an IP router after its basic mechanism: IP routing. A gateway or router may be used to qualify a host which interconnects two or more IP networks, and which performs the task of routing datagrams from one network to another.

In the Internet protocol, outgoing IP datagrams pass through the IP routing algorithm. This determines:

- On which physical adapter the datagram should be sent.
- If the destination host is on the local network or not. If it is on the local network, the datagram is sent to the physical address of the destination host;

otherwise, the datagram is sent to the physical address of the gateway, which is actually the next router.

The IP routing is based on the network address part of the destination′s host IP address (target IP address). The local part of the target IP address is of no importance at this stage.

This base algorithm, needed on all IP implementations, is sufficient to perform the base gateway function. Incoming datagrams (that is, datagrams directed to one of the physical adapters, using the physical network address of these adapters to reach this host) will be checked to see if the local host is the IP destination host:

yes The datagram is passed to the higher-level protocols.

no The datagram is treated as an outgoing datagram and the IP routing algorithm is used to determine on which network the datagram is to be forwarded (next hop).

Thus, an Internet gateway is basically a normal host running TCP/IP, as the gateway functionality is included in the base IP protocol. However, some configurations require more than just the base gateway (called routers with partial information); they will have additional protocols for gateway-to-gateway communications.

Figure 51 shows an IP datagram, going from IP address A (consisting of Network address X, Host A) to IP address B (consisting of Network address Y, Host B), through two physical networks.

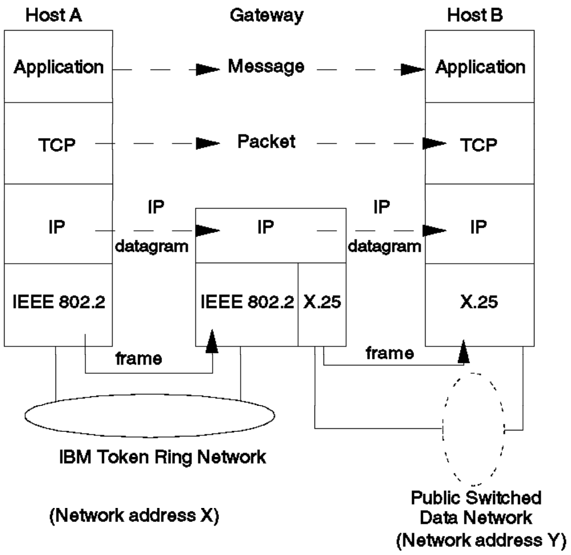


Figure 51. Internetworking with TCP/IP

4.1.3.5 TCP/IP Summary

The fastest growing and most popular set of products in the networking world are TCP/IP-based products. Originally, TCP/IP was not so popular. In fact, in the late 1970s and early 1980s, many thought that Xerox Network Systems (XNS) was a better solution, and that Xerox could have preempted TCP/IP if they had made all of XNS (including upper layers such as Clearinghouse, Courier, and Virtual Terminal) public domain. Please refer to 4.1.8, “Xerox Network Systems (XNS)” on page 177 for more information.

TCP/IP′s principal limitations include:

- IP′s limited addressing capabilities, which are fully described in several RFCs, such as RFC 1296 and RFC 1347. However, the project generally known as IPng (IP: The Next Generation) is addressing this limitation. This is considered to be one of the biggest challenges that the Internet is facing to date. Basically, due to the incredible growth of the Internet, address space is being depleted. This is analogous to “running out of phone numbers.” It is also analogous to running out of SNA addresses (which IBM had to deal with back in the 1980s). This project is trying to resolve this address space problem without impacting the entire Internet community, and also address how new features, such as security, support for mobile hosts, and real-time resource reservations, can be included. RFC 1550 was released so that anyone from the Internet community could submit white papers detailing any specific requirements that they felt an IPng must fulfill or any factors that they felt might sway the IPng selection.

  Recently, Scott Bradner of Harvard, who is a co-director of the IPng project, mentioned that by changing how the NIC assigns IP addresses, the address range may not be depleted until after the year 2000. If this interim solution were to be implemented, this would provide additional time for IPng to decide upon the best solution to resolve this problem.

  Incidentally, a bigger problem for the Internet is managing huge routing tables on backbone Internet routers. This problem is also being resolved by changing addressing and routing practices.

- Difficulty supporting communications like bit synchronous communications.
- Limited upper-layer support, which causes more interoperability issues since network application developers need to provide more function.
- The necessity for foreign countries to follow a U.S. standard, which is essentially beyond their control.
- The Internet, which was TCP/IP′s main test-bed, is now a production network, which has made innovation more difficult. As a result, alternative test-beds are being developed, such as OSI Intermediate System to Intermediate System (IS-IS), which is the ISO-specific protocol for routing within a single routing domain.

In spite of these limitations, TCP/IP is very successful with much widespread support and many implementations. Reasons which helped make TCP/IP successful in the past include:

- Funding from the U.S. Department of Defense (DoD)
- Free development and support of the Internet by college students and others
- University of California at Berkeley developing UNIX and making it a public domain at almost no charge

- Free use of ARPANET (Advanced Research Projects Agency Network), which was an experimental wide-area network created by the U.S. Department of Defense in 1969, and which first adopted TCP/IP as its standard and has now evolved to become the Internet connecting many universities, research projects, and businesses

Reasons for TCP/IP′s continued current success and popularity include:

- No commercial vendor controls its specifications.
- Simple model providing error-free data communication across many heterogeneous transmission systems (such as, multivendor interoperability).
- Stable specification since both IP and TCP have been implemented many times.
- Rapid growth in commercial use of the Internet.

4.1.4 Novell′s Network Protocol (IPX/SPX)

Novell′s network protocol (IPX/SPX) is an implementation of Xerox′s Internetwork Datagram Packet (IDP) protocol. IPX allows applications running on DOS, Macintosh, Windows, or OS/2 workstations to access the NetWare network drivers and communicate directly with workstations, servers, or other devices.

Novell′s network protocol uses a variety of peer protocols such as:

- Internetwork Packet Exchange (IPX)
- Sequenced Packet Exchange (SPX)
- Service Advertising Protocol (SAP)
- NetWare Core Protocol (NCP)
- Routing Information Protocol (RIP)
- ERROR and ECHO Protocols

To give applications that use NetBIOS the ability to use IPX as a network protocol, Novell has also implemented a NetBIOS emulator. Figure 52 illustrates the NetWare protocol stack.

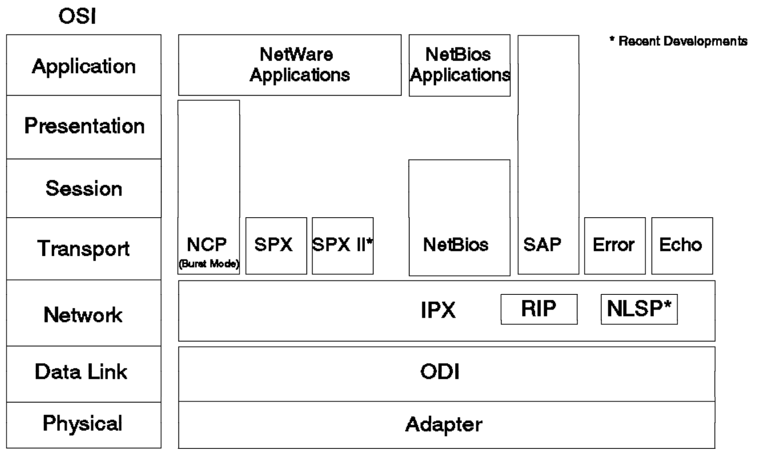


Figure 52. NetWare Protocol Stack

4.1.4.1 Open Data Link Interface (ODI)

Novell′s ODI standard provides the function of the OSI data link layer. The purpose of ODI is to allow single or multiple protocol drivers to use single or multiple adapter drivers. Network protocol drivers such as IPX/SPX written to the ODI standard, are independent of physical media and media protocol.

4.1.4.2 Internetwork Packet Exchange (IPX)

The Internet Packet Exchange (IPX) protocol is prevalent in many networks today. This proprietary protocol is used by Novell′s NetWare servers and requesters. IPX is a true datagram protocol. It makes a best-effort attempt to send data packets but does not guarantee the delivery of the data. The term datagram means that each packet is treated as an individual entity, having no logical or sequential relationship to any other packet.

IPX is a connectionless, full-duplex protocol; that is, it transmits data to a remote node, but does not wait for a response or acknowledgment indicating that the data has been received successfully. It is left to the higher level protocols to provide this guaranteed data transmission. In the NetWare environment, this is implemented within the NetWare shell.

4.1.4.3 IPX Addressing

When sending a NetWare IPX packet from one node to another, the station sending the data must know the correct 12-byte address of the receiving station. The NetWare IPX packet, therefore, contains two 12-byte addresses, one for the destination and one for the source. Each address is comprised of the following:

- **Network Address**

  This is an 8-digit (4 bytes) hexadecimal number which uniquely identifies a network and is used to address individual workgroups. Each network address has to be unique in the entire internetwork. (Note: Internetwork in this case refers to the entire IPX/SPX network.)

- **Node Address**

  This is a 12-digit (6 bytes) hexadecimal number that uniquely identifies the adapter card, or in NetWare terms, a network interface card (NIC), in a LAN segment. (This is analogous to the token-ring adapter address.)

- **Socket Number**

  This is a 4-digit (2 bytes) hexadecimal number that contains the number of the socket on which the server will receive requests. This number may be used to multiplex between functions within a node.

4.1.4.4 IPX Internetworking

Sockets enable multiple, higher layer protocols to use IPX services concurrently. A socket is an address used by the IPX layer in identifying the higher protocol to which the packet should be passed.

In order for an IPX client protocol to communicate with its peer partner, it needs to know the socket number of its remote partner application. To ensure the IPX packet travels through the network to the right partner system, the client program also needs to know its partner′s network address and node address. Figure 53 on page 153 illustrates how IPX Sockets work.

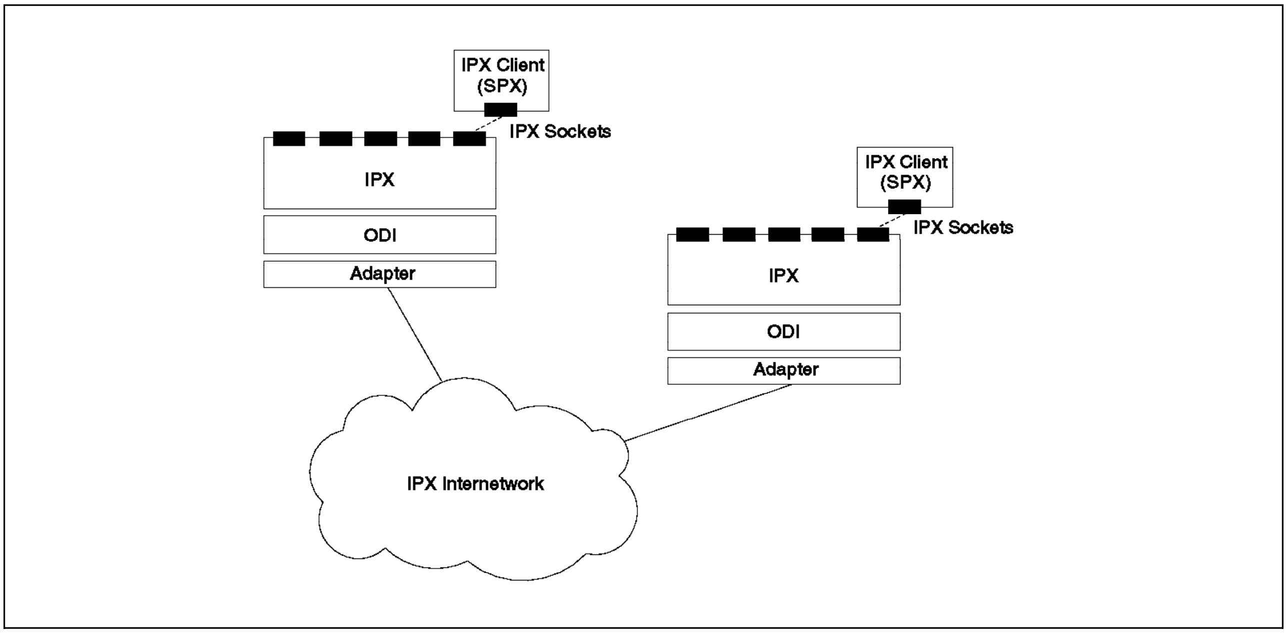


Figure 53. IPX Sockets

Nodes in the Same Network Segment: Figure 54 provides an example of a network station sending data to another. Both nodes are connected to the same network segment. In this case, both nodes will have the same network address. The sender addresses the packet the following way:

- The node address of the destination node is placed in the MAC header destination address field. The node address of the sending node is placed in the MAC header source address field.
- The full internetwork address (node address, network number, socket number) of the destination node is placed in the IPX header destination address field. The full internetwork address of the sending node is placed in the IPX header source address field.

  The MAC node address and the IPX node address are the same.

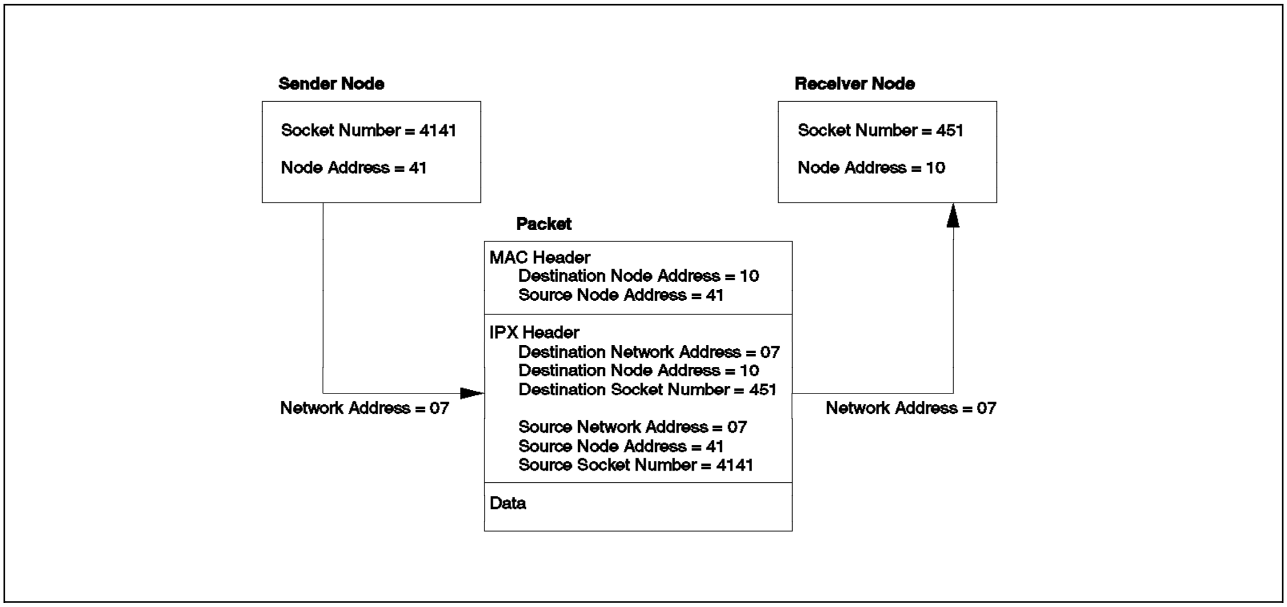


Figure 54. Addressing without Routing

Note: In the example above, the number of characters in the node address and the network address have been modified.

Nodes Separated by a Router: Figure 55 on page 155 shows an example of a node sending data to another node. However, in this case, the two are not on the same network segment. A router is used to connect the two networks. The nodes now have different network addresses.

The sending node must find a router on its own segment that can forward packets to the destination node′s network segment. To find this router, the workstation broadcasts a RIP packet requesting the fastest route to the destination node′s network number. Once the sender receives the router′s node address, it has enough information to send packets to the destination node. The packet is addressed in the following way:

- The node address of the router is placed in the MAC header destination address field. The node address of the sender is placed in the MAC header source address field.
- The full internetwork address (node address, network number, socket number) of the destination node is placed in the IPX header destination address field. The full internetwork address of the sending node is placed in the IPX header source address field.

When a router receives a packet to be routed, it first checks in its Routing Information table to see if the destination node is on a network segment to which it is directly connected.

If so, the router places the destination node address from the IPX header in the destination node address field of the MAC header. It places its own node address into the source address field of the MAC header and transmits the packet.

If the router is *not* directly connected to the destination network number segment, it passes the packet to the next router. It puts the node address of the next router into the destination node address field of the MAC header and places its own node address into the source node address field from the MAC header.

In this case, the MAC node address is not the same as the IPX node address.

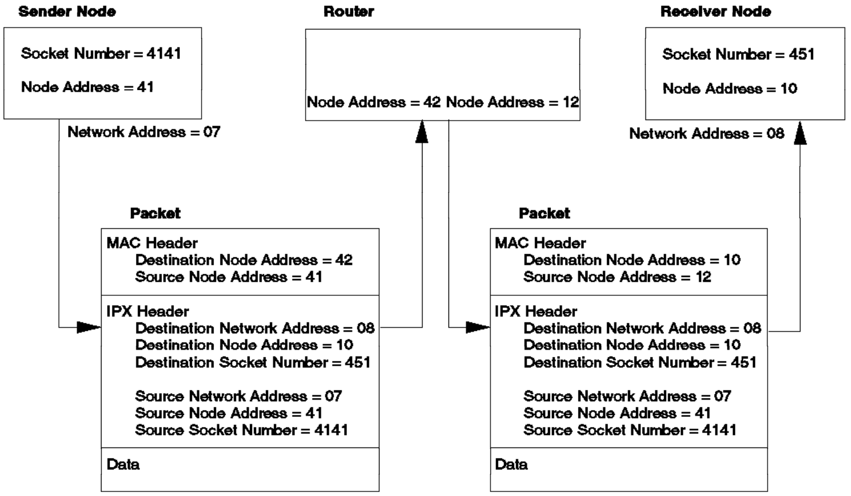

Sender Node
Socket Number = 4141
Node Address = 41
Network Address = 07
Router
Node Address = 42 Node Address = 12
Receiver Node
Socket Number = 451
Node Address = 10
Network Address = 08
Packet
MAC Header
Destination Node Address = 42
Source Node Address = 41
IPX Header
Destination Network Address = 08
Destination Node Address = 10
Destination Socket Number = 451
Source Network Address = 07
Source Node Address = 41
Source Socket Number = 4141
Data
Packet
MAC Header
Destination Node Address = 10
Source Node Address = 12
IPX Header
Destination Network Address = 08
Destination Node Address = 10
Destination Socket Number = 451
Source Network Address = 07
Source Node Address = 41
Source Socket Number = 4141
Data

Figure 55. Addressing with Routing

Note: In the example above, the number of characters in the node address and the network address have been modified.

4.1.4.5 Review of Other Novell Protocols

Sequenced Packet Exchange (SPX) and SPX II: The SPX protocol is intended to be used as a foundation upon which a variety of sophisticated applications may be built, including communication servers, PC-to-host gateways, and direct inter-workstation messaging systems.

SPX is a connection-oriented communications protocol built on top of IPX. When an application program makes a call to SPX to send a packet, SPX does some housekeeping-type work on the packet, then calls IPX to send the packet. SPX guarantees packet delivery, whereas IPX delivers packets on a best effort basis. This feature of SPX has obvious advantages, but also adds overhead to the data transfer cycle.

One of the functions SPX performs is the task of recovery from duplicate data and lost data errors. To achieve this, each side of the SPX connection maintains a sequence number for each packet it sends out. A receiving station does not need to acknowledge every single packet immediately. SPX allows several requests to be outstanding. Instead of acknowledging every single packet, it waits until the “window” (the allowable number of outstanding packets) is reached. It then sends a single acknowledgement for all of the packets received.

SPX II is an enhancement over the original SPX available in NetWare Version 4. One of the major advantages of SPX II is that it implements a true windowing algorithm and therefore improves packet throughput and reduces load on the network. SPX II receives positive and negative acknowledgments which indicate if packets have been successfully received by the partner station, or if some of the packets must be retransmitted. If, after sending a window, some packets have been lost, only the packets which are missing will be retransmitted by the originating station.

In contrast to SPX, which can only send packets of up to 576 bytes (the packet size that can be carried on any Novell router network), SPX II can negotiate the largest packet size both stations are capable of handling with its partner station during the establishment of a connection. SPX II is able to renegotiate the packet size automatically if an established connection fails, and a new route is determined. This gives SPX II the opportunity to send the largest possible packet size.

SPX II also provides the facility which allows both connection partners to release the session without losing data. That is, SPX II ensures that data currently transmitted will not be discarded. (This operates much like the de-allocate "flush" function found in SNA APPC.)

NetWare Core Protocol (NCP): In the NetBIOS environment, the LAN requester and LAN server communicate, pass data, and issue commands to one another via a protocol called the Server Message Block (SMB). NetWare has a similar protocol, which it calls the NetWare Core Protocol or NCP. This protocol defines the procedures that the file server′s operating system follows to accept and respond to workstation requests. NCP service protocols exist for every service that a workstation is able to request from a file server.

Every service available from a NetWare file server has been assigned a number. When a client, needs to submit a request to a server, it places the number in the request code field of the NCP packet. In addition, a connection number is also assigned for every session established with a file server so that the file server can keep track of clients making requests.

Originally, each packet transmitted by NCP had to be acknowledged causing poor wide area network performance for NCP read and write requests, especially for large files. A subsequent enhancement to NCP technology, called Burst Mode Protocol, addresses this problem. Burst Mode uses a *sliding window* algorithm, with an adaptive, self-tuning flow control mechanism.

Without Burst Mode, each read/write request (up to 512 bytes) is answered with a response acknowledgment from the server. This has been acceptable in simple LAN environments since the delay for the acknowledgement is low, but over a WAN link, this NCP request/response relationship may be very high and, therefore, impact performance. Also in a large LAN environment when the client and server are separated by several bridge and router hops, Burst Mode can improve large file transfer time.

With Burst Mode, a client is now able to issue a single read or write request for a block of data up to 64 KB. Basically, Burst Mode divides the file into packets, which are transmitted to the client, and no reply packets are required until the last burst packet is received by the client.

Service Advertising Protocol (SAP): The Service Advertising Protocol allows service-providing nodes, such as file servers, print servers, and gateway servers to advertise their services and addresses, in order for clients to access these services.

A service-providing node broadcasts an SAP packet or packets (each SAP packet can contain up to seven services) containing its service information every 60 seconds. This time is the same for all such nodes.

In large bridged networks, this broadcast traffic can pose congestion problems, especially for low-speed WAN links. This problem can be reduced by using IPX routers, which will collect all broadcasts from one subnetwork and send a single consolidated SAP broadcast (which could be multiple packets) to other attached networks. In addition, because these broadcasts are sent to the LAN "all stations" broadcast address, adapter congestion can occur in devices that have nothing to do with IPX/SPX (such as IBM 3745 Token-Ring interfaces).

An SAP *agent* exists in every service-providing node. These agents collect this service and address information and store them in a table called the Server Information table. If all SAP agents on the internetwork exchange SAP information properly, each agent′s Server Information table should have all the servers on the internetwork.

Clients (requesters) are able to contact an SAP agent or file server for server information. The clients receive information on the service type, and the internetwork address of the machine providing this service. With this address, the client is able to initiate a session with that specific server.

Routing Information Protocol (RIP): The Routing Information Protocol (RIP) facilitates the exchange of routing information on a NetWare internetwork. Weighting on network routes based on circuit delay or bandwidth is provided by a "cost" metric. The following information is exchanged:

1. Route Request broadcast by a workstation.
2. Request for routing information from other routers to update their own internal routing tables.
3. Response to route requests from routers or workstations.
4. Periodic broadcasts to ensure that all other routers are aware of the internetwork configuration (every 60 seconds).

   Note: Again, in large bridged environments, these broadcasts can cause network congestion.
5. Broadcast whenever a change in the internetwork occurs.
6. Final broadcast when the router is brought down.

ERROR and ECHO Protocols: ERROR and ECHO are two other protocols used by other NetWare peer protocols for internal maintenance. For example, an application may wish to communicate with a node in the internetwork . It sends an NCP request to IPX, indicating the destination address and the remote node. Normally, NCP expects to get a packet back from its peer. When this packet is received at the router, it checks its Routing Information table. If the destination network is unreachable or nonexistent, it cannot send the packet. It then sends an ERROR packet back to the sending node.

ECHO may be used by an application to test the viability of the path to a given destination. If a destination node receives an ECHO packet, it simply echoes it back to the sending application.

4.1.4.6 Novell Summary

NetWare Link Services Protocol (NLSP) was developed by Novell to replace the Service Advertising Protocol (SAP) and Router Information Protocol (RIP). SAP and RIP were designed at a time when internetworks were local and relatively small. As previously mentioned, both protocols have some limitations and characteristics which are not suited to large internetworks.

NLSP offers the following major benefits over RIP and SAP:

Improved routing

Each NLSP router stores a complete map of the network. In this way, NLSP is able to make more intelligent routing decisions.

Reduced network overhead

NLSP only transmits routing information and service information once something has changed.

Better manageability

Any management console which uses the Simple Network Management Protocol can monitor and control the operation of NLSP routers.

NLSP is backward compatible. NLSP-based and RIP-based routers can be used in the same network. A smooth upgrade from existing networks is possible.

Novell makes NLSP available for NetWare V3.11 and above. It will also be included in future releases of the NetWare Multiprotocol Router software.

4.1.5 Network Basic Input/Output System (NetBIOS)

When IBM developed the IBM PC Network and NetBIOS was introduced, there were no protocol standards available in the LAN networking environment. NetBIOS has become a de facto standard between application programs and a local area network (LAN). Most networking vendors have implemented the specification given by IBM that allows almost any application written to the NetBIOS interface to run unmodified on any network operating system that uses NetBIOS, regardless of the hardware or transport protocols used in the LAN. Therefore, the NetBIOS rules do not specify what hardware, software, protocols or physical media are to be used.

The NetBIOS protocol is used on networks by applications such as IBM LAN Server, Microsoft LAN Manager, and Lotus Notes. NetBIOS is one of the LAN oriented protocols, like XNS of Xerox, AppleTalk of Apple Computer Co., and NetWare (SPX/IPX) of Novell. NetBIOS was developed almost independently of the wide area network protocols. It is supported in Ethernet, token-ring, and IBM PC network environments. Originally it was designed as an interface between the application program and the network adapter of LANs, but some transport-like functions have been added. Currently NetBIOS contains selected functions of the transport layer (layer 4) and of the session layer (layer 5).

A common problem with the NetBIOS specification, however, is that it only deals with the upper layer functions of the interface. It does not specify what communications protocol should be used underneath it. As a result, almost every networking vendor has a NetBIOS API on top of their own proprietary communications protocol, which cannot communicate with other vendors′ protocols.

Novell′s NetBIOS emulator is implemented on top of NetWare IPX in the same way as SPX. Novell calls their implementation of NetBIOS an emulator, and it does not generate frames compatible with the IBM NetBIOS.

IBM uses NetBIOS to provide the session services between the IBM LAN Requester and the IBM LAN Server. IBM LAN Server provides users the capability to access shared printers and shared files across the network.

Figure 56 compares the IBM and NetWare NetBIOS implementations in a DOS environment.

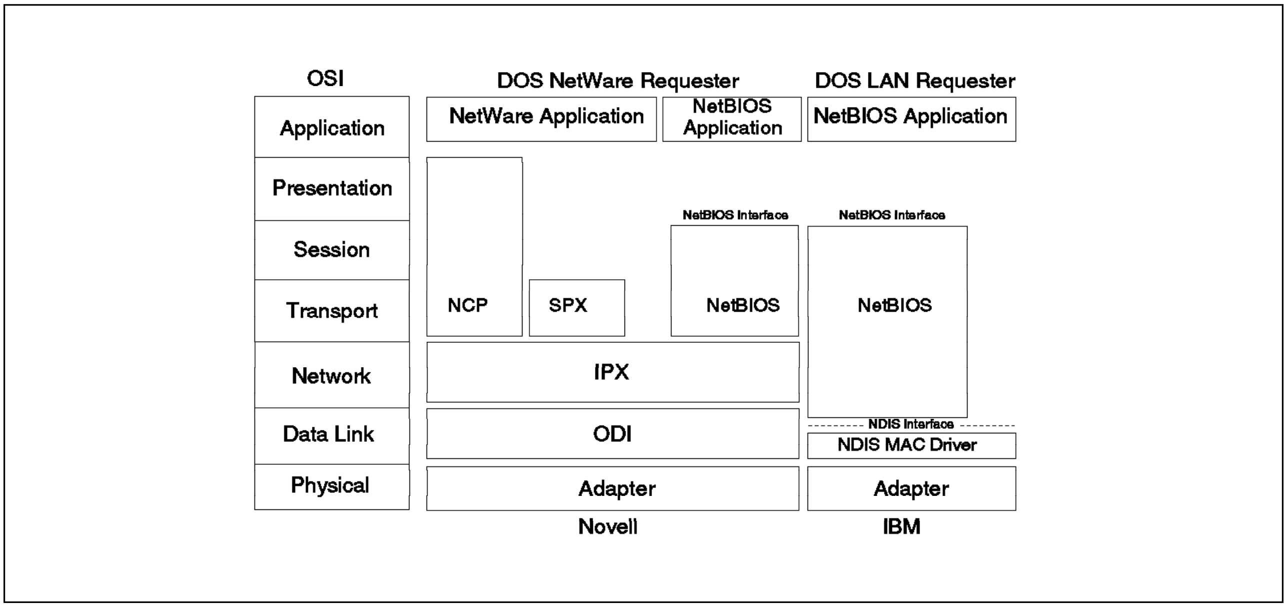


Figure 56. Examples of NetBIOS Implementations

An application using the IBM NetBIOS environment is designed to communicate with other devices using the IBM NetBIOS protocol. Novell has provided a NetBIOS API that allows NetBIOS applications to be ported into the NetWare environment by using the IPX protocol. In this environment, IPX becomes the transport protocol to link these applications. Because the protocol stacks are different, between IPX and IBM NetBIOS, these two cannot communicate with each other.

4.1.5.1 NetBIOS Functional Layers

NetBIOS supports a layered communications architecture. Names rather than network addresses are used by NetBIOS applications for session support. NetBIOS support provides services to locate and associate these names with MAC network addresses.

NetBIOS functional layers are:

- **Data Link Layer**

  At the link layer, NetBIOS can use connectionless (user datagram) or connection-oriented services. This layer is responsible for the assembly of data units, and the delivery of these data units to the physical media for transmission. This type of data transfer is best effort and receipt of data is not guaranteed.

- **Network Layer**

The network layer is implemented as a “null” layer.

Note: This lack of layer 3 function makes NetBIOS (as provided by IBM and Microsoft) unique in the world of LAN protocols. All of the other LAN-oriented protocols have some notion of intermediate node routing; thus, they can be routed across networks. To support NetBIOS over WANs requires bridging or some level of support, which provides a mapping to a routed protocol, such as using:

- Data Link Switching (DLSW)
- Multiprotocol Transport Networking (MPTN)
- LAN-to-LAN-over-WAN (LTLW) program, which is a LAN router that maps NetBIOS to an APPC connection
- Internet RFC 1001/1002, which supports NetBIOS over TCP/IP

- **Transport Layer**

  At this level, point-to-point connections are created supporting data transmission and acknowledgements. Any required flow control or pacing is handled at this level.

- **Session Layer**

  Through the session layer, NetBIOS establishes *name* pairs. A name is an identifier for a logical entity in which all session-level communication activity is centered. A *session* is a logical connection between two named resources supporting peer-to-peer communications. When a session has been established with another named resource, information can be exchanged over the session between the two resources. This reliable data transfer is provided by the session layer. A program may refer to multiple names and therefore sustain multiple sessions between itself and other stations on the LAN.

The NetBIOS interface provides an application programming interface offering commands for session control, name support, datagram support, and debugging facilities.

4.1.5.2 NetBIOS Addressing

The NetBIOS environment is particularly suitable for client/server relationships because clients and servers can use a common interface. Workstation sharing of print servers or disk servers are popular client/server applications.

As previously mentioned, communication between NetBIOS applications is based on resource names and not network addresses. The NetBIOS function within each workstation maintains a table of names that a node and its session partners are known by on the network. These names are provided to NetBIOS by the application program, and are up to 16 alphanumeric characters long. A name can be a *unique name*, or a *group name*. NetBIOS checks the network to verify that a unique name is not already in use at another adapter. A group name can be used by several adapters. This enables resources to be defined for functions that may exist in multiple workstations in a distributed application environment. For example, a department could have a group name invoked for sending data to all department members.

Only one NetBIOS unique name can be active on a NetBIOS network. However, no permanent relationship exists between a name and a specific workstation. Once a user is finished using a name, another workstation can use it. For example, an application named “sales report” could be used by several salesmen on different workstations to report a sale to a central database. When one salesman has completed his or her report, another at a different workstation could use the same name to send in his or her report.

Names are used as the basis for communication between application programs. NetBIOS provides services to ensure uniqueness of names. If a name is in the NetBIOS name table, a session can be established from this named resource with another named resource. NetBIOS maintains up to 254 selectable names plus one permanent node name (10 bytes of binary zeros followed by the adapter′s burned-in address).

4.1.5.3 NetBIOS Operation

The NetBIOS interface consists of five basic services, although some vendors may provide proprietary extensions. These services are:

- General Control, which includes resetting the NetBIOS interface, cancelling commands, and finding out the status of the network hardware adapter.
- Name Support, which includes adding unique and group names to the local name table and deleting local names from the table when they are no longer needed
- Datagram Support, which includes sending and receiving short messages addressed to specific NetBIOS unique names or group names, and sending and receiving broadcasts addressed to all machines.
- Session Control, which allows applications to establish sessions, check their status, and bring them down.
- Session Data Transfer, which includes sending and receiving data on already established sessions, and breaking larger messages into smaller messages to compensate for any physical network limitations.

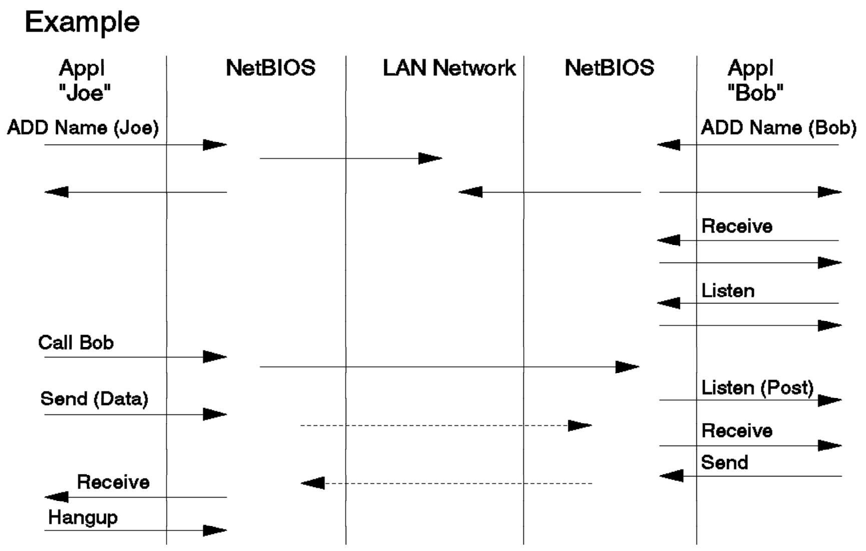


Figure 57. An Example of NetBIOS Operation

Figure 57 shows an example of a NetBIOS session-level data transfer sequence. A NetBIOS session is a logical connection between any two names on the network. All of NetBIOS commands are passed to the NetBIOS interface via a

Network Control Block (NCB), which contains several fields. Some fields are used to pass input values to NetBIOS and others are used by NetBIOS to return results from the command execution. No setup or BIND information is necessary like with other protocols. Both stations must already "know" each other and the receiver must have set the "listen" state for NetBIOS communications to start. Once a session is established, two-way guaranteed-delivery communication is possible between the two names.

In Figure 57 on page 161, each step on a system must have a complementary step performed on the other system.

1. Add "Joe," a unique name to the NetBIOS local name table using the ADD NAME command. "Bob" is also added to its NetBIOS local name table.

   Note: To ensure that a name is unique on the network, the ADD NAME QUERY is broadcasted to all NetBIOS stations. To make certain that each station receives the query and to avoid timing problems, this broadcast will be sent up to ten times on half-second intervals.

2. If Joe does not already know about Bob, Joe will broadcast a NAME QUERY looking for Bob (again, up to ten times on half-second intervals). As the network size gets large, this type of broadcast traffic can cause congestion problems, especially on wide area links.

3. Joe initiates a session with Bob by using the CALL command. This can be done since Bob had already issued a LISTEN command. If Bob had not issued a LISTEN command first, Joe′s CALL command would fail.

4. Transfer messages using the SEND and RECEIVE commands.

5. Terminate the session with the HANGUP command.

4.1.5.4 NetBIOS Summary

The popularity of NetBIOS is due to several factors:

- The NetBIOS interface isolates the application from the communications process.
- NetBIOS interface calls are independent of the physical hardware and of the communications protocol.
- The commands for token-ring are the same for Ethernet or PC Network.
- The application program only concerns itself with the NetBIOS interface and lets NetBIOS perform the communications tasks. For this reason, NetBIOS is popular since it hides the complexities of communications from the user.
- NetBIOS applications are easily ported from one system to another. For example, an application developed on a PC can be moved to a more powerful workstation which uses a different communications protocol. The NetBIOS support programs in each system take care of the differences.

However, NetBIOS, was not designed for wide area networks (WANs). Its name space is limited. It lacks support for subnetworks other than LAN, and recovery is not as good as that those found in other protocols, such as APPC (Advanced Program-to-Program Communication). Most importantly, NetBIOS has no concept of intermediate node routing.

4.1.6 Apple Computer′s Network Protocol (AppleTalk)

Apple Computer′s network protocol, known as AppleTalk, is used to connect Apple Macintosh Computers together to form a local area network. AppleTalk is a proprietary protocol stack, designed and developed by Apple Computer, Inc.

AppleTalk Phase 1: AppleTalk Phase 1 was the first release of AppleTalk. It was designed to fulfill the communication needs of workgroups and departments. With AppleTalk Phase 1, a node is identified through an 8-bit node ID. This limited the number of nodes to 254 (two IDs are reserved). In a LocalTalk environment, 254 network stations may be sufficient, but in a token-ring or Ethernet network, 254 addresses may be restrictive.

AppleTalk Phase 1 networks are also known as *nonextended networks*. A nonextended network has only one zone name per physical network. A zone is a grouping mechanism providing the ability to group the internetwork into logical user groupings.

AppleTalk Phase 2: AppleTalk Phase 2 is the new release of AppleTalk, and it extends the network addressing capability of Phase 1. AppleTalk Phase 2 networks are also known as *extended networks*.

With AppleTalk Phase 2, a single physical network may have more than 254 nodes. A node is identified through a 16-bit network number and an 8-bit node ID. AppleTalk Phase 2 allows multiple zones and multiple network numbers per physical network.

4.1.6.1 AppleTalk Protocol Stack

Figure 58 on page 164 illustrates the AppleTalk Protocol Stack, which is a layered suite of protocols that easily fit into the OSI Reference Model. These protocols can be further classified into five functions:

- Physical and data link functions
- End-to-end data flow functions
- Named entities functions
- Reliable data delivery functions
- End-user services

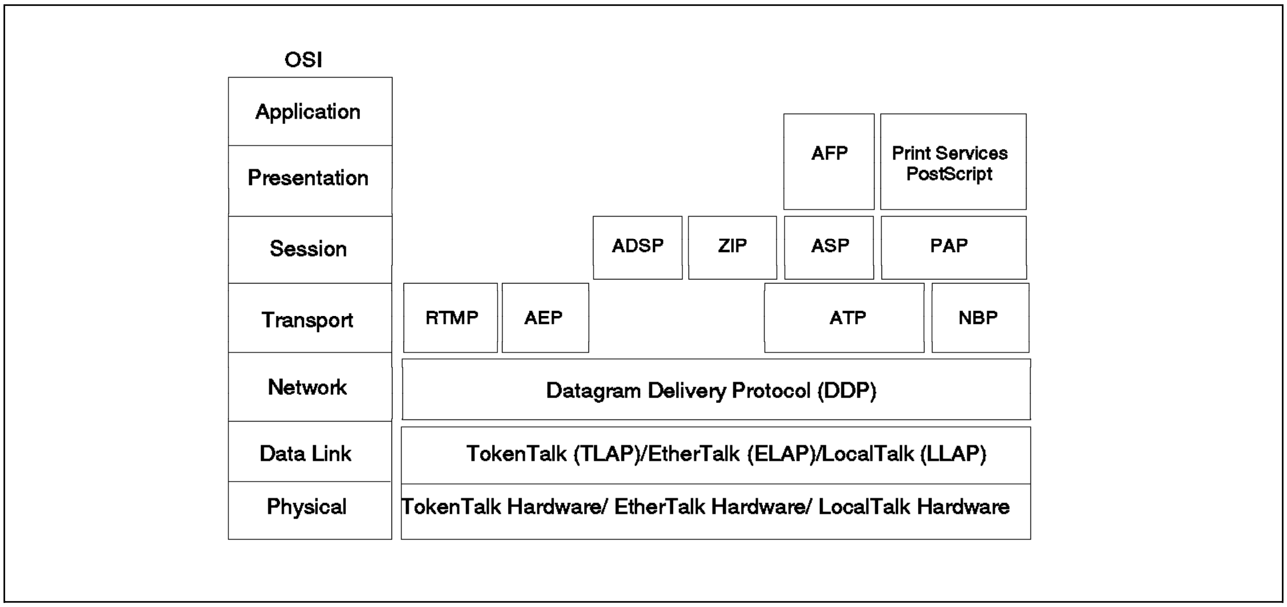

OSI
Application
Presentation
Session
Transport
Network
Data Link
Physical
AFP
Print Services PostScript
ADSP
ZIP
ASP
PAP
RTMP
AEP
ATP
NBP
Datagram Delivery Protocol (DDP)
TokenTalk (TLAP)/EtherTalk (ELAP)/LocalTalk (LLAP)
TokenTalk Hardware/ EtherTalk Hardware/ LocalTalk Hardware

Figure 58. AppleTalk Protocol Stack

Physical and Data Link Functions: The physical layer is responsible for handling the network hardware. As shown in Figure 58, standard network hardware, such as Ethernet and token-ring, can be used with AppleTalk. LocalTalk is Apple′s own network hardware and it uses a synchronous RS422A bus for communications.

The data link layer is responsible for interfacing to the network hardware. There are three protocols called Local Access Protocols (LAPs), upon which AppleTalk networks may be established:

- **EtherTalk Link Access Protocol (ELAP)**

  Apple first introduced EtherTalk as an extension of the AppleTalk Phase I protocols. Ethernet Version 2 was used as defined by DEC, Intel, and Xerox. When Apple updated ELAP to support AppleTalk Phase 2, Apple adopted Ethernet (IEEE 802.3) standards. Apple also changed the address and frame formats for ELAP between Phase 1 and Phase 2. As a result, a Macintosh running Phase 2 cannot directly talk to a Macintosh running Phase 1 and vice versa.

  ELAP uses the *AppleTalk Address Resolution Protocol (AARP)* to translate or map the AppleTalk node ID into an appropriate data link address used on an Ethernet LAN, which is 48 bits long. AARP also belongs at the data link layer but is not shown in Figure 58 since it is usually included within the definition of each LAP.

- **LocalTalk Link Access Protocol (LLAP)**

  LocalTalk is an easily configurable cabling system to connect workstations and other computer devices (for example, printers, scanners, etc.) to an AppleTalk network environment. Transmitter and receiver hardware for LocalTalk is built into every Apple device. LocalTalk is laid out in a bus topology, and the operation of a LocalTalk network is managed by the LocalTalk Link Access Protocol (LLAP).

  LLAP provides the basic service of packet transmission between the nodes of a LocalTalk or compatible network. LLAP manages access to the shared link by using an access method known as Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA).

LLAP has a node addressing mechanism that allows the protocol to dynamically assign a node ID. LLAP uses an 8-bit node identifier number.

- **TokenTalk Link Access Protocol (TLAP)**

  TokenTalk LAP did not exist before AppleTalk Phase 2. Since there is only one definition of token-ring packets, it doesn′t have the packet-typing problems that EtherTalk has.

  TokenTalk uses IEEE 802.5 standard for token-ring packets. By supporting 802.5, 802.2 Logical Link Control, and 802.3 Ethernet standards, AppleTalk can transmit data over an internetwork composed of all three networks: Ethernet, LocalTalk, and token-ring.

  TLAP′s address node ID is 48 bits long. Again, to communicate on a token-ring, the AppleTalk node ID must be translated by AARP into an address used by the token-ring.

End-to-End Data Flow Functions: The end-to-end data flow functions fall into two layers:

- Layer 3

  Layer 3, the network layer, is responsible for accepting data from the layers above it and dividing the data into packets that can be sent over the network. Since the network layer creates the packets, it is also responsible to define how they should be treated if the packets get damaged as they are transmitted.

- Layer 4

  Layer 4, the transport layer, is responsible for controlling and testing the route that the data packets take to their destination. There are four protocols at this layer as shown in Figure 58 on page 164, but only two perform the end-to-end data flow functions. These are RTMP and AEP. The other two protocols, ATP and NBP, perform other functions and will be discussed next.

The major function of the protocols in this category is to make sure that data is transmitted to the right destination. These protocols are not concerned with reliability in transmitting the data on the network, but just in getting the data from one network socket to another. There are three protocols that fall into this category:

- **Datagram Delivery Protocol (DDP)**

  The Datagram Delivery Protocol is responsible for network routing, actually transmitting and delivering application data, and basic data packaging into datagrams. Datagrams are packets of data carried between the sockets over an AppleTalk internetwork. DDP is connectionless; it routes a packet to the destination node, but it does not require acknowledgement of receipt of the packet. The datagram defined by DDP is limited to 586 bytes of data and includes a mechanism for detecting errors in data transmission, known as *checksum*. With checksum, the bytes comprising the data are added together and the sum is appended to the end of the data packet. The receiver recomputes the sum of the bytes in the packet and compares it to the sum received from the sender to determine if any data has been garbled.

  There are two types of DDP headers:

  - Short DDP header

    This was used in AppleTalk Phase 1 and uses 8-bit socket numbers and 8-bit node numbers as a means to address source and destination sockets.

  - Long DDP header

    Introduced in AppleTalk Phase 2 and uses an extended address of the 8-bit socket number, 8-bit node ID, and 16-bit network number.

  In AppleTalk, a packet cannot traverse more than 15 routers. To track this, each DDP packet contains a *hop count*, which gets incremented every time a packet passes a router.

- **Routing Table Maintenance Protocol (RTMP)**

  The Routing Table Maintenance Protocol is used to establish and maintain information about internetwork addresses and connections between the various networks, as stored in routing tables. Routing tables are central to the process of forwarding datagrams from any source socket to any destination socket on an internetwork. RTMP allows routers to exchange their routing table information periodically. If a router receives this information from another router, it compares and updates its own routing table to record the shortest path for each destination network.

- **AppleTalk Echo Protocol (AEP)**

  The AppleTalk Echo Protocol is a special test protocol which may be used by DDP clients (ATP, ASP, etc.) to determine whether a particular node is accessible over the internetwork. AEP may also be used to obtain an estimate of the time taken for a packet to reach a node in the internetwork.

  This protocol is used to send a datagram from one node to another and cause the destination node to return, or *echo*, the datagram back to the sender. To obtain an echo from a particular node, each AppleTalk node must have an echoer socket installed.

Named Entities Functions: There are two protocols that are responsible for handling and assigning names to every network device or service so that they can be found and are easy to use. One of these protocols, NBP, is defined at the transport level (layer 4), and the other, ZIP, operates at the session level (layer 5). The main purpose of the session layer is to manage the conversations among end users and to synchronize communications between applications located throughout the network.

- **Name Binding Protocol (NBP)**

  The Name Binding Protocol (NBP) is used to convert or translate numeric internetwork addresses, which the network uses to perform communications, into alphanumeric named entities or nickname addresses, which is what AppleTalk uses. Each file server, printer, or other service provider in the AppleTalk internetwork is given a unique name or nickname. A user uses this name as an address for the service providing nodes. If an application refers to a service by this name, the name must be converted into a network address.

  NBP maintains this table of mappings between internetwork addresses and named entities. Because each node maintains its own list of named entities, the names directory within an AppleTalk network is not centralized, but is a distributed database of all nodes on the internetwork.

- **Zone Information Protocol (ZIP)**

  AppleTalk uses logical groupings of networks to simplify routing. Each logical grouping can be identified by a name, or "zone."

  The main function of the Zone Information Protocol (ZIP) is to map the AppleTalk internetwork network numbers with the appropriate zone name so that a zone can be assigned during startup of a node. This includes helping routers maintain a mapping of network numbers to zones for the entire internetwork. This information is stored in a Zone Information Table (ZIT) in each router. ZIP is primarily implemented by routers, since non-router nodes use a small subset of ZIP during their startup process.

Reliable Data Delivery Functions: There are four protocols that fall into this category. In general, they are layered above the end-to-end data flow protocols, which means that they ensure reliable data delivery without knowing the source and destination of the data. These protocols pass on their data to the lower protocols, which take the responsibility for finding the right destination.

Three out of the four protocols are defined at the session layer since they are used to manage the end user conversations. Only ATP resides at the transport layer.

- **AppleTalk Transaction Protocol (ATP)**

  ATP is a connection-oriented protocol. The main purpose of ATP is to guarantee delivery of client data packets. To perform this delivery, ATP requires an acknowledgement once the packet is received from the destination node. This means that every time that ATP is requested to send a packet (*transaction request*), the receiver socket must report the outcome of the transfer (*transaction response*).

  Since ATP maintains a conversation between the two sockets by pairing transaction requests with transaction responses, ATP attempts recovery if the transaction request is lost, if the transaction receiver is lost or delayed, and if a responder becomes unreachable.

  ATP uses a timer to determine if any of the above conditions have occurred. If the timer expires, ATP retransmits the original request and it continues to retransmit until a maximum retry count is reached.

  For better performance, ATP is able to send up to eight packets, and only requires a response when all packets are received by the partner station.

- **AppleTalk Data Stream Protocol (ADSP)**

  ADSP is a connection-oriented protocol which provides an interface to an AppleTalk internetwork. ADSP services establish maintenance of full-duplex streams of data between two sockets in the internetwork, which means that a conversation between two computers can take place in both directions at the same time.

  ADSP also controls the rate at which data is sent from one node to another so that a fast sender does not overwhelm a slow receiver. This is accomplished by the receiving node periodically informing the sending node how much available buffer space is left. This also guarantees the correct delivery of data.

- **AppleTalk Session Protocol (ASP)**

  AppleTalk Session Protocol (ASP) is a client of ATP. ASP provides a reliable session service, which may be used to transport workstation commands to the server. This means that ASP makes sure that the commands received are in the same order as when they were sent.

A *session* is a logical relationship between two network entities. Sessions are identified by a unique session identifier. This makes it possible for more than one workstation to establish a session with the server at the same time. ASP is responsible for opening and closing a session, handling the commands and replies to and from a workstation and a server, and assuring that both ends are operational.

- **Printer Access Protocol (PAP)**

  Printer Access Protocol is a connection-oriented protocol. Its responsibility is to maintain communications between a workstation and a printer or print service. Functions include setting up and maintaining a connection, as well as transferring the data and releasing the connection upon job completion.

  Print manager and printer software are just two examples of applications which may use PAP. PAP relies on the Apple Transaction Protocol (ATP) and the Name Binding Protocol (NBP). ATP and NBP use the Datagram Delivery Protocol (DDP).

End-User Services: This category defines what the user wants to do with the network, which normally is either file sharing or printing. There are two protocols that fall into this category of functions and they are both defined at layer 6, presentation layer, which handles issues related to data files and formats. In addition, data encryption and data compression are also part of this layer.

- **AppleTalk Filing Protocol (AFP)**

  A native workstation file system command is able to manipulate local resources, that is, the hard disk and diskette drive and other memory resources. AFP enables the user to access the memory resources on a remote system with these same commands.

  If a workstation application sends a file system command to the local file system, it identifies whether the requested file resides locally or remotely by looking at this data structure. If the local file system discovers that the file is not local, it reroutes the command to the AFP translator. The AFP translator translates the file system command into AFP calls and sends them through the AFP interface to the file server.

  It is also possible for an application to directly access the AFP interface.

  AFP calls provide the following services:

  - The ability to obtain information about the file server and other parts of the file system structure
  - The ability to modify this information
  - Creation and deletion of files and directories
  - Reading and writing information in individual files

- **PostScript**

  PostScript is the well-known page description language used by many printers. Although PostScript is owned by Adobe Systems, it is included in the protocol stack as illustrated in Figure 58 on page 164 because it is a common interface for networked printer output on Apple Macintosh computers.

The Application Layer: All of AppleTalk′s protocols fall into one of the seven layers as illustrated in Figure 58 on page 164. However, as shown, there are no specific protocols mentioned at layer 7, the application layer. This is the layer at which all applications execute and it is at this layer where decisions on which application the AppleTalk user needs to use to fulfill his/her needs.

4.1.6.2 AppleTalk Addressing

An AppleTalk node in an internetwork is identified by an address. This address is composed of:

- A 16-bit network number, assigned from within the range for the node′s network
- An 8-bit, dynamically assigned AppleTalk node ID

This node address offers the ability to address 16 million nodes in one AppleTalk internetwork.

As discussed previously, LocalTalk networks are assigned one network number and are one zone. Such networks are called *nonextended networks*. Because of the 8-bit length of the network number, no more than 254 nodes may be concurrently active.

EtherTalk and TokenTalk are examples of *extended networks.* Such network nodes may be identified by the unique network number/node ID combination. Up to 16 million nodes may be concurrently active on such a network.

The node address in a nonextended network is simply a unique 8-bit node ID. The first time the node is brought up, it chooses a node ID randomly within the allowable range and broadcasts the ID on the network. Only a workstation with the same ID responds. If there is no response, the workstation uses this ID as its node address. If there *is* a response, the node tries another one. The ID is stored, and the next time the node is brought up it uses this old ID the first time.

The node address in an extended network is a 24-bit, unique network address/node ID combination. The process of assigning a node address is different in extended networks.

The acquisition of a node address in an extended network takes place in two steps:

1. The node takes a provisional node address to communicate with a router, thereby identifying the network number range valid for the network to which the node is connected.
2. A valid network number is chosen from the network number range, and a node ID is chosen randomly.

   This address is then broadcast to the network to certify that this node ID is not already in use by another node. If the network address is in use by another node, another valid node ID with the same network number is tried. If all node IDs are in use, another network number is taken from the range originally received by the router. The final network address is stored and used in the next initialization.

Sockets are logical entities within the nodes connected to an AppleTalk internetwork. Sockets are the identification numbers for partner applications to exchange information between nodes. Every program in an Apple environment has a socket number assigned to enable the exchange of information with a partner process in another node.

Each socket within a given node is identified by an 8-bit socket number. There can be up to 254 different socket numbers in a node. Sockets are owned by socket clients, which are typically processes (or functions in a process) implemented within the software in a node.

4.1.6.3 AppleTalk Networking

An AppleTalk internetwork consists of one or more AppleTalk networks connected together by intelligent nodes referred to as internetwork routers (IR). These need not be dedicated machines, but rather may perform additional functions such as file serving.

Internet routers are datagram (packet) forwarding agents. Datagrams can be sent between any two nodes of an Internet using a store-and-forward process through a series of Internet routers. An IR may consist of a single node connected to two or more AppleTalk networks. An IR might also consist of two nodes connected to each other through a communication channel (backbone network). See Figure 59 for an illustration of this.

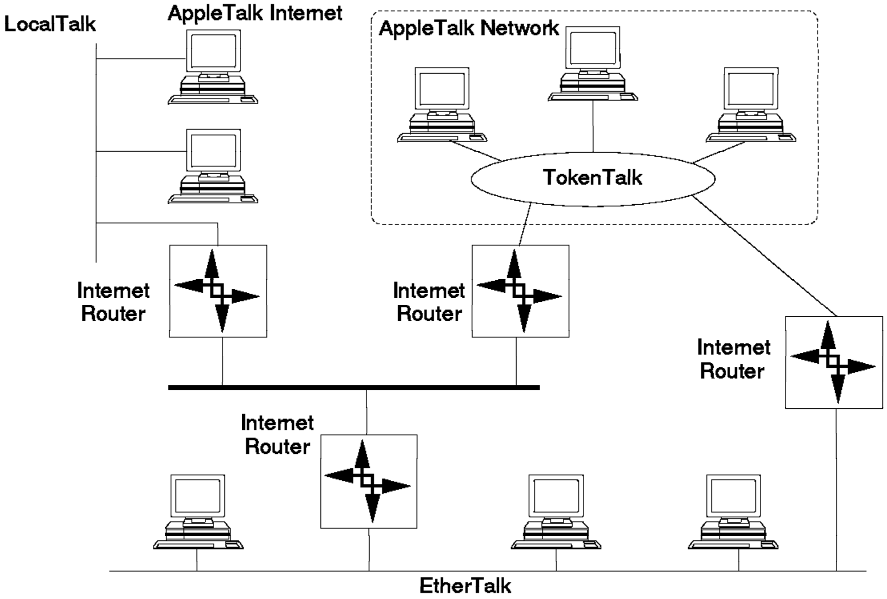


Figure 59. AppleTalk Internet and Internet Routers

4.1.6.4 AppleTalk Summary

A Macintosh can be connected to many other networking systems, such as DECnet, TCP/IP, and SNA. They can coexist with standard LAN operating systems, such as Novell NetWare, LAN Manager, 3Com 3+, and Banyan VINES (Virtual Network System). Basically, a Macintosh can be linked to any of these networking protocols in one of two ways:

- Run the protocols native to the other networking system on the Macintosh.
- Use a gateway between AppleTalk protocols and the non-AppleTalk protocol suite.

The major difference between these two implementations is performance. Running the foreign protocol on the Macintosh gives better performance than running the same protocols through a gateway since there would not be any competition with other users for the gateway resources. However, if there is only an occasional need to access these other protocols, then a gateway approach would probably be the most logical.

AppleTalk is one of the easiest protocols to administer because many networking functions are performed automatically via broadcast mechanisms. In large bridged networks, this broadcast traffic could cause some congestion especially in networks with low-speed wide area links. One solution to relieve this congestion is to install AppleTalk routers, which will reduce the broadcast traffic by segmenting the network.

4.1.7 Digital Equipment Company′s Network Protocol (DECnet)

Digital Network Architecture (DNA) was developed in the early 1970s, when networking had just begun. DECnet is the set of networking products based on DNA and in the 1980s, Digital added TCP/IP and OSI to DNA. Digital published its DNA specifications at about the same time that IBM announced Systems Network Architecture (SNA). In 1991, Digital introduced ADVANTAGE-NETWORKS, also known as Phase V, which integrates OSI and DECnet products and allows them to coexist and interoperate with TCP/IP.

DECnet has evolved through five phases since its introduction. Below is a brief summary of the five phases:

- **DNA Phase I**
  - Introduced in 1974
  - Supported only PDP-11 mini-computers running the RSX-11 operating system
  - Defined standards for point-to-point network communications between pairs of processors
  - Introduced first release of electronic mail to remote systems on any computer processor
- **DNA Phase II**
  - Introduced in 1976
  - Supported many other Digital operating systems: RSX-11M, RSX-11D, IAS, RSTS, and RT
  - Guaranteed backward compatibility so that changes made from one phase to the next would not be incompatible
  - Provided first step in solving interoperability problems by defining precisely how different DECnet implementations could be supported
- **DNA Phase III**
  - Introduced in 1980
  - Supported major Digital operating systems: RSX-11M, RSX-11M+, RSX-11D, IAS, RSTS, RT, VMS, Tops 10 and Tops 20
  - Introduced many networking enhancements, which included:
    - Incrementing network size to allow users to connect up to 255 processors in any configuration

        - Introducing adaptive routing capability so that each node could locate all the other nodes to route messages across the network to them
        - Support of routing through intermediate nodes
        - Introduction of a network management architecture to monitor and manage larger complex networks
        - Introduction of gateways from DECnet to other networks such as IBM SNA and X.25
    - Became a leader in network integration

- **DNA Phase IV**
    - Introduced in 1982, and is what most customers use today.
    - Supports major Digital operating systems: RSX-11M, RSX-11M+, RSX-11D, IAS, VMS, ULTRIX, VAXELN, DOS, OS/2, and P/OS.
    - Defines new standard for 16-bit addressing, which allows users to construct networks of up to 64,000 nodes.
    - Introduces many networking enhancements, which include:
        - Support for Ethernet, which was jointly developed by Digital, Intel and Xerox, and which provides high-speed communications and allows many devices to connect to a local area network
        - Expands Phase III adaptive routing standard to include support for hierarchical routing, which basically divides a network into areas to describe logical groupings of nodes which allows for network segmentation
    - Is layered after the ISO model, although the ISO standards were not finalized when this phase was introduced. (ISO network layer is very similar to the DECnet network layer.)

- **DNA Phase V**
    - Introduced in 1987
    - Is also known as ADVANTAGE-NETWORKS
    - Developed DECnet/OSI products from 1987-1991
    - Is available on VMS and ULTRIX operating systems
    - Incorporates the ISO standards into DECnet
    - Supports millions of interconnected nodes, which is an enhancement over DNA Phase IV
    - Provides improvements on performance
    - Introduces many networking enhancements, which include:
        - Link state routing
        - Enterprise Management Architecture (EMA) for network management

In addition to OSI and DNA Phase IV, Digital has been developing and is considered to be a leading supplier of TCP/IP products. They are, therefore, committed to all three protocol families. In addition, at this time, not all three protocols are sold together. Although TCP/IP is offered on both VMS and ULTRIX, only OSI and DECnet are sold as one. Digital is now considering selling all three together.

4.1.7.1 DECnet Functional Layers

The architecture standard for DECnet communication products is Digital Network Architecture (DNA). This architecture describes how the protocols and interfaces at each layer interrelate. Figure 60 shows how similar DNA is to the ISO model. Digital′s PHASE V model is shown in Figure 61.

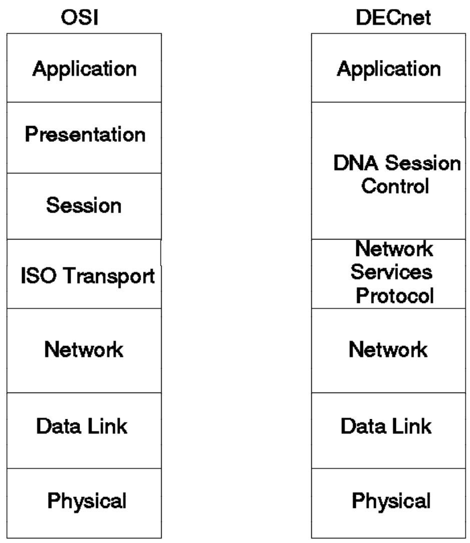


Figure 60. DECnet Protocol Stack

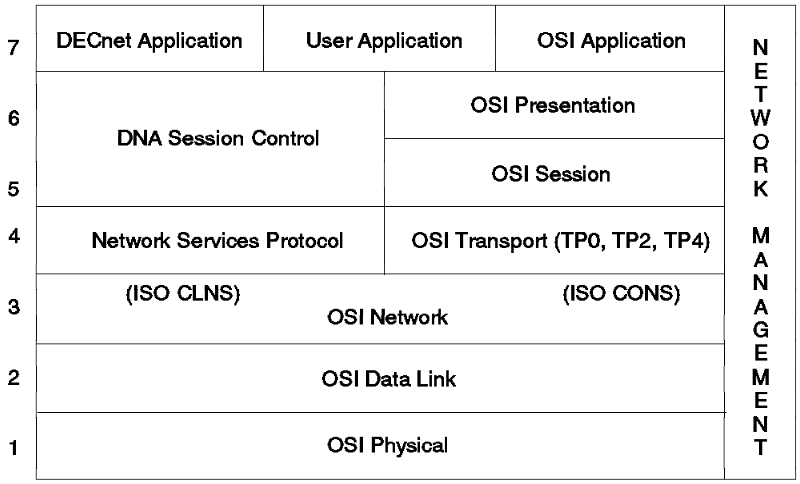


Figure 61. Digital′s Fifth Generation Protocol Stack

A brief description of a few of DECnet main protocols are:

- **Digital Data Communications Message Protocol (DDCMP)**

  This protocol, which is located at the data link layer, is used in wide area networks between two nodes as a point-to-point connection. It is also used in multidrop connections.

- **DNA Network Services Protocol (NSP)**

  This is a transport protocol defined in DNA Phase IV, but has been enhanced and refined by ISO so that it is also in DNA Phase V. This protocol establishes connections with a peer NSP entity and sends data, control, and acknowledgement messages.

- **Data Access Protocol (DAP)**

  This protocol is used for file transfer in DECnet, via the VMS COPY command, ULTRIX dcp command, or the programming READ and WRITE statements. This protocol also allows arbitary access to records within files.

- **Command Terminal Protocol (CTERM)**

  This protocol provides terminal emulation for local and remote DECnet terminals.

- **Mail-11**

  This protocol is the mail protocol in DECnet, upon which VMSmail is based. ALL-IN-ONE mail and other Digital mail applications are not related to this protocol.

- **Directory Services**

  In DECnet Phase IV, directory services is provided by a node database, while in DECnet Phase V, DECdns provides this sevice.

- **Connectionless Network Service Protocol (CLNS)**

  Connectionless assumes that no reliability is provided by the services at a particular layer.

- **Connection-Oriented Network Service Protocol (CONS)**

  As previously mentioned, connection-oriented protocols assume that, within a given layer of a protocol suite, there are mechanisms that ensure reliable communications with the same peer layer at another node running the same protocol suite.

- **OSI Transport Protocols**

  OSI specifies five classes of transport protocols, TP0 to TP4. Digital only supports TP0, TP2, and TP4, since both TP1 and TP3 are not very popular in the industry and in fact even GOSIP does not make it a requirement.

  - **TP0**

    TPO is the simplest of the protocols and it assumes that most requirements for supplying a reliable transport connection are handled by the network layer.

  - **TP1**

    TP1 provides minimal error recovery for errors signaled by the network, and it was designed for X.25.

  - **TP2**

    TP2 is an enhancement of TP0 and supports multiplexing, or the creation of multiple transport connections using a single network connection. It assumes a highly reliable network without a need for error recovery.

  - **TP3**

    TP3 provides multiplexing and basic error recovery.

  - **TP4**

    TP4 is the most popular of the protocols since it provides a reliable transport connection and runs on top of either CLNS or CONS. It provides retransmission, duplicate detection, flow control, connection establishment and termination, and recovery from crashes.

- **Network Time Protocol (NTP)**

  This protocol provides accurate, dependable, and synchronized time for hosts on both wide area networks, like the Internet, and local area networks.

- **Local Area Transport Protocol (LAT)**

  This protocol supports character-oriented devices that operate on a local area network to permit communication between nodes and other devices such as terminals, printers, and modems. LAT also supports arbitrary byte transfers. For example, Digital supplies X-Windows terminals that can use either LAT or TCP/IP transport connections to host applications.

- **Maintenance Operations Protocol (MOP)**

  This protocol supports Digital′s diskless workstation for installations and maintenance. In fact, this protocol can be used to support any type of workstation, terminal server, and X-Windows terminal for software installations, such as downloading operating systems, and for testing maintenance.

4.1.7.2 DECnet Addressing

Each DECnet node has a unique node address. The Phase IV addressing scheme is hierarchical in that it identifies each node as a member of an area. The node address must be configured for a node. A node address consists of:

- A 6-bit area number in the range 1 to 63
- A 10-bit node number in the range 1 to 1023

4.1.7.3 DECnet Internetworking

Like OSI and TCP/IP, DECnet nodes also have a peer relationship to each other. Any node can communicate with any other node without the need for a central controlling node. Digital′s peer-to-peer relationship helped make DECnet networks distributed while maintaining good performance. It also avoided many problems generated when the central controlling node failed.

A DECnet network is logically divided into *areas*. A large network can be divided into 63 areas, each containing up to 1023 nodes. A *node* is a computer that is connected to a network and supports DECnet software. Each node may be a member of only one area. A sample DECnet network is shown in Figure 62 on page 176.

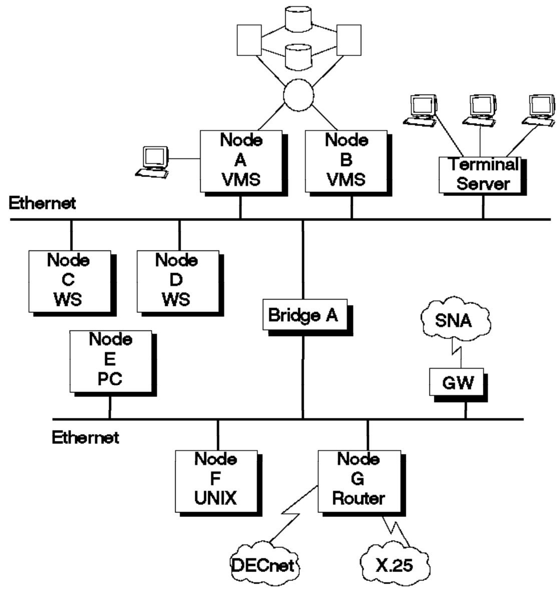


Figure 62. Digital Network

Nodes A and B are two large VAXs connected in a VAX cluster. The cluster can serve hundreds of users. Both nodes A and B are connected to an Ethernet cable. Nodes C and D are workstations. Node E is a PC and node F is a UNIX workstation.

On the Ethernet cable to the right of node B there is a terminal server. A terminal server connects many user terminals to the network which gives the users access to any node in the network. Terminal servers are another entity in the DEC network. The term entity is used to describe the hardware and software resources that are used cooperatively to form a DEC network. An entity can be a physical resource, such as a modem, or a logical resource, such as a circuit. Entities are classified into entity types, which include nodes, bridges, printer, circuits, links, and lines, to mention a few.

Bridges are entities in the DEC network that connect different Ethernet segments to each other. In Figure 62, Bridge A connects two Ethernet segments.

Node G is a router. Routers allow other networks and other DECnet areas to be connected to each other. In Figure 62, Node G is connected to another DECnet network and an X.25 network.

Figure 62 shows physical entities in a DEC network. In addition to physical entities, there are many logical entities that provide services in a DEC network. A circuit is an example of a logical entity that provides a virtual connection between two nodes.

There are two kinds of routing, direct and indirect. All routing is based on the network number of the Internet address and the use of Internet Protocol (IP) routing tables. All hosts, including routers, use routing tables to route and deliver data. Direct routing occurs between hosts on the same physical network. If the network number on the packet being routed is different than that of the sending host, then the host will use indirect routing to send it. In this case, the sending host sends the packet to a router on its own network to send the packet to its final destination.

A well-known and frequently referenced product solution for DEC networking is PATHWORKS, which is based on the Digital′s Personal Computing Systems Architecture (PCSA), which is an extension of Digital′s DNA. It provides file sharing, printer sharing, and similar services that integrate personal computers with Digital′s large systems. Possible clients include OS/2, DOS, and Macintosh. Possible servers include VMS, ULTRIX, OSF/1 and OS/2.

4.1.7.4 DECnet Summary

Digital is committed to developing open systems products that are reliable, standards-based, and willing to conform. It has been performing rigorous testing to ensure the interoperability of its networking products as early as prior to the introduction of DNA Phase III. Since Digital has allowed other vendors to build DECnet protocols for their own systems, it requires that they too perform tests of their implementations to demonstrate that they conform.

Digital′s release of DEC OSF/1 1.2 for Alpha AXP platforms, which is Digital′s newest architecture, represents a significant step in DEC′s effort in developing RISC/UNIX products. DEC OSF/1 is a direct implementation of the Open Software Foundation′s (OSF) product. Although OSF/1 1.2 represents a second-generation UNIX implementation with good functionality, efficiency, modularity, and room for growth, it does lack in applications compared to other UNIX platforms, due to it′s immaturity. In fact, OSF/1′s immaturity is seen also when compared to the maturity and depth of layered products in Digital′s VMS operating system. Digital has plans to make OSF/1 support DEC′s LAT (Local Area Transport) protocol. LAT is the alternative to TCP/IP′s Telnet for network login that optimizes bandwidth for small-packet traffic by bundling together many small logical packets into one network packet, similar to the way that X.25 does. Digital is also planning for DEC OSF/1 to participate in heterogeneous networks, which includes linking to other environments like AppleTalk, LAN Manager, and Novell.

4.1.8 Xerox Network Systems (XNS)

Xerox Network Systems (XNS) was developed by a group of people working at Xerox Corporation′s Palo Alto Research Center (PARC) and Office Systems Division (OSD) in the late 1970s and early 1980s for integrating their office products and computer systems. XNS is still one of the most commonly used protocol suites in distributed networks composed of multiple LANs.

Although XNS was developed by Xerox for their office systems products, most XNS networks were developed by other vendors. Xerox has published and made available its XNS protocols so that other vendors can write interfaces or applications for them. For example, in the late 1970s, Ungermann-Bass adopted the XNS protocols as the foundation of what would become its Net One line of products. 3Com Corporation adopted XNS in the early 1980s for its 3Com 3+ and 3+ Open product lines. Banyan VINES implements a proprietary protocol derived from XNS, which has a substantial number of modifications. And lastly, Novell (NetWare) also uses some XNS transport protocols in their own networking.

Basically, XNS is a network architecture that supports distributed computing. What this means is that a user from his/her workstation has the capability to manipulate local data, as well as access, via a network connection, any other shared data stored in large databases. In addition, the user would be able to send and receive electronic mail or use any printer anywhere on the network.

4.1.8.1 XNS Functional Layers

XNS is very similar to OSI and TCP/IP. Figure 63 illustrates XNS′s relationship to the seven-layer OSI Reference Model.

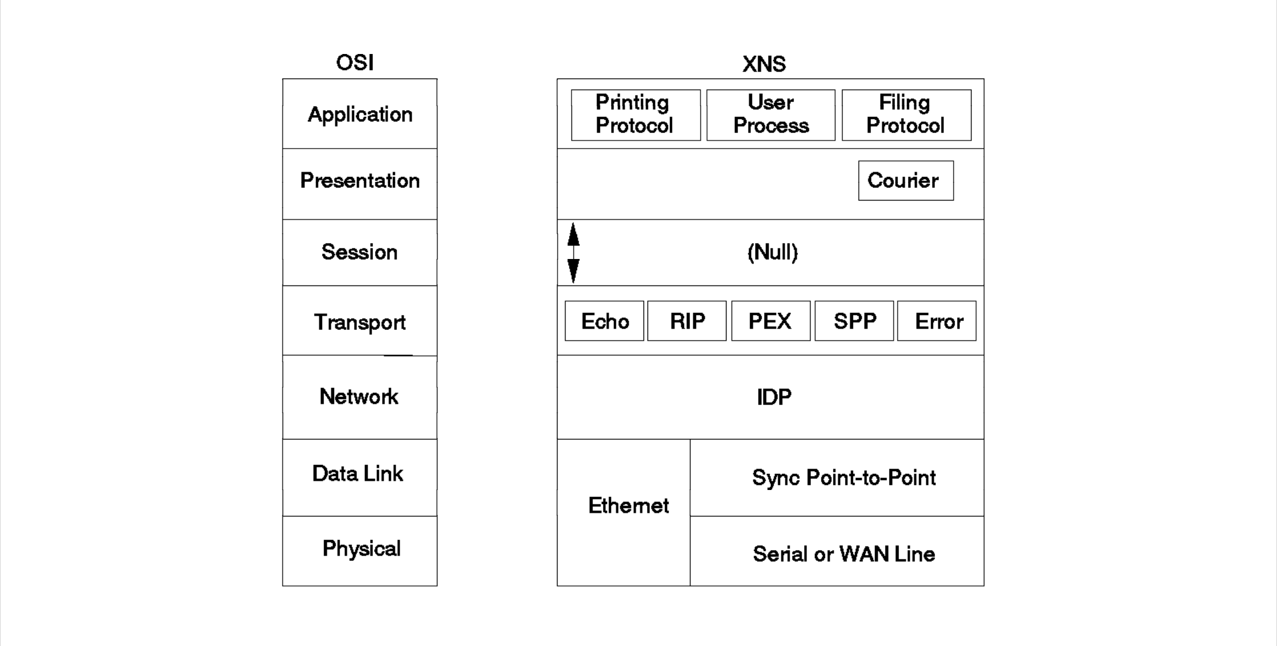


Figure 63. XNS Functional Layers

- Layers 1 and 2 (Physical and Data Link)

  There are three protocols supported at these layers. The most widely used and typical XNS network is Ethernet. Synchronous Point-to-Point protocol is also supported over serial or WAN line.

- Layer 3 (Network)

  XNS uses Internet Datagram Protocol (IDP) to route protocols through the network. Everything in XNS is transmitted in an IDP packet. IDP provides a connectionless and unreliable delivery service. Each packet contains all required addressing information and is transferred independently of all other packets on the internetwork. Since Ethernet is connectionless, it is easier and cleaner to use a connectionless protocol above it rather than a connection-oriented one. XNS does have a connection-oriented protocol, known as XNS Sequenced Packet Protocol (SPP), and this is implemented at the transport layer, which will be discussed next.

- Layer 4 (Transport)

  As Figure 63 illustrates, there are several protocols at this layer:

  – **Error Protocol**

    This protocol provides a means for any agent in the system to report that it has noticed an error and as a result, has discarded a packet. This protocol is intended as a diagnostic tool as well as a means to improve performance. Its use is optional.

  – **Echo Protocol**

    This protocol is executed when a host or router detects an inconsistency or network fault that prohibits packet delivery to the next hop or destination (not end-to-end recovery). The Echo protocol is used to verify the existence and correct the operation of a host and the path to it. Most XNS implementations support this protocol.

  - **Routing Information Protocol (RIP)**

    This is the only routing exchange protocol defined for XNS. It is the predecessor to TCP/IP′s RIP, which is similar in function. RIP maintains a routing database for use on a host in the forwarding of IDP packets to another host.

  - **Sequenced Packet Protocol (SPP)**

    This connection-oriented protocol provides reliable transmission of successive Internet packets on behalf of client processes. It provides a byte stream for the user process with optional message boundaries.

  - **Packet Exchange Protocol (PEX)**

    This protocol is used to transmit a packet and receive a response with reliability greater than that achievable by transmitting Internet packets directly as datagrams, but with less reliability than that achievable through use of SPX. It is therefore considered an unreliable, connectionless, datagram protocol for user processes. This protocol, however, supports retransmission, but does not perform duplication detection.

- Layer 5 (Session)

  XNS has no session layer. All session management is done at the XNS application layer.

- Layer 6 (Presentation)

  The most important protocol at this layer is the Courier remote procedure call, which is both a protocol and a specification language. It is used mainly to communicate with Xerox print and file servers. Courier only uses SPP for the network connection between the client and server. Since SPP is a connection-oriented protocol, there is no limit on the amount of data that can be exchanged between the client and server. Courier is very similar to ISO Abstract Syntax Notation (ASN) One, in that it also provides a way of representing data exchanged by higher-level protocols.

- Layer 7 (Application)

  Most of the applications use the Courier remote procedure facility. The most used protocols include the Printing Protocol, which is used to send files to a Xerox printer, and the Filing Protocol, which is used to send and receive files from a Xerox file server. Other applications at this layer include Mail, Clearinghouse, which provides a directory service function, Authentication, which enables processes to identify themselves to other processes as a matter of security, and Virtual Circuit Terminal, which is used primarily by workstations emulating terminals communicating to remote hosts via the XNS network.

4.1.8.2 XNS Addressing

An XNS address is 12 bytes and consists of a 32-bit network ID, a 48-bit host ID, and a 16-bit port number (known as a socket).

- Network ID

  This is a number that uniquely identifies the Ethernet on which the addressee resides. Network IDs are used by routers to identify the Ethernet to which a packet should be sent. There is a registration process which occurs to make sure that the network IDs are globally unique so that other networks can be connected easily.

- Host ID

  This number is also unique and it identifies the addressee. It is also registered to keep its uniqueness so that the router can transfer the packet to the destination host.

- Socket Number

  This is a logical identifier of a process within a host system. The association between a socket and a specific process is actually performed at the transport layer. There are some well-known socket numbers that are associated with a specific function or type of processing. For example, routing information is exchanged between routers through Socket Number 1 on all routers.

4.1.8.3 XNS Internetworking

Figure 64 shows a small typical XNS configuration. XNS supports a distributed system that is composed of a certain number of user workstations and shared servers which are all connected to an Ethernet LAN. The LANs are then interconnected via Internetwork Routing Systems (IRS).

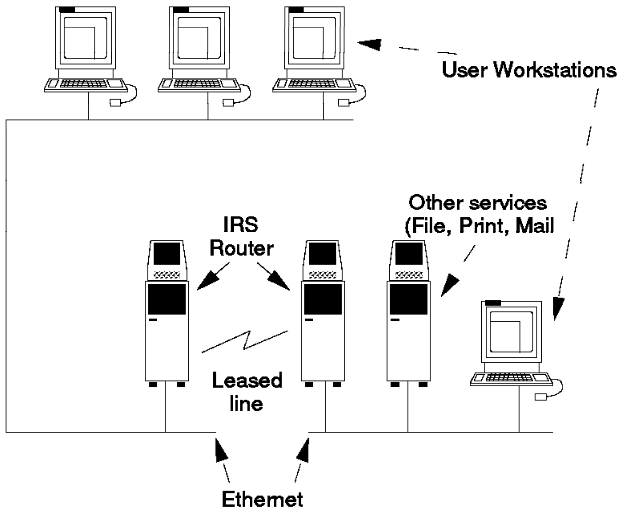


Figure 64. XNS Internetworking

XNS provides both connectivity and application functionality throughout an extended network of Ethernets interconnected by point-to-point leased lines. XNS calls this type of extended network an *internet*. There is no mainframe computer that controls the network like in SNA, so XNS is very similar to TCP/IP and OSI in this respect.

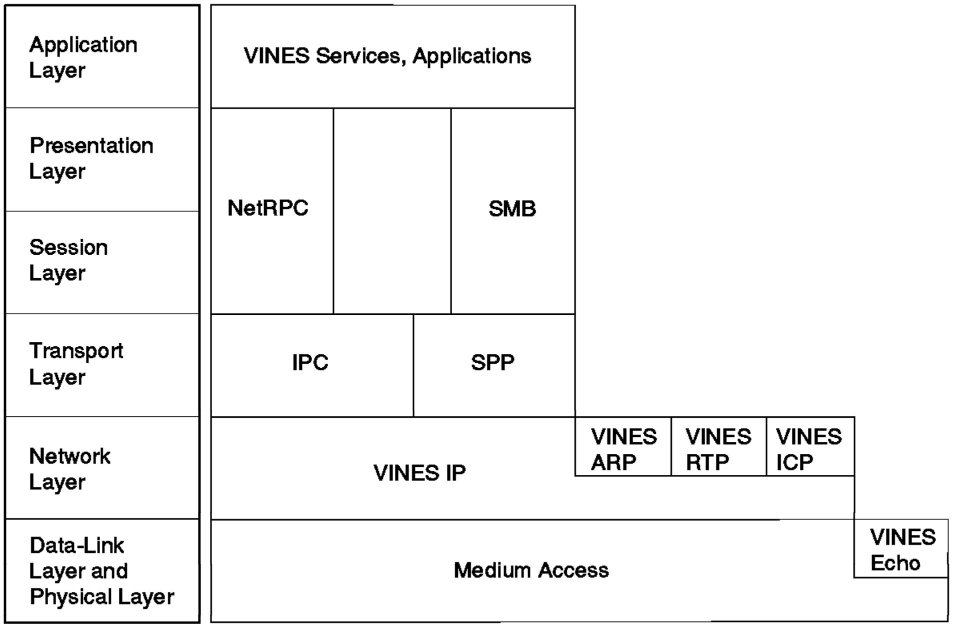
XNS supports two point-to-point wide area network (WAN) protocols, so as to enable the IRS routers to communicate with each other over longer distances. These routers use dedicated leased lines because they must frequently communicate with each other. In fact, switch (dial-up) lines are not supported. The routers can also communicate through packet-switched networks using X.25 protocols.

4.1.8.4 XNS Summary

In the last few years, the importance of XNS has lessened due to the growth of non-proprietary networking architectures, as TCP/IP and OSI. XNS derivatives, however, continue to exist in LAN segments running protocols, such as NetWare, Banyan VINES, and 3+ Open.

4.1.9 Banyan VINES

Banyan VINES is a service-based network operating system which has a seven-layer architecture that consists of industry standard and proprietary protocols. A comparison of the VINES seven-layer architecture to the OSI architecture is shown in Figure 65.



Legend:

ARP	=	Address Resolution Protocol
ICP	=	Internet Control Protocol
IP	=	Internet Protocol
IPC	=	InterProcess Control Protocol
NetRPC	=	Network Remote Procedure Calls
RTP	=	RouTing Update Protocol
SMB	=	Server Message Block Protocol
SPP	=	Sequenced Packet Protocol

Figure 65. VINES Protocol Stack

4.1.9.1 Medium Access Protocols

Medium access is the combination of the physical and data link layers. At the physical layer, VINES packets on the transmission media are supported on token-ring, Ethernet, and serial ports.

VINES Echo Protocol: LAN-attached VINES clients send echo frames to their routing servers using the VINES Echo Protocol.

After it receives an echo frame from a client, a routing server switches the source and destination MAC addresses in the frame and returns the frame to the client. If the original destination address is the broadcast address, the server replaces the broadcast address with its own MAC address before it returns the frame.

The VINES Echo Protocol actually resides above the LLC sublayer within the data link layer. Therefore, the VINES architecture provides a protocol ID for the data link layer echo frames in each of its encapsulation methods for LAN data link protocols.

4.1.9.2 Network Layer Protocols

VINES network layer protocols:

- VINES Internet Protocol (VINES IP)
- VINES RouTing Update Protocol (VINES RTP)
- VINES Address Resolution Protocol (VINES ARP)
- VINES Internet Control Protocol (VINES ICP)

VINES Internet Protocol (VINES IP): VINES IP routes the transport layer′s reliable messages and unreliable datagrams through the VINES internetwork. It provides a connectionless delivery service that guarantees only to provide the received messages and datagrams to the transport layer without any size or bit errors.

VINES IP does not guarantee that the messages or datagrams will reach the destination node nor that the messages or datagrams will arrive in the order in which they were sent by the source node. It is also possible that more than one copy of a packet may arrive at the destination.

VINES RouTing Update Protocol (VINES RTP): VINES RTP provides the routing database from which VINES IP selects the appropriate route on which to forward a packet. It builds the routing database from the network topology information shared among routing nodes.

The database contains the least cost route to each known destination network node. A route′s cost is determined by its routing metric, which is an indication of the round trip delay between the source and the destination network nodes. The metric unit is 200 milliseconds.

The routing metric is the sum of the interface metrics along the path from the source network node to the destination network node. The VINES architecture defines an interface metric for each type of interface based on the cost of communicating over that type of interface. The interface metric is also known as the neighbor metric.

VINES releases prior to 5.50 contain the original VINES RTP, now known as nonsequenced RTP. VINES releases starting with 5.50 contain an updated VINES RTP, known as sequenced RTP. Sequenced RTP is compatible with nonsequenced RTP.

VINES Address Resolution Protocol (VINES ARP): VINES servers and routers have preassigned VINES internet addresses, whereas VINES clients do not. VINES ARP allows a client to receive a unique VINES internet address.

When a client starts its VINES software, the client sends a query request to find a VINES server or router. The first server or router that responds to the query becomes that client′s routing server. The client then sends an assignment request to its routing server who then responds with the client′s address.

VINES releases prior to 5.50 contain the original VINES ARP, now known as nonsequenced ARP. VINES releases starting with 5.50 contain an updated VINES ARP, known as sequenced ARP. Sequenced ARP is compatible with nonsequenced ARP.

VINES Internet Control Protocol (VINES ICP): VINES ICP provides notifications of certain conditions and a network layer echo.

The transport layer protocols at the source need to know the metric of the interface between the destination network node and the final destination. They find this interface metric by setting the metric subfield bit in a packet. When VINES IP, at the destination network node, routes this packet to the neighboring destination, VINES ICP sends a metric notification containing the value of the interface metric to the source of the packet.

The connection-oriented protocols of the transport layer need to be informed when a packet cannot reach the intended destination. Consequently, the error subfield bit is set in all packets of Interprocess Communications Protocol (IPC) reliable messages and Sequenced Packet Protocol (SPP) data streams. If VINES IP, in a routing node along the route to the destination, cannot route nor receive one of these packets, VINES ICP sends an error notification to the source of the packet.

VINES ICP also provides the ICP echo, a network layer echo that verifies that a usable route exists between two network nodes.

4.1.9.3 Transport Layer Protocols

The VINES architecture includes these transport layer services:

- IPC unreliable datagram service
- IPC reliable message service
- SPP data stream service

4.1.9.4 VINES Address Definition

The VINES internet address is 48 bits. The high-order 32 bits are the network number. It is a logical network number that identifies a specific VINES server or VINES router. Each network number must be unique within a VINES internetwork.

A VINES server obtains its network number from a VINES server key that unlocks the VINES software on that server. A VINES server′s network number is globally unique.

Banyan Systems, Inc. assigns ranges of network numbers to VINES routers.

A VINES client obtains its network number from a neighboring VINES server or VINES router through the use of the VINES ARP Protocol.

The low-order 16 bits of the VINES internet address are the subnetwork number. It identifies each node within the logical network. Each subnetwork number must be unique within the logical network.

VINES servers and VINES routers have a subnetwork number equal to X′ 0001′. VINES clients have a subnetwork number in the range X′ 8000′ to X′ FFFE′.

4.1.9.5 The VINES IP Packet Format

The VINES IP packet format is shown in Figure 66.

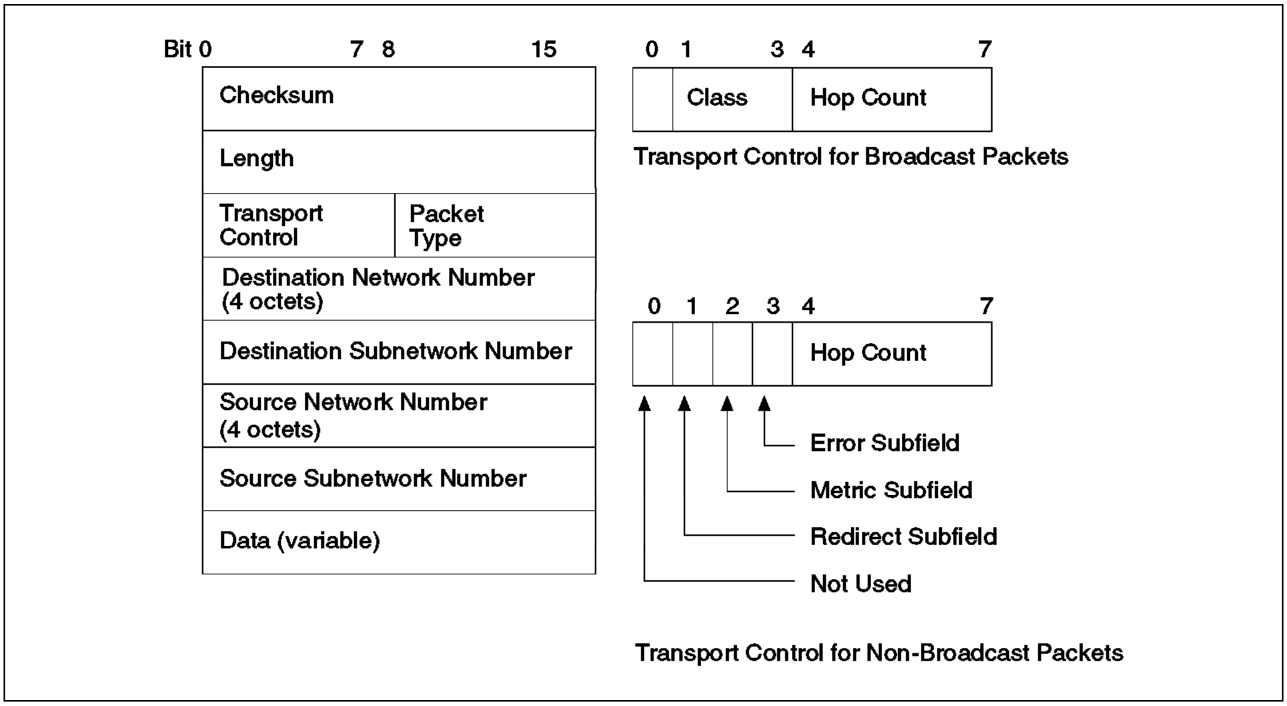


Figure 66. VINES IP Packet Format

Field — **Explanation**

Checksum
This 2-octet field is used by VINES IP to detect bit-error corruption of the packet.

Length
This 2-octet field contains the length of the packet, including the length of the VINES IP header.

Transport Control
This 1-octet field depends on whether the packet is a broadcast packet. For broadcast packets, VINES IP uses bits 1 through 3 for the class subfield. For other packets, VINES IP uses bits 1, 2, and 3 for the error, metric, and redirect subfields, respectively.

Class Subfield

VINES IP uses the class subfield to filter broadcasts on transmission. VINES IP uses the class subfield with the hop count to qualify the range that a broadcast packet can travel. This reduces the number of extraneous broadcast packets that a node receives.

Error Subfield

The error subfield indicates that the packet's originator wants to receive exception notification. If a VINES IP entity cannot route a packet for any reason, and the error subfield is set to one, the entity requests ICP to send an exception notification packet to the transport layer protocol entity from which the packet originated. The packet tells the entity that the packet cannot be routed.

Metric Subfield

Transport layer protocol entities on routers set the metric subfield to one when they need to learn the routing cost of moving packets between another router and one of its neighbor client nodes. The routing cost is referred to as a metric.

Redirect Subfield

When the redirect subfield is set to one, the last router that forwarded the packet is capable of sending and accepting redirect packets.

Hop Count Subfield

The hop count subfield can contain any value from 0x0 to 0xF. The source node initializes the hop count subfield to the maximum number of hops or routers that can handle the packet.

Packet Type

This 1-octet field indicates the network or transport layer protocol entity destination.

Destination and Source Address Fields

The rest of the VINES IP header fields contain the VINES internet addresses of the source of the packet and the destination of the packet. The source address follows the destination address in the order of transmission.

4.1.9.6 Flow Control and Acknowledgments

In the VINES architecture, flow control and acknowledgements are provided by the transport layer protocols. The SPP data stream service provides both flow control and acknowledgements. The IPC reliable message service provides acknowledgements, but not flow control. The IPC datagram service provides neither.

4.1.9.7 Features Routers Provide for VINES

A router provides the following features for VINES:

- Configurable interface metrics
- Periodic full routing updates on serial links
- Serverless option

Configurable Interface Metrics: A router allows you to override a port′s default interface metric. You may wish to choose your own interface metric in several different circumstances. On Ethernet ports, you may want to increase the interface metric to compensate for the delay through transparent bridges. On any port, you may want to adjust an interface metric in order to favor one route over another.

Routing Updates on Serial Links: A router provides two methods for sending routing updates on serial links.

The default method adheres to the RTP specification in the VINES Protocol Definition for sending routing updates on serial links. In accordance with this specification, a router sends three full routing updates at 90-second intervals when a serial link is first opened. After these initial updates, the link becomes permanent. Thereafter, the router asynchronously sends delta routing updates on the serial link only after it detects a change in the network topology. A delta update contains information only about the change to the network topology.

The optional method does not mark the serial link as permanent and continues to send periodic full routing updates on the link every 90 seconds, even if the network topology has not changed.

Serverless Option: The VINES architecture allows for a client to be no more than one hop away from its designated server. A router provides a serverless option that allows a client to be more than one hop away.

You enable the serverless option on any port that receives the broadcasts a client sends in search of a server. Then you specify the number of hops you want the broadcasts to travel using the Distance to Server parameter. The number of hops includes a router as one of the hops.

With the serverless option enabled, a router looks for the client′s broadcasts. When it receives one of these broadcasts, the router:

- Sets the packet′s hop count to one less than the value configured for the Distance to Server parameter.
- Changes the packet′s class subfield value to ″all reachable routers, regardless of cost″.
- Rebroadcasts the packet onto every LAN and serial port enabled for VINES that has a VINES router or VINES server neighbor (except the port over which the packet was received).

If a router receives a packet that does not meet the serverless option criteria, the router handles the packet as a normal broadcast in the manner specified by VINES IP.

If you attach a router to the client′s LAN, you need to enable the serverless option only on the router port attached to that LAN. In the example shown in Figure 67 on page 187, the router 6611 has the serverless option enabled on its token-ring port with a serverless hop count of two.

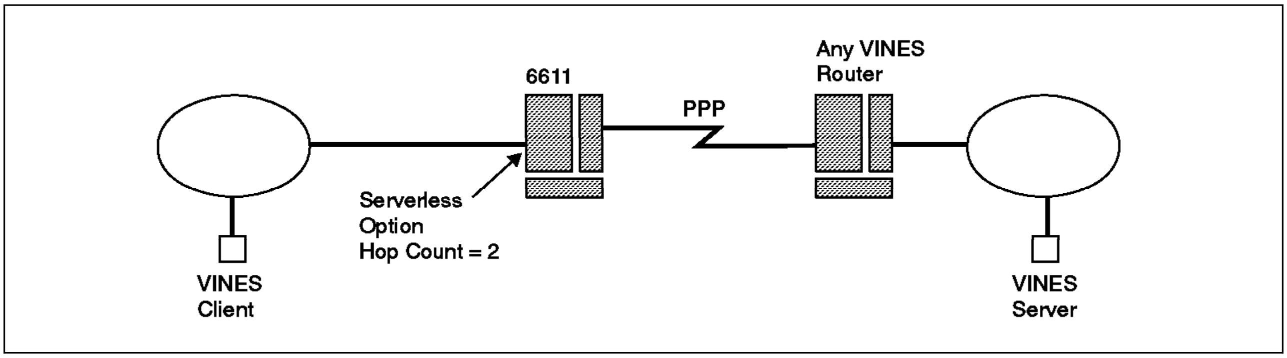


Figure 67. Serverless Client Attached to the 6611

You may also be able to interconnect the router 6611 when another vendor′s router is attached to the client. In the example shown in Figure 68, the router 6611 has the serverless option enabled on its serial port with a serverless hop count of one. The non-6611 router would have its serverless function enabled according to its manufacturer′s instructions.

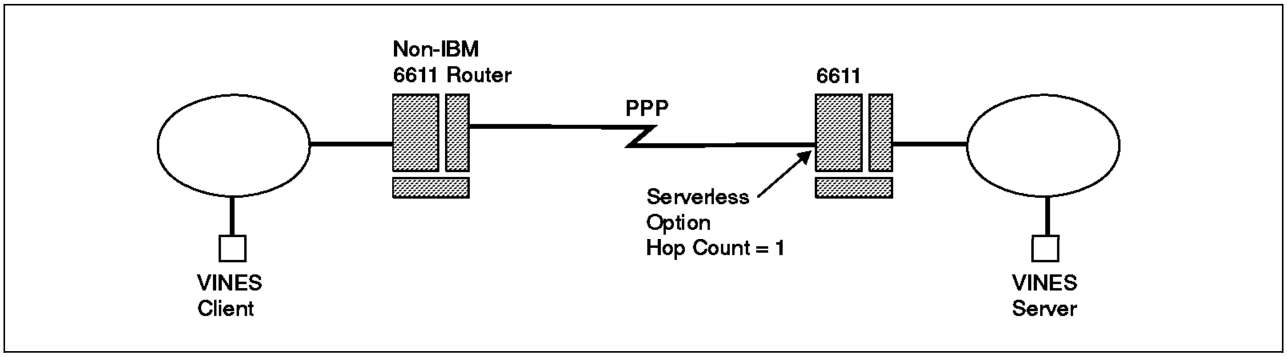


Figure 68. Serverless Client Not Attached to the 6611

For the 6611′s serverless option to work with the non-6611′s equivalent, the non-6611 router must transfer the client′s zero hop count broadcasts unaltered to the serial link. (Although, the class subfield may be altered to be any value greater than four.)

4.1.9.8 VINES Filtering

On the system level, VINES allows you to define filters to control RTP packets. On the port level, VINES allows you to define filters which act on the VINES IP header information or the transport layer header information contained in VINES packets.

System-Level Filters: VINES provides the following system-level filters:

- RTP router filters

  RTP router filters determine from which VINES network nodes a router accepts RTP packets. RTP router filters operate on the source network number field in the VINES IP header of the VINES RTP update, response, and redirect packets.

- Inbound RTP filters

  Inbound RTP filters determine which networks received in an RTP packet are retained in a router′s routing database. Inbound RTP filters operate on the topological entries contained in the VINES RTP update, response, and redirect packets.

- Outbound RTP filters

  Outbound RTP filters determine which networks a router advertises in RTP packets. Outbound RTP filters operate on the topological entries contained in the VINES RTP update and response packets.

The VINES system-level filters are composed of the following elements:

- Filter ID

  The filter ID contains a unique number from 1 to 128 that identifies a specific filter. This filter ID does not determine the order in which filters are applied.

- Filter type

  VINES system-level filters are always singular.

- Filtering mode

  The filtering mode determines whether a router processes or discards filtered traffic.

  If you select Permit as the filtering mode, then all traffic that matches the defined filters will be processed. All other traffic will be discarded.

  If you select Deny as the filtering mode, then all traffic that matches the defined filters will be discarded. All other traffic will be processed.

Table 16. Algorithm Used for VINES System-Level Filters

A Packet with Applied with This Result...	Yields This Outcome...	Filtering Mode of...
Deny	Filter and packet contents match.	Packet is discarded; no other filters are applied to it.
Deny	Filter and packet contents do not match.	Packet is passed to the next filter, until all filters have been applied. If the packet does not match any filters, the packet is then processed.
Permit	Filter and packet contents match.	Packet is processed; no other filters are applied to it.
Permit	Filter and packet contents do not match.	Packet is passed to the next filter, until all filters have been applied. If the packet does not match any filters, the packet is then discarded.

- Filter address

  VINES system-level filters use the eight digit hexidecimal VINES network number. The subnetwork number is not used.

- Filter mask

  VINES system-level filters allow you to specify either all network numbers or one network number.

VINES inbound and outbound RTP filters also allow you to specific the port to which the filter is applied.

Port-Level Filters: VINES provides the following port-level filters:

- Inbound port filters

  Inbound port filters are applied to traffic entering a router from either a LAN or a WAN interface.

- Outbound port filters

  Outbound port filters are applied to traffic leaving a router through either a LAN or a WAN interface.

VINES port-level filters are composed of the following elements:

- Filter ID

  The filter ID contains a unique number from 1 to 10 that identifies a specific filter. This filter ID does not determine the order in which filters are applied.

- Filter type

  VINES port-level filters are always singular.

- Filtering mode

  The filtering mode determines whether a router processes or discards filtered traffic.

  If you select Permit as the filtering mode, then all traffic that matches the defined filters will be processed. All other traffic will be discarded.

  If you select Deny as the filtering mode, then all traffic that matches the defined filters will be discarded. All other traffic will be processed.

Table 17. Algorithm Used for VINES Port-Level Filters

A Packet with Applied with This Result...	**Yields This Outcome...**	**Filtering Mode of...**
Deny	Filter and packet contents match.	Packet is discarded; no other filters are applied to it.
Deny	Filter and packet contents do not match.	Packet is passed to the next filter, until all filters have been applied. If the packet does not match any filters, the packet is then processed.
Permit	Filter and packet contents match.	Packet is processed; no other filters are applied to it.
Permit	Filter and packet contents do not match.	Packet is passed to the next filter, until all filters have been applied. If the packet does not match any filters, the packet is then discarded.

- Filter address VINES port filters are based on the following fields in the:
  - VINES IP header:
    - Hop count
    - Protocol type
    - Destination network
    - Destination subnetwork

  - Source network
  - Source subnetwork
- Transport layer header:
  - Destination port number
  - Source port number

4.2 Multiprotocol Transport Networking (MPTN)

With the growth of networking in general and local area networks in particular, it is not uncommon to see configurations with many different networks using four or five different protocols, such as TCP/IP, NetBIOS, SNA, IPX/SPX or AppleTalk. One of the problems in these environments is interoperability, since applications that run on one network often do not run with applications on other networks. The interoperability problem is commonly circumvented by writing application gateways. However, each gateway handles only one specific application. If applications are transport-independent, installing general transport gateways between networks would solve the problem. To provide this solution, you need the capability to write an application once and have it run on different networks. Also, you need the capability to create a logical connection across heterogeneous networks that matching applications use. Besides interoperability, applications need to be transport independent because of the increased use of single-protocol endpoints and migration to new transport networks. In most workstations, it is desirable to install one efficient communications protocol and be able to run all applications.

The Networking Blueprint addresses these problems by structuring applications, applications interfaces and support protocols, transport network and subnetworking layers. Support for multiple protocols can be achieved in a number of ways. The Networking Blueprint provides the context for discussing them. The Multiprotocol Transport Network (MPTN) is a solution framed within the blueprint. It is defining an interface to a set of transport services that concatenate connections across multiple networking protocols. In addition to logical connections across heterogeneous networks, MPTN separates applications from networks by providing a common transport interface so that messages from the application can be transported over any protocol under the interface. Thus, MPTN is like a group of single-protocol transport networks (those that have a group of nodes, as in a NetBIOS LAN, that are physically connected and implement the same network protocol) each of which has its own transport protocol. MPTN appears as a single logical network having a single protocol. This appearance is provided by two aspects of MPTN: The Common Transport Semantics (CTS) and MPTN gateways.

In order to build a multiprotocol network, several functions are needed:

- Function Compensation

  With MPTN, the transport user selects the transport services that it wants without concern for what the underlying transport provider supports. If a transport provider does not support a service requested by a transport user, MPTN defines function compensation to provide the transport user′s service. The compensation bridges the gap between the needs of the transport user and the services provided by the underlying transport provider.

- Address Mapping

One MPTN function allows transport users to continue to use accustomed transport address formats, even if a transport provider uses a different address format. MPTN provides various techniques for mapping from the transport addresses provided by the transport users to those required by non-matching transport providers. An example of a mapping is SNA LU name to TCP/IP address.

- Transport Gateways

  The MPTN transport gateway allows two different transport networks to be concatenated so that they appear to the user as one logical network. The transport gateway relays data received from one network into the other network to provide end-to-end service. A key function of an MPTN transport gateway is to connect individual native transport networks to an integrated heterogeneous network. Another key function is the ability to interconnect more than two transport networks.

4.2.1 The Common Transport Semantics (CTS)

The CTS provides a semantic interface so that higher-level protocols or application interfaces written for a particular transport protocol can be transported over another protocol with no apparent change.

4.2.1.1 Function Compensation

The Common Transport Semantics (CTS) achieves multiprotocol networking at layer 4, the transport layer. CTS includes all of the functions in the underlying transport providers in the Networking Blueprint. If needed functions are missing from any of the transport providers, CTS itself provides those functions. CTS functions can be delivered in different ways depending upon the situation:

- Same protocol

  CTS includes the traditional case where the transport user protocol matches the transport provider protocol. The transport user protocol is the protocol, that an application wants to use. The transport provider protocol is the protocol which is provided to the CTS. When the characteristics supported by the transport provider match exactly those required by the transport user, the connection is said to be native. No changes are made to the protocol flows. A same protocol flow can be thought of as an invisible line through the CTS boundary.

- Mixed protocols

  – Existing standards

    CTS function can be delivered in specific situations by accepted industry standards, such as RFC 1006 for OSI over TCP/IP or RFCs 1001 and 1002 for NetBIOS over TCP/IP.

    With the RFC implementations, specific compensations are provided when the transport user does not match the transport provider. For example, with RFC 1006 OSI over TCP/IP, one set of compensations are provided, and with RFCs 1001 and 1002 for NetBIOS over TCP/IP a second set of compensations are provided. Some of the compensations in these two sets may be for the same missing function, but they are implemented differently and specifically.

  – MPTN

    Multiprotocol Transport Networking (MPTN) delivers CTS function in situations where the transport user protocol does not match the transport provider protocol. MPTN provides a standard common set of compensations for missing functions which can be used by any protocols.

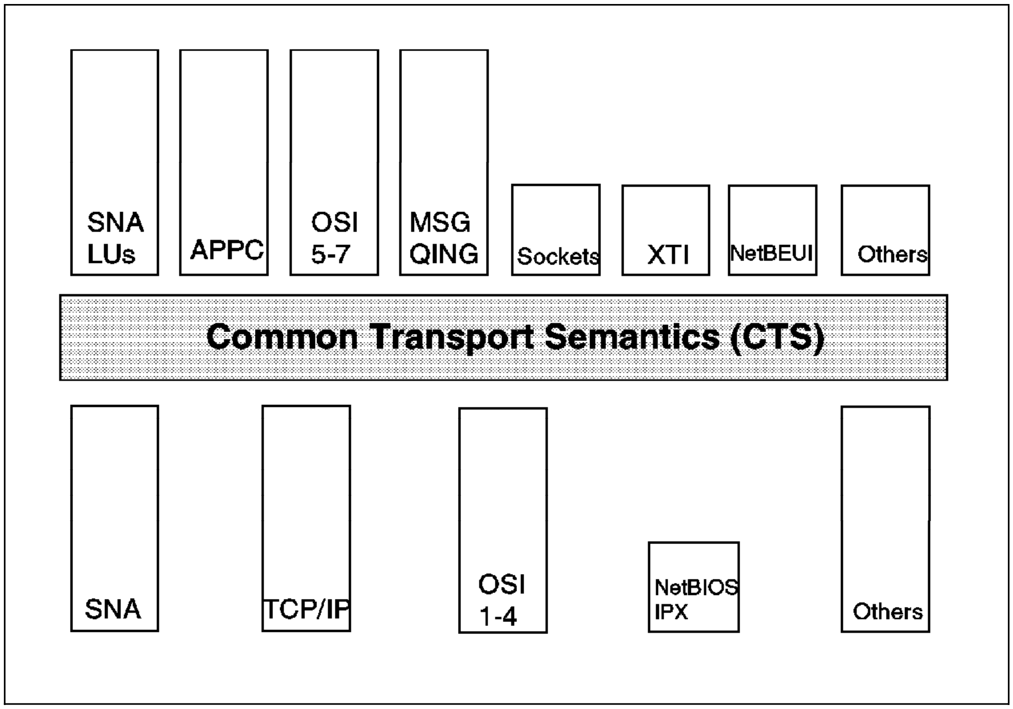


Figure 69. Common Transport Semantics

Support for multiple protocols can also be achieved in different ways outside of CTS, such as through multiprotocol routers like the IBM Multiprotocol Network Processor or encapsulation (enveloping) like IBM′s LAN-to-LAN via WAN program (LTLW).

Multiprotocol routers operate in the transport portion of the Networking Blueprint. Routers are meant to adapt to whatever protocols exist in the networks being connected. The difficulty with this type of solution is that routers do not reduce the number of protocols in the overall network. Even though they can enable consolidation of physical resources in the backbone network, (for example, reduce the number of lines and boxes) the exterior networks (and their workstations or end nodes) continue to require full support for all protocols that their applications require.

Encapsulation is the generic technique typically used to ″tunnel″ traffic from one network type through another; it involves sending link-level packets as user data. With encapsulation, two full networks are present, and data is transported through two entire protocol stacks. The set of headers from one protocol is imbedded in a full set of headers from a second protocol. In contrast, with MPTN two full networks are not required. Only the upper layers of one protocol need be present along with the lower layers of another protocol. The data is transported using the upper layer headers and the non-matching lower layer headers, along with the MPTN header with compensation information. MPTN has a potential performance advantage since two full protocol stacks do not have to be traversed. The following figure shows native SNA and SNA over TCP/IP using MPTN.

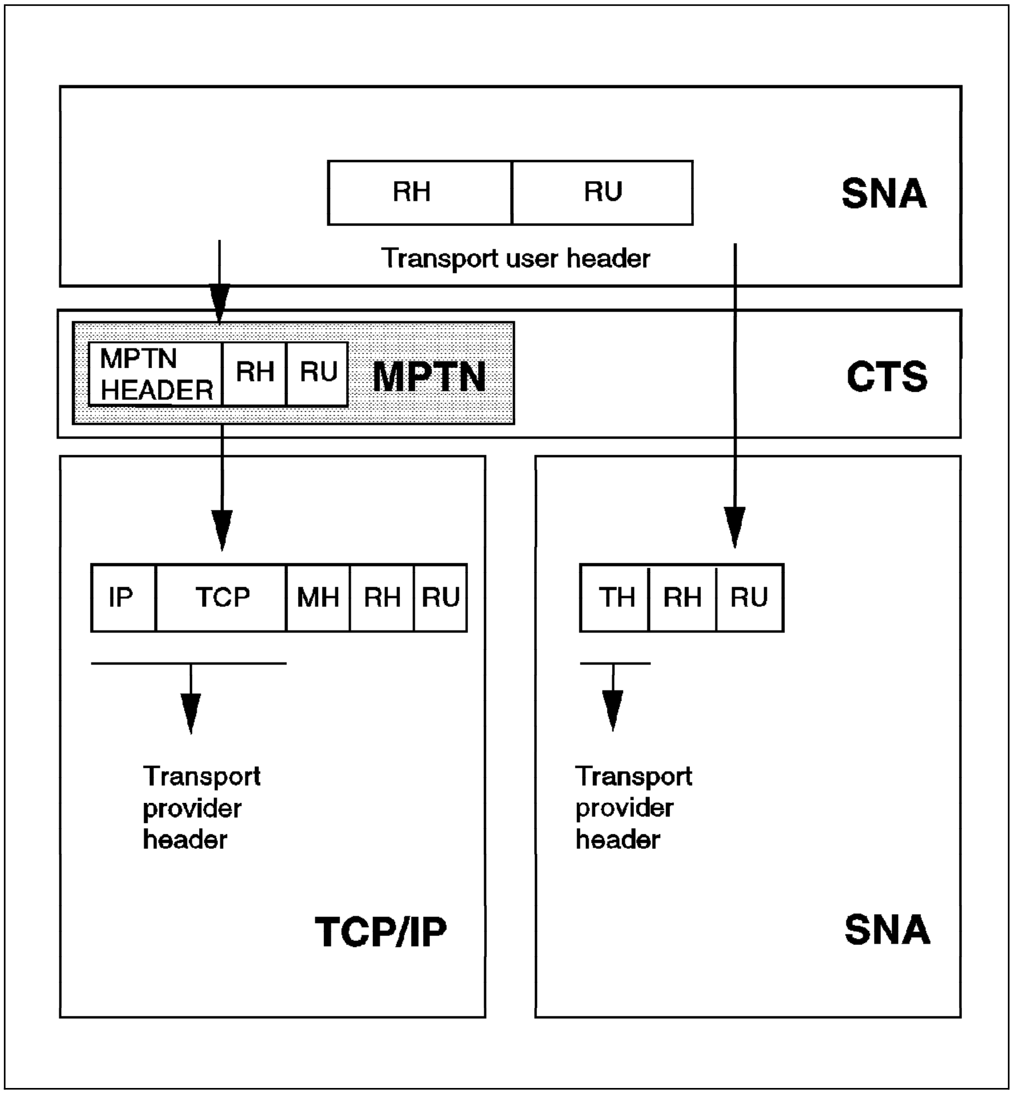


Figure 70. Data Transport (Native and Non-Native)

4.2.1.2 Address Mapping

If the transport user and transport provider have different addressing schemes, address mapping is required. Since MPTN supports connectivity across heterogeneous networks it must deal with different address formats of the network protocols it supports. In order to avoid disruption to either transport users or transport providers, MPTN allows a transport user to use its existing transport address format, while the serving transport provider uses transport addresses that its network expects. MPTN provides an address mapping between a transport user′s address and a transport provider′s address when different addressing schemes are used. For example, TCP/IP uses a 32-bit internetwork address and SNA uses a 17 character name.

MPTN has a component that maps between the transport addresses used by the transport users and those understood by non-matching transport providers. MPTN performs these address mapping functions dynamically:

- A transport user registers an address, making itself accessible to other transport users.

- MPTN locates a partner, given a transport user address. This allows a user to set up a non-native connection with another transport user.

The MPTN architecture uses three alternative approaches to address mapping as the lower cost solutions do not work work for all cases:

- With algorithmic mapping, the transport provider implements an algorithm that generates a native address in the provider′s address space from the transport user′s non-native address. This method is appropriate when the user address space is smaller than the provider′s address space. An example is internetwork addresses mapped to SNA names, since IP address space is smaller than the SNA name space.
- Another alternative is to enhance protocol-specific directories to handle transport addresses of various formats. The protocol-specific directory of the transport provider holds a mapping that correlates the provider address to the user address and user protocol. This method is appropriate when the transport provider′s directory supports the registration of different address types. An example is registering an SNA LU name in TCP′s domain name server.
- When dynamic directories are required, or if multiprotocol combinations are used, then the address mapper option can be used. The address mapper is a database which holds the transport user to transport provider mappings. An example is mapping an SNA name to an IPX MAC address.

4.2.1.3 Transport Gateways

MPTN Transport Gateways support different gateway functions:

- Interconnection of two networks with different transport protocols
- Support of native systems
- Support of parallel gateways

If a network consists of only a single protocol transport network, only MPTN access nodes and address mapping services are needed to run applications written for other transport protocols. In environments that include multiple networking protocols, the MPTN transport gateway function allows networks with two different network protocols to be concatenated so that they appear to the user as one logical network. A single connection from one protocol maps to a single connection over a different protocol. No multiplexing of connections is done which makes network management easier.

The MPTN transport gateway takes care of resource location in the MPTN network, routes connections and datagrams through the MPTN network, sets up connections through the MPTN network, and relays data from one network to the next network. The gateway relay function transports data received from the incoming network to the outgoing network in order to provide end-to-end service.

An MPTN transport gateway is able to exchange data between native and non-native systems. It is possible to transport data from an MPTN node to a non-MPTN node over the gateway with no changes to the non- MPTN node. Connectivity is provided from a transport user running on a non-native protocol stack in the MPTN access node to a matching transport user in a native protocol stack. For example a node can run SNA over TCP/IP (using MPTN) to the MPTN gateway and native SNA from the gateway to a node with a native SNA protocol stack.

4.2.2 MPTN Implementations

IBM has different implementation of the MPTN architecture. Products differ in what protocols they support.

4.2.2.1 AnyNet

The AnyNet product family is IBM′s implementation of the MPTN architecture. AnyNet products benefit users of mainframes and personal computers on interconnected LANs and WANs by accommodating a variety of combinations of application programs and transport networks.

4.2.2.2 Multiprotocol Transport Service (MPTS)

An example for a MPTN implementation is Socket/MPTS (Multiprotocol Transport Service) that comes with the IBM OS/2 LAN Server 4.0 and Requester. Socket/MPTS supports TCP/IP and NetBIOS protocols, but does not include the MPTN gateways. Also, it is imperative that the underlying protocols are the same, that is, communicating peers must both use the same protocol.

IBM′s Multiprotocol Transport Network (MPTN) architecture is aimed at providing a general solution to interconnect applications. The protocol-independent transport interface defined in the MPTN architecture decouples higher layer protocols and application programming interfaces from protocols at the transport layer and below. The common transport semantics thus need two features: protocol compensation, which adjusts for discrepancies between services required by the common transport interface and those provided by individual transport protocols; and address mapping, which resolves the difference when the application and the protocol used to transport data use different address types.

Socket is an abstract object that is used to send and receive messages. MPTS supplies a transport framework that provides common transport semantics across TCP/IP and NetBIOS protocols. The framework is accessible through the Socket API. Socket/MPTS supplies drivers for:

- Common transport semantics
- IBM OS/2 TCP/IP protocol drivers
- Protocol compensation and address mapping for NetBIOS
- Local interprocess communication

The OS/2 transport-independent interface is provided using the socket programming interface. The following figure describes the common transport interface used to access multiple transport networks.

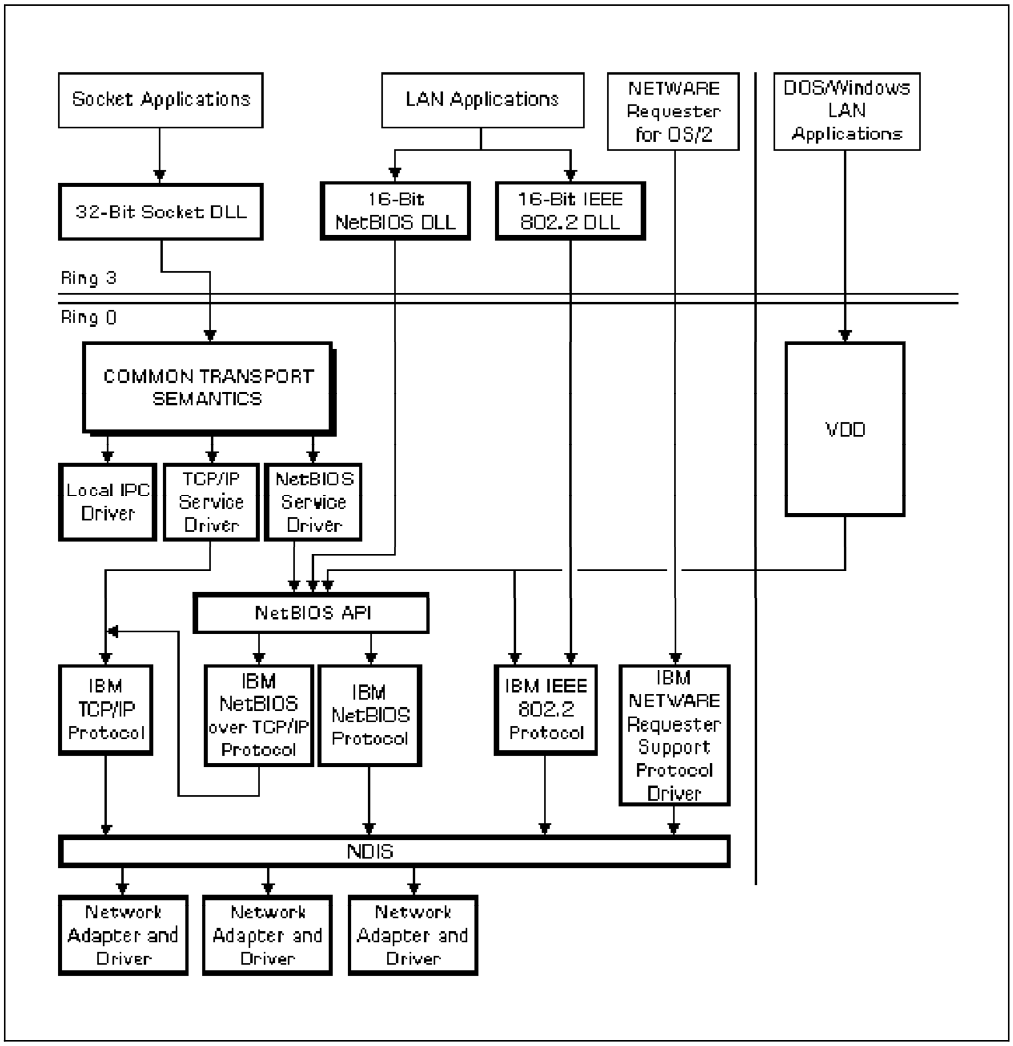


Figure 71. Multiprotolcol Transport Service (MPTS)

Socket/MPTS also provides the 32-bit socket interface.

The Socket API can be used to communicate over TCP/IP and NetBIOS protocols *natively*, which means that both the transport user and the protocol used to transport data are from the same protocol architecture. For example, if a socket created for a domain uses TCP/IP protocol, then it is termed a native communication. If, however, the socket created for a domain uses the NetBIOS protocol for communications, then it is termed INET non-native over NetBIOS. MPTS provides non-native INET support over NetBIOS, which allows TCP/UDP applications to run atop NetBIOS.

4.3 Router Protocols

Router protocols are used to exchange routing information between routers in order to maintain their routing tables. Routing tables are used to determine the correct route of data packets.

For IP routing several protocols exist to exchange IP routing information with other routers. The IP routing protocols currently supported include the following:

- RIP

- OSPF
- HELLO
- EGP
- BGP

The first three of these protocols are referred to as *interior gateway protocols* (IGPs), while the remainder are referred to as *exterior gateways protocols* (EGPs).

The term gateway is used here to mean router. IP routers are commonly called gateways in the TCP/IP world. It is not being used in the strict ISO OSI sense.

4.3.1 Interior and Exterior Gateway Protocols

Gateway protocols are referred to as interior or exterior depending on whether they are used within, or between autonomous systems (ASs).

Interior gateway protocols allow routers to exchange routing information *within* an AS.

Exterior gateway protocols allow the exchange of summary reachability information *between* separately administered ASs.

ASs are defined as logical portions of larger IP networks that are administered by single authorities. An AS would normally comprise the internetwork within an organization, and would be designated as such to allow communication over public IP networks with ASs belonging to other organizations. It is mandatory to register an organization′s internetwork as an AS in order to use these public IP services.

Figure 72 on page 198 illustrates two ASs interconnected by routers. It shows how IGPs are used within the ASs, and an EGP between them.

ASs must be registered publicly. If you require an AS number, or an IP network address, to allow your network to connect to public IP services you should contact:

Network Information Center
GSI
14200 Park Meadow Drive, Suite 200
Chantilly, VA 22021
USA
Tel: (703) 802-4535

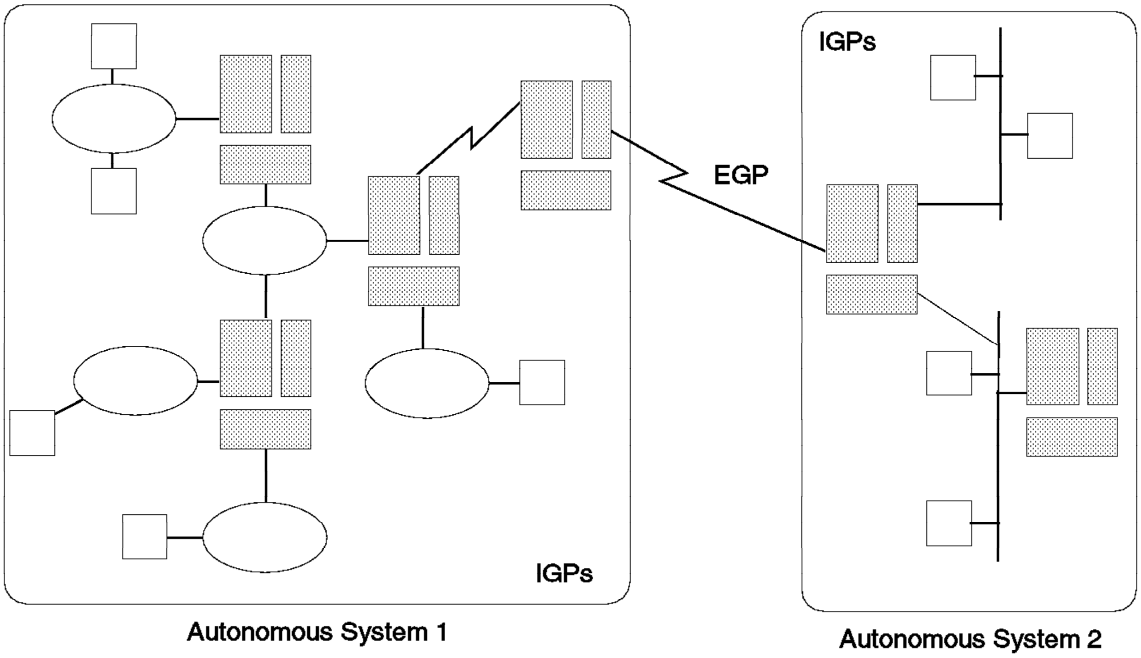


Figure 72. Autonomous Systems

4.3.2 Choosing Gateway Protocols

Within an AS (or if you are building a private IP network) you are free to choose the interior gateway protocol, or combination of protocols, that best meets your needs.

However, each interior gateway protocol has different characteristics and selection must be carried out carefully to meet internetwork design requirements.

If you wish to communicate with other ASs you are once again, in principle, free to choose the exterior gateway protocol that best meets your needs. The interior gateway protocol used within an AS is not constrained by the choice of exterior gateway protocol. However, there is synergy between some interior and exterior gateway protocols (for example, OSPF and BGP).

In practice AS to AS communication is governed by rules set by the administrators of the public Internet, or by private IP service providers. An internetwork design must accommodate the exterior gateway protocols required by the IP service provider.

4.3.3 Routing Algorithms

Interior and exterior gateway protocols currently implemented use one of two generic classes of dynamic routing algorithm.

These are known as *distance vector* and *link state* routing algorithms.

Dynamic routing algorithms allow routers to exchange route or link information, from which are calculated the best paths to reachable destinations in an internetwork.

Static routing may also be used to supplement dynamic routing.

4.3.4 Routing Information Protocol (RIP)

The Routing Information Protocol (RIP) is an interior gateway protocol defined in RFC 1058.

It is an IAB standard protocol and its status is elective. This means that it is one of several interior gateway protocols available, and it may or may not be implemented on a system. However, if a system does implement it, the implementation should be in line with the RFC.

RIP is based on the Xerox PUP and XNS routing protocols. The RFC was issued after many RIP implementations had been completed. For that reason some do not include all the enhancements to the basic distance vector routing protocol (such as poison reverse and triggered updates).

RIP is very widely used because the code (known as *ROUTED*) was incorporated on the Berkeley Software Distribution (BSD) UNIX operating system, and in other UNIX systems based on it.

4.3.4.1 Protocol Description

RIP is a standard distance vector routing protocol. RIP packets are transmitted onto a network in User Datagram Protocol (UDP) datagrams, which in turn are carried in IP datagrams. RIP sends and receives datagrams using UDP port 520. RIP datagrams have a maximum size of 512 octets. Tables larger than this must be sent in multiple UDP datagrams.

RIP datagrams are normally broadcast onto LANs using the LAN MAC All-Stations broadcast address and the IP network or subnetwork broadcast address. They are specifically addressed on point-to-point and multiaccess non-broadcast networks, using the destination router IP address.

Routers normally run RIP in *active mode*; that is, they advertise their own distance vector tables and update them based on advertisements from neighbors. End nodes, if they run RIP, normally operate in *passive (or silent) mode*; that is, they update their distance vector tables on the basis of advertisements from neighbors, but do not advertise them.

RIP specifies two packet types: *request* and *response*.

A request packet is sent by routers to ask neighbors to send part of their distance vector table (if the packet contains destinations), or their entire table (if no destinations are specified).

A response packet is sent by routers to advertise their distance vector table in the following circumstances:

- Every 30 seconds
- In response to a request packet
- When distance vector tables change (if triggered updates are supported)

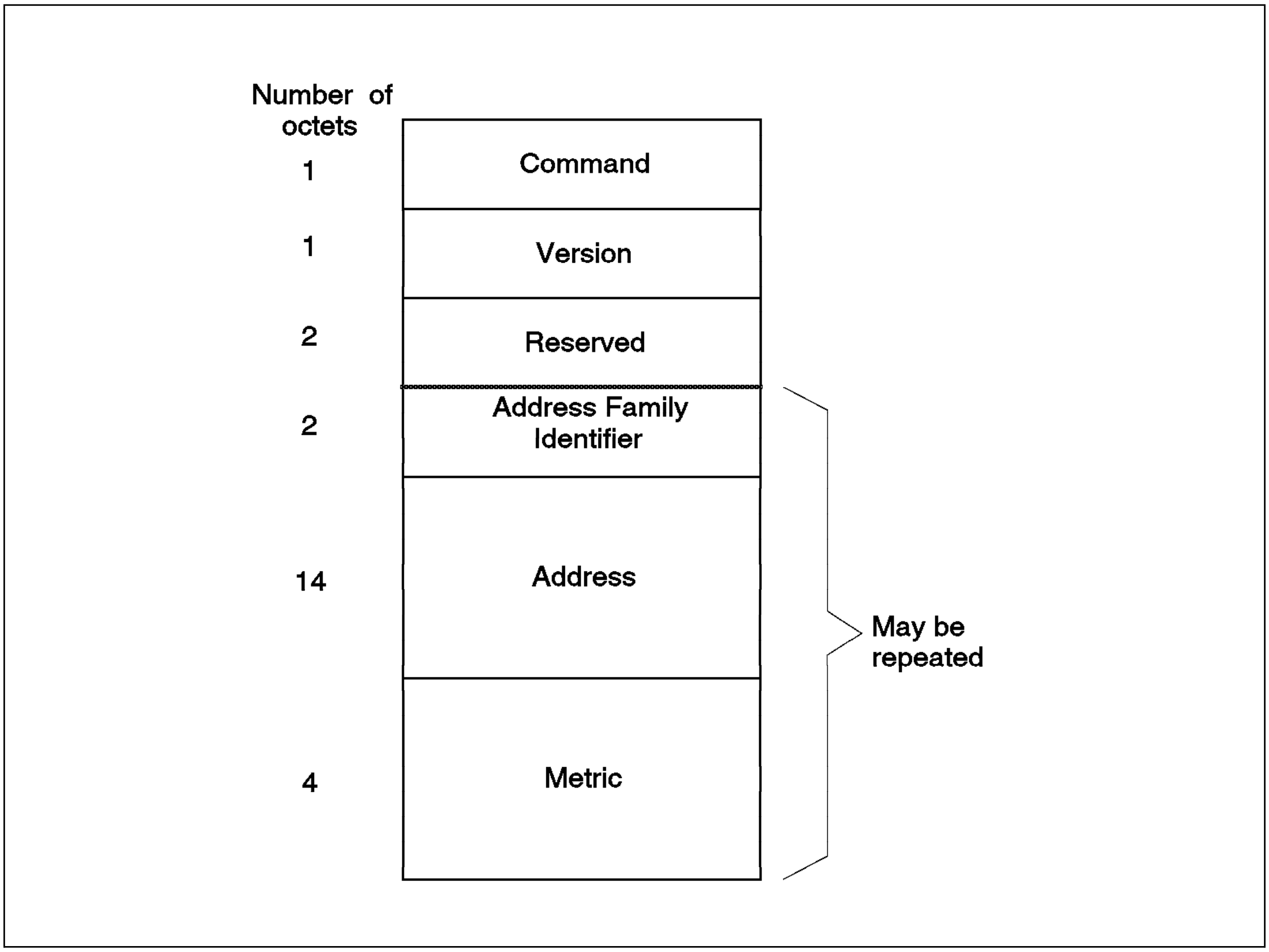


Figure 73. Generalized RIP Packet Format

Active and passive systems listen for all response packets and update their distance vector tables accordingly. A route to a destination, computed from a neighbor′s distance vector table, is kept until an alternate is found with lower cost, or it is not readvertised in six consecutive RIP responses. In this case the route is timed out and deleted.

The RFC defines a packet format that can be used with different network protocols. It does this by specifying an *address family identifier* which defines the type (and hence the length) of the network address. The generalized format of RIP packets is shown in Figure 73.

RIP may be used, therefore, for protocols other than IP, simply by setting the address family identifier. The RFC requires that RIP response handling discards entries for unsupported address families but processes entries for supported address families in the normal way.

When RIP is used with IP the address family identifier is 2, and the address fields are 4 octets. To reduce problems of counting to infinity the maximum metric is 16 (unreachable) and directly connected networks are defined as having a metric of one.

The RIP packet format for IP is shown in Figure 74 on page 201.

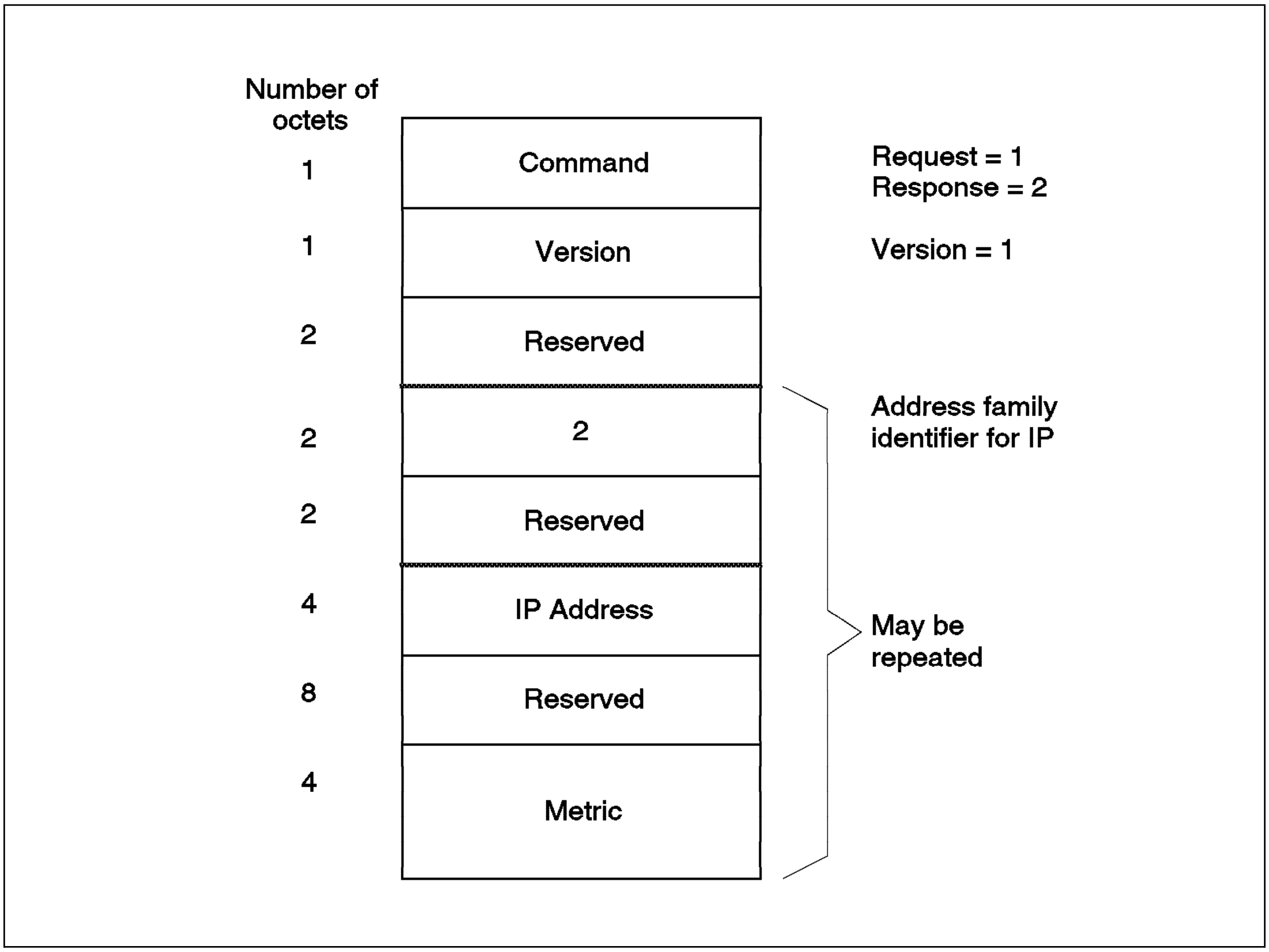


Figure 74. IP-Specific RIP Packet Format

RIP makes no provision for passing subnet masks with its distance vector tables. A router receiving a RIP response must already have subnet mask information to allow it to interpret the network identifier and host identifier portions of the IP address correctly.

In the absence of subnet mask information a router will interpret routes as best it can. If it knows an IP network has a specific subnet mask it will interpret all other route information for that network on the basis of that single mask. If it receives a packet with bits set in the field that it regards as the host field it will interpret it as a route to a host with a mask of *255.255.255.255*.

The above makes it impossible for RIP to be used in an internetwork with variable length subnet masks.

4.3.5 Open Shortest Path First (OSPF)

The Open Shortest Path First (OSPF) V2 protocol is an interior gateway protocol defined in RFC 1247. A report on the use of OSPF V2 is contained in RFC 1246 *Experience with the OSPF Protocol*.

It is an IAB standard protocol; its status is elective.

OSPF is important because it has a number of features not found in other interior gateway protocols. Support for these additional features makes OSPF the preferred choice for new IP internetwork implementations:

- Variable length subnet masks
- Alternate routes based on IP Type of Service (TOS)
- Equal cost multipath routes

OSPF uses specific terminology which must be understood before the protocol can be described.

Areas: OSPF internetworks are organized into *areas*.

An OSPF area consists of a number of networks and routers that are logically grouped together. Areas may be defined on a per location or a per region basis, or they may be based on administrative boundaries.

All OSPF networks consist of at least one area, the backbone, plus as many additional areas as are demanded by network topology and other design criteria.

Within an OSPF area all routers maintain the same topology database, exchanging link state information to maintain their synchronization. This ensures that all routers calculate the same network map for the area.

Information about networks outside an area is summarized by *area border* or *AS boundary routers* (see below “Intra-Area, Area Border and AS Boundary Routers”) and flooded into the area. Routers within an area have no knowledge of the topology of networks outside the area, only of routes to destinations provided by area border and AS boundary routers.

The importance of the area concept is that it limits the size of the topology database that must be held by routers. This has direct impact on the processing to be carried out by each router, and on the amount of link state information that must be flooded into individual networks.

The OSPF Backbone: All OSPF networks must contain at least one area, the *backbone*, which is assigned an area identifier of *0.0.0.0*.

The backbone has all the properties of an area, but has the additional responsibility of distributing routing information between areas attached to it.

Normally an OSPF backbone should be contiguous, that is, with all backbone routers attached one to another. This may not be possible because of network topology, in which case backbone continuity must be maintained by the use of *virtual links*.

Virtual links are backbone router-to-backbone router connections that traverse a non-backbone area.

Routers within the backbone operate identically to other intra-area routers and maintain full topology databases for the backbone area.

Intra-Area, Area Border and AS Boundary Routers: There are three possible types of routers in an OSPF network.

Routers that are situated entirely within an OSPF area are called *intra-area routers*. All intra-area routers flood router link advertisements into the area to define the links to which they are attached. If elected designated or backup designated router (see “Designated and Backup Designated Router” on page 204) they also flood network link advertisements to define the identity of all routers attached to the network. Intra-area routers maintain a topology database for the area in which they are situated.

Routers that connect two or more areas are referred to as *area border routers*. Area border routers maintain topology databases for each area to which they are attached, and exchange link state information with other routers in those areas. Area border routers also flood summary link state advertisements into each area to inform them of inter-area routes.

Routers that are situated at the periphery of an OSPF internetwork and exchange reachability information with routers in other ASs using exterior gateway protocols are called *AS boundary routers*. Routers that import static routes or routes from other IGPs, such as RIP or HELLO, into an OSPF network are also AS boundary routers. AS boundary routers are responsible for flooding AS external link state advertisements into all areas within the AS to inform them of external routes.

Figure 75 shows the location of intra-area, area border and AS boundary routers within an OSPF internetwork.

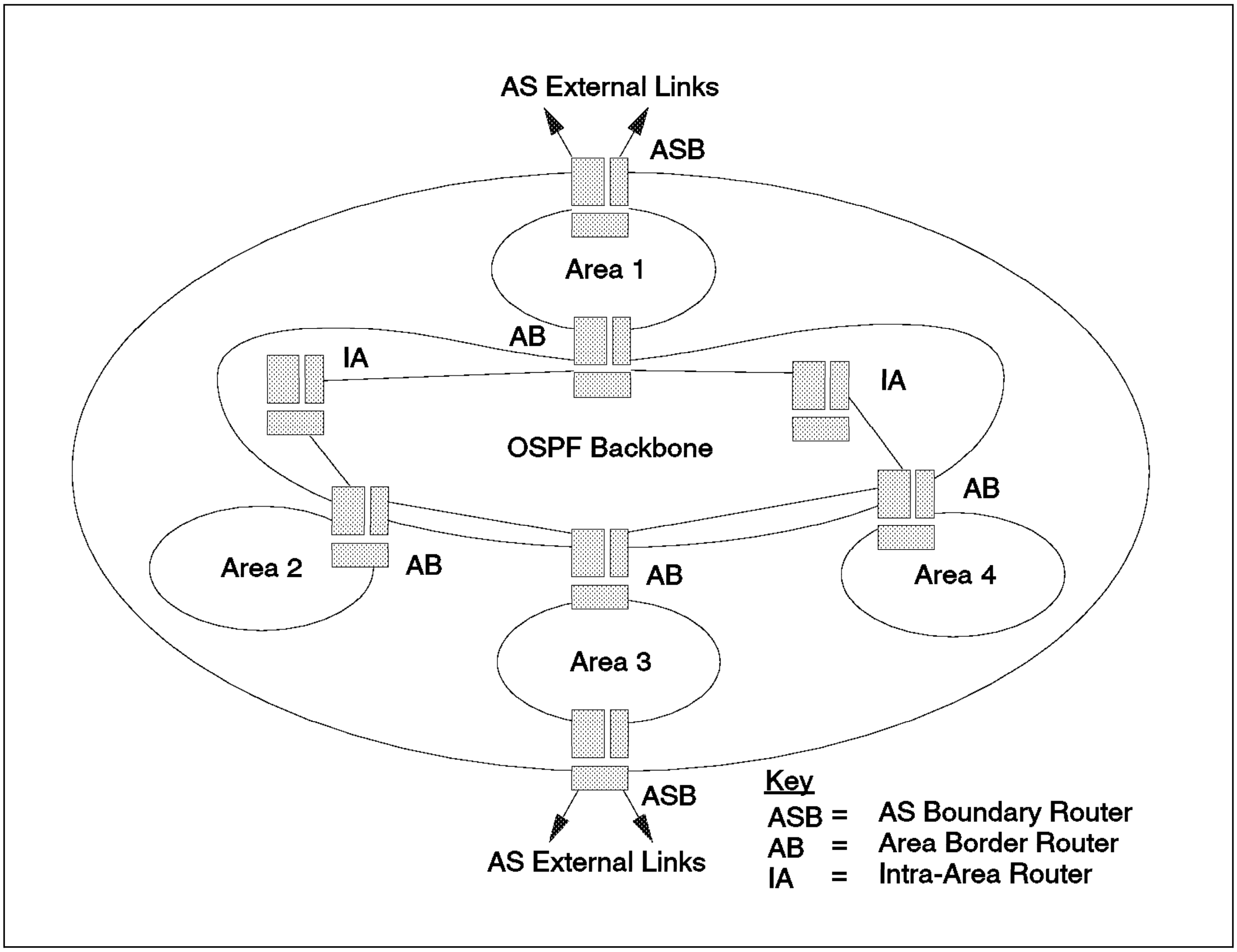


Figure 75. OSPF Network

Neighbor Router: Two routers that have interfaces to a common network are said to be *neighbors*.

Neighbor routers are discovered by the OSPF Hello protocol, which is described in “Discovering Neighbors - The OSPF Hello Protocol” on page 207.

Adjacent Router: Neighbor routers may become *adjacent*. They are said to be adjacent when they have synchronized their topology databases through the exchange of link state information.

Link state information is exchanged only between adjacent routers, not between neighbor routers.

Not all neighbor routers become adjacent. Neighbors on point-to-point links do so; on multiaccess networks, adjacencies are only formed between individual routers and the designated and backup designated routers.

The exchange of link state information between neighbors can create significant amounts of network traffic. Limiting the number of adjacencies on multiaccess networks in this way achieves considerable reductions in network traffic.

Designated and Backup Designated Router: All multiaccess networks have a *designated* and a *backup designated* router.

These routers are elected automatically for each network once neighbor routers have been discovered by the HELLO protocol.

The designated router performs two key roles for a network:

- It generates network link advertisements that list the routers attached to a multiaccess network.
- It forms adjacencies with all routers on a multiaccess network and therefore becomes the focal point for forwarding all link state advertisements.

The backup designated router forms the same adjacencies as the designated router. Therefore, it has the same topology database and is able to assume designated router functions should it detect that the designated router has failed.

Point-to-Point and Multiaccess Networks: All OSPF areas consist of aggregates of networks linked by routers. OSPF categorizes networks into two different types: point-to-point and multiaccess.

Point-to-point networks directly link two routers. OSPF packets on a point-to-point network are multicast to the neighbor router. *Multicasting* is the term used for transmitting IP datagrams to a functional rather than a specific IP address. A functional address will typically be recognized by a number of systems and can be considered a form of limited broadcast. The OSPF RFC defines the use of two multicast addresses for OSPF router interactions.

Multiaccess networks are those which support the attachment of more than two routers. They are further subdivided into two types:

- Broadcast
- Non-broadcast

Broadcast networks have the capability of directing OSPF packets to all attached routers, using an address that is recognized by all of them. An Ethernet LAN is an example of a broadcast multiaccess network.

Non-broadcast networks do not have this capability and all packets must be specifically addressed to routers on the network. This requires that routers on a non-broadcast network be configured with the addresses of neighbors. An X.25 public data network is an example of a non-broadcast multiaccess network.

Link State Advertisements: Link state information is exchanged by adjacent OSPF routers to allow area topology databases to be maintained, and inter-area and inter-AS routes to be advertised.

Link state information consists of five types of *link state advertisement*. Together these provide all the information needed to describe an OSPF network and its external environment:

1. Router links
2. Network links
3. Summary links (type 3 and 4)
4. AS external links

Router links advertisements are generated by all OSPF routers and describe the state of the router′s interfaces (links) within the area. They are flooded throughout a single area only.

Network links advertisements are generated by the designated router on a multiaccess network and list the routers connected to the network. They are flooded throughout a single area only.

Summary links advertisements are generated by area border routers. There are two types; one describes routes to destinations in other areas, and the other routes to AS boundary routers. They are flooded throughout a single area only.

AS external links advertisements are generated by AS boundary routers and describe routes to destinations external to the OSPF network. They are flooded throughout all areas in the OSPF network.

4.3.5.1 Protocol Description

The OSPF protocol is an implementation of a *link state* routing protocol. OSPF packets are transmitted directly in IP datagrams. IP datagrams containing OSPF packets can be distinguished by their use of *protocol identifier 89* in the IP header. Therefore, OSPF packets are not contained in TCP or UDP headers. OSPF packets are always sent with IP *type of service* set to *0*, and the IP *precedence field* set to internetwork control. This is to aid them in getting preference over normal IP traffic.

Further details of IP protocol identifiers, type of service and precedence can be found in RFC 791 *Internet Protocol*.

OSPF packets are sent to a standard multicast IP address on point-to-point and broadcast networks. This address is *224.0.0.5*, referred to as *All SPF Routers* in the RFC. They are sent to specific IP addresses on non-broadcast networks using neighbor network address information that must be configured for each router.

All OSPF packets share a common header which is shown in Figure 76 on page 206.

This header provides general information such as area identifier and originating router identifier, and also includes a checksum and authentication information. A type field defines each OSPF packet as one of five possible types:

1. Hello
2. Database Description
3. Link State Request
4. Link State Update
5. Link State Acknowledgment

The router identifier, area identifier, and authentication information are configurable for each OSPF router.

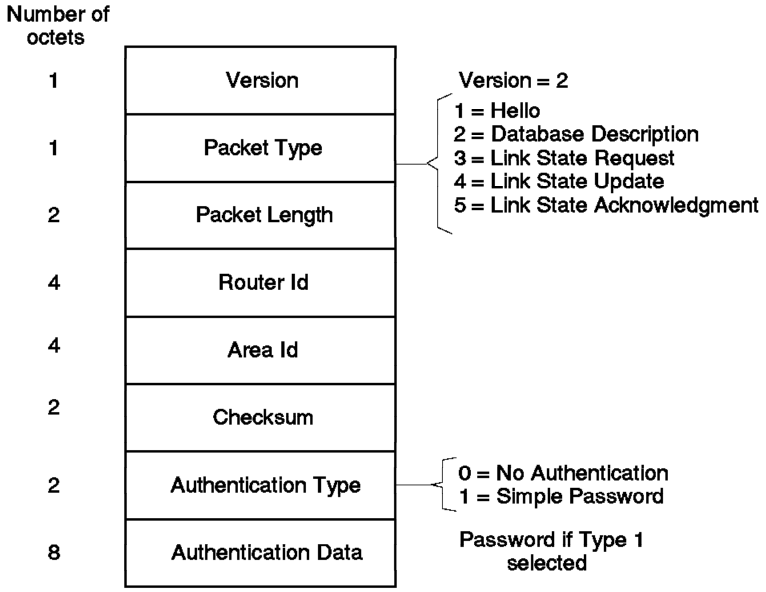


Figure 76. OSPF Common Header

The OSPF protocol defines a number of stages that must be executed by individual routers. They are as follows:

- Discovering neighbors
- Electing the designated router
- Initializing neighbors
- Propagating link state information
- Calculating routing tables

The use of the five OSPF packet types to implement stages of the OSPF protocol is described in the following subsection.

During OSPF operation, a router cycles each of its interfaces through a number of states from *Down*, through *Waiting*, to *DR Other*, *BackupDR* or *DR* (DR stands for *designated router*) depending on the status of each attached network and the identity of the designated router elected for each of them. A detailed description of these states is outside the scope of this document, but can be found in RFC 1247.

At the same time, a router cycles each neighbor interface (interaction) through a number of states as it discovers them and then becomes adjacent. These states are *Down*, *Attempt*, *Init*, *2-Way*, *ExStart*, *Exchange*, *Loading* and *Full*. Once again a description of these is outside the scope of this document but can be found in the RFC.

Discovering Neighbors - The OSPF Hello Protocol: The hello protocol is responsible for discovering neighbor routers on a network, and establishing and maintaining relationships with them.

Note

The OSPF hello protocol is entirely separate from the HELLO protocol described in 4.3.6, “HELLO” on page 216. They should not be confused.

Hello packets are sent out periodically on all router interfaces. The format of these is shown in Figure 77.

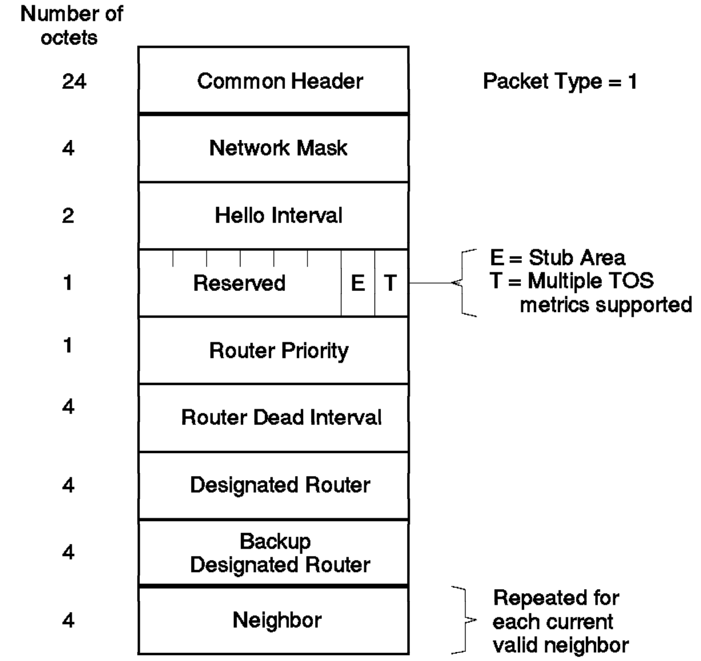


Figure 77. OSPF Hello Packet

Hello packets contain the identities of neighbor routers whose hello packets have already been received over a specific interface. They also contain the network mask, router priority, designated router identifier and backup designated router identifier. The final three parameters are used to elect the designated router on multiaccess networks.

The network mask, router priority, hello interval and router dead interval are configurable for each interface on an OSPF router.

A router interface changes state from Down to Point-to-Point (if the network is point-to-point), to DR Other (if the router is ineligible to become designated router), or to Waiting as soon as hello packets are sent over it.

A router receives hello packets from neighbor routers via its network interfaces. When this happens the neighbor interface state changes from Down to Init. Bidirectional communication is established between neighbors when a router sees itself listed in a hello packet received from another router. Only at this point are the two routers defined as true neighbors, and the neighbor interface changes state from Init to 2-Way.

Electing the Designated Router: All multiaccess networks have a designated router. There is also a backup designated router that takes over in the event that the designated router fails.

The use of a backup, which maintains an identical set of adjacencies and an identical topology database to the designated router, ensures there is no extended loss of routing capability if the designated router fails.

The designated router performs two major functions on a network:

- It originates network links advertisements on behalf of the network.
- It establishes adjacencies with all other routers on the network. Only routers with adjacencies exchange link state information and synchronize their databases.

The designated router and backup designated router are elected on the basis of the router identifier, router priority, designated router and backup designated router fields in hello packets. Router priority is a single octet field that defines the priority of a router on a network. The lower the value of the priority field the more likely the router is to become the designated router, hence the higher its priority. A zero value means the router is ineligible to become designated or backup designated router.

The process of designated router election is as follows:

1. The current values for designated router and backup designated router on the network are initialized to 0.0.0.0.
2. The current values for router identifier, router priority, designated router and backup designated router in hello packets from neighbor routers are noted. Local router values are included.
3. Backup designated router election:

   Routers that have been declared as designated routers are ineligible to become backup designated routers.

   The backup designated router will be declared to be:

   - The highest priority router that has been declared as backup designated router
   - The highest priority router if no backup designated router has been declared

   If equal priority routers are eligible, the one with the highest router identifier is chosen.

4. Designated Router Election:

   The designated router will be declared to be:

   - The highest priority router that has been declared designated router
   - The highest priority router if no designated router has been declared

5. If the router carrying out the above determination is declared the designated or backup designated router then the above steps are re-executed. This ensures that no router can declare itself both designated and backup designated router.

Once designated and backup designated routers have been elected for a network, they proceed to establish adjacencies with all routers on the network.

Completion of the election process for a network causes the router interface to change state from Waiting to DR, BackupDR, or DR Other depending on whether the router is elected the designated router, the backup designated router or neither of these.

Establishing Adjacencies - Database Exchange: A router establishes adjacencies with a subset of neighbor routers on a network.

Routers connected by point-to-point networks and virtual links always become *adjacent*. Routers on multiaccess networks form adjacencies with the designated and backup designated routers only.

Link state information flows only between adjacent routers. Before this can happen it is necessary for them to have the same topology database and to be synchronized.

This is achieved in OSPF by a process called *database exchange*.

Database exchange between two neighbor routers occurs as soon as they attempt to bring up an adjacency. It consists of the exchange of a number of database description packets that define the set of link state information present in the database of each router. The link state information in the database is defined by the list of link state headers for all link state advertisements in the database (see Figure 82 on page 213 for information on the link state header).

The format of database description packets is shown in Figure 78 on page 210.

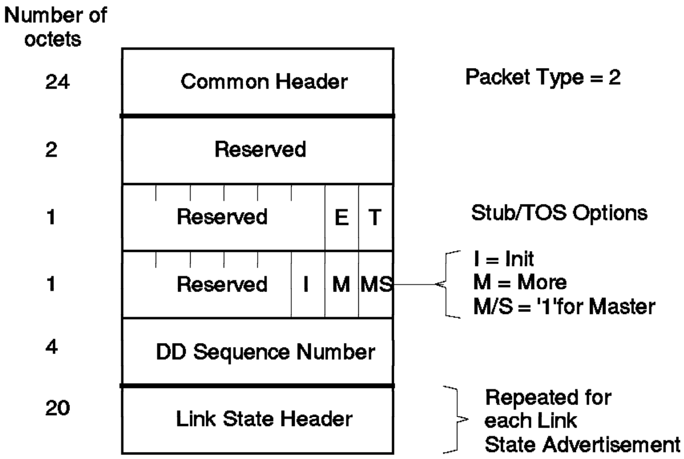


Figure 78. OSPF Database Description Packet

During the database exchange process, the routers form a *master/slave* relationship, the master being the first to transmit. The master sends database description packets to the slave to describe its database of link state information. Each packet is identified by a sequence number and contains a list of the link state headers in the master′s database. The slave acknowledges each packet by sequence number and includes its own database of headers in the acknowledgments.

Flags in database description packets indicate whether they are from a master or slave (the M/S bit), the first such packet (the I bit) and if there are more packets to come (the M bit). Database exchange is complete when a router receives a database description packet from its neighbor with the M bit off.

During database exchange, each router makes a list of the link state advertisements for which the adjacent neighbor has a more up-to-date instance. (All advertisements are sequenced and time stamped.) Once the process is complete, each router requests these more up-to-date instances of advertisements using link state requests.

The format of link state request packets is shown in Figure 79 on page 211.

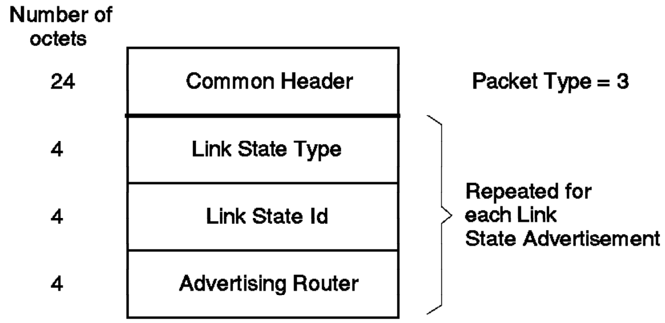


Figure 79. OSPF Link State Request Packet

The database exchange process sequences the neighbor interface state from *2-Way* through:

- *ExStart* as the adjacency is created and the master agreed upon
- *Exchange* as the topology databases are being described
- *Loading* as the link state requests are being sent and responded to
- *Full* when the neighbors are fully adjacent

In this way, the two routers synchronize their topology databases and are able to calculate identical network maps for their OSPF area.

Link State Propagation: Information about the topology of an OSPF network is passed from router to router in link state advertisements.

Link state advertisements pass between adjacent routers in the form of *link state update* packets, the format of which is shown in Figure 80.

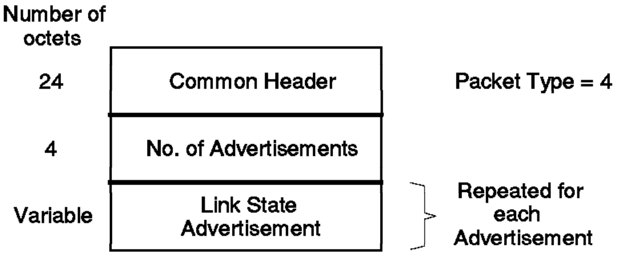


Figure 80. OSPF Link State Update Packet

Link state advertisements consist of five types: router links, network links, summary links (two types) and AS external links as noted earlier in this section.

Link state updates pass as a result of link state requests during database exchange, and also in the normal course of events when routers wish to indicate a change of network topology. Individual link state update packets can contain multiple link state advertisements.

It is essential that each OSPF router in an area has the same network topology database, and hence the integrity of link state information must be maintained.

For that reason link state update packets must be passed without loss or corruption throughout an area. The process by which this is done is called *flooding*.

A link state update packet floods one or more link state advertisements one hop further away from their originator. To make the flooding procedure reliable each link state advertisement must be acknowledged separately. Multiple acknowledgments can be grouped together into a single *link state acknowledgment packet*. The format of the link state acknowledgment packet is shown in Figure 81.

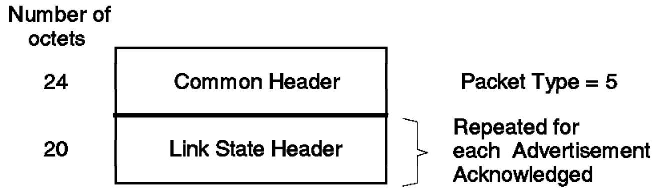


Figure 81. OSPF Link State Acknowledgment Packet

In order to maintain database integrity, it is essential that all link state advertisements are rigorously checked to ensure validity.

The following checks are applied and the advertisement discarded if:

- The link state checksum is incorrect.
- The link state type is invalid.
- The advertisement′s age has reached its maximum.
- The advertisement is older than or the same as one already in the database.

If an advertisement passes the above checks, then an acknowledgment is sent back to the originator. If no acknowledgment is received by the originator then the original link state update packet is retransmitted after a timer has expired.

Once accepted, an advertisement is flooded onward over the router′s other interfaces until it has been received by all routers within an area.

Advertisements are identified by their *link state type*, *link state ID* and the *advertising router*. They are further qualified by their *link state sequence number*, *link state age* and *link state checksum number*.

The age of a link state advertisement must be calculated to determine whether it should be installed into a router′s database. Only a more recent advertisement should be accepted and installed. Advertisements are only considered more recent if they have a newer sequence number, or if sequence numbers are equal if they have the larger checksum, or if checksums are equal if they have their age set to *max age*.

Valid link state advertisements are installed into the topology database of the router. This causes the topology map or graph to be recalculated and the routing table to be updated.

Link state advertisements all have a common header. This is shown in Figure 82. The five link state advertisement types are shown in Figure 83 on page 214, in Figure 84 on page 214, in Figure 85 on page 215, and in Figure 86 on page 215.

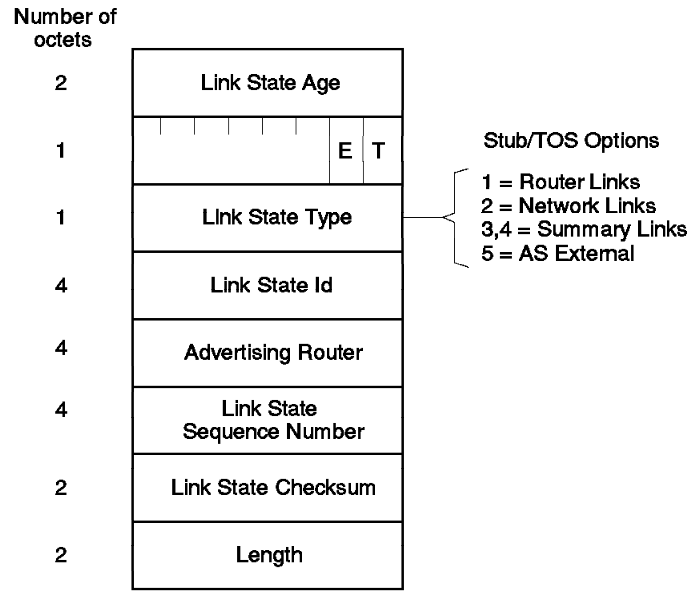


Figure 82. OSPF Link State Header

Routing Table Calculation: Each router in an OSPF area builds up a topology database of validated link state advertisements and uses them to calculate the network map for the area. From this map the router is able to determine the best route for each destination and insert it into its routing table.

Each advertisement contains an age field which is incremented while the advertisement is held in the database. An advertisement′s age is never incremented past *max age*. When max age is reached it is excluded from routing table calculation, and reflooded through the area as a newly originated advertisement.

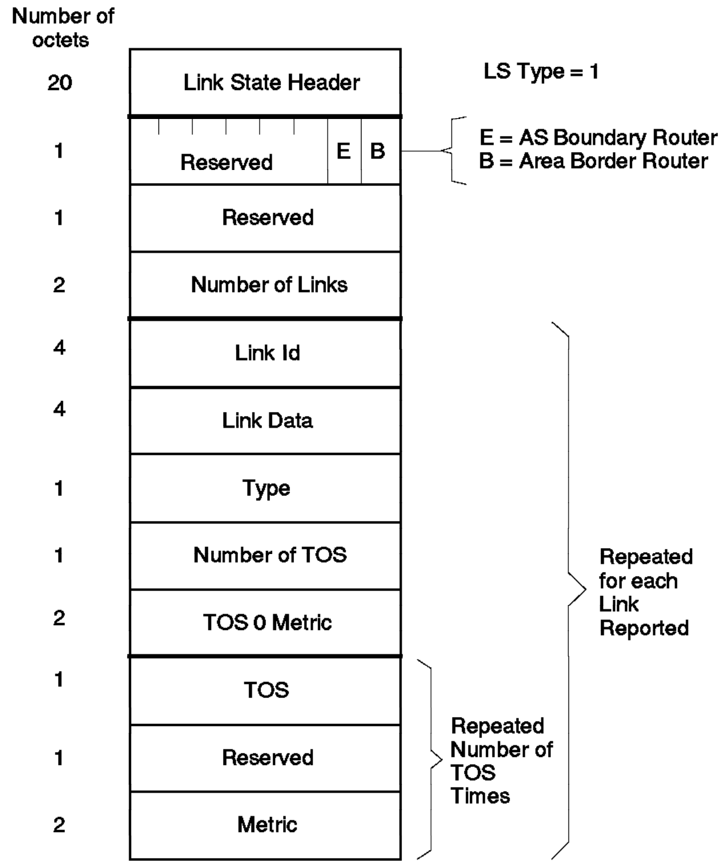


Figure 83. OSPF Router Links Advertisement

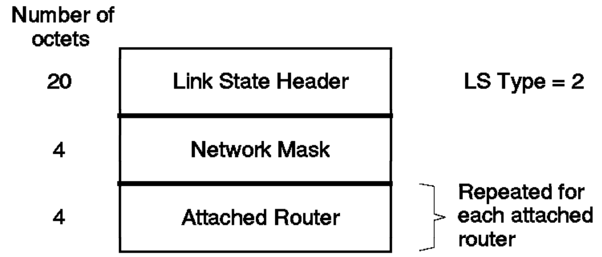


Figure 84. OSPF Network Links Advertisement

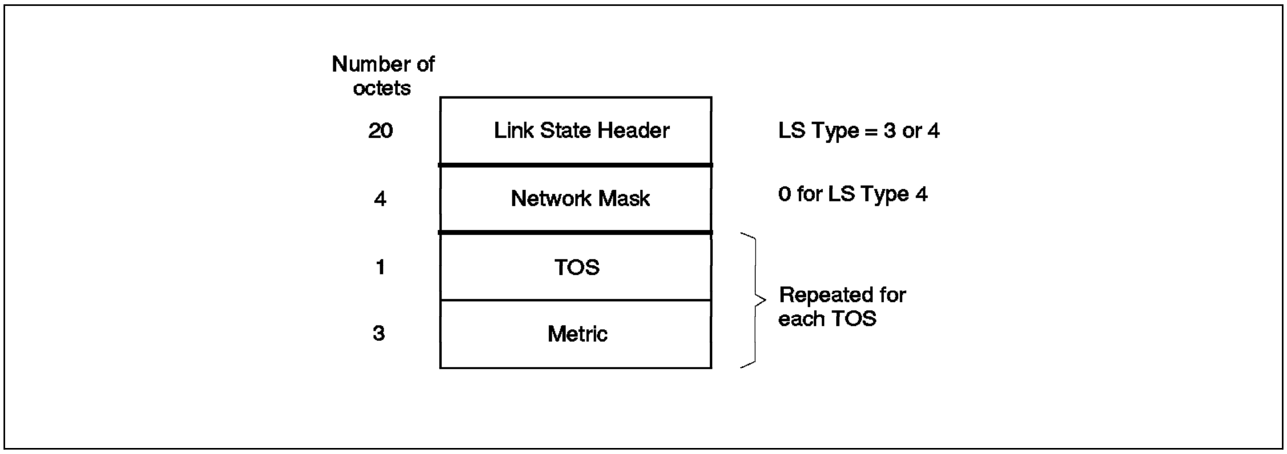


Figure 85. OSPF Summary Links Advertisement

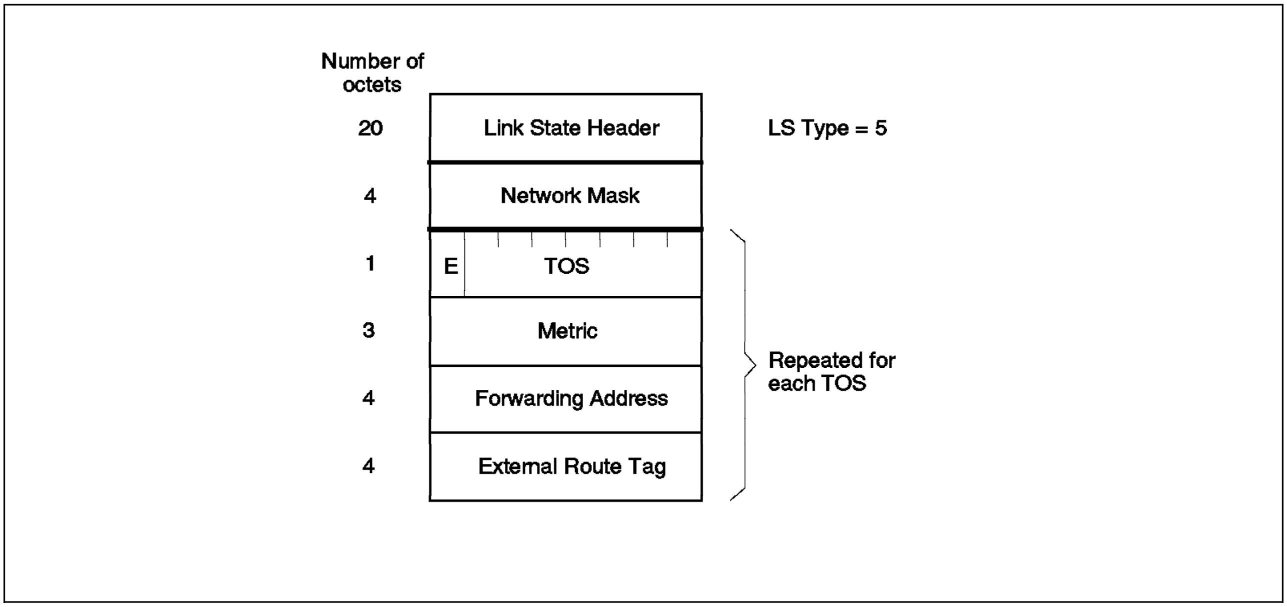


Figure 86. OSPF External Links Advertisement

Routers build up their routing table from the database of link state advertisements in the following sequence:

1. The shortest path tree is calculated from router and network link advertisements allowing the best routes within the area to be determined.
2. Inter-area routes are added by examination of summary link advertisements.
3. AS external routes are added by examination of AS external link advertisements.

The topology graph or map constructed from the above process is used to update the routing table. The routing table is recalculated each time a new advertisement is received.

4.3.6 HELLO

The HELLO protocol is an interior gateway protocol defined in RFC 891. It is an IAB standard protocol; its status is elective.

The HELLO protocol was used in the *Fuzzball* software that formed part of early Internet experience with the original NSFNET backbone. It should not be used for new network implementations.

HELLO should not be confused with the OSPF Hello protocol described in "Discovering Neighbors - The OSPF Hello Protocol" on page 207. It is *not* necessary to configure the HELLO protocol to run OSPF on a router.

HELLO is significant because it provides an example of a distance vector protocol that does not use hop counts as its metric. Instead it uses round-trip delays.

4.3.6.1 Protocol Description

HELLO provides functions to synchronize clocks among a set of routers, and then to use shortest delay to select the best routes to destination routers and networks.

In the basic HELLO protocol every router is given a unique *host identifier*, conventionally its IP address. Every router maintains two tables:

- One contains entries for every network connected to the local network, for certain remote networks and for the routers that should be used to access them. This table must be configured into the router.
- A second contains the best estimate of the round trip delay to and the logical clock offset of every router on the network. This table is maintained by the router.

Routers maintain a best estimate of the clock in every router from which they will receive routing information. All routing table updates include a time stamp from the originating router from which the receiving router can calculate the delay across the link from the originating router by subtracting the time stamp from the current estimate of the originating router′s clock. Periodically routers poll their neighbors to re-establish clock synchronization.

HELLO messages are sent by routers to define destinations they can reach and the delay and clock offset associated with each of them. Overall round-trip delays are calculated to each destination by the receiving router adding the delay across the link to the originating router to the values in the HELLO message. This gives overall delays to all destinations, allowing calculation of the best routes to each destination.

This is a very simplified description of the protocol but describes the principle by which *delay* can be used as a routing metric.

4.3.7 Exterior Gateway Protocol (EGP)

The *Exterior Gateway Protocol* (EGP) is an exterior gateway protocol defined in RFC 904.

It is an IAB standard protocol; its status is recommended. Recommended status means that a system should implement it.

EGP is a protocol used between routers in separate autonomous systems (ASs) to advertise reachability information about networks within their local ASs. Routers exchanging EGP information are said to be *exterior neighbors* and EGP an *inter-AS* routing protocol.

Reachability information consists of the natural network numbers of networks accessible within an AS, and the identity of the router through which they can be reached.

It is fundamental to the operation of EGP that only information about networks belonging to the local AS is forwarded. This is a different approach to that of other router protocols, which usually forward all routing table information.

EGP itself has no ability to determine networks or routes within an AS. It must rely on an interior gateway protocol, such as RIP or OSPF, to do this.

4.3.7.1 Protocol Description

EGP packets are transmitted in IP datagrams. IP datagrams containing EGP packets can be distinguished by the use of *protocol identifier 8* in the IP datagram header. Therefore, EGP packets are not contained in TCP or UDP headers.

EGP packets are specifically addressed to neighbors whose identities must be configured into each router running EGP.

The EGP protocol consists of three main elements:

- A *Neighbor Acquisition* protocol to request a neighbor to enter into an exchange of reachability information
- A *Neighbor Reachability* protocol to confirm that a neighbor is alive
- A *Network Reachability* protocol to request reachability information from a neighbor

Neighbor Acquisition Protocol: A router sends neighbor acquisition messages when it wishes to establish EGP communications with a neighbor. A neighbor would normally be acquired as a result of it being configured as an exterior neighbor in the router running EGP.

Neighbor acquisition messages, like all EGP messages, include a standard header with *sequence number*, *version number*, *AS number* and *checksum*.

A neighbor is acquired by sending a *neighbor acquisition request* message, to which the neighbor router should respond with a *neighbor acquisition confirm*. In the event that the request was invalid, a *neighbor acquisition refuse* should be sent.

The format of neighbor acquisition messages is shown in Figure 87 on page 218.

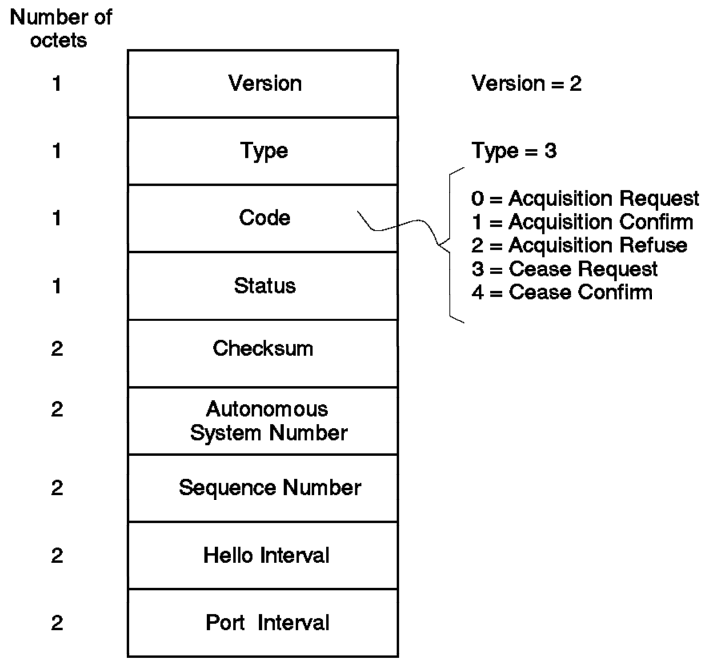


Figure 87. EGP Neighbor Acquisition Packet

Once the request/confirm exchange has been completed, EGP communications can commence, and will continue until either neighbor wishes to terminate. In this case a *cease request* message is sent, which should be acknowledged by a *cease confirm* from the receiving neighbor.

Neighbor Reachability Protocol: The initial exchange of acquisition messages defines a *hello interval*. This interval tells a neighbor how often it should send hello packets to check that the originating router is alive. The originating router will also send hello packets to ensure that its neighbor is alive.

Neighbor reachability is confirmed by a simple exchange of *Hello* requests and *I Heard You* responses at intervals defined by the hello interval. The format of these messages is shown in Figure 88 on page 219.

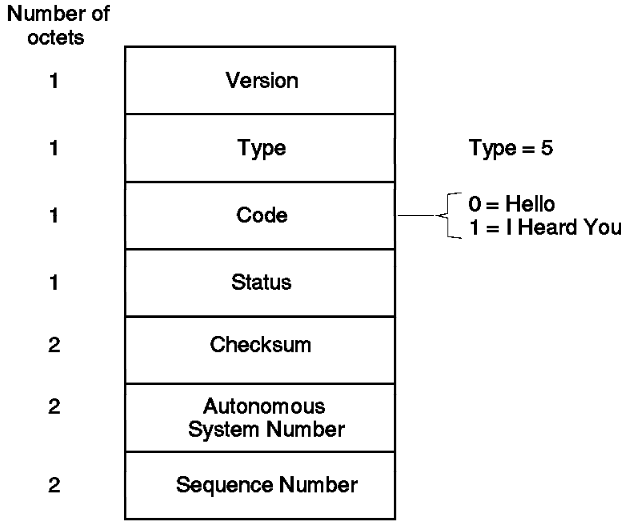


Figure 88. EGP Neighbor Reachability Packet

Separating the process of neighbor reachability from that of advertising reachability information is important because it leads to reduced network traffic over what may be relatively low-speed, high cost point-to-point network links. This advantage can only be realized because reachability information about networks within ASs changes relatively infrequently.

It is possible, of course, that Hello and I Heard You messages may be lost between neighbors. To ensure that this does not cause neighbors to be marked down, EGP specifies that *n* out of *m* message exchanges must fail before a neighbor is marked down. In addition, the EGP standard recommends that two successive exchanges must fail before a neighbor is marked down.

Network Reachability Protocol: An EGP network reachability poll message allows a router to request reachability information from a neighbor. This should not be done more frequently than once per minute.

The poll message consists of the standard EGP *header* along with the common *IP source network number* to which both neighbors attach. This network is used as the reference point for all routing information. Its format is shown in Figure 89 on page 220.

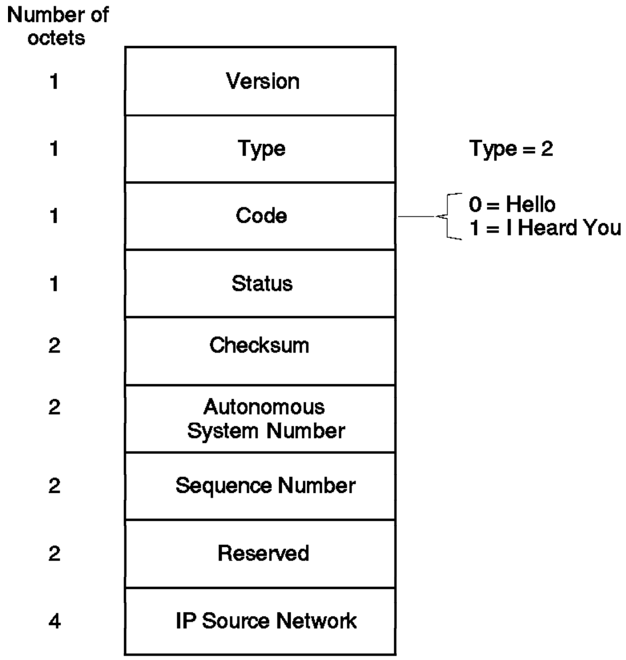


Figure 89. EGP Poll Packet

The response to a poll is a *network reachability update*, which contains a list of destinations with their distances measured relative to the specified source network, and the identity of routers through which they may be reached. The reachability information sent in network reachability messages must only be for destinations within the AS to which the EGP router is attached. The router, through which an intra-AS network may be reached, must be on the same common source network as the neighbor EGP routers. The format of this message is shown in Figure 90 on page 221.

EGP routers may also send routing update messages to neighbors at any time, without the need to wait for a poll. However, at most *one* unsolicited update may be sent between successive polls.

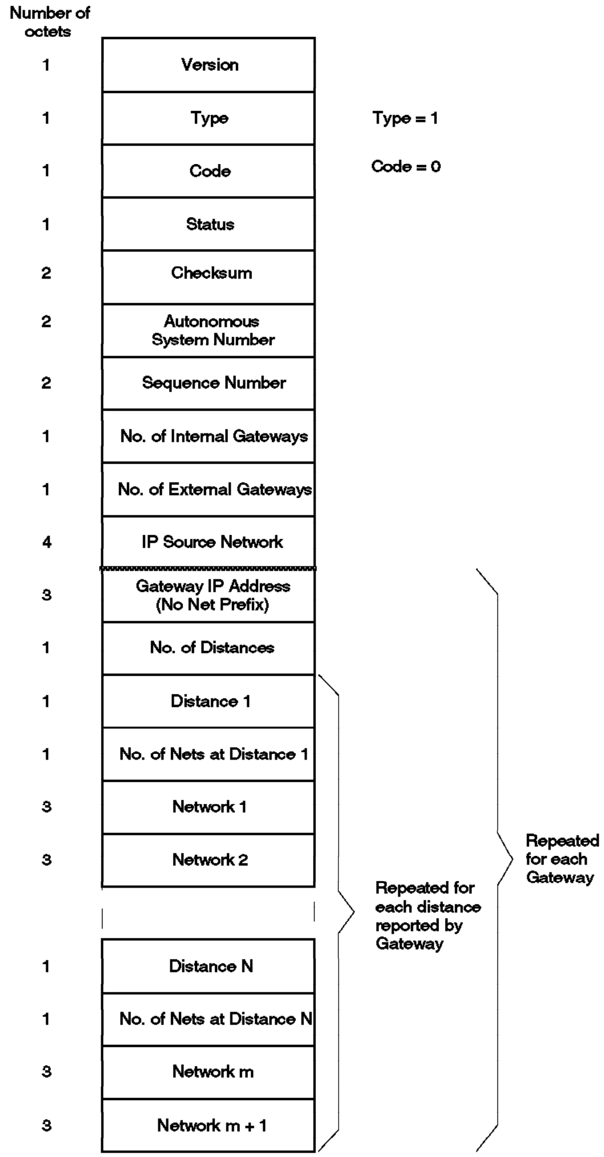

Number of octets
1 Version
1 Type
Type = 1
1 Code
Code = 0
1 Status
2 Checksum
2 Autonomous System Number
2 Sequence Number
1 No. of Internal Gateways
1 No. of External Gateways
4 IP Source Network
3 Gateway IP Address (No Net Prefix)
1 No. of Distances
1 Distance 1
1 No. of Nets at Distance 1
3 Network 1
3 Network 2
1 Distance N
1 No. of Nets at Distance N
3 Network m
3 Network m + 1
Repeated for each distance reported by Gateway
Repeated for each Gateway

Figure 90. EGP Network Reachability Update Packet

4.3.8 Border Gateway Protocol (BGP)

The *Border Gateway Protocol* (BGP) Version 3 is an exterior gateway protocol defined in RFC 1267. A companion document, RFC 1266 *Experience with the BGP Protocol*, details experience of its use in the Internet.

It is an IAB standard protocol; its status is recommended. BGP was built upon experience with EGP, and is intended as a functionally superior replacement for it.

BGP is an *inter-AS* routing protocol designed to exchange reachability information with BGP neighbors. BGP only advertises routes that it uses itself, but unlike EGP, they can be routes outside the local AS. In this case the route consists of the path (the AS_Path) through ASs to the destination, along with the next hop router. As with EGP, all routes advertised are for natural network numbers only; that is, no subnetworks are advertised.

BGP, like EGP, has no ability to determine routes within an AS. It must rely on an interior gateway protocol, such as RIP or OSPF, to do this.

4.3.8.1 Protocol Description

BGP runs over a reliable transport layer connection between neighbor routers. BGP relies on the transport connection for fragmentation, retransmission, acknowledgment and sequencing. It assumes that the transport connection will close in an orderly fashion, delivering all data, in the event of an error notification.

Practical implementations of BGP use TCP as the transport mechanism. Therefore, BGP protocol data units are contained within TCP packets. Connections to the BGP service on a router use TCP port 179.

The BGP protocol comprises four main stages:

- Opening and confirming a BGP connection with a neighbor router
- Maintaining the BGP connection
- Sending reachability information
- Notification of error conditions

Opening and Confirming a BGP Connection: BGP communications between two routers commences with the TCP transport protocol connection being established. Description of this is outside the scope of this document, but can be found in *TCP/IP Tutorial and Technical Overview*, GG24-3376.

Once the connection is established, each router sends an *open* message to its neighbor.

The BGP open message, like all BGP messages, consists of a standard header plus packet-type specific contents. The standard header consists of a *16-octet marker* field, which is set to all 1s, the *length* of the total BGP packet, and a *type* field that specifies the packet to be one of four possible types:

1. Open

2. Update

3. Notification

4. Keep alive

The format of the BGP header is shown in Figure 91 on page 223.

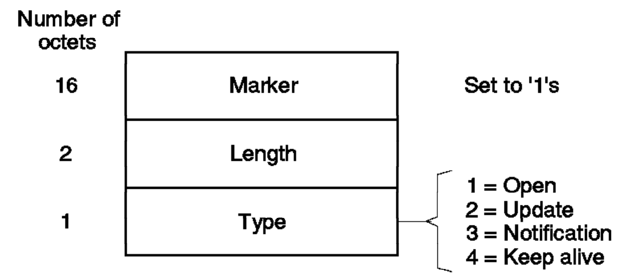


Figure 91. BGP Message Header

The open message defines the originating router's *AS number*, its BGP *router identifier* and the *hold time* for the connection. If no keep alive, update or notification messages are received for a period of hold time, the originating router assumes an error, sends a notification message, and closes the connection.

The open message also provides an *authentication code* and *authentication data*. The use of these fields is not fully defined in the RFC and current BGP implementations use authentication code *0* with authentication data of all *0*s.

The format of the open message is shown in Figure 92.

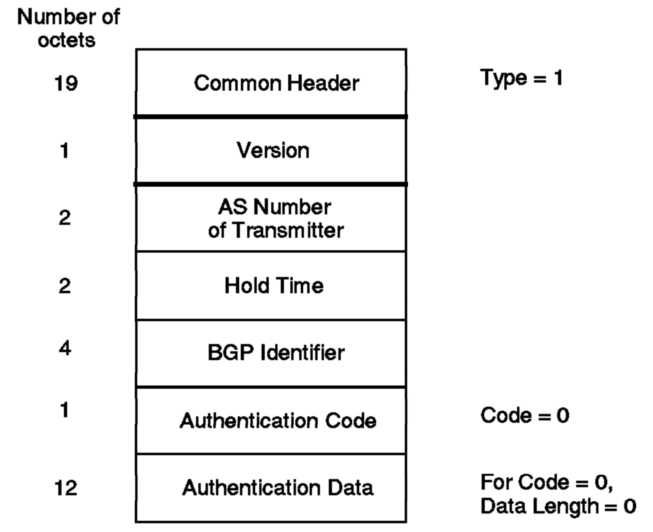


Figure 92. BGP Open Message

An acceptable open message is acknowledged by a *keep alive* message. Once neighbor routers have sent keep alives in response to opens, they can proceed to exchange further keep alives, notifications and updates.

Maintaining the BGP Connection: BGP messages must be exchanged periodically between neighbors. If no messages are received for a period defined by hold time in the open message, then an error on the connection is assumed.

BGP uses keep alive messages to maintain the connection between neighbors. Keep alive messages consist of the BGP packet header only with no data. The RFC recommends that they should be sent at intervals of approximately one third of the hold time.

Sending Reachability Information: Reachability information is exchanged between BGP neighbors in update messages.

Update messages provide network advertisements using a series of path attributes for each advertised network. The format of these is shown in Figure 93.

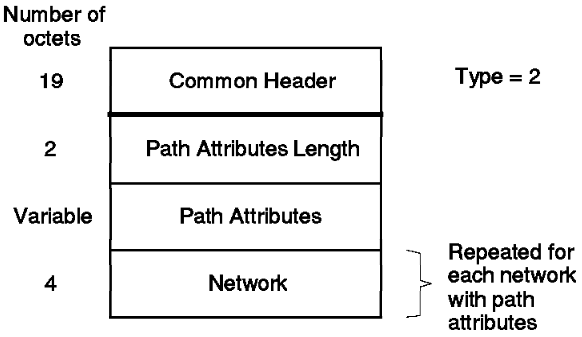


Figure 93. BGP Update Message

Each *path attribute* consists of a triple set of values: *attribute flags*, *attribute type* and *attribute value*. Three of the attribute flags provide information about the status of the attribute types, and may be *optional* or *well-known*, *transitive* or *non-transitive and partial* or *complete*.

Attribute flags must be read in conjunction with their associated attribute types. There are five attribute types which together define an advertised route:

- *Origin*, which must be well-known, and defines the origin of the route as an interior gateway protocol, an exterior gateway protocol or other (for example a static route).
- *AS Path*, which must be well-known, and defines the ASs which must be crossed to reach the network being advertised.
- *Next Hop*, which must be well-known, and defines the next hop to the network being advertised.
- *Unreachable*, which must be well-known, and indicates that a previously advertised route is unreachable.
- *Inter AS Metric*, which must be optional and non-transitive, and is used to define the metric of a route within its local AS. This can be used to prioritize between multiple exit points from a network.

The first three of these attributes are mandatory and must be supplied with all network advertisements. Each attribute type sent in a BGP update message must have its attribute flags set as indicated; otherwise, an error occurs.

The third element of the triple assigns values to the five attributes listed above.

The format of BGP path attributes is shown in Figure 94 on page 225.

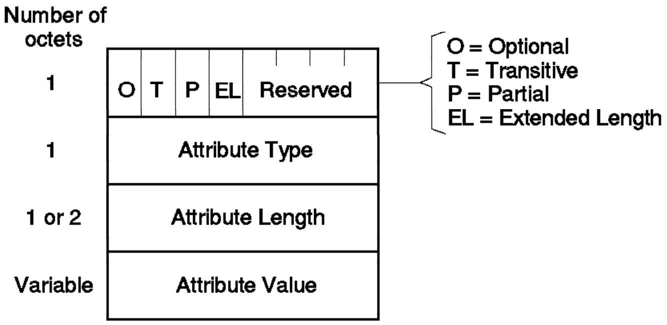


Figure 94. BGP Path Attributes

Multiple networks with the same path attributes can be advertised in a single EGP update. Networks with different path attributes must be advertised in separate update messages.

Notifying Errors: Notification messages are sent to a neighbor router when error conditions are detected. The BGP transport connection is closed immediately after a notification message has been sent.

Notification messages consist of an *error code* and an *error subcode* which further qualifies the main error. The format of notification messages is shown in Figure 95.

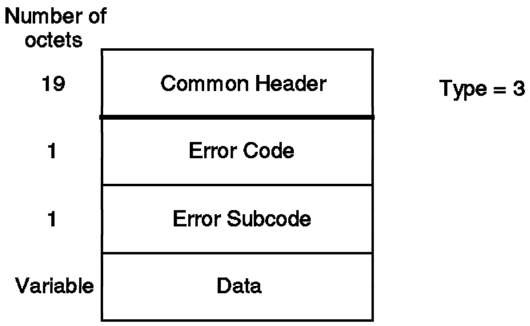


Figure 95. BGP Notification Message

Error codes that are provided by BGP are as follows:

- Message Header Error
- Open Message Error
- Update Message Error
- Hold Timer Expired
- Finite State Machine Error
- Cease

A *data field* is included in the notification message to provide additional diagnostic information.

4.4 LAN Management Protocols

This section describes the major protocols used in LAN management products. For more details on LAN management products refer to Chapter 5 of *Local Area Network Concepts and Products: LAN Adapters, Hubs and ATM*, SG24-4754.

4.4.1 Internet Control Message Protocol (ICMP)

The ICMP protocol is used for host-to-host datagram services to provide problem feedback. ICMP allows gateways to communicate between themselves for control purposes. ICMP is an integral part of IP and must be implemented by every IP module. ICMP provides feedback about problems in the communication environment, but does not make IP reliable.

The messages have a common header of type and code and 64 bits of the original data datagram if applicable.

- *Destination unreachable* can indicate whether the net, host, protocol, or port is unreachable or whether fragmentation is needed or the source route failed.
- *Time exceeded* can indicate if time to live was exceeded in transit or if fragment reassembly time was exceeded.
- *Parameter problem* indicates header parameter problems such that it cannot complete processing of the datagram. This may include incorrect arguments in an option.
- *Source quench* is sent when the gateway discards a datagram due to:
  - Insufficient buffer space to queue datagram on its router
  - If the datagrams are arriving too fast to be processed
  - The capacity limit is being approached and the host is requested to reduce the rate at which it is sending traffic
- *Redirect* is sent when the gateway recognizes a shorter path.
- *Echo/echo reply* allows return of information to verify paths.
- *Timestamp* returns the time the sender last touched the message before sending it, the time the echoer first touched it on receipt and the transmit time when the echoer last touched the message on sending it.
- *Information* returns the number of the network it is on.

4.4.1.1 Simple Network Management Protocol (SNMP)

SNMP is a TCP/IP network management protocol. It is based on a manager-agent interaction. Agents gather management data and store it, while managers solicit this data and process it. The base for SNMP communication is the Management Information Base (MIB) stores information about an SNMP-managed resource. The MIB is usually made up of two components:

- **MIB II:** This is a standard definition that defines the data layout (such as the length of fields and what the field is to contain) for the management data for resource, such as the resource name and address.

- **MIB Extension:** This incorporates unique information about a resource. It is defined by the manufacturer of the resource that is being managed. These are usually unique and proprietary in nature.

4.4.1.2 What Is an SNMP Agent?

The SNMP agent is responsible for managed resources and keeps data about the resources in an MIB. The SNMP agent has two responsibilities:

1. To place data into the MIB fields
2. To react to changes in certain fields made by the manager

4.4.1.3 SNMP Agent Added Features

- Proxy agent - SNMP agent for another entity that does not support its own MIB values.
- Security - No security has been standardized for SNMP.
- Enterprise-specific information - MIBs and traps.

4.4.1.4 What Is an SNMP Manager (Client)?

An SNMP manager has the ability to issue the SNMP commands and be the end point for traps being sent by the agent. Commands are sent to the agent using the MIB as a communication vehicle.

4.4.1.5 SNMP Manager (Client) Added Features

- Trouble ticketing
- Network statistics
- Conversion of traps to alerts for centralized management
- Reporting capabilities
- Database support
- Graphics interface
- Auto-discovery
- Dynamic MIB updates
- UNIX′s PING, Telnet support

The manager uses commands to either get data from the MIB in the agent (GET command), or to place data into the MIB on the agent (SET command). An agent may notify a manager of an event using the TRAP command. The manager polls the agents for information on a regular basis to keep track of the status of the resources. Polling is fundamental to the operation of SNMP.

MIB Enterprise Specific Definitions: The MIB II definitions are all aligned to a prescribed format; however, the extensions are unique. In the past, many vendors did not publish their MIB extensions and produced their own proprietary manager for the resource. Through the use of MIB compilers, products such as AIX/NetView 6000 produce the mapping required to read, and make sense of, the MIB data.

4.4.1.6 MIB Support

The MIB defines the following groups and the information stored for each:

- System Group - textual description of the entity being managed.
- Interface Group - tabular description of the network interfaces.
- Address Translation Group - translation tables for physical addresses.
- IP Group - addresses, indicators, and counters for IP decisions on datagrams and routing information for these datagrams.
- ICMP Group - the error input and output statistics.
- TCP Group - information of TCP connections, transmission.
- UDP Group - information on unguaranteed datagrams.
- EGP Group - information on the exterior gateway neighbors.

- Transmission Group - information on different types of transmission media. (This group was added by MIB II.)
- SNMP Group - Information on SNMP for use by applications when using SNMP statistical information. (This group was added by MIB II.)

4.4.2 Common Management Information Protocol (CMIP)

CMIP is part of a complete set of management definitions for an OSI network. There are two major components: Common Management Protocol (CMIP) and Common Management Information Service (CMIS). CMIS defines the services used and CMIP defines the protocol that carries the services. The OSI management model uses an object-oriented approach. Managed objects all exhibit the same exterior appearance and accept the same set of commands. Part of the OSI standards is the GDMO (Guidelines for the Definition of Managed Objects). This standard provides a common way to define the objects being managed by a manager.

4.4.2.1 CMIP Manager Functions

CMIP Managers manage CMIP agents and can communicate with many different types of objects. The CMIP applications fall into five categories:

1. Fault
2. Configuration
3. Accounting
4. Performance
5. Security

4.4.2.2 CMIP Agent Functions

CMIP agents execute the CMIS services requested by the manager and maintain data in the objects. When an object detects a problem, the agent notifies the manager. The agent can perform the management services (CMIS) required by the manager by invoking the methods associated with the object definition.

CMIP was initially defined for use across an X.25 network. Subsequently, there have been standards to use CMIP protocols using other transports. CMIP over TCP/IP (CMOT) and CMIP over LAN (CMOL) are examples. In both cases, the transport protocol simply envelopes the underlying CMIP protocol.

4.4.3 X/Open Management Protocol API (XMP)

The XMP API is an important industry standard and is on the SystemView interfaces. The XMP API defines a means to request management information services. Its main benefit is to allow generalized applications to control multiple, unique resource types regardless of the underlying management protocols used. The interface uses the generic commands defined by the International Standards Organization (ISO) as the basis for LAN systems management. The service elements of the API correspond to the abstract services of CMIS and SNMP. It also uses the GDMO object definitions. This API has been selected by SystemView, OSF and NMF as the key API for managing open system environments.

Appendix A. IEEE 802 Documents Summary

These tables describe the currently available IEEE 802 documents. More information about how to order these documents may be obtained by calling the IEEE at 1-800-678-IEEE.

Table 18. IEEE 802.0 Executive Committee Docs

IEEE 802.0	Description
P802.0-91/170	Operating rules of IEEE Project 802.Draft 11/91
P802.0-93/00A	Executive Committee Member & Liaison list 3/93 *
P802.0-91	IEEE 802 Functional Requirements Rev 6.10 11/91
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 19. IEEE 802.1 Higher Layer Interface (HILI) Working Group Docs

IEEE 802.1	Description
IEEE Std 802.1B-1992	LAN/MAN Management 11/92 (SH15701) ***
IEEE Std 802.1E-1990	System load protocol 3/91 (SH13573) ***
IEEE Std 802-1990	Overview & Architecture 12/90 (SH13557)
IEEE Std 802.1D-1990	MAC Bridges 3/91 (SH13565) ***
P802.1F/D13	Common definitions & procedures for 802 Mgmt info 11/92 *
IEEE Std 802.1i-1992	Supplement to MAC Bridges : FDDI 8/92 (SH15198)
IEEE 802.1	IEEE 802.1 Working group meeting minutes 11/92 *
P802.1G/D7	Remote MAC bridging 12/92 *
P802.1H/D4	MAC Bridging Ethernet V2.0 in 802 LANs 1/93 *
P802.1K/D6	LAN/MAN mgmt Discovery & Dynamic control of event fwd. 8/92
P802.1m/D4	System load protocol managed object defns. & PICS proforma 8/92
P802.1-92/14	IEEE 802 Protocol Identifier Assignments 3/92
P802.1J/D2	Managed objects for MAC bridges 12/92 *
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 20 (Page 1 of 2). IEEE 802.2 Logical Link Control (LLC) Working Group Docs

IEEE 802.2	Description
P802.2-90/47	Type 3 operation, ISO 8802-2:1989/AM2
ISO 8802-2	1989 LLC Standard, Revised, (SH12930) ***

Table 20 (Page 2 of 2). IEEE 802.2 Logical Link Control (LLC) Working Group Docs

IEEE 802.2	Description
P802.2-91/3	ISO/IEC DTR 10178 Structure & Coding of LSAPs, 3/91

Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE

Table 21. IEEE 802.3 CSMA/CD Working Group Docs

IEEE 802.3	Description
ISO 8802/3-1990	ANSI/IEEE 802.3-1990 CSMA/CD AMC & PHY (SH13482) ***
IEEE Std 802.3h-1990	Layer Management, Section 5, 3/91 (SH13755) ***
P802.3-92/3	Meeting minutes , Minnesota, 7/92
IEEE Std 1802.3	Conformance test for CSMA/CD***
IEEE St 802.3i-1990	Sect.13-14, type 10BaseT, 12/90 (SH13763) ***
P802.3j/D13-92	Fiber Optic CSMA/CD LAN, 10BaseF, 3/92
IEEE Std 802.3k-1992	Layer Management for 10 Mbps Baseband Repeaters ***
IEEE Std 802.3l-1992	10BaseT PICS Proforma, #14.10 (SH15727) ***
P802.3G/D5-91	MAU ATS, MAU Conformance Test , Draft 5 6/91
P802.3M&N/D1	Consolidated 2nd and 3rd Maintenance Packages, 3/92
P1802.3D/D5-92	10BaseT MAU Conformance Testing, 6/92
P802.3P/D6-92	Layer Management for MAUs, 11/92
P802.3Q/D5-92	GDMO Revisions to Layer mgmt for DTEs, 11/92

Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE

Table 22. IEEE 802.4 Token Bus (TBUS) Working Group Docs

IEEE 802.4	Description
	Interim Conformance Test specification document set, 11/91
IEEE Std 804-1991-	Token Passing Bus- MAC and PHY ***
IEEE Std 802.4b-1992	Physical Layer Diversity (SH15735) ***

Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE

Table 23. IEEE 802.5 Token-Ring Working Group Docs

IEEE 802.5	Description
ISO/IEC 8802-5/1992	Token-Ring MAC & PHY Layer Std. (SH15131) ***
P802.5J/D24	Fiber Optic station attachment, Draft 24, 2/26/93 *
P802.5M/D7	Source Route/Transparent Bridge Annex to 802.1d, 6/92
P802.5P/D6	Route Determination Entity Change to ISO 8802-2, 2/93 *
P802.5Q/D2A	802.5 MAC & STP/UTP PHY Revisions, Draft 2A ,2/26/93*
P802.5N/11-01	PHY Working Group Presentations
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 24. IEEE 802.6 Metropolitan Area Network (MAN) Working Group Docs

IEEE 802.6	Description
IEEE Std 802.6-1990	DQDB Subnetwork of a MAN, 7/3/91, (SH13961) ***
P802.6C-90/D2	PLCP for DS1, Draft 2, 11/90
P802.6i/D5-93/01	Remote LAN Bridging Using 802.6 MAN, 3/93 *
P802.6K/D2-91/18	DQDB Supplement to IEEE Std 802.1D, 5/28/91
P802.6F/D3-91/100	PICS proforma, 11/91
P802.6G/D2	Layer Management, 6/92
P802.6D-92/01&02	PLCP or Sonet, 1/92
P802.6H/D1-92/42	Isochronous Service, 11/92
P802.6-92/17	Minutes 3/92
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 25. IEEE 802.7 Broadband Technical Advisory Group Working Docs

IEEE 802.7	Description
IEEE Std 802.7-1989	Broadband LAN Recommended Practices, (SH12955) ***
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 26 (Page 1 of 2). IEEE 802.8 Fiber Optic Technical Advisory Group Docs

IEEE 802.8	Description
P802-87*8.12	Fiberoptic CSMA/CD Network Approaches

Table 26 (Page 2 of 2). IEEE 802.8 Fiber Optic Technical Advisory Group Docs

IEEE 802.8	**Description**
P802.8-92/4	Status of Various FO Connectors Tutorial, 3/92
P802.8-92/5	″Why IBM Selected the SC Connector-End User View″, 3/92
P802.8-92/7	Minutes, Irvine,3/92
P802.8-92/9	Minutes, Bloomington, 7/92
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 27. IEEE 802.9 Integrated Services LAN Group Docs

IEEE 802.9	**Description**
P802.9-88/22	Architectural Tutorial Document, Draft Issue 4 3/88
P802.9/D19	IEEE 802.9 ISLAN Draft Standard, Draft 19, 3/93 *
P802.9/4	Minutes, Austin, 1/92
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 28. IEEE 802.10 Interoperable LAN Security

IEEE 802.10	**Description**
IEEE 802.10-1992	SILS - Secure Data Exchange (SH15610) ***
P802.10C/92-3	Part C- Key Management, NOV 23, 1992
P802.10-92/1	Minutes, Salt Lake City, 1/92
P802.10-92/2	Minutes, Irvine 3/92
P802.10-92/6	Minutes, Bloomington 7/92
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Table 29. IEEE 802.11 Wireless LANs Working Group Docs

IEEE 802.11	**Description**
P802.4L/89-14A	Factory RF Channel Modeling 2/89
P802.4L/89-14B	UHF Characterisation in Factory, 6/88
P802.4L/90-08A	Microwave Oven Interference Measurement 2/90
P802.11-93/20	Draft Std, Wireless LAN MAC/PHY Specification
P802.11-92/134	Minutes of Nov 92 and Jan 93 Meetings and papers
Note: * = New or revised document, *** = Can be ordered by calling 1-800-678-IEEE	

Appendix B. Special Notices

This publication is intended to help customers, systems engineers, services specialists, and marketing specialists understand LANs and IBM LAN solutions and architectures for planning and support purposes. The information in this publication is not intended as the specification of any programming interfaces that are provided by the products mentioned in this book. See the PUBLICATIONS section of the IBM Programming Announcement for IBM LAN products for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM′s product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM′s intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The information about non-IBM (″vendor″) products in this manual has been supplied by the vendor and IBM assumes no responsibility for its accuracy or completeness. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer′s ability to evaluate and integrate them into the customer′s operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

ACF/VTAM
Advanced Peer-to-Peer Networking
AFP
AIX
AIX/6000
AnyNet
APPN
ARTour
AS/400
CICS
DatagLANce
IBM
LAN Distance
LANStreamer
MVS/ESA
NetFinity
NetView
Nways
Operating System/2
OS/2
PS/2
RT
S/390
SP
SystemView
Trouble Ticket
VM/ESA
VTAM

400

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc.

PC Direct is a trademark of Ziff Communications Company and is used by IBM Corporation under license.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

Microsoft, Windows, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

Adobe	Adobe Systems Incorporated
Advantis	Advantis
Alpha AXP, DEC, DECnet, Digital, PATHWORKS, ULTRIX, VAX	Digital Equipment Corporation
Apollo, NCS, Network Computing System	Apollo Computer, Incorporate
Apple, AppleTalk, EtherTalk, LocalTalk, Macintosh, TokenTalk	Apple Computer, Incorporated
ARCnet	Datapoint Corporation
ARDIS	Ardis Company
AT&T	American Telephone and Telegraph Company
Attachmate	Attachmate Corporation
Banyan, VINES	Banyan Systems, Incorporated
Cisco	Cisco Systems, Incorporated
Clearinghouse, Xerox, Xerox Network Systems, XNS	Xerox Corporation
Compaq	Compaq Computer Corporation
EtherLink, 3Com	3Com Corporation
Hewlett-Packard, HP, OpenView	Hewlett-Packard Company
HYPERchannel	Network Systems Corporation
IDNX	Network Equipment Technologies, Incorporated
Intel, Pentium, 386	Intel Corporation
IPX, LANalyzer, NetWare, Novell	Novell, Incorporated
LANtastic	Artisoft, Incorporated
Lotus, Lotus Notes	Lotus Development Corporation
Mobitex	Televerket
Motorola	Motorola, Incorporated
MS-DOS, Windows NT	Microsoft Corporation
Network File System, NFS, SunOS	Sun Microsystems, Incorporated
Open Software Foundation, OSF, OSF/1	Open Software Foundation, Incorporated
PostScript	Adobe Systems Incorporated
Proteon	Proteon, Incorporated
Qualitas, 386MAX	Qualitas
Quarterdeck	Quarterdeck Corporation
Sniffer Network Analyzer	Network General Corporation
SPARCstation	SPARC International, Incorporated
SynOptics	SynOptics Communication Incorporated
Toshiba	Toshiba Corporation
Ungermann-Bass	Ungermann-Bass Corporation
Wellfleet	Wellfleet Communications, Incorporated
PEX, X Window System, X-Windows	Massachusetts Institute of Technology
X/Open	X/Open Company Limited

Other trademarks are trademarks of their respective companies.

Appendix C. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

C.1 International Technical Support Organization Publications

For information on ordering these ITSO publications see "How To Get ITSO Redbooks" on page 237.

- *FDDI Concepts and Products*, GG24-3865-00
- *High-Speed Networking Technology: An Introductory Survey*, GG24-3816-01 (available on CD-ROM, SK2T-6022)
- *Asynchronous Transfer Mode (ATM) Technical Overview*, SG24-4625-00
- *TCP/IP Tutorial and Technical Overview,* GG24-3376

A complete list of International Technical Support Organization publications, known as redbooks, with a brief description of each, may be found in:

International Technical Support Organization Bibliography of Redbooks, GG24-3070.

C.2 Other Publications

These publications are also relevant as further information sources:

- *IBM Token-Ring Network Introduction and Planning Guide,* GA27-3677
- *F-Coupler Planning Guide,* GA27-3949
- *IBM Cabling System Planning and Installation Guide,* GA27-3361
- *IBM Cabling System Optical Fiber Planning and Installation Guide,* GA27-3943
- *FDDI Introduction and Planning Guide,* GA27-3892
- *IBM Token-Ring Network Architecture Reference,* SC30-3374
- *Systems Network Architecture Technical Overview,* GC30-3073-04
- *Internetworking with TCP/IP Volume I: Principles, Protocols, and Architecture*, SC31-6144

How To Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at URL http://www.redbooks.ibm.com/redbooks.

How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **PUBORDER** — to order hardcopies in United States
- **GOPHER link to the Internet**

  Type GOPHER.WTSCPOK.ITSO.IBM.COM
- **Tools disks**

  To get LIST3820s of redbooks, type one of the following commands:

  ```
  TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
  TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
  ```

  To get lists of redbooks:

  ```
  TOOLS SENDTO WTSCPOK TOOLS REDBOOKS GET REDBOOKS CATALOG
  TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
  TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET LISTSERV PACKAGE
  ```

  To register for information on workshops, residencies, and redbooks:

  ```
  TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1996
  ```

  For a list of product area specialists in the ITSO:

  ```
  TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ORGCARD PACKAGE
  ```
- **Redbooks Home Page on the World Wide Web**

  http://w3.itso.ibm.com/redbooks/redbooks.html
- **IBM Direct Publications Catalog on the World Wide Web**

  http://www.elink.ibmlink.ibm.com/pbl/pbl

  IBM employees may obtain LIST3820s of redbooks from this page.
- **ITSO4USA category on INEWS**
- **IBM Bookshop** — send orders to:

  USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL
- **Internet Listserver**

  With an Internet E-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an E-mail note to announce@webster.ibmlink.ibm.com with the keyword subscribe in the body of the note (leave the subject line blank). A category form and detailed instructions will be sent to you.

How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **IBMLINK**

  Registered customers have access to PUBORDER to order hardcopy, to REDPRINT to obtain BookManager BOOKs

- **IBM Bookshop** — send orders to:

  usib6fpl@ibmmail.com (United States)
  bookshop@dk.ibm.com (Outside United States)

- **Telephone orders**

1-800-879-2755	Toll free, United States only
(45) 4810-1500	Long-distance charge to Denmark, answered in English
(45) 4810-1200	Long-distance charge to Denmark, answered in French
(45) 4810-1000	Long-distance charge to Denmark, answered in German
(45) 4810-1600	Long-distance charge to Denmark, answered in Italian
(45) 4810-1100	Long-distance charge to Denmark, answered in Spanish

- **Mail Orders** — send orders to:

  IBM Publications
  P.O. Box 9046
  Boulder, CO 80301-9191
  USA

  IBM Direct Services
  Sortemosevej 21
  DK-3450 Allerød
  Denmark

- **Fax** — send orders to:

1-800-445-9269	Toll-free, United States only
45-4814-2207	Long distance to Denmark

- **1-800-IBM-4FAX (United States only)** — ask for:

  Index # 4421 Abstracts of new redbooks
  Index # 4422 IBM redbooks
  Index # 4420 Redbooks for last six months

- **Direct Services**

  Send note to softwareshop@vnet.ibm.com

- **Redbooks Home Page on the World Wide Web**

  http://www.redbooks.ibm.com/redbooks

- **IBM Direct Publications Catalog on the World Wide Web**

  http://www.elink.ibmlink.ibm.com/pbl/pbl

- **Internet Listserver**

  With an Internet E-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an E-mail note to announce@webster.ibmlink.ibm.com with the keyword subscribe in the body of the note (leave the subject line blank). A category form and detailed instructions will be sent to you.

IBM Redbook Order Form

Please send me the following:

Title	**Order Number**	**Quantity**

- **Please put me on the mailing list for updated versions of the IBM Redbook Catalog.**

First name Last name

Company

Address

City Postal code Country

Telephone number Telefax number VAT number

- Invoice to customer number

- Credit card number

Credit card expiration date Card issued to Signature

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

DO NOT SEND CREDIT CARD INFORMATION OVER THE INTERNET.

Index

Printed in U.S.A.

SG24-4753-00