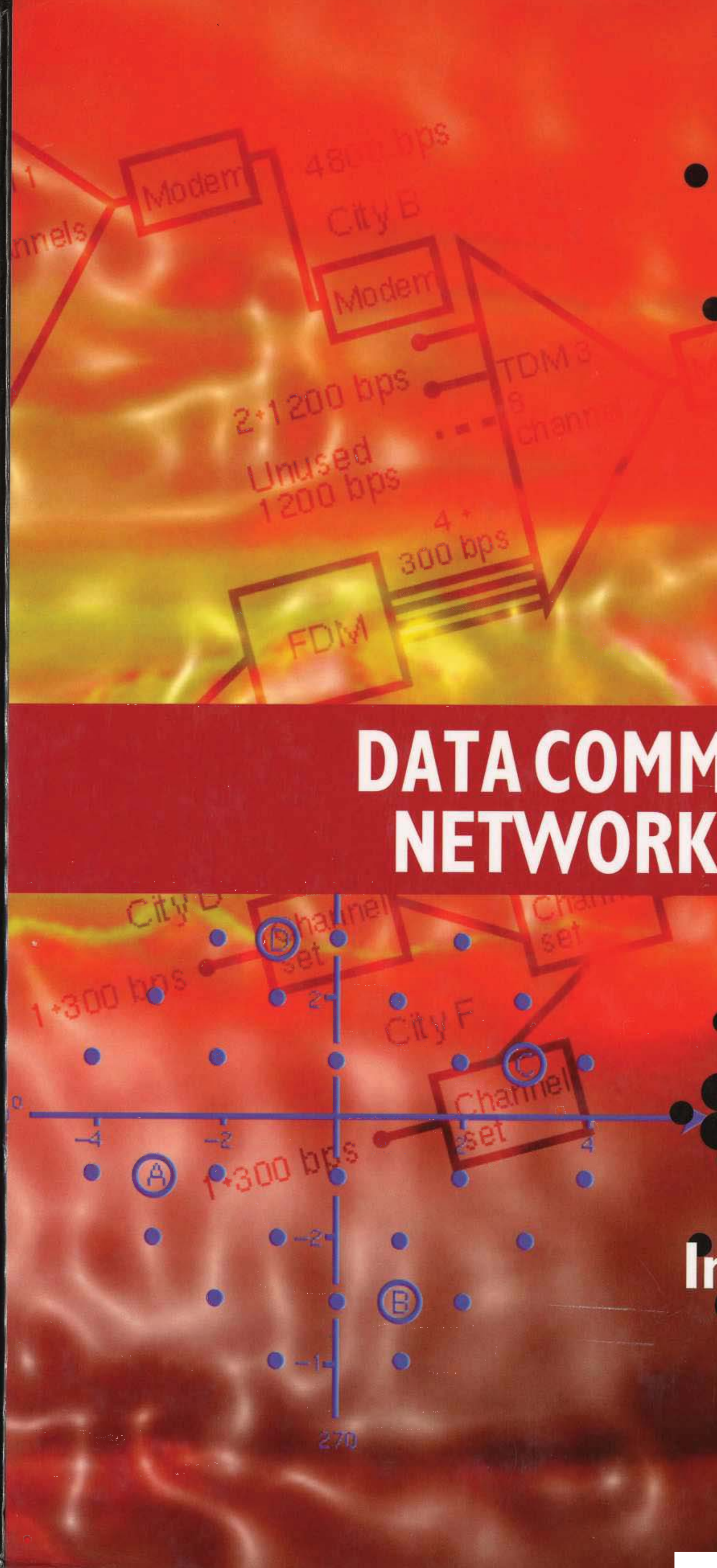
GILBERT HELD



DATA COMMUNICATIONS NETWORKING DEVICES

Fourth Edition

Operation, Utilization and LAN and WAN Internetworking

WILEY

ARRIS EX. 1006

BOOKS BY GILBERT HELD, PUBLISHED BY WILEY

DATA COMMUNICATIONS

UNDERSTANDING DATA COMMUNICATIONS
From Fundamentals to Networking
Second Edition

This comprehensive, introductory text offers a fundamental understanding of the role, operation and utilization of communication devices and facilities. Ideal for both the data communications student and the professional telecommunications engineer, this foundation text will appeal to anyone who wishes to gain a detailed knowledge of the evolution and future of communications.

1996 0 471 96820 X

DATA COMMUNICATIONS NETWORKING DEVICES
Third Edition

This popular and authoritative text evaluates the networking products that can be used in the design, operation and optimization of a data communications network. Reflecting the many technological advances in the field, this unique book covers over 30 networking devices.

1992 0 471 93072 5

DATA AND IMAGE COMPRESSION
Tools and Techniques
Fourth Edition

Considerably expanded, this fourth edition explains how practical data and image compression techniques are vital for efficient, low-cost transmission and data storage requirements. The book takes into account the advances in this area, including image and fax compression.

1996 0 471 95247 8

FRAME RELAY NETWORKING
Covers all aspects of Frame Relay Networking, from the evolution and rationale to architecture and equipment.
1999 0 471 98578 3

INTERNETWORKING LANs AND WANs
Concepts, Techniques and Methods
Second Edition

Internetworking is one of the fastest growing markets in the field of computer communications. However, the interconnection of LANs and WANs tends to cause significant technological and administrative difficulties. This updated version provides valuable guidance, enabling the reader to avoid the pitfalls and achieve successful connection.
1998 0 471 97514 1

THE MULTIPLEXER REFERENCE MANUAL

Designed to provide the reader with a detailed insight into the operation, utilization and networking of six distinct types of multiplexers, this book will appeal to practising electrical, electronic and communications engineers, students in electronics, network analysts and designers.
1993 0 471 93484 4

PRACTICAL NETWORK DESIGN TECHNIQUES

Many network design problems are addressed and solved in this informative volume. Gil Held confronts a range of issues including through-put problems, line facilities, economic trade-offs and multiplexers. Readers are also shown how to determine the numbers of ports, dial-in lines and channels to install on communications equipment in order to provide a defined level of service.
1991 0 471 92938 7 (Set)

Please refer to the back of the book for further details

DATA COMMUNICATIONS NETWORKING DEVICES: OPERATION, UTILIZATION AND LAN AND WAN INTERNETWORKING

Fourth Edition

DATA COMMUNICATIONS NETWORKING DEVICES: OPERATION, UTILIZATION AND LAN AND WAN INTERNETWORKING

Fourth Edition

Gilbert Held
4-Degree Consulting
Macon, Georgia
USA

JOHN WILEY & SONS
Chichester · New York · Weinheim · Brisbane · Singapore · Toronto

Baffins Lane, Chichester,
West Sussex PO19 1UD, England

National 01243 779777
International (+44) 1243 779777
e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on http://www.wiley.co.uk or http://www.wiley.com

Other Wiley Editorial Offices

John Wiley & Sons, Inc., 605 Third Avenue,
New York, NY 10158-0012, USA

WILEY-VCH Verlag GmbH, Pappelallee 3,
D-69469 Weinheim, Germany

Jacaranda Wiley Ltd, 33 Park Road, Milton,
Queensland 4064, Australia

John Wiley & Sons (Canada) Ltd, 22 Worcester Road,
Rexdale, Ontario M9W 1L1, Canada

John Wiley & Sons (Asia) Pte Ltd, Clementi Loop #02-01,
Jin Xing Distripark, Singapore 129809

Libaray of Congress Cataloging-in-Publication Data

Held, Gilbert, 1943–
Data communications networking devices : operation, utilization, and LAN and WAN internetworking / Gilbert Held. — 4th ed.
p. cm.
Includes index.
ISBN 0-471-97515-X (alk. paper)
1. Computer networks. 2. Computer networks—Equipment and supplies. 3. Data transmission systems. I. Title.
TK5105.5.H44 1998
004.6—dc21 98-27200
CIP

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0 471 97515-X

Typeset in 10/12pt Imprint by Thomson Press (India) Ltd, New Delhi, India
Printed and bound in Great Britain by Bookcraft (Bath) Ltd
This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

To Beverly, Jonathan and Jessica
for their patience understanding and support—
I love you all

To Dr Alexander Ioffe and family of Moscow—
congratulations on next year in Jerusalem being each year!

CONTENTS

PREFACE

Over fifteen years ago I introduced the first edition of this book with the statement 'data communications networking devices are the building blocks upon which networks are constructed.' Although networking technology has made significant advances, that statement retains its validity. Today you can use devices such as bridges and routers that were non-existent in the late 1970s to link local and wide area networks together, while boosting LAN productivity and access through the use of switches and remote access servers that represent products of the 1990s. Thus, the basic rationale and goal of this fourth edition, which is to provide readers with an intimate awareness of the operation and utilization of important networking products that can be used in the design, modification, or optimization of a data communications network, has not changed from the rationale and goal of the first edition. What has changed is the scope and depth of the material included in this book.

In developing this new edition I have taken into consideration and acted upon comments received from both individuals and professors who used the book for a college course on networking. Major changes include an expansion and subdivision of the Fundamental Concepts chapter, which now covers both WANs and LANs in a series of separate chapters focused upon fundamental concepts and advanced networking topics. Other significant changes in this new edition include a chapter covering Wide Area Networks as a separate entity and another covering LAN internetworking devices. In addition, a significant amount of material was revised and updated to provide detailed information covering the operation and utilization of additional networking devices and the updating of information concerning the operating characteristics of other devices. To facilitate the use of this book as a text as well as for reader review purposes, the questions at the end of each chapter reference the sections in each chapter. Through the use of a numbering scheme, students can easily reference an appropriate section in the book for assistance in answering a question while instructors can easily reference the assignment of questions to reading assignments based upon specific sections within chapters.

The expansion of the Fundamental Concepts chapter followed by the addition of two new chapters covering wide area networks and local area networks provides readers new to the field of data communications with the ability to use these chapters as a detailed introduction to this field. For more experienced readers the information in these chapters can be used as a reference to the many facets of data communications.

The new chapter covering wide area networks first explains the different types of networks and then examines network architecture and the flow of data in several popular networks. Similarly, the new chapter on LANs provides a solid foundation concerning the topology, access methods, and operation of several types of popular local area networks, laying the groundwork for detailed information concerning the operation of WAN and LAN internetworking devices presented in later chapters in this book.

Similar to prior editions of this book, this edition was structured for a two semester course at a high level undergraduate or first-year graduate course level. In addition, this book can be used as a comprehensive reference to the operation and utilization of different networking devices and as a self-study guide for individuals who wish to pace themselves at their leisure.

As I once again rewrote this book, I again focused attention upon explaining communications concepts which required an expansion of an already comprehensive introductory chapter into a series of three chapters in order to cover the fundamental concepts common to all phases of data communications. All three chapters should be read first by those new to this field and can be used as a review mechanism for readers with a background in communications concepts. Thereafter, each chapter is written to cover a group of devices based upon a common function.

Through the use of numerous illustrations and schematic diagrams, I believe readers will easily be able to see how different devices can be integrated into networks, and some examples should stimulate new ideas for even the most experienced person. At the end of each chapter I have included a comprehensive series of questions that cover many of the important concepts covered in the chapter. These questions can be used by the reader as a review mechanism prior to going forward in the book.

For those readers actually involved in the sizing of network devices I have include several appendices in this book that cover this area. Since the mathematics involved in the sizing process can result in a considerable effort to obtain the required data, I have enclosed computer program listings that readers can use to generate a series of sizing tables. Then after reading the appendices and executing the computer programs, you can reduce many sizing problems to a table lookup procedure. As always I look forward to receiving reader comments, either though my publisher whose address is on the back cover of this book or via email to 235-8068@meimail.com

Gilbert Held
Macon, Georgia

ACKNOWLEDGEMENTS

The preparation of a manuscript that gives birth to a book requires the cooperation and assistance of many persons.

First and foremost, I must thank my family for enduring those long nights and missing weekends while I drafted and redrafted the manuscript to correspond to each of the editions of this book. The preparation of the first edition was truly a family affair, since both my wife and my son typed significant portions of the manuscript on our mobile Macintosh, with both my family and the Macintosh traveling a considerable distance during the preparation of the manuscript. For the preparation of the second and third editions of this book I am grateful for the efforts of Mrs Carol Ferrell who turned my handwritten inserts and drawings into a legible manuscript. As a frequent traveler I write the old-fashioned way—with pen and paper—to avoid battery drain and electrical outlet incompatibilities affecting my productivity. However, it still requires a talent to decipher my handwriting, especially since aircraft turbulence periodically affects my writing effort. Thus, I am most appreciative of Mrs Linda Hayes's efforts in turning my latest manuscript into typed pages that resulted in the book you are reading. In addition, I would also like to thank Auerbach Publishers, Inc., for permitting me to use portions of articles I previously wrote for their *Data Communications Management* publication. Excerpts from these articles were used for developing the section covering integrated services digital network (ISDN) presented in Chapter 1, for expanding the statistical and T1 multiplexing in Chapter 5, and for the voice digitization, data compression and fiber optic transmission systems presented in Chapter 7.

Last but not least, one's publishing editor, editorial supervisor and desk editor are the critical link in converting the author's manuscript into the book you are now reading. To Ian Shelley, who enthusiastically backed the first edition of this book, I would like to take the opportunity to thank you again for your efforts. To Ann-Marie Halligan and Ian McIntosh who provided me with the opportunity to produce the third and fourth editions, I would again like to acknowledge your efforts in a multinational way. Cheers! To Stuart Gale, Robert Hambrook, and Sarah Lock who moved my manuscripts through proofs and into each edition of this book, many thanks for your fine effort.

1

FUNDAMENTAL WIDE AREA NETWORKING CONCEPTS

The main purpose of this chapter is to provide readers with a common level of knowledge concerning wide area networking (WAN) communications concepts. To achieve this goal we will examine the fundamental concepts associated with wide area network communications. Commencing with a description of the three components necessary to establish communications, we will expand our base of knowledge by discussing the types of line connections available for use, different types of transmission services and transmission devices, carrier offerings, transmission modes and techniques, and other key concepts. In doing so we will obtain a base of knowledge that will allow readers to better understand how devices and transmission facilities are interconnected to establish networks and interconnect geographically separated local area networks which is the focus of Chapters 2 and 3. In addition, the material presented in this chapter will enable readers to better understand the operation and utilization of devices explained in subsequent chapters.

While the transmission of data may appear to be a simple process, many factors govern the success or failure of a communications session. In addition, the exponential increase in the utilization of personal computers and a corresponding increase in communications between personal computers and other personal computers and large-scale computers had enlarged the number of hardware and software parameters you must consider. Although frequently we will use the terms 'terminal' and 'personal computers' interchangeably and refer to them collectively as 'terminals' in this book, in certain instances we will focus our attention upon personal computers in order to denote certain hardware and software characteristics unique to such devices. In these instances we will use the term 'personal computer' to explicitly reference this terminal device. In other instances we will use the term 'workstation' to refer to any computational device from a personal computer to a mainframe that is connected to a local area network. Such general use of this term should not be confused with its usage to represent a specialized powerful computer designed to facilitate the mathematical operations that are required to generate 3-D graphics, perform computer-aided design or similar compute-intensive operations, a topic beyond the scope of this book.

1.1 COMMUNICATIONS SYSTEM COMPONENTS

To transmit information between two locations it is necessary to have a transmitter, a receiver, and a transmission medium which provides a path or link between the transmitter and the receiver. In addition to transmitting signals, a transmitter must be capable of translating information from a form created by humans or machines into a signal suitable for transmission over the transmission medium. The transmission medium provides a path to convey the information to the receiver without introducing a prohibitive amount of signal distortion that could change the meaning of the transmitted signal. The receiver then converts the signal from its transmitted form into a form intelligible to humans or machines.

1.2 LINE CONNECTIONS

Three basic types of line connections are available to connect terminal devices to computers or to other terminals via a wide area network: dedicated, switched, and leased lines.

Dedicated line

A dedicated line is similar to a leased line in that the terminal is always connected to the device on the distant end, transmission always occurs on the same path, and, if required, the line may be able to be tuned to increase transmission performance.

The key difference between a dedicated and a leased line is that a dedicated line refers to a transmission medium internal to a user's facility, where the customer has the right of way for cable laying, whereas a leased line provides an interconnection between separate facilities. The term facility is usually employed to denote a building, office, or industrial plant. Dedicated lines are also referred to as direct connect lines and normally link a terminal or business machine on a direct path through the facility to another terminal or computer located at that facility. The dedicated line can be a wire conductor installed by the employees of a company or by the computer manufacturer's personnel, or it can be a local line installed by the telephone company.

Normally, the only cost associated with a dedicated line in addition to its installation cost is the cost of the cable required to connect the devices that are to communicate with one another.

Leased line

A leased line is commonly called a private line and is obtained from a communications company to provide a transmission medium between two facilities which could be in separate buildings in one city or in distant cities. In addition to a one-time installation charge, the communications carrier will normally bill the user on a monthly basis for the leased line, with the cost of the line usually based upon the distance between the locations to be connected.

Switched line

A switched line, often referred to as a dial-up line, permits contact with all parties having access to the analog public switched telephone network (PSTN) or the digital switched network. If the operator of a terminal device wants access to a computer, he or she dials the telephone number of a telephone which is connected to the computer. In using switched or dial-up transmission, telephone company switching centers establish a connection between the dialing party and the dialed party. After the connection is set up, the terminal and the computer conduct their communications. When communications are completed, the switching centers disconnect the path that was established for the connection and restore all paths used so they become available for other connections.

The cost of a call on the PSTN is based upon many factors which include the time of day when the call was made, the distance between called and calling parties, the duration of the call and whether or not operator assistance was required in placing the call. Direct dial calls made from a residence or business telephone without operator assistance are billed at a lower rate than calls requiring operator assistance. In addition, most telephone companies have three categories of rates: 'weekday', 'evening' and 'night and weekend'. Typically, calls made between 8 a.m. and 5 p.m. Monday to Friday are normally billed at a 'weekday' rate, while calls between 5 p.m. and 10 p.m. on weekdays are usually billed at an 'evening' rate, which reflects a discount of approximately 25% over the 'weekday' rate. The last rate category, 'night and weekend', is applicable to calls made between 10 p.m. and 8 a.m. on weekdays as well as anytime on weekends and holidays. Calls during this rate period are usually discounted 50% from the 'weekday' rate.

Table 1.1 contains a sample PSTN rate table which is included for illustrative purposes but which should not be used by readers for determining the actual cost of a PSTN call as the cost of intrastate calls by state and interstate calls varies. In addition, the cost of using different communications carriers to place a call between similar locations will typically vary from vendor to vendor and readers should obtain a current interstate and/or state schedule from the vendor they plan to use in order to determine or project the cost of using PSTN facilities.

Table 1.1 Sample PSTN rate table (cost per minute in cents)

	Rate category					
	Weekend		Evening		Night and weekend	
Mileage between location	First minute	Each additional minute	First minute	Each additional minute	First minute	Each additional minute
1–100	0.31	0.19	0.23	0.15	0.15	0.10
101–200	0.35	0.23	0.26	0.18	0.17	0.12
201–400	0.48	0.30	0.36	0.23	0.24	0.15

Cost trends

Although many vendors continue to maintain a rate table similar to the one shown in Table 1.1, other vendors have established a variety of flat-rate billing schemes in which calls made anywhere within a country are billed at a uniform cost per minute regardless of distance. During 1996 Sprint introduced a 10 cents per minute long-distance charge for calls made between 7 p.m. and 7 a.m. Monday through Friday and all day at weekends. During 1997 AT&T introduced a flat 15 cents per minute charge for calls made anywhere in the United States at any time. Both offerings require the selection of one communications carrier as your primary long-distance carrier and the selection of an appropriate calling plan to obtain flat-rate billing.

Factors to consider

Cost, speed of transmission, and degradation of transmission are the primary factors used in the selection process between leased and switched lines.

As an example of the economics associated with comparing the cost of PSTN and leased line usage, assume that a personal computer user located 50 miles from a mainframe needs to communicate between 8 a.m. and 5 p.m. with the mainframe once each business day for a period of 30 minutes. Using the data in Table 1.1, each call would cost $0.31 \times 1 + 0.19 \times 29$ or \$5.82. Assuming there are 22 working days each month, the monthly PSTN cost for communications between the PC and the mainframe would be \$5.82 × 22 or \$128.04. If the monthly cost of a leased line between the two locations was \$250, it is obviously less expensive to use the PSTN for communications. Suppose the communications application lengthened in duration to 2 hours per day. Then, from Table 1.1, the cost per call would become $0.31 \times 1 + 0.19 \times 119$ or \$22.92. Again assuming 22 working days per month, the monthly PSTN charge would increase to \$504.24, making the leased line more economical.

Thus, if data communications requirements involve occasional random contact from a number of terminals at different locations and each call is of short duration, dial-up service is normally employed. If a large amount of transmission occurs between a computer and a few terminals, leased lines are usually installed between the terminal and the computer.

Since a leased line is fixed as to its routing, it can be conditioned to reduce errors in transmission as well as permit ease in determining the location of error conditions since its routing is known. Normally, analog switched circuits are used for transmission at speeds up to 33 600 bits per second (bps); however, in certain situations data rates as high as 56 000 bps are achievable when transmission on the PSTN occurs through telephone company offices equipped with modern electronic switches.

Some of the limiting factors involved in determining the type of line to use for transmission between terminal devices and computers are listed in Table 1.2. Information in this table is applicable to both analog and digital transmission facilities and as such was generalized. For more specific information concerning the speed of transmission obtainable on analog and digital transmission facilities, readers are referred to the analog facilities and digital facilities subsections in Section 1.3.

Table 1.2 General line selection guide

Line type	Distance between transmission points	Speed of transmission	Use of transmission
Dedicated (direct connect)	Local	Limited by conductor	Short or long duration
Switched (dial-up)	Limited by telephone access availability	Normally up to 33 600 bps (analog), 1.544 Mbps (digital)	Short-duration transmission
Leased (private)	Limited by telephone company availability	Limited by type of facility	Long duration or numerous short duration calls

1.3 TYPES OF SERVICES AND TRANSMISSION DEVICES

Digital devices which include terminals, mainframe computers, and personal computers transmit data as unipolar digital signals, as indicated in Figure 1.1(a). When the distance between a terminal device and a computer is relatively short, the transmission of digital information between the two devices may be obtained by cabling the devices together. As the distance between the two devices increases, the pulses of the digital signals become distorted because of the resistance, inductance, and capacitance of the cable used as a transmission medium. At a certain distance between the two devices the pulses of the digital data will distort, such that they are unrecognizable by the receiver, as illustrated in Figure 1.1(b). To extend the transmission distance between devices, specialized equipment must be employed, with the type of equipment used dependent upon the type of transmission medium employed.

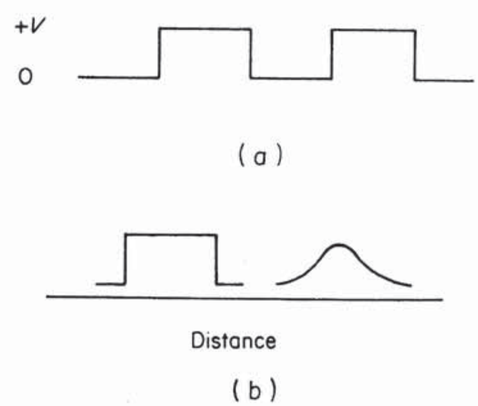


Figure 1.1 (a) Digital signaling. Digital devices to include terminals and computers transmit data as unipolar digital signals. (b) Digital signal distortion. As the distance between the transmitter and receiver increases digital signals become distorted because of the resistance, inductance, and capacitance of the cable used as a transmission medium

Digital repeaters

You can transmit data in a digital or analog form. To transmit data long distances in digital form requires repeaters to be placed on the line at selected intervals to reconstruct the digital signals. The repeater is a device that essentially scans the line looking for the occurrence of a pulse and then regenerates the pulse into its original form. Thus, another name for the repeater is a data regenerator. As illustrated in Figure 1.2, a repeater extends the communications distance between terminal devices to include personal computers and mainframe computers.

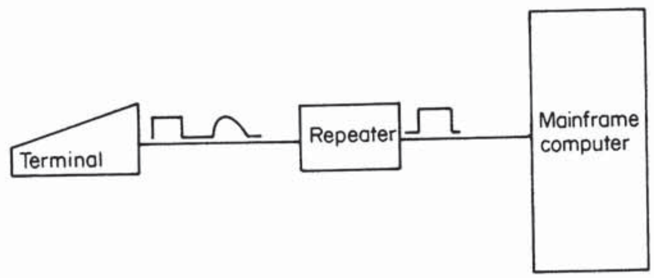


Figure 1.2 Transmitting data in digital format. To transmit data long distances in digital format requires repeaters to be placed on the line to reconstruct the digital signals

Unipolar and bipolar signaling

Since unipolar signaling results in a dc voltage buildup when transmitting over long distance, digital networks require unipolar signals to be converted into a modified bipolar format for transmission on this type of network. This requires the installation at each end of the circuit of a device known as a data service unit (DSU) in the United States and a network terminating unit (NTU) in the United Kingdom. The utilization of DSUs for transmission of data on a digital network is illustrated in Figure 1.3. Although not shown, readers should note that repeaters are placed on the path between DSUs to regenerate the bipolar signals. Later in this chapter we will examine digital facilities in more detail.

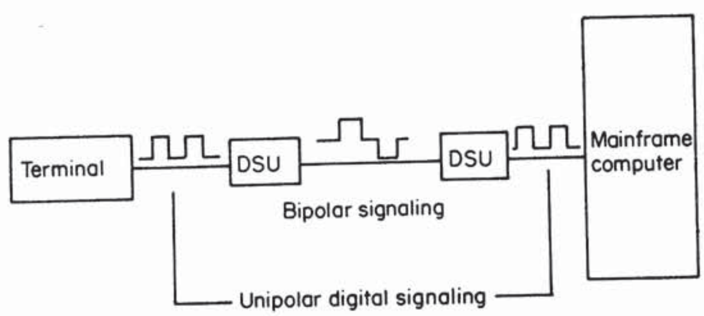


Figure 1.3 Transmitting data on a digital network. To transmit data on a digital network, the unipolar digital signals of terminal devices and computers must be converted into a bipolar signal

Repeaters are primarily used on wide area network digital transmission facilities at distances of approximately 6000 feet from one another on lines connecting subscribers to telephone company offices serving those subscribers. From local telephone company offices data will travel either by microwave or via fiber optic cable to a higher level telephone company office for routing through the telephone network hierarchy. By the late 1990s, over 99.9% of long-distance transmission was being carried in digital form via fiber optic cable. A vast majority of connections between telephone company subscribers and the local office serving those subscribers were, however, over twisted-pair copper cables that have amplifiers inserted to boost the strength of analog signals. Such connections require the conversion of digital signals into an analog form to enable the signal to be carried over the analog transmission facility.

Other digital signaling methods

In a LAN environment the full bandwidth of the cable is usually available for use. In comparison, the communications carrier commonly uses filters to limit the bandwidth usable on the local loop between a telephone company office and a subscriber's premises to 4 kHz or less. Although the absence of filters enables LAN designers to obtain a much higher data rate than that obtainable on a local loop, other operational considerations, to include the potential buildup of dc voltage and the cost of constructing equipment to operate at a high signaling rate to provide a high data transmission rate, resulted in the development of several digital signaling techniques used on LANs. Two of the more popular techniques are Manchester and Differential Manchester which are used on Ethernet and Token–Ring networks, respectively. In Chapter 3, when we focus our attention on local area networks, we will also turn our attention to the digital signaling methods used by different types of LANs.

Modems

Since telephone lines were originally designed to carry analog or voice signals, the digital signals transmitted from a terminal to another digital device must be converted into a signal that is acceptable for transmission by the telephone line. To effect transmission between distant points, a data set or modem is used. A modem is a contraction of the compound term modulator–demodulator and is an electronic device used to convert digital signals generated by computers and terminal devices into analog tones for transmission over telephone network analog facilities. At the receiving end, a similar device accepts the transmitted tones, reconverts them to digital signals, and delivers these signals to the connected device.

Signal conversion

Signal conversion performed by modems is illustrated in Figure 1.4. This illustration shows the interrelationship of terminals, mainframe computers, and

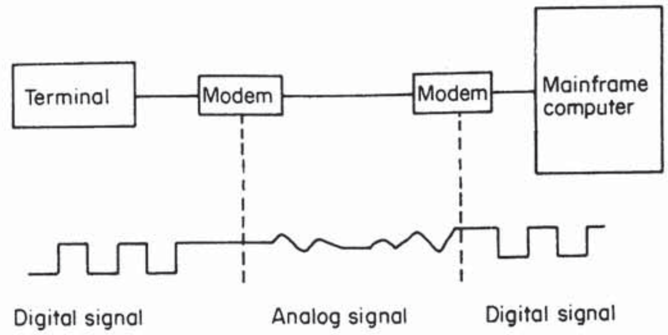


Figure 1.4 Signal conversion performed by modems. A modem converts (modulates) the digital signal produced by a terminal into an analog tone for transmission over an analog facility

transmission lines when an analog transmission service is used. Analog transmission facilities include both leased lines and switched lines; therefore, modems can be used for transmission of data over both types of analog line connections. Although an analog transmission medium used to provide a transmission path between modems can be a direct connect, a leased, or a switched line, modems are connected (hard-wired) to direct connect and leased lines, whereas they are interfaced to a switched facility; thus, a terminal user can communicate only with one distant location on a leased line, but with many devices when there is access to a switched line.

Acoustic couplers

Although popular with data terminal users in the 1970s, today only a very small percentage of persons use acoustic couplers for communications. The acoustic coupler is a modem whose connection to the telephone line is obtained by acoustically coupling the telephone headset to the coupler. The primary advantage of the acoustic coupler was the fact that it required no hard-wired connection to the switched telephone network, enabling terminals and personal computers to be portable with respect to their data transmission capability. Owing to the growth in modular telephone jacks, modems that interface the switched telephone network via a plug, in effect, are portable devices. Since many hotels and older office buildings still have hard-wired telephones, the acoustic coupler permits terminal and personal computer users to communicate regardless of the method used to connect a telephone set to the telephone network.

Signal conversion

The acoustic coupler converts the signals generated by a terminal device into a series of audible tones, which are then passed to the mouthpiece or transmitter of the telephone and in turn onto the switched telephone network. Information transmitted from the device at the other end of the data link is converted into audible

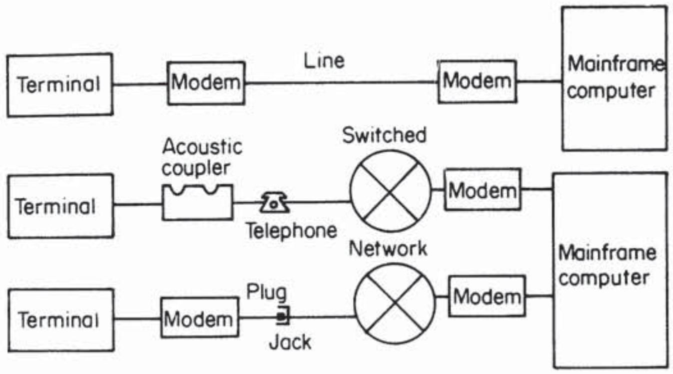


Figure 1.5 Interrelationship of terminals, modems, acoustic couplers, computers and analog transmission mediums. When using modems on an analog transmission medium, the line can be a dedicated, leased, or switched facility. Terminal devices can use modems or acoustic couplers to transmit via the switched network

tones at the earpiece of the telephone connected to the terminal's acoustic coupler. The coupler then converts those tones into the appropriate electrical signals recognized by the attached terminal. The interrelationship of terminals, acoustic couplers, modems and analog transmission media is illustrated in Figure 1.5.

In examining Figure 1.5, you will note that a circle subdivided into four equal parts by two intersecting lines is used as the symbol to denote the public switched telephone network or PSTN. This symbol will be used in the remainder of the book to illustrate communications occurring over this type of line connection.

Analog facilities

Several types of analog switched and leased line facilities are offered by communications carriers. Each type of facility has its own set of characteristics and rate structure. Normally, for extensive communications requirements, an analytic study is conducted to determine which type or types of service should be utilized to provide an optimum cost-effective service for the user. Common types of analog switched facilities are direct distance dialing, wide area telephone service, and foreign exchange service. The most common type of analog line is a voice grade private line.

DDD

Direct distance dialing (DDD) permits a person to dial directly any telephone connected to the public switched telephone network. The dialed telephone may be connected to another terminal device or mainframe computer. The charge for this service, in addition to installation costs, may be a fixed monthly fee if no long-distance calls are made, a message unit rate based upon the number and duration of local calls, or a fixed fee plus any long-distance charges incurred. Depending upon the time of day a long-distance call is initiated and its destination (intrastate or interstate), discounts from normal long-distance tolls are available for selected calls made without operator assistance.

WATS

Introduced by AT&T for interstate use in 1961, wide area telephone service (WATS) is now offered by most long-distance communications carriers. Its scope of coverage has been extended from the continental United States to Hawaii, Alaska, Puerto Rico, the US Virgin Islands, and Europe, as well as selected Pacific and Asian countries.

Wide area telephone service (WATS) may be obtained in two different forms, each designed for a particular type of communications requirement. Outward WATS is used when a specific location requires placing a large number of outgoing calls to geographically distributed locations. Inward WATS service provides the reverse capability, permitting a number of geographically distributed locations to communicate with a common facility. Calls on WATS are initiated in the same manner as a call placed on the public switched telephone network. However, instead of being charged on an individual call basis, the user of WATS facilities pays a flat rate per block of communications hours per month occurring during weekday, evening, and night and weekend time periods.

A voice-band trunk called an access line is provided to the WATS users. This line links the facility to a telephone company central office. Other than cost considerations and certain geographical calling restrictions which are a function of the service area of the WATS line, the user may place as many calls as desired on this trunk if the service is outward WATS or receive as many calls as desired if the service is inward. Inward WATS, the well-known '800' area code which was extended to the '888' area code during 1996, permits remotely located personnel to call your facility toll-free from the service area provided by the particular inward WATS-type of service selected. The charge for WATS is a function of the service area. This can be intrastate WATS, a group of states bordering the user's state where the user's main facility is located, a grouping of distant states, or International WATs which extends inbound 800 service to the United States from selected overseas locations. Another service very similar to WATS is AT&T's 800 READYLINE^{SM} service. This service is essentially similar to WATS; however, calls can originate or be directed to an existing telephone in place of the access line required for WATS service.

Figure 1.6 illustrates the AT&T WATS service area one for the state of Georgia. If this service area is selected and a user in Georgia requires inward WATS service, he or she will pay for toll-free calls originating in the states surrounding Georgia–Florida, Alabama, Mississippi, Tennessee, Kentucky, South Carolina, and North Carolina. Similarly, if outward WATS service is selected for service area one, a person in Georgia connected to the WATS access line will be able to dial all telephones in the states previously mentioned. The states comprising a service area vary based upon the state in which the WATS access line is installed. Thus, the states in service area one when an access line is in New York would obviously differ from the states in a WATS service area one when the access line is in Georgia. Fortunately, AT&T publishes a comprehensive book which includes 50 maps of

Figure 1.6 AT&T WATS service area one for an access line located in Georgia

the United States, illustrating the composition of the service areas for each state. Similarly, a time-of-day rate schedule for each state based upon state service areas is also published by AT&T.

In general, since WATS is a service based upon volume usage its cost per hour is less than the cost associated with the use of the PSTN for long-distance calls. Thus, one common application for the use of WATS facilities is to install one or more inward WATS access lines at a data processing center. Then, terminal and personal computer users distributed over a wide geographical area can use the inward WATS facilities to access the computers at the data processing center.

Since International 800 service enables employees and customers of US companies to call them toll-free from foreign locations, this service may experience a considerable amount of data communications usage. This usage can be expected to include applications requiring access to such databases as hotel and travel reservation information as well as order entry and catalog sales data updating. Persons traveling overseas with portable personal computers as well as office personnel using terminals and personal computers in foreign countries who desire access to computational facilities and information utilities located in the United States represent common International 800 service users. Due to the business advantages of WATS its concept has been implemented in several foreign countries, with inward WATS in the United Kingdom marketed under the term Freefone.

FX

Foreign exchange (FX) service may provide a method of transmission from a group of terminal devices remotely located from a central computer facility at less

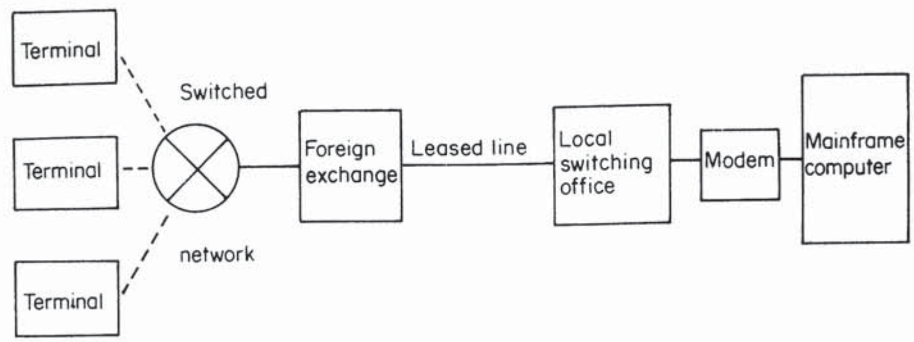


Figure 1.7 Foreign exchange (FX) service. A foreign exchange line permits many terminal devices to use the facility on a scheduled or on a contention basis

than the cost of direct distance dialing. An FX line can be viewed as a mixed analog switched and leased line. To use an FX line, a user dials a local number which is answered if the FX line is not in use. From the FX, the information is transmitted via a dedicated voice line to a permanent connection in the switching office of a communications carrier near the facility with which communication is desired. A line from the local switching office which terminates at the user's home office is included in the basic FX service. This is illustrated in Figure 1.7.

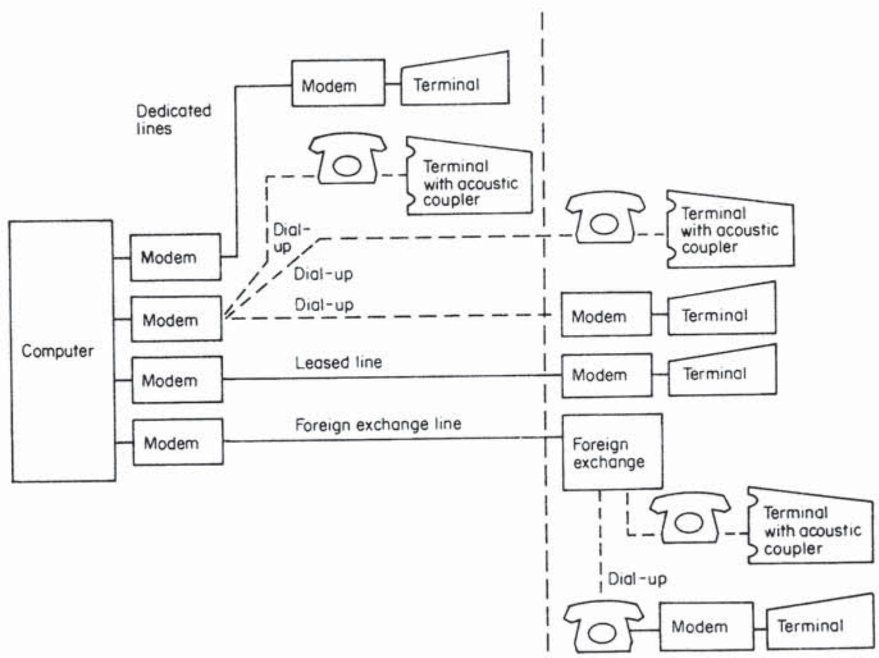


Figure 1.8 Terminal-to-computer connections via analog mediums. A mixture of dedicated, dialup, leased and foreign exchange lines can be exployed to connect local and remote terminals to a central computer facility

The use of an FX line permits the elimination of long-distance charges that would be incurred by users directly dialing a distant computer facility. Since only one person at a time may use the FX line, normally only groups of users whose usage can be scheduled are suitable for FX utilization. Figure 1.8 illustrates the possible connections between remotely located terminal devices and a central computer where transmission occurs over an analog facility.

The major difference between an FX line and a leased line is that any terminal dialing the FX line provides the second modem required for the transmission of data over the line; whereas a leased line used for data transmission normally has a fixed modem attached at both ends of the circuit.

Leased lines

The most common type of analog leased line is a voice grade private line. This line obtains its name from its ability to permit one voice conversation with frequencies between 300 and 3300 Hz to be carried on the line. In actuality, the bandwidth or range of frequencies that can be transmitted over a twisted-pair analog switched or analog leased line extends from 0 to approximately 1 MHz. However, to economize on the transmission of multiple voice conversations routed between telephone company offices, the initial design of the telephone company cable infrastructure resulted in the use of filters to remove frequencies below 300 Hz and above 3300 Hz, resulting in a 3000 Hz bandwidth for voice conversations. At telephone company offices voice conversations destined to another office are multiplexed onto a trunk or high speed line by frequency, requiring only 300 Hz of bandwidth per conversation, enabling one trunk to transport a large number of voice conversations shifted in frequency from one another. At the distant office other frequency division multiplexing equipment shifts each conversation back into its original frequency range as well as routing the call to its destination. Although the use of filters has considerably economized on the cost of routing multiple calls on trunks connecting telephone company offices, they have resulted in a bandwidth limit of 3000 Hz which makes high speed transmission on an analog loop most difficult to obtain. As we turn our attention to the operation of different types of modems later in this book, we will also obtain an appreciation of how the 3000 Hz bandwidth of analog lines limits the communications rate to most homes and many offices.

Figure 1.9 illustrates the typical routing of a leased line in the United States. The routing from each subscriber location to a telephone company central office serving the subscriber is known as a local loop. Normally the local loop is a two-wire or four-wire copper single or dual twisted-pair cable with amplifiers inserted

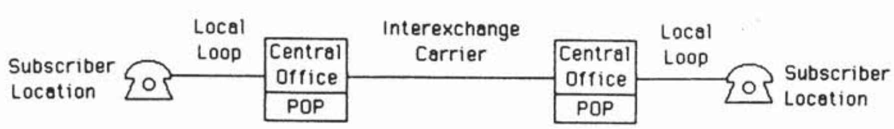


Figure 1.9 Leased line routing. Leased lines are routed from a local telephone company serving a subscriber to an interexchange carrier at the point of presence (POP)

on the local loop to boost the strength of the signal. Both the local loop and the central office are operated by the telephone company serving each subscriber location. If the leased line is routed outside the local telephone company's serving area it must be connected to an interexchange carrier (IXC), such as AT&T, MCI, or Sprint. The location where this interconnection takes place is called the point of presence (POP), which is normally located in the central office of the local telephone company. Although data on an analog leased line flows in an analog format, by the early 1990s most interexchange carriers digitized analog signals at the POP. Thus, between POPs most analog data is actually carried in digital form. Since the local loop is still an analog medium, it is the local loop which limits the data transmission rate obtainable through the use of an analog leased line. By 1998 modems permitting a 33.6 kbps data transmission rate on leased lines were commonly available, and some vendors had introduced products that allow data rates of up to 56 kbps in one direction and up to 33.6 kbps in the opposite direction, a technique referred to as asymmetrical transmission.

Digital facilities

In addition to the analog service, numerous digital service offerings have been implemented by communications carriers over the last decade. Using a digital service, data is transmitted from source to destination in digital form without the necessity of converting the signal into an analog form for transmission over analog facilities as is the case when modems or acoustic couplers are interfaced to analog facilities.

To understand the ability of digital transmission facilities to transport data requires an understanding of digital signaling techniques. Those techniques provide a mechanism to transport data end-to-end in modified digital form on LANs as well as on wide area networks that can connect locations hundreds or thousands of miles apart.

Digital signaling

Digital signaling techniques have evolved from use in early telegraph systems to provide communications for different types of modern technology, ranging in scope from the data transfer between a terminal and a modem to the flow of data on a LAN and the transport of information on high speed wide area network digital communications lines. Instead of one signaling technique numerous techniques are used, each technique having been developed to satisfy different communications requirements. In this section we will focus our attention upon digital signaling used on wide area network transmission facilities, deferring a discussion of LAN signaling until later in this book.

Unipolar non-return to zero

Unipolar non-return to zero (NRZ) is a simple type of digital signaling which was originally used in telegraphy. Today, unipolar non-return to zero signaling is used

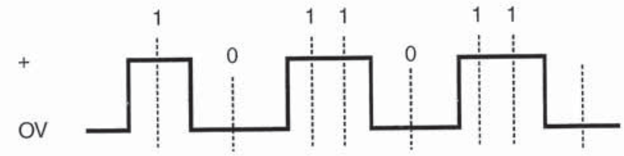


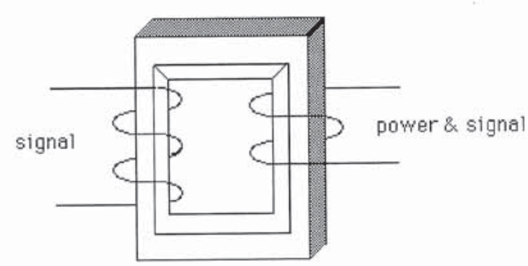
Figure 1.10 Unipolar non-return to zero signaling

in computers as well as by the common RS–232/V.24 interface between data terminal equipment and data communications equipment.

Figure 1.10 illustrates an example of unipolar non-return to zero signaling. In this signaling scheme, a dc current or voltage represents a mark, while the absence of current or voltage represents a space. Since voltage or current does not return to zero between adjacent set bits, this signaling technique is called non-return to zero. Since voltage or current is only varied from 0 to some positive level the pulses are unipolar, hence the name unipolar non-return to zero.

There are several problems associated with unipolar non-return to zero signaling which make it unsuitable for use as a signaling mechanism on wide area network digital transmission facilities. These problems include the need to sample the signal and the fact that it provides residual dc voltage buildup.

Since two or more repeated marks or spaces can stay at the same voltage or current level, sampling is required to distinguish one bit value from another. The ability to sample requires clocking circuitry which drives up the cost of communications. A second problem related to the fact that a sequence of marks or set bits can occur is that this condition results in the presence of residual dc levels. Residual dc requires the direct attachment of transmission components, while the absence of residual dc permits ac coupling via the use of a transformer. When communications carriers engineered their early digital networks they were based upon the use of copper conductors, as fiber optics did not exist. Communications carriers attempt to do things in an economical manner. Rather than install a separate line to power repeaters, they examined the possibility of carrying both power and data on a common line, removing the data from the power at the distant end as illustrated in Figure 1.11. This required transformer coupling at the distant end,



A signaling technique which does not produce residual dc enables the use of transformer coupling to separate the signal from power.

Figure 1.11 Transformer coupling

which is only possible if residual dc is eliminated. Thus, communications carriers began to search for an alternative signaling method.

Unipolar return to zero

One of the first alternative signaling methods examined was unipolar return to zero (RTZ). Under this signaling technique, which is illustrated in Figure 1.12, the current or voltage always returns to zero after every '1' bit. While this signal is easier to sample since each mark has a pulse rise, it still results in residual dc buildup and was unsuitable for use as the signaling mechanism on communications carrier digital transmission facilities.

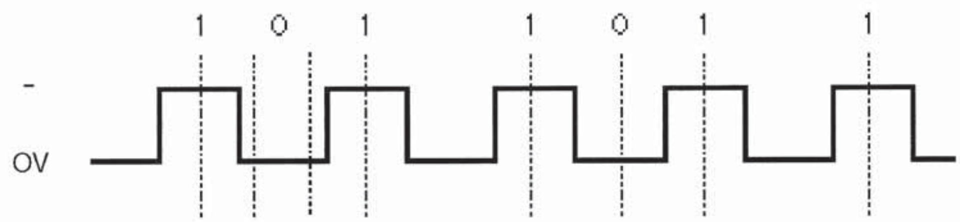


Figure 1.12 Unipolar return to zero signaling

Bipolar return to zero

After examining a variety of signaling methods communications carriers focused their attention upon a technique known as bipolar return to zero. Under bipolar signaling alternating polarity pulses are used to represent marks, while a zero level pulse is used to represent a space. In the bipolar return to zero signaling method the bipolar signal returns to zero after each mark. Figure 1.13 illustrates an example of bipolar return to zero signaling.

The key advantage of bipolar return to zero signaling is the fact that it precludes dc voltage buildup on the line. This enables both power and data to be carried on the same line, enabling repeaters to be powered by a common line. In addition, repeaters can be placed relatively far apart in comparison to other signaling techniques, which reduces the cost of developing the digital transmission infrastructure.

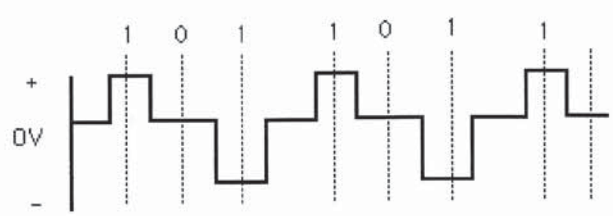


Figure 1.13 Bipolar return to zero signaling

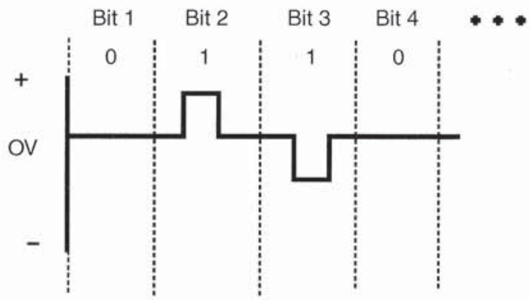


Figure 1.14 Bipolar (AMI) RTZ 50 percent duty cycle

On modern wide area network digital transmission facilities a modified form of bipolar return to zero signaling is employed. That modification involves the placement of pulses in their bit interval so that they occupy 50% of the interval, with the pulse centered at the center of the interval. This positioning eliminates high frequency components of the signal that can interfere with other transmissions and results in a bipolar pulse known as a 50% duty cycle alternate mark inversion (AMI). An example of this pulse is illustrated in Figure 1.14.

Now that we have a general appreciation for the type of digital signaling used on digital transmission facilities and the rationale for the use of that type of signaling, we will discuss some of the types of digital transmission facilities available for use. In doing so we will first consider the manner by which the digitization of voice conversations was performed, as the voice digitization effort had a significant effect upon the evolution of digital service offerings.

Evolution of service offerings

The evolution of digital service offerings can be traced to the manner by which telephone company equipment was developed to digitize voice and the initial development of digital multiplexing equipment to combine multiple voice conversations for transmission between telephone company offices. The digitization of voice was based upon the use of a technique referred to as Pulse Code Modulation (PCM) which requires an analog voice conversation to be sampled 8000 times per second. Each sample is encoded using an 8-bit byte, resulting in a digital data stream of 64 kbps to carry one digitized voice conversation.

The first high speed digital transmission circuit used in North America was designed to transport 24 digitized voice conversations. This circuit, which is referred to as a T1 line, transports a Digital Signal Level 1 (DS1) signal. That signal represents 24 digitized voice samples plus one framing bit repeated 8000 times per second. Thus, the operating rate of a T1 circuit becomes $(24 \times 8 + 1)$ bits/sample $\times$ 8000 samples/second, or 1.544 Mbps.

Each individual time slot within a DS1 signal is referred to as DS0 or Digital Signal Level 0 and represents a single digitized voice conversation transported at 64 kbps. Figure 1.15 illustrates the relationship between DS0s and a DS1. In examining Figure 1.15 note that the DS1 frame is repeated 8000 times per second,

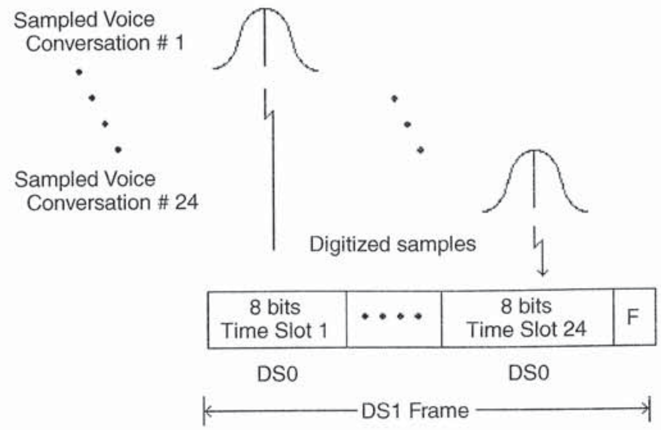


Figure 1.15 Relationship between DS0s and a DS1 signal (F = frame bit)

resulting in 8000 bps of framing information transmitted between telephone company central offices. Initially the framing bits were used for synchronization and the transmission of certain types of alarm indications. Later the sequence of framing bits was altered to enable control information to be transmitted as part of the framing. Later in this book when we investigate the operation and utilization of T1 multiplexers, we will examine the use of framing bits in detail.

Due to the necessity of transporting certain types of telephone information such as on-hook and off-hook information with each DS0, a technique referred to as bit-robbing was developed. Under bit-robbing one of the bits used to represent the height of a digitized sample was periodically 'stolen' for use to convey telephone set information. This bit-robbing process only occurred in the 6th and 12th frames in each continuous sequence of 12 frames, resulting in the inability of the human ear to recognize that a few digitized samples were encoded in seven bits instead of eight. However, when early digital transmission facilities were developed, the bit-robbing process limited data transmission to seven bits per eight-bit byte, which explains why switched 56 service operates at 56 kbps instead of 64 kbps. Later, the development of a separate network by telephone companies to convey signaling information enabled the full transmission capacity of DS0s to be used for data transmission. Today many communications carriers offer switched 56 and switched 64 kbps digital transmission as well as digital fractional T1 leased lines that operate in increments of 56 or 64 kbps.

AT&T offerings

In the United States, AT&T offers several digital transmission facilities under the ACCUNET^{SM} Digital Service mark. Dataphone$^{®}$ Digital Service was the charter member of the ACCUNET family and is deployed in over 100 major metropolitan cities in the United States as well as having an interconnection to Canada's digital network. Dataphone Digital Service operates at synchronous data transfer rates of 2.4, 4.8, 9.6, 19.2 and 56 kbps, providing users of this service with dedicated, two-way simultaneous transmission capability.

Originally all AT&T digital offerings were leased line services where a digital leased line is similar to a leased analog line in that it is dedicated for full-time use to one customer. In the late 1980s, AT&T introduced its Accunet Switched 56 service, a dial-up 56 kbps digital data transmission service. This service enables users to maintain a dial-up backup for previously installed 56 kbps AT&T Dataphone Digital Services leased lines or a partial backup for ACCUNET T1.5 service which is described later in this section. In addition, this service can be used to supplement existing services during peak transmission periods or for applications that only require a minimal amount of transmission time per day since the service is billed on a per minute basis.

Access to Switched 56 service is obtained by dialing area code 700 numbers available in approximately 100 cities in the United States. All numbers are 10-digit, of the form 700-56X-XXXX.

Transmission on both leased line and switched Dataphone Digital Service requires the use of a Data Service Unit and Channel Service Unit (DSU/CSU) in comparison to the use of modems when transmission occurs on an analog transmission facility. Originally separate devices, most vendors now market combined DSU/CSU products that are commonly and collectively referred to as DSUs. The operation of DSUs is described later in this section and in significantly more detail in Chapter 5.

Although DDS was a very popular digital transmission service during the 1980s, the expansion of communications carriers' digital infrastructure based upon the installation of tens of thousands of miles of fiber cable during the late 1980s and early 1990s resulted in a range of new digital offerings. Those offerings are based upon the use of portions of, or entire, T1 circuits, with the former referred to as fractional T1, and are considerably more cost-effective than DDS. Thus, although DDS was still in use during 1998 its future days of use are probably limited.

Another offering from AT&T, ACCUNET T1.5 Service is a high capacity 1.544 Mbps terrestrial digital service which permits 24 voice-grade channels or a mixture of voice and data to be transmitted in digital form. This service was originally only obtainable as a leased line and is more commonly known as a T1 channel or circuit. Transmission on a T1 circuit also requires the use of a DSU. However, the DSU portion of the DSU/CSU is commonly built into terminal equipment connected to T1 lines. Separate CSUs are required, therefore, to interface T1 circuits. Channel Service Units manufactured for use on T1 lines are described in detail in Chapter 5. Readers should note that field trials of switched 384 kbps and 1.544 Mbps services during the mid-1990s resulted in their availability for commercial usage.

Until 1989 there was a significant gap in the transmission rates obtainable on digital lines. DDS users could transmit at data rates up to 56 kbps, while the use of a T1 line resulted in the transmission of data at 1.544 Mbps. Recognizing the requirements of many organizations for the transmission of data at rates above that obtainable on DDS but below the T1 rate, several communications carriers introduced fractional T1 digital service. AT&T's fractional T1 service is called ACCUNET Spectrum of Digital Service (ASDS). Under ASDS digital transmission is furnished via leased lines at data rates ranging from 9.6, 56, or 64 kbps up to 1.544 Mbps, in increments of 64 kbps from 64 kbps to 1.544 Mbps. In actuality, 9.6 and 56 kbps ASDS services are not fractional T1 as they represent special digital services that allow DDS to be carried on a fraction of a T1 circuit at a considerable reduction in cost.

Data rates of 64, 128, 256, 384, 512 and 768 kbps available under ASDS can be considered as true fractional T1 as they represent distinct fractions of a T1 circuit. A 64 kbps data rate represents 1/24th of a T1 circuit since 64 kbps is the data rate of one digitized voice channel on a T1 circuit and that circuit carries 24 digitized voice channels.

The majority of access to a fractional T1 line requires the use of a T1 local loop. Although the data transmission rate on the local loop is 1.544 Mbps, the fractional T1 subscriber in actuality uses one or more 64 kbps channels on the local loop which is routed to the telephone company central office serving the subscriber. At that location the fractions of the T1 local loop used by the subscriber are removed from the T1 line and input into an interexchange carrier's equipment at the point of presence. The interexchange carrier combines the fractions of T1 circuits used by many subscribers into a full T1 circuit operating at 1.544 Mbps which is then routed through the carrier's transmission facilities to the point of presence serving the distant location. At that point of presence the 64 kbps channels representing the fractional T1 channel used by the subscriber are passed to another telephone company which, more than likely, routes the transmission via a T1 line to the subscriber.

In addition to T1 lines, AT&T and other communications carriers offer T3 digital circuits operating at 44.736 Mbps. A T3 circuit transports a DS3 signal which is formed via the multiplexing of 28 DS1 signals. Similar to the recognition that many organizations cannot use a full T1, AT&T and other communications carriers recognized that only a limited number of organizations can use the capacity of a full T3 digital circuit. As you might expect, this realization resulted in the development of fractional T3 (FT3) offerings.

European offerings

In Europe, a number of countries have established digital transmission facilities. One example of such offerings is British Telecom's KiloStream service. KiloStream provides synchronous data transmission at 2.4, 4.8, 9.6, 48 and 64 kbps and is very similar to AT&T's Dataphone Digital Service. Each KiloStream circuit is terminated by British Telecom with a network terminating unit (NTU), which is the digital equivalent of the modem required on an analog circuit. In comparison with the T1 circuit used in North America which was based upon a design for carrying 24 digitized voice conversations, European countries use E1 circuits. Such circuits were constructed based upon the placement of 32 channels on one circuit and operate at 2.048 Mbps. In the United Kingdom British Telecom's E1 service is marketed as MegaStream.

DSUs

A data service unit (DSU) provides a standard interface to a digital transmission service and handles such functions as signal translation, regeneration, reformatting, and timing. Most DSUs are designed to operate at one of four speeds—2.4, 4.8, 9.6, and 56 kbps—while some DSUs also support 19.2 kbps operations. The transmitting portion of the DSU processes the customer's signal into bipolar pulses suitable for transmission over the digital facility. The receiving portion of the DSU is used both to extract timing information and to regenerate mark and space information from the received bipolar signal. The second interface arrangement originally developed for AT&T's Dataphone Digital Service is called a channel service unit (CSU) and was provided by the communication carrier to those customers who wish to perform the signal processing to and from the bipolar line, as well as to retime and regenerate the incoming line signals through the utilization of their own equipment. Originally marketed as separate devices, almost all DSUs and CSUs designed for use on AT&T Dataphone Digital Service and equivalent facilities from other carriers are now manufactured as one integrated device which is commonly referred to as a DSU or a DSU/CSU. Since most terminal devices connected to T1 lines contain a built-in data service unit, a separate channel service unit is required for transmission on that type of digital transmission facility.

As data is transmitted over digital facilities, the signal is regenerated by the communications carrier numerous times prior to its arrival at its destination. In general, digital service gives data communications users improved performance and reliability when compared to analog service, owing to the nature of digital transmission and the design of digital networks. This improved performance and reliability is due to the fact that digital signals are regenerated whereas, when analog signals are amplified, any distortion to the analog signal is also being amplified.

Although a digital service is offered in many locations, for those locations outside the serving area of a digital facility the user will have to employ analog devices as an extension in order to interface to the digital facility. The utilization of digital service via an analog extension is illustrated in Figure 1.16. As depicted in Figure 1.16, if the closest city to the terminal located in city 2 that offers digital

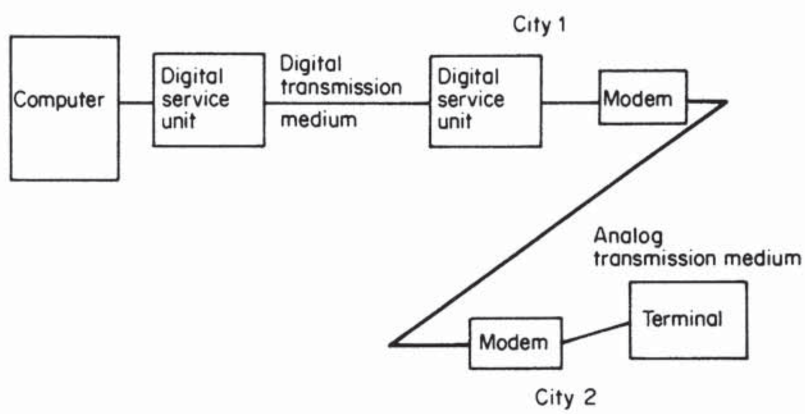


Figure 1.16 Analog extension to digital service. Although data is transmitted in digital form from the computer to city 1, it must be modulated by the modern at that location for transmission over the analog extension

service is city 1, then to use digital service to communicate with the computer an analog extension must be installed between the terminal location in city 2 and city 1. In such cases, the performance, reliability, and possible cost advantages of using digital service may be completely dissipated.

1.4 TRANSMISSION MODE

One method of characterizing lines, terminal devices, mainframe computers, and modems is by their transmission or communications mode. The three classes of transmission modes are simplex, half-duplex, and full-duplex.

Simplex transmission

Simplex transmission occurs in one direction only, disallowing the receiver of information a means of responding to the transmission. A home AM radio which receives a signal transmitted from a radio station is an example of a simplex communications mode. In a data transmission environment, simplex transmission might be used to turn on or off specific devices at a certain time of the day or when a certain event occurs. An example of this would be a computer-controlled environmental system where a furnace is turned on or off depending upon the thermostat setting and the current temperature in various parts of a building. Normally, simplex transmission is not utilized where human-machine interaction is required, owing to the inability to turn the transmitter around so that the receiver can reply to the originator.

Half-duplex transmission

Half-duplex transmission permits transmission in either direction; however, transmission can occur in only one direction at a time. Half-duplex transmission is used in citizen band (CB) radio transmission where the operator can either transmit or receive but cannot perform both functions at the same time on the same channel. When the operator has completed a transmission, the other party must be advised that they have finished transmitting and is ready to receive by saying the term 'over'. Then the other operator can begin transmission.

When data is transmitted over the telephone network, the transmitter and the receiver of the modem or acoustic coupler must be appropriately turned on and off as the direction of the transmission varies. Both simplex and half-duplex transmission require two wires to complete an electrical circuit. The top of Figure 1.17 illustrates a half-duplex modem interconnection while the lower portion of that illustration shows a typical sequence of events in the terminal's sign-on process to access a computer. In the sign-on process, the user first transmits the word NEWUSER to inform the computer that a new user wishes a connection to the computer. The computer responds by asking for the user's password, which is then furnished by the user.

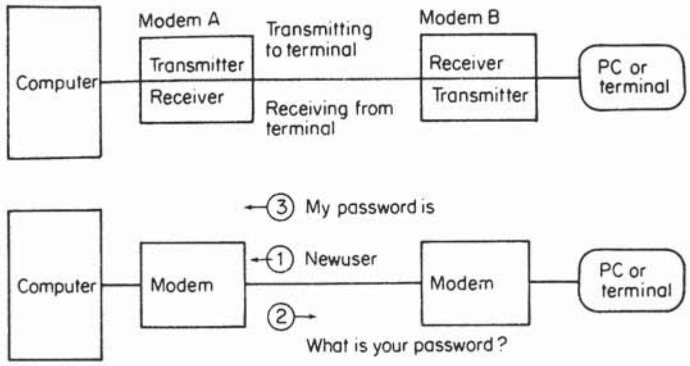


Figure 1.17 Half-duplex transmission. Top: control signals from the mainframe computer and terminal operate the transmitter and receiver sections of the attached modems. When the transmitter of modem A is operating, the receiver of modem B operates; when the transmitter of modem B operates; the reveiver of modem A operates. However, only one transmitter operates at any one time in the half-duplex mode of transmission. Bottom: during the sign-on sequence, transmission is turned around several times

In the top portion of Figure 1.17, when data is transmitted from a computer to a terminal, control signals are sent from the computer to modem A which turns on the modem A transmitter and causes the modem B receiver to respond. When data is transmitted from the terminal to the computer, the modem B receiver is disabled and its transmitter is turned on while the modem A transmitter is disabled and its receiver becomes active. The time necessary to effect these changes is called a transmission turnaround time, and during this interval transmission is temporarily halted. Half-duplex transmission can occur on either a two-wire or four-wire circuit. The switched network is a two-wire circuit, whereas leased lines can be obtained as either two-wire or four-wire links. A four-wire circuit is essentially a pair of two-wire links which can be used for transmission in both directions simultaneously. This type of transmission is called full-duplex.

Full-duplex transmission

Although you would normally expect full-duplex transmission to be accomplished over a four-wire connection that provides two two-wire paths, full-duplex transmission can also occur on a two-wire connection. This is accomplished by the use of modems that subdivide the frequency bandwidth of the two-wire connection into two distinct channels, permitting simultaneous data flow in both directions on a two-wire circuit. This technique will be examined and explained in more detail later in this book, when the operating characteristics of modems are examined in detail.

Full-duplex transmission is often used when large amounts of alternate traffic must be transmitted and received within a fixed time period. If two channels were used in our CB example, one for transmission and another for reception, two simultaneous transmissions could be effected. While full-duplex transmission

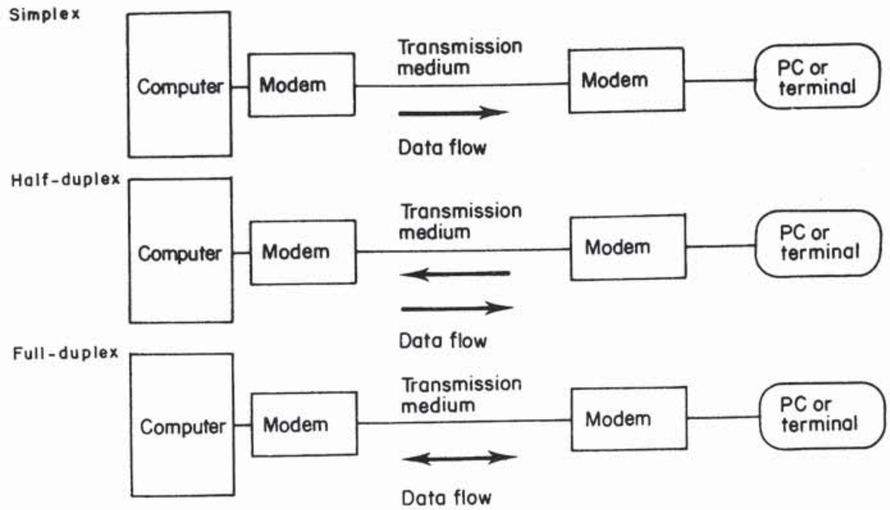


Figure 1.18 Transmission modes. Top: simplex transmission is in one direction only; transmission cannot reverse direction. Center: half-duplex transmission permits transmission in both directions but only one way at a time. Bottom: full-duplex transmission permits transmission in both directions simultaneously

provides more efficient throughput, this efficiency was originally negated by the cost of two-way lines and more complex equipment required by this mode of transmission. The development of low cost digital signal processor chips enabled high speed modems to operate in a full-duplex transmission mode on a two-wire circuit through a technique referred to as echo cancellation. This technique will be described when we discuss the operation of modems in detail in Chapter 5.

In Figure 1.18, the three types of transmission modes are illustrated, while Table 1.3 summarizes the three transmission modes previously discussed.

In Table 1.3, the column headed ITU refers to the International Telecommunications Union, a United Nations agency headquartered in Geneva, Switzerland. Previously, the standardization body of the ITU was known as the Consultative Committee on International Telephone and Telegraph (CCITT) and, even though the renaming occurred several years ago, many people still use the term CCITT when referencing certain standards.

Table 1.3 Transmission mode comparison

Symbol	ANSI	US telecommunications industry	ITU	Historical physical line requirement
←	One-way only	Simplex		Two-wire
← →	Two-way alternate	Half-duplex (HDX)	Simplex	Two-wire
⇄	Two-way simultaneous	Full-duplex (FDX)	Duplex	Four-wire

Until the 1980s most ITU standards were primarily followed in Europe. Since then, ITU modem standards have achieved a worldwide audience of followers, and enable true global communications compatibility. Although most modern modems are compatible with ITU standards, a large base of modems are still being used that were designed to Bell System standards that were popular in North America through the 1980s. In Chapter 5 we will examine some of the more common ITU modem standards as well as a few Bell System standards that will provide us with an understanding of some of the compatibility problems that can occur when communications are attempted between Bell System and ITU compatible modems.

Terminal and mainframe computer operations

When referring solely to terminal operations, the terms half-duplex and full-duplex operation take on meanings different from the communications mode of the transmission medium. Vendors commonly use the term half-duplex to denote that the terminal device is in a local copy mode of operation. This means that each time a character is pressed on the keyboard it is printed or displayed on the local terminal as well as transmitted. Thus, a terminal device operated in a half-duplex mode would have each character printed or displayed on its monitor as it is transmitted.

When one says a terminal is in a full-duplex mode of operation this means that each character pressed on the keyboard is transmitted but not immediately displayed or printed. Here the device on the distant end of the transmission path must 'echo' the character back to the originator, which, upon receipt, displays or prints the character. Thus, a terminal in a full-duplex mode of operation would only print or display the characters pressed on the keyboard after the character is echoed back by the device at the other end of the line. Figure 1.19 illustrates the terms full- and half-duplex as they apply to terminal devices. Note that although most conventional terminals have a switch to control the duplex setting of the device, personal computer users normally obtain their duplex setting via the

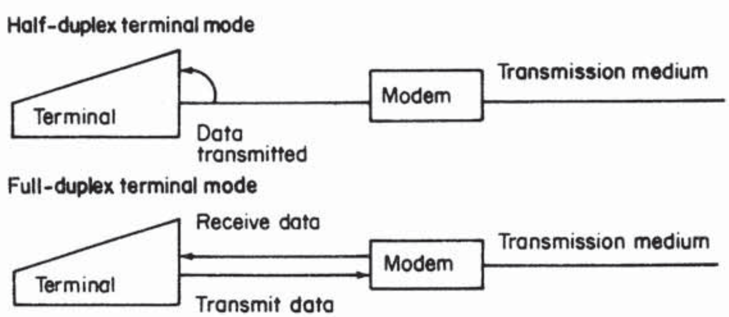


Figure 1.19 Terminal operation modes. Top: the term half-duplex terminal mode implies that data transmitted is also printed on the local terminal. This is known as local copy. Bottom: the term full-duplex terminal mode implies that no local copy is provided

software program they are using. Thus, the term 'echo on' during the initialization of a communications software program would refer to the process of displaying each character on the user's screen as it is transmitted.

When we refer to half- and full-duplex with respect to mainframe computer systems we are normally specifying whether or not they echo received characters back to the originator. A half-duplex computer system does not echo characters back, while a full-duplex computer system echoes each character it receives.

Different character displays

When considering the operating mode of the terminal device, the transmission medium, and the operating mode of the mainframe computer on the distant end of the transmission path as an entity, three things could occur in response to each character you press on a keyboard. Assuming a transmission medium is employed that can be used for either half- or full-duplex communications, your terminal device could print or display no character for each character transmitted, one character for each character transmitted, or two characters for each character transmitted. Here the resulting character printed or displayed would be dependent upon the operating mode of the terminal device and the host computer you are connected to as indicated in Table 1.4.

Table 1.4 Operating mode and character display

Operating mode		
Terminal device	Host computer	Character display
Half-duplex	Half-duplex	1 character
Half-duplex	Full-duplex	2 characters
Full-duplex	Half-duplex	No characters
Full-duplex	Full-duplex	1 character

To understand the character display column in Table 1.4, let us examine the two-character display result caused by the terminal device operating in a half-duplex mode while the host computer operates in a full-duplex mode.

When the terminal is in a half-duplex mode it echoes each transmitted character onto its printer or display. At the other end of the communications path, if the computer is in a full-duplex mode of operation it will echo the received character back to the terminal, causing a second copy of the transmitted character to be printed or displayed. Thus, two characters would appear on your printer or display for each character transmitted. To alleviate this situation, you would change the transmission mode of your terminal to full-duplex. This would normally be accomplished by turning 'echo' off during the initialization of a communications software program, if using a personal computer; or you would turn a switch to half-duplex if operating a conventional terminal.

1.5 TRANSMISSION TECHNIQUES

Data can be transmitted either synchronously or asynchronously. Asynchronous transmission is commonly referred to as a start–stop transmission where one character at a time is transmitted or received. Start and stop bits are used to separate characters and synchronize the receiver with the transmitter, thus providing a method of reducing the possibility that data becomes garbled.

Most devices designed for human–machine interaction that are teletype compatible transmit data asynchronously. By teletype compatible, we refer to terminals and personal computers that operate similarly to the Teletype® terminal manufactured by Western Electric, originally a subsidiary of AT&T and now known as Lucent Technology after this equipment manufacturing arm of AT&T was spun off during 1996. Various versions of this popular terminal have been manufactured for over 30 years and an installed base of approximately one million such terminals at one time during the late 1970s were in operation worldwide. As characters are depressed on the device's keyboard they are transmitted, with idle time occurring between the transmission of characters. This is illustrated in Figure 1.20(b). Although many teletype terminals have been replaced by personal computers, asynchronous transmission pioneered by those terminals remains very popular in use.

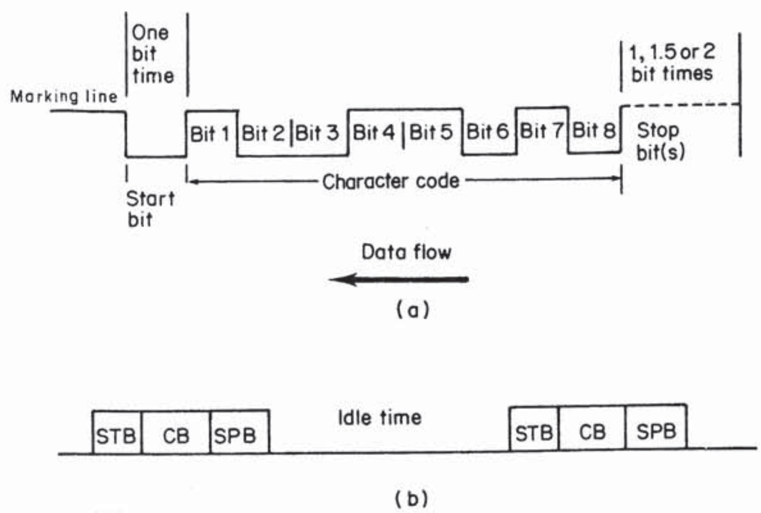


Figure 1.20 Asynchronous (start-stop) transmission. (a) Transmission of one 8-bit character. (b) Transmission of many characters. STB = start bit; CB = character bits; SPB = stop bit(s); idle time is time between character transmission

Asynchronous transmission

In asynchronous transmission, each character to be transmitted is encoded into a series of pulses. The transmission of the character is started by a start pulse equal in length to a code pulse. The encoded character (series of pulses) is followed by a stop pulse that may be equal to or longer than the code pulse, depending upon the transmission code used.

The start bit represents a transition from a mark to a space. Since in an idle condition when no data are transmitted the line is held in a marking condition, the start bit serves as an indicator to the receiving device that a character of data follows. Similarly, the stop bit causes the line to be placed back into its previous 'marking' condition, signifying to the receiver that the data character is completed.

As illustrated in Figure 1.20(a), the transmission of an 8-bit character requires either 10 or 11 bits, depending upon the length of the stop bit. In actuality the eighth bit may be used as a parity bit for error detection and correction purposes. The use of the parity bit is described in detail in Section 1.12.

In the start–stop mode of transmission, transmission starts anew on each character and stops after each character. This is indicated in Figure 1.20(b). Since synchronization starts anew with each character, any timing discrepancy is cleared at the end of each character, and synchronization is maintained on a character-by-character basis. Asynchronous transmission normally is used for transmission at speeds at or under 33 600 bps over the switched telephone network or on leased lines, while data rates up to 115 200 bps are possible over a direct connect cable whose distance is normally limited to approximately 50 feet.

The term asynchronous TTY, or TTY compatible, refers to the asynchronous start–stop protocol employed originally by Teletype® terminals and is the protocol where data are transmitted on a line-by-line basis between a terminal device and a mainframe computer. In comparison, more modern terminals with cathode ray tube (CRT) displays are usually designed to transfer data on a full screen basis.

Personal computer users only require an asynchronous communications adapter and a software program that transmits and receives data on a line-by-line basis to connect to a mainframe that supports asynchronous TTY compatible terminals. Here the software program that transmits and receives data on a line-by-line basis is normally referred to as a TTY emulator program and is the most common type of communications program written for use with personal computers.

When a personal computer is used to transmit and receive data on a full screen basis a specific terminal emulator program is required. Most terminal emulator programs emulate asynchronous terminals; however, some programs emulate synchronous terminals. Concerning the latter, such programs usually require the installation of a synchronous communications adapter in the personal computer and the use of a synchronous modem. Since a personal computer includes a video display onto which characters and graphics can be positioned, the PC can be used to emulate a full-screen addressable terminal. Thus, with appropriate software or a combination of hardware and software you can use a personal computer as a replacement for proprietary terminals manufactured to operate with a specific type of mainframe computer, as well as to perform such local processing as spreadsheet analysis and word processing functions. In fact, during the early 1990s a large majority of conventional terminal devices in business use were replaced by PC-operating terminal emulation software.

Synchronous transmission

A second type of transmission involves sending a grouping of characters in a continuous bit stream. This type of transmission is referred to as synchronous or bit-stream synchronization. In the synchronous mode of transmission, modems located at each end of the transmission medium normally provide a timing signal or clock that is used to establish the data transmission rate and enable the devices attached to the modems to identify the appropriate characters as they are being transmitted or received. In some instances, timing may be provided by the terminal device itself or a communication component, such as a multiplexer or front-end processor channel. No matter what timing source is used, prior to beginning the transmission of data the transmitting and receiving devices must establish synchronization among themselves. In order to keep the receiving clock in step with the transmitting clock for the duration of a stream of bits that may represent a large number of consecutive characters, the transmission of the data is preceded by the transmission of one or more special characters. These special synchronization or 'syn' characters are at the same code level (number of bits per character) as the coded information to be transmitted. They have, however, a unique configuration of zero and one bits which are interpreted as the syn character. Once a group of syn characters is transmitted, the receiver recognizes and synchronizes itself onto a stream of those syn characters.

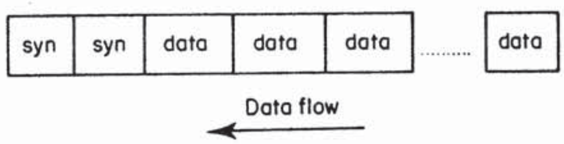


Figure 1.21 Synchronous transmission. In synchronous transmission, one or more syn characters are transmitted to establish clocking prior to the transmission of data

After synchronization is achieved, the actual data transmission can proceed. Synchronous transmission is illustrated in Figure 1.21. In synchronous transmission, characters are grouped or blocked into groups of characters, requiring a buffer or memory area so characters can be grouped together. In addition to having a buffer area, more complex circuitry is required for synchronous transmission since the receiving device must remain in phase with the transmitter for the duration of the transmitted block of information. Synchronous transmission is normally used for data transmission rates in excess of 2000 bps. The major characteristics of asynchronous and synchronous transmission are denoted in Table 1.5.

In examining the entries in Table 1.5 a word of explanation is in order concerning the fifth entry for asynchronous transmission. The ability to transmit at 56 000 bps is based upon a relatively new modem technology that allows one analog to digital conversion to occur. This means that one end of a dialed circuit path must directly connect to digital equipment. Normally this can only be accomplished in one direction when a call is originated over the switched telephone network.

Table 1.5 Transmission technique characteristics

Asynchronous

1. Each character is prefixed by a start bit and followed by one or more stop bits.
2. Idle time (period of inactivity) can exist between transmitted characters.
3. Bits within a character are transmitted at prescribed time intervals.
4. Timing is established independtly in the computer and terminal.
5. Transmission speeds normally do not exceed 33 600 bps bidirectional (or 56 000 in one direction and 33 600 in the opposite direction) over switched facilities or leased lines and 56 000 bps over analog dedicated links and leased lines.

Synchronous

1. Syn characters prefix transmitted data.
2. Syn characters are transmitted between blocks of data to maintain line synchronization.
3. No gaps exist between characters.
4. Timing is established and maintained by the transmitting and receving modems, the terminal, or other devices.
5. Terminals must have buffers.
6. Transmission speeds normally are in excess of 2000 bps.

However, when dedicated or leased lines are used it becomes theoretically possible to obtain a 56 kbps transmission capability in both directions. Although modem equipment marketed during 1998 only enabled 56 kbps transmission in one direction via the switched network, it is quite possible that by the time you read this book a bidirectional 56 kbps transmission capability will be available for dedicated and leased analog lines.

1.6 TYPES OF TRANSMISSION

The two types of data transmission one can consider are serial and parallel. For serial transmission the bits which comprise a character are transmitted in sequence over one line, whereas in parallel transmission characters are transmitted serially but the bits that represent the character are transmitted in parallel. If a character consists of eight bits, then parallel transmission would require a minimum of eight lines. Additional lines may be necessary for control signals and for the transmission of a parity bit. Although parallel transmission is used extensively in computer-to-peripheral unit transmission, it is not normally employed other than in dedicated data transmission usage over relatively short distances owing to the cost of the extra circuits required.

A typical use of parallel transmission is the in-plant connection of badge readers and similar devices to a computer in that facility. Parallel transmission can reduce the cost of terminal circuitry since the terminal does not have to convert the internal character representation to a serial data stream for transmission. The cost of the transmission medium and interface will, however, increase because of the additional number of conductors required. Since the total character can be transmitted at the same moment in time using parallel transmission, higher data transfer rates can be obtained than are possible with serial transmission facilities.

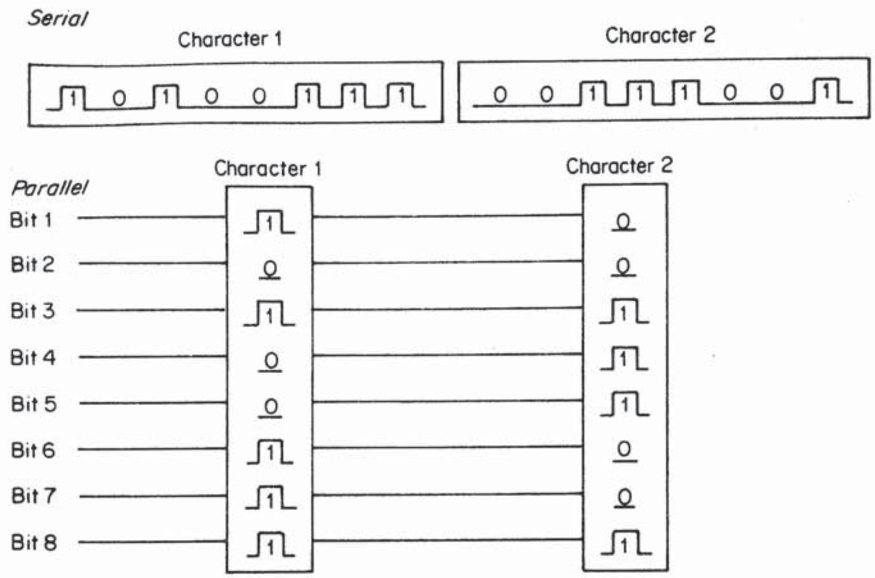


Figure 1.22 Types of data transmission. In serial transmission, the bits that comprise the character to be transmitted are sent in sequence over one line. In parallel transmission, the characters are transmitted serially but the bits that represent the character are transmitted in parallel

For this reason, most local facility communications between computers and their peripheral devices are accomplished using parallel transmission. In comparison, communications between terminal devices and computers normally occur serially, since this requires only one line to interconnect the two devices that need to communicate with one another. Figure 1.22 illustrates serial and parallel transmission.

1.7 LINE STRUCTURE

The geographical distribution of terminal devices and the distance between each device and the device it transmits to are important parameters that must be considered in developing a wide area network configuration. The method used to interconnect personal computers and terminals to mainframe computers or to other devices is known as line structure and results in a computer's network configuration.

Types of line structure

The two types of line structure used in networks are point-to-point and multipoint, the latter also commonly referred to as multidrop lines.

Point-to-point

Communications lines that only connect two points are point-to-point lines. An example of this line structure is depicted in Figure 1.23(a). As illustrated, each terminal transmits and receives data to and from a computer via an individual connection that links a specific terminal to the computer. The point-to-point connection can utilize a dedicated circuit or a leased line, or can be obtained via a connection initiated over the switched (dial-up) telephone network.

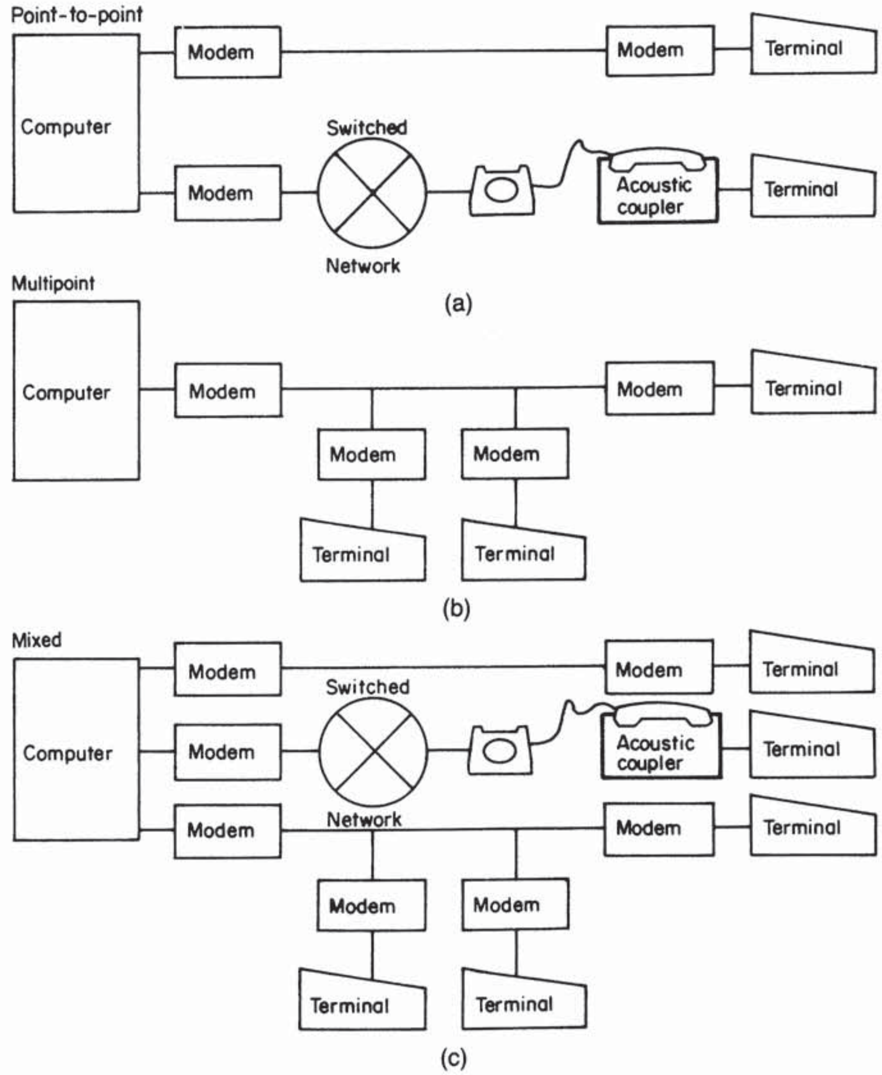


Figure 1.23 Line structures in networks. Top: point-to-point line structure. Center: multipoint (multidrop) line structure. Bottom: mixed network line structure

Multipoint

When two or more terminal locations share portions of a common line, the line is a multipoint or multidrop line. Although no two devices on such a line can transmit data at the same time, two or more devices may receive a message at the same time. The number of devices receiving such a message is dependent upon the addresses assigned to the message recipients. In some systems a 'broadcast' address permits all devices connected to the same multidrop line to receive a message at the same time. When multidrop lines are employed, overall line costs may be reduced since common portions of the line are shared for use by all devices connected to that line. To prevent data transmitted from one device from interfering with data transmitted from another device, a line discipline or control must be established for such a link. This discipline controls transmission so no two devices transmit data at the same time. A multidrop line structure is depicted in Figure 1.23(b).

Although modems are shown on point-to-point and multipoint line structures illustrated in Figure 1.23, both line structures are also applicable to digital transmission services. Dataphone Digital Service facilities support both point-to-point and multipoint line structures, with the use of modems and acoustic couplers shown in Figure 1.23 replaced by the use of DSUs. T1 and fractional T1 lines, however, are only available as a point-to-point line structure and require the use of DSUs/CSUs in place of the modems shown at the top of Figure 1.23.

Depending upon the type of transmission facilities used, it may be possible to intermix both point-to-point and multipoint lines in developing a network. For example, analog facilities and the use of DDS support both line structures which enable the development of a mixed network line structure as shown in Figure 1.23(c).

1.8 LINE DISCIPLINE

When several devices share the use of a common, multipoint communications line, only one device may transmit at any one time, although one or more devices may receive information simultaneously. To prevent two or more devices from transmitting at the same time, a technique known as 'poll and select' is utilized as the method of line discipline for multidrop lines. To utilize poll and select, each device on the line must have a unique address of one or more characters as well as circuitry to recognize a message sent from the computer to that address. When the computer polls a line, in effect it asks each device in a predefined sequence if it has data to transmit. If the device has no data to transmit, it informs the computer of this fact and the computer continues its polling sequence until it encounters a device on the line that has data to send. Then the computer acts on that data transfer.

As the computer polls each device, the other devices in the line must wait until they are polled before they can be serviced. Conversely, transmission of data from the computer to each device on a multidrop line is accomplished by the computer selecting the device address to which those data are to be transferred, informing the device that data are to be transferred to it, and then transmitting data to the

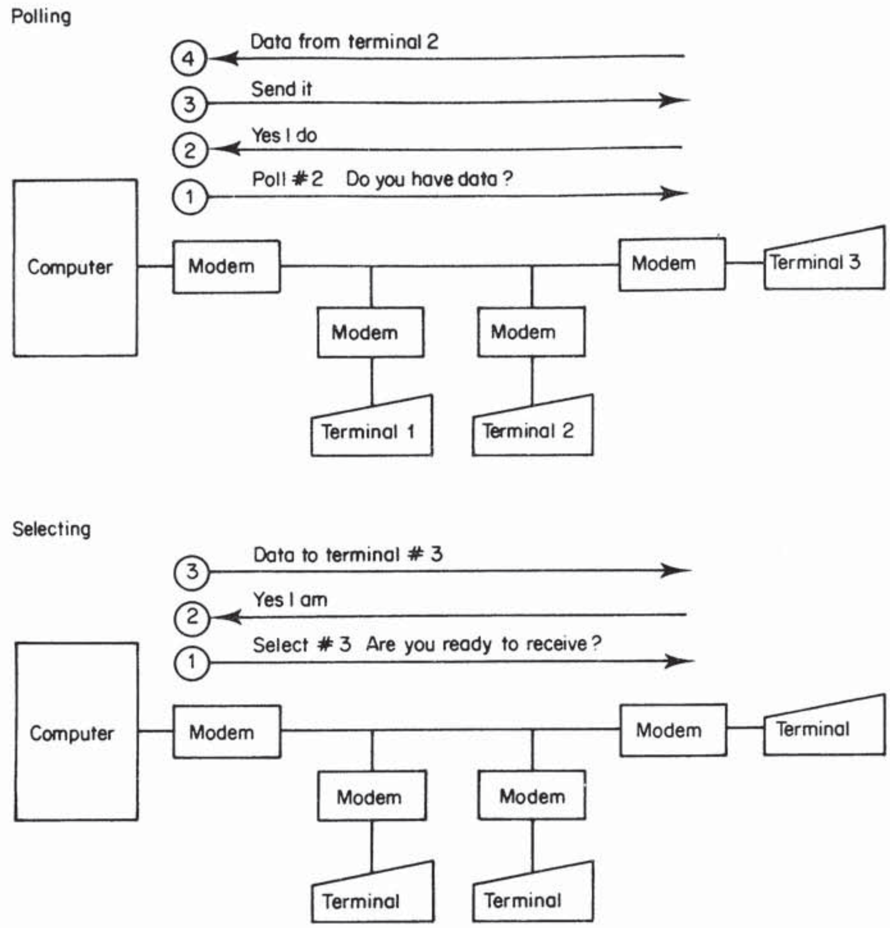


Figure 1.24 Poll and select line discipline. Poll and select is a line discipline which permits several devices to use a common line facility in an orderly manner

selected device. Polling and selecting can be used to service both asynchronous or synchronous operating terminal devices connected to independent multidrop lines. Owing to the control overhead of polling and selecting, synchronous devices are normally serviced in this type of environment. By the use of signals and procedures, polling and selecting line control ensures the orderly and efficient utilization of multidrop lines. An example of a computer polling the second terminal on a multipoint line and then receiving data from that device is shown at the top of Figure 1.24. At the bottom of that illustration, the computer first selects the third terminal on the line and then transfers a block of data to that device.

When terminals transmit data on a point-to-point line to a computer or another terminal, the transmission of that data occurs at the discretion of the terminal operator. This method of line control is known as 'non-poll-and-select' or 'free-wheeling' transmission.

1.9 NETWORK TOPOLOGY

The topology or structure of a wide area network is based upon the interconnection of point-to-point and/or multipoint lines used to develop the network. Although an infinite number of network topologies can be developed by connecting lines to one another, there are several basic types of topologies that are commonly used to construct wide area networks or parts of those networks. These topologies include the point-to-point, 'v', star, and mesh.

Figure 1.25(a) illustrates the four basic WAN topology structures, with dots used to reference end points or network nodes. These structures can be interconnected to one another to link organizational locations that represent different types of clusters of locations. Figure 1.25(b) illustrates the creation of two examples of different network topologies based upon connecting two 'v' structures and joining a 'v' to a star structure.

The design of a wide area network topology is commonly based upon economics and reliability. From an economic perspective the WAN should use one or more network structures that minimize the distance of circuits used to connect geographically separated locations, since the monthly cost of a circuit is in general proportional to its distance. From a reliability perspective, the ability to have multiple paths between nodes enables an alternative path to be used in the event of a transmission impairment making the primary path inoperative. Since redundancy requires additional communications paths, a trade-off occurs between the need to obtain an economically efficient network and a highly reliable network. For example, the use of a mesh structure permits multiple paths between nodes but requires more circuits than a star topology.

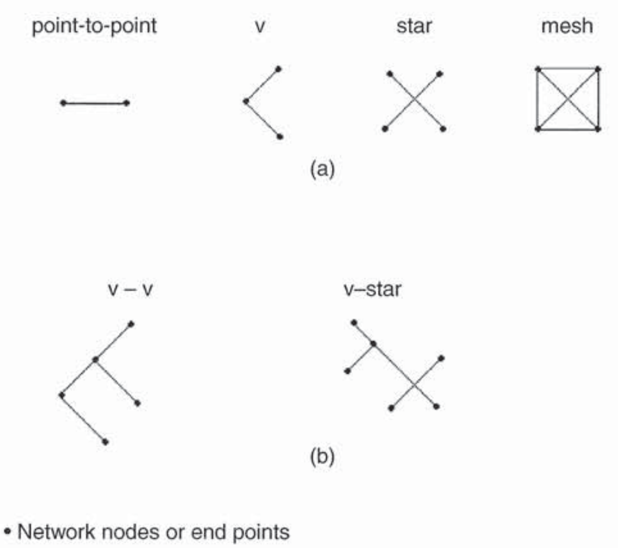


Figure 1.25 Wide Area Network Topologies: (a) basic topology structures; (b) connecting different structures

1.10 TRANSMISSION RATE

Many factors can affect the transmission rate at which data is transferred on a wide area network. The types of modems and acoustic couplers used on analog facilities or the types of DSUs or CSUs used on digital facilities, as well as the line discipline and the type of computer channel to which a terminal is connected via a transmission medium, play governing roles that affect transmission rates. However, the transmission medium itself is a most important factor in determining transmission rates.

Data transmission services offered by communications carriers such as AT&T, MCI, Sprint, and British Telecom are based on their available plant facilities. Depending upon terminal and computer locations, two types of transmission services may be available. The first type of service, analog transmission, is readily available in the form of switched and leased telephone lines. Digital transmission is available in most large cities; however, the type of digital service offered can vary from city to city as well as between different locations within a city. For example, AT&T's DDS service, while available in over 100 cities in North America, is only supported by certain telephone company offices in each city. Thus, an analog extension is required to connect to this service from non-digital service locations as previously illustrated in Figure 1.16. Within each type of service several grades of transmission are available for consideration.

Analog service

In general, analog service offers the user three grades of transmission: narrowband, voice-band and wideband. The data transmission rates capable on each of these grades of service depend upon the bandwidth and electrical properties of each type of circuit offered within each grade of service. Basically, transmission speed is a function of the bandwidth of the communications line: the greater the bandwidth, the higher the possible speed of transmission.

Narrowband facilities are obtained by the carrier subdividing a voice-band circuit or by grouping a number of transmissions from different users onto a single portion of a circuit by time. Transmission rates obtained on narrowband facilities range between 45 and 300 bps. Teletype® terminals that connect to message switching systems are the primary example of the use of narrowband facilities.

The primary use of analog narrowband transmission facilities is by information distribution organizations, such as news services and weather bureaux. Although the utilization of analog narrowband transmission facilities was popular through the 1970s, since that time many organizations that used that medium have migrated to transmission facilities that permit higher data transmission rates.

While narrowband facilities have a bandwidth in the range of 200 to 400 Hz, voice-band facilities have a bandwidth in the range of 3000 Hz. Data transmission speeds obtainable on voice-band facilities are differentiated by the type of voice-band facility utilized—switched dial-up transmission or transmission via a leased line. For transmission over the switched telephone network data rates up to 56 kbps may be obtainable in one direction and up to 33.6 kbps in the opposite direction when communicating with an information utility or corporate remote access device that is directly connected to the switched network, resulting in only one analog–digital conversion in the downstream direction. Although the fact that leased lines can be conditioned enabled higher speed transmission to occur on analog leased lines than the switched telephone network through the 1980s, since then most modems have been designed to operate on both transmission facilities at equal operating rates. When operating on the switched network a high speed modem must use echo cancellation to obtain a full-duplex transmission capability on a two-wire circuit. In comparison, when used on a four-wire leased line a modem can use each two-wire pair for transmission in different directions.

Although low data speeds can be transmitted on both narrowband and voice-band circuits, one should not confuse the two, since a low data speed on a voice circuit is transmission at a rate far less than the maximum permitted by that type of circuit, whereas a low rate on a narrowband facility is at or near the maximum transmission rate permitted by that type of circuit.

Analog facilities which have a higher bandwidth than voice-band are termed wideband or group-band facilities since they provide a wider bandwidth through the grouping of a number of voice-band circuits. Wideband facilities are available only on leased lines and permit transmission rates at or in excess of 40 800 bps. Transmission rates on wideband facilities vary with the offerings of communications carriers. Speeds normally available include 40.8, 50 and 230.4 kbps. The growth in the availability of high speed digital transmission facilities during the late 1980s resulted in a corresponding decrease in the use of high speed analog transmission facilities. By the late 1990s, most wideband analog transmission facilities were replaced by the use of high speed digital transmission facilities.

For direct connect circuits, transmission rates are a function of the distance between the terminal and the computer as well as the gauge of the conductor used.

Digital service

In the area of digital service, several offerings are currently available for user consideration, that can be categorized by the data transmission rates they support—low, medium, and high data rates.

Low speed digital transmission services include AT&T's Dataphone Digital Service (DDS) and equivalent offerings from other communications carriers. DDS provides full-duplex, point-to-point, and multipoint synchronous transmission via leased lines at speeds of 2.4, 4.8, 9.6, 19.2 and 56 kbps as well as switched 56 kbps service.

Medium speed digital transmission service is represented by AT&T's Accunet Spectrum of Digital Services (ASDS) and equivalent offerings from other communication carriers. As discussed earlier in this chapter, ASDS provides a fractional T1 transmission capability from 64 kbps to 1.544 Mbps in increments of 64 kbps. Common data transmission rates supported by ASDS and equivalent offerings from other carriers include 64, 128, 256, 512 and 768 kbps.

High speed digital transmission service is represented by T1 facilities operating at 1.544 Mbps through T3 facilities that operate at 44.736 Mbps. A relatively new

Table 1.6 Common wide area network transmission facilities

Facility	Transmission speed	Typical use
Analog		
Narrowband	45–300 pbs	Message switching
Voice-band		
Switched	Up to 56 000 bps	Time sharing; remote job entry; information utility access; file transfer
Leased	Up to 56 000 bps	Computer-to-computer; remote job entry; tape-to-tape transmission
Wideband	At or over 40 800 bps	High-speed terminal to high-speed terminal
Digital		
Switched	56 kbps, 384 kbps, 1.544 Mbps	LAN internetworking, Videoconferencing, computer-to-computer
Leased line	2.4, 4.8, 9.6, 19.2 kbps	Remote job entry; computer-to-computer
	56 kbps, $n \times 56/64$ kbps	High speed fax; LAN internetworking
	1.544 Mbps	Integrated voice and data
	44 Mbps	Integrated voice and data

digital offering known as a synchronous optical network (SONET) in North America, and as the Synchronous Digital Hierarchy (SDH) in Europe, provides a mechanism for interconnecting high speed networks via optical media. SONET speeds start at 51.84 Mbps and increase in multiples of that data rate to 2.488 Gbps.

Table 1.6 lists the main analog and digital facilities, the range of transmission speeds over those facilities, and the general use of such facilities. The entry $n \times 56/64$ represents fractional service, with n varying from 1 to 12 on some carrier offerings while other carriers support n varying up to 24, the latter providing a full T1 service. When we discuss the T1 carrier later in this book we will also discuss in additional detail the reason why the fractional T1 operating rate increments at either 56 or 64 kbps intervals.

1.11 TRANSMISSION CODES

Data within a computer or terminal device is structured according to the architecture of the device. The internal representation of data in either device is seldom suitable for transmission other than to peripheral units attached to terminals or computers. In most cases, to effect data transmission, internal formatted data must be redesigned or translated into a suitable transmission code. This transmission code creates a correspondence between the bit encoding of data for transmission or internal device representation and printed symbols. The end result of the translation is usually dictated by the character code that the devices communicating with one another mutually support. Frequently available codes include Baudot, which is a 5-level (5 bits per character) code; binary-coded decimal (BCD), which is a 6-level code; American Standard Code for Information Interchange (ASCII), which is normally a 7-level code; and the extended binary-coded decimal interchange code (EBCDIC), which is an 8-level code.

In addition to information being encoded into a certain number of bits based upon the transmission code used, the unique configuration of those bits to represent certain control characters can be considered as a code that can be used to effect line discipline. These control characters may be used to indicate the acknowledgement of the receipt of a block of data without errors (ACK), the start of a message (SOH), or the end of a message (ETX), with the number of permissible control characters standardized according to the code employed. With the growth of computer-to-computer data transmission, a large amount of processing can be avoided by transferring data in the format used by the computer for internal processing. Such transmission is known as binary mode transmission, transparent data transfer, code-independent transmission, or native mode transmission.

Morse code

One of the most commonly known codes, the Morse code, is not practical for utilization in a computer communications environment. This code consists of a series of dots and dashes, which, while easy for the human ear to decode, are of unequal length and not practical for data transmission implementation. In addition, since each character in the Morse code is not prefixed with a start bit and terminated with a stop bit, it was initially not possible to construct a machine to automatically translate received Morse transmissions into their appropriate characters.

Baudot code

The Baudot code, which is a 5-level (5 bits per character) code, was the first code to provide a mechanism for encoding characters by an equal number of bits, in this case five. The 5-level Baudot code was devised by Emil Baudot to permit teletypewriters to operator faster and more accurately than relays used to transmit information via telegraph.

Since the number of different characters which can be derived from a code having two different (binary) states is 2^m, where m is the number of positions in the code, the 5-level Baudot code permits 32 unique character bit combinations. Although 32 characters could be represented normally with such a code, the necessity of transmitting digits, letters of the alphabet, and punctuation marks made it necessary to devise a mechanism to extend the capacity of the code to include additional character representations. The extension mechanism was accomplished by the use of two 'shift' characters: 'letters shift' and 'figures shift'. The transmission of a shift character informs the receiver that the characters which follow the shift character should be interpreted as characters from a symbol and numeric set or from the alphabetic set of characters.

Table 1.7 Five-level Baudot code

		Bit Selection				
Letters	Figures	1	2	3	4	5
Characters						
A	-	1	1			
B	?	1			1	1
C	:		1	1	1	
D	$	1			1	
E	3	1				
F	!	1		1	1	
G	&		1		1	1
H				1		
I	8		1	1		
J	'	1	1		1	
K	(	1	1	1	1	
L	)		1			1
M	.			1	1	1
N	,			1	1	
O	9					1
P	Ø		1	1		1
Q	1	1	1	1		1
R	4		1		1	
S		1		1		
T	5				1	
U	7	1	1	1		
V	;		1	1	1	1
W	2	1	1			1
X	/	1		1	1	1
Y	6	1		1		1
Z	"	1				1
Functions						
Carriage return					1	
Line feed			1			
Space				1		
Lettershift	<	1	1	1	1	1
Figures shift	=	1	1		1	1

The 5-level Baudot code is illustrated in Table 1.7 for one particular terminal pallet arrangement. A transmission of all ones in bit positions 1 to 5 indicates a letter shift, and the characters following the transmission of that character are interpreted as letters. Similarly, the transmission of ones in bit positions 1, 2, 4 and 5 would indicate a figures shift, and the following characters would be interpreted as numerals or symbols based upon their code structure. Although the Baudot code is quite old in comparison to the age of personal computers, it is the transmission code used by the Telex network which at one time was extensively employed in the business community to send messages throughout the world.

BCD code

The development of computer systems required the implementation of coding systems to convert alphanumeric characters into binary notation and the binary notation of computers into alphanumeric characters. The BCD system was one of the earliest codes used to convert data to a computer-acceptable form. This coding technique permits decimal numeric information to be represented by four binary bits and permits an alphanumeric character set to be represented through the use of six bits of information. This code is illustrated in Table 1.8. An advantage of this code is that two decimal digits can be stored in an 8-bit computer word and

Table 1.8 Binary-coded decimal (BCD) system

Bit position						
b_6	b_5	b_4	b_3	b_2	b_1	Character
0	0	0	0	0	1	A
0	0	0	0	1	0	B
0	0	0	0	1	1	C
0	0	0	1	0	0	D
0	0	0	1	0	1	E
0	0	0	1	1	0	F
0	0	0	1	1	1	G
0	0	1	0	0	0	H
0	0	1	0	0	1	I
0	1	0	0	0	1	J
0	1	0	0	1	0	K
0	1	0	0	1	1	L
0	1	0	1	0	0	M
0	1	0	1	0	1	N
0	1	0	1	1	0	O
0	1	0	1	1	1	P
0	1	1	0	0	0	Q
1	1	1	0	0	1	R
1	0	0	0	1	0	S
1	0	0	0	1	1	T
1	0	0	1	0	0	U
1	0	0	1	0	1	V
1	0	0	1	1	0	W
1	0	0	1	1	1	X
1	0	1	0	0	0	Y
1	0	1	0	0	1	Z
1	1	0	0	0	0	0
1	1	0	0	0	1	1
1	1	0	0	1	0	2
1	1	0	0	1	1	3
1	1	0	1	0	0	4
1	1	0	1	0	1	5
1	1	0	1	1	0	6
1	1	0	1	1	1	7
1	1	1	0	0	0	8
1	1	1	0	0	1	9

manipulated with appropriate computer instructions. Although only 36 characters are shown for illustrative purposes, a BCD code is capable of containing a set of 2^6 or 64 different characters.

EBCDIC code

In addition to transmitting letters, numerals, and punctuation marks, a considerable number of control characters may be required to promote line discipline. These control characters may be used to switch on and off devices which are connected to the communications line, control the actual transmission of data, manipulate message formats, and perform additional functions. Thus, an extended character set is usually required for data communications. One such character set is EBCDIC code. The extended binary coded decimal interchange code (EBCDIC) is an extension of the BCD system and uses 8 bits for character representation. This code permits 2^8 or 256 unique characters to be represented, although currently a lesser number is assigned meanings. This code is primarily used for transmission by byte-oriented computers, where a byte is normally a grouping of eight consecutive binary digits operated on as a unit by the computer. The use of this code by computers may alleviate the necessity of the computer performing code conversion if the connected terminals operate with the same character set.

One of the more interesting examples of computer terminology is the use of the terms bytes and octets. In the early days of computer development machines were constructed using different groupings of bits that were operated upon as an entity and referred to as a byte. At various times during the 1960s groupings of 4, 6, 8, 12 and 16 bits operated upon as an entity were referred to as bytes. Owing to this ambiguity, the term octet became commonly used in communications to refer to an 8-bit byte. However, since virtually all modern computers use 8-bit bytes, this author will use the term byte to refer to a grouping of 8 bits.

Several subsets of EBCDIC exist that have been tailored for use with certain devices. As an example, IBM 3270 type terminal products would not use a paper feed and its character representation is omitted in the EBCDIC character subset used to operate that type of device, as indicated in Table 1.9.

ASCII code

As a result of the proliferation of data transmission codes, several attempts to develop standardized codes for data transmission were made. One such code is the American Standard Code for Information Interchange (ASCII). This 7-level code is based upon a 7-bit code developed by the International Standards Organization (ISO) and permits 128 possible combinations or character assignments, to include 96 graphic characters that are printable or displayable and 32 control characters to include device control and information transfer control characters. Table 1.10 lists the ASCII character set while Table 1.11 lists the ASCII control characters by position and their meaning. A more detailed explanation of these control characters is contained in the section covering protocols in this chapter.

Table 1.9 EBCIDIC code implemented for the IBM 3270 information display system

Bits 4567	Hex1	00				01				10				11				← Bits 0, 1
		00	01	10	11	00	01	10	11	00	01	10	11	00	01	10	11	← 2, 3
		0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	← Hex 0
0000	0	NUL	DLE			SP	&	-									0	
0001	1	SOH	SBA					/		a	j			A	J		1	
0010	2	STX	EUA		SYN					b	k	s		B	K	S	2	
0011	3	ETX	IC							c	l	t		C	L	T	3	
0100	4									d	m	u		D	M	U	4	
0101	5	PT	NL							e	n	v		E	N	V	5	
0110	6			ETB						f	o	w		F	O	W	6	
0111	7			ESC	EOT					g	p	x		G	P	X	7	
1000	8									h	q	y		H	Q	Y	8	
1001	9		EM							i	r	z		I	R	Z	9	
1010	A					c	!	1p	:									
1011	B					.	$	,	#									
1100	C		DUP		RA	<	*	%	@									
1101	D		SF	ENQ	NAK	(	)	—	'									
1110	E		FM			+	;	>	=									
1111	F		ITB		SUB	|	—ar	?	"									

The primary difference between the ASCII character set listed in Table 1.10 and other versions of the ITU International Alphabet Number 5 is the currency symbol. Although the bit sequence 0 1 0 0 0 1 0 is used to generate the dollar ($) currency symbol in the United States, in the United Kingdom that bit sequence results in the generation of the pound sign (£). Similarly, this bit sequence generates other currency symbols when the ITU International Alphabet Number 5 is used in other countries.

Extended ASCII

Members of the IBM PC series and compatible computers use an extended ASCII character set which is represented as an 8-level code. The first 128 characters in the character set, ASCII values 0 through 127, correspond to the ASCII character set listed in Table 1.10 while the next 128 characters can be viewed as an extension of that character set since they require an 8-bit representation.

Caution is advised when transferring IBM PC files since characters with ASCII values greater than 127 will be received in error when they are transmitted using 7

Table 1.10 The ASCII character set. This coded character set is to be used for the general interchange of information among information processing systems, communications systems, and associated equipment

Bits												
b_7					0	0	0	0	1	1	1	1
b_6					0	0	1	1	0	0	1	1
b_5					0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1	COLUMN / ROW	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	DLE	SP	0	@	P	\	p
0	0	0	1	1	SOH	DC1	!	1	A	Q	a	q
0	0	1	0	2	STX	DC2	"	2	B	R	b	r
0	0	1	1	3	ETX	DC3	#	3	C	S	c	s
0	1	0	0	4	EOT	DC4	$	4	D	T	d	t
0	1	0	1	5	ENQ	NAK	%	5	E	U	e	u
0	1	1	0	6	ACK	SYN	&	6	F	V	f	v
0	1	1	1	7	BEL	ETB	/	7	G	W	g	w
1	0	0	0	8	BS	CAN	(	8	H	X	h	x
1	0	0	1	9	HT	EM	)	9	I	Y	i	y
1	0	1	0	10	LF	SUB	★	:	J	Z	j	z
1	0	1	1	11	VT	ESC	+	;	K	[	k	{
1	1	0	0	12	FF	FS	,	<	L	\	l	∫
1	1	0	1	13	CR	GS	–	=	M	j	m	}
1	1	1	0	14	SO	RS	.	>	N	∧	n	~
1	1	1	1	15	SI	US	/	?	O	—	o	DEL

Note that b_7 is the higher order bit and b_1 is the low order bit as indicated by the following example for coding the letter C.

b_7	b_6	b_5	b_4	b_3	b_2	b_1
1	0	0	0	0	1	1

data bits. This is because the ASCII values of these characters will be truncated to values in the range 0 to 127 when transmitted with 7 bits from their actual range of 0 to 255. To alleviate this problem from occurring you can initialize your communications software for 8-bit data transfer; however, the receiving device must also be capable of supporting 8-bits ASCII data.

Table 1.11 ASCII control characters

Column/row	Control character		Mnemonic and meaning
0/0	^@	NUL	Null (CC)
0/1	^A	SOH	Start of heading (CC)
0/2	^B	STX	Start of text (CC)
0/3	^C	ETX	End of text (CC)
0/4	^D	EOT	End of transmission (CC)
0/5	^E	ENQ	Enquiry (CC)
0/6	^F	ACK	Acknowledgement (CC)
0/7	^G	BEL	Bell
0/8	^H	BS	Backspace (FE)
0/9	^I	HT	Horizonatal tabulation (FE)
0/10	^J	LF	Line feed (FE)
0/11	^K	VT	Vertical tabulation (FE)
0/12	^L	FF	Form feed (FE)
0/13	^M	CR	Carriage return (FE)
0/14	^N	SO	Shift out
0/15	^O	SI	Shift in
1/0	^P	DLE	Date link escape (CC)
1/1	^Q	DC1	Device control 1
1/2	^R	DC2	Device control 2
1/3	^S	DC3	Device control 3
1/4	^T	DC4	Device control 4
1/5	^U	NAK	Negative acknowledge (CC)
1/6	^V	SYN	Synchronous idle (CC)
1/7	^W	ETB	End of transmission block (CC)
1/8	^X	CAN	Cancel
1/9	^Y	EM	End of medium
1/10	^Z	SUB	Substitute
1/11	^[	ESC	Escape
1/12	^/	FS	File separator (IS)
1/13	^]	GS	Group separator (IS)
1/14	~	RS	Record separator (IS)
1/15	,	US	Unit separator (IS)
7/15	^-	DEL	Delete

(CC) communications control; (FE) format effector; (IS) information separator.

Although conventional ASCII files can be transmitted in a 7-bit format, many word processing and computer programs contain text graphics represented by ASCII characters whose values exceed 127. In addition, EXE and COM files which are produced by assemblers and compilers contain binary data that must also be transmitted in 8-bit ASCII to be accurately received. While most communications programs can transmit 7- or 8-bit ASCII data, some programs, especially those developed in the early 1980s and handed down with antiquated computers, may not be able to transmit binary files accurately. This is due to the fact that communications programs that use the control Z character (ASCII SUB) to identify the end of a file transfer will misinterpret a group of 8 bits in an EXE or COM file being transmitted when they have the same 8-bit format as a control Z, and upon detection prematurely close the file. To avoid this situation you should obtain a communications software program that transfers files by blocks of bits or converts the data into a hexadecimal or octal ASCII equivalent prior to transmission if this type of data transfer will be required.

Another communications problem you can encounter occurs when attempting to transmit files using some electronic mail services that are limited to transferring 7-bit ASCII. If you created a file using a word processor which employs 8-bit ASCII codes to indicate special character settings, such as bold and underlined text, attempting to transmit the file via a 7-bit ASCII electronic mail service will result in the proverbial gobbledygook being received at the distant end. Instead, you should first save the file as a 'text' file prior to transmission. Although you will lose any embedded 8-bit ASCII control codes, you will be able to transmit the document over an electronic mail system limited to the transfer of 7-bit ASCII.

Code conversion

A frequent problem in data communications is that of code conversion. Consider what must be done to enable a computer with an EBCDIC character set to transmit and receive information from a terminal with an ASCII character set. When that terminal transmits a character, that character is encoded according to the ASCII character code. Upon receipt of that character, the computer must convert the bits of information of the ASCII character into an equivalent EBCDIC character. Conversely, when data is to be transmitted to the terminal, it must be converted from EBCDIC to ASCII so the terminal will be able to decode and act according to the information in the character that the terminal is built to interpret.

One of the most frequent applications of code conversion occurs when personal computers are used to communicate with IBM mainframe computers.

Normally, ASCII to EBCDIC code conversion is implemented when an IBM PC or compatible personal computer is required to operate as a 3270 type terminal. This type of terminal is typically connected to an IBM or IBM compatible mainframe computer and the terminal's replacement by an IBM PC requires the PC's ASCII coded data to be translated into EBCDIC. There are many ways to obtain this conversion, including emulation boards that are inserted into the system unit of a PC and protocol converters that are connected between the PC and the mainframe computer. Later in this book, we will explore these and other methods that enable the PC to communicate with mainframe computers that transmit data coded in EBCDIC.

Table 1.12 lists the ASCII and EBCDIC code character value for the 10 digits for comparison purposes. In examining the difference between ASCII and EBCDIC coded digits you will note that each EBCDIC coded digit has a value precisely Hex C0 (decimal 192) higher than its ASCII equivalent. Although this might appear to make code conversion a simple process of adding or subtracting a fixed quantity depending upon which way the code conversion takes place, in reality many of the same ASCII and EBCDIC coded characters differ by varying quantities. As an example, the slash (/) character is Hex 2F in ASCII and Hex 61 in EBCDIC, a difference of Hex 92 (decimal 146). In comparison, other characters such as the carriage return and form feed have the same coded value in ASCII and EBCDIC, while other characters are displaced by different amounts in these two

Table 1.12 ASCII and EBCDIC digits comparison

ASCII				
Dec	Oct	Hex	EBCDIC	Digit
048	060	30	F0	0
049	061	31	F1	1
050	062	32	F2	2
051	063	33	F3	3
052	064	34	F4	4
053	065	35	F5	5
054	066	36	F6	6
055	067	37	F7	7
056	070	38	F8	8
057	071	39	F9	9

codes. Due to this, code conversion is typically performed as a table lookup process, with two buffer areas used to convert between codes in each of the two conversion directions. Thus, one buffer area might have the ASCII character set in Hex order in one field of a two-field buffer area, with the equivalent EBCDIC Hex values in a second field in the buffer area. Then, upon receipt of an ASCII character its Hex value is obtained and matched to the equivalent value in the first field of the buffer area, with the value of the second field containing the equivalent EBCDIC Hex value which is then extracted to perform the code conversion.

1.12 ERROR DETECTION AND CORRECTION

As a signal propagates down a transmission medium several factors can cause it to be received in error: the transmission medium employed and impairments caused by nature and machinery.

The transmission medium will have a certain level of resistance to current flow that will cause signals to attenuate. In addition, inductance and capacitance will distort the transmitted signals and there will be a degree of leakage which is the loss in a transmission line due to current flowing across, though insulators, or changes in the magnetic field.

Transmission impairments result from numerous sources. First, Gaussian or white noise is always present as it is the noise level that exists due to the thermal motions of electrons in a circuit. Next, impulse can occur from line hits due to atmospheric static or poor contacts in a telephone system.

Regardless of the cause of a transmission disturbance, its duration and the operating rate of your communications session are the governing factors that determine the effect of the disturbance. For example, consider a noise burst of 0.01 s which is not uncommon and which sounds like a short click during a voice conversation. At a 1200 bps transmission rate the noise burst will result in 12 bits having the potential to be received in error. At 2400 bps the number of bits having the potential to be received in error is increased to 24. Similarly, increasing the duration of the noise burst increases the number of bits that have the potential to be received in error. For example, a noise burst of 0.1 s would result in 120 bits having the potential to be received in error when the data rate is 1200 bps, while 240 bits would have the potential to be received in error when the data rate is 2400 bps.

Asynchronous transmission

In asynchronous transmission the most common form of error control is the use of a single bit, known as a parity bit, for the detection of errors. Owing to the proliferation of personal computer communications, more sophisticated error detection methods have been developed which resemble the methods employed with synchronous transmission.

Parity checking

Character parity checking, also known as vertical redundancy checking (VRC), requires an extra bit to be added to each character in order to make the total quantity of 1 s in the character either odd or even, depending upon whether you are employing odd parity checking or even parity checking. When odd parity checking is employed, the parity bit is set to 1 if the number of 1 s in the character's data bits is even; or it is set at 0 if the number of 1 s in the character's data bits is odd. When even parity checking is used, the parity bit is set to 0 if the number of 1 s in the character's data bits is even; or if it is set to 1 if the number of 1 s in the character's data bits is odd.

Two additional terms used to reference parity settings are 'mark' and 'space'. When the parity bit is set to a mark condition the parity bit is always 1 while space parity results in the parity bit always set to 0. Although not actually a parity setting, parity can be set to none, in which case no parity checking will occur. When transmitting binary data asynchronously, such as between personal computers via a wide area network transmission facility, parity checking must be set to none or off. This enables all 8 bits to be used to represent a character. Table 1.13 summarizes the effect of five types of parity checking upon the eighth data bit in asynchronous transmission.

Table 1.13 Parity effect upon eighth data bit

Parity type	Parity effect
Odd	Eighth bit is logical zero if the total number of logical 1s in the first seven data bits is odd
Even	Eighth bit is logical zero if the total number of logical 1s in the first seven data bits is even.
Mark	Eighth bit is always logical 1.
Space	Eighth bit is always logical zero.
None/Off	Eighth bit is ignored.

For an example of parity checking, let us examine the ASCII character R whose bit composition is 1 0 1 0 0 1 0. Since there are three 1 bits in the character R, a 0 bit would be added if odd parity checking is used or a 1 bit would be added as the parity bit if even parity checking is employed. Thus, the ASCII character R would appear as follows:

data bits	parity bit	
1 0 1 0 0 1 0	0	odd parity check
1 0 1 0 0 1 0	1	even parity check
1 0 1 0 0 1 0	1	mark parity check
1 0 1 0 0 1 0	0	space parity check

Since there are three bits set in the character R, a 0 bit is added if odd parity checking is employed while a 1 bit is added if even parity checking is used. Similarly, mark parity results in the parity bit being set to 1 regardless of the composition of the data bits in the character, while space parity results in the parity bit always being set to 0.

Undetected errors

Although parity checking is a simple mechanism to investigate if a single bit error has occurred, it can fail when multiple bit errors occur. This can be visualized by returning to the ASCII R character example and examining the effect of additional bits erroneously being transformed as indicated in Table 1.14. Here the ASCII R character has three set bits and a one-bit error could transform the number of set bits to four. If parity checking is employed, the received set parity bit would result in the character containing five set bits, which is obviously an error since even parity checking is employed. Now suppose two bits are transformed in error as indicated in Table 1.14. This would result in the reception of a character containing six set bits, which would appear to be correct under even parity checking. Thus, two bit errors in this situation would not be detected by a parity error detection technique.

Table 1.14 Character parity cannot detect an even number of bit errors

ASCII character R	1 0 1 0 0 1 0
Adding an even parity bit	1 0 1 0 0 1 0 1
	1
1 bit in error	1 ∅ 1 0 0 1 0 1
	1 1
2 bits in error	1 ∅ 1 ∅ 0 1 0 1
	1 0 1
3 bits in error	1 ∅ ✗ ∅ 0 1 0 1
	0 1 0 1
4 bits in error	✗ ∅ ✗ ∅ 0 1 0 1

Now consider the effect of three bit errors shown in Table 1.14. In this example, the number of set bits was charged to five which is a detectable error when even parity checking is employed. Let us, however, examine the effect of four bits being in error. In this example at the bottom of Table 1.14 the number of set bits was changed to four, which does not result in the detection of errors when even parity is employed. From the preceding we should recognize that parity can detect the occurrence of an odd number of bits errors within a character. Errors that effect an even number of bits will not be detected. This problem exists regardless of the method of parity checking used.

In addition to the potential of undetected errors, parity checking has several additional limitations. First, the response to parity errors will vary based upon the type of computer with which you are communicating. Certain mainframes will issue a 'Retransmit' message upon detection of a parity error. Some mainframes will transmit a character that will appear as a 'fuzzy box' on your screen in response to detecting a parity error, while other computers will completely ignore parity errors.

When transmitting data asynchronously on a personal computer, most communications programs permit the user to set parity to odd, even, off, space, or mark. Off or no parity would be used if the system with which your are communicating does not check the parity bit for transmission errors. No parity would be used when you are transmitting 8-bit EBCDIC or an extended 8-bit ASCII coded data, such as that available on the IBM PC and similar personal computers. Mark parity means that the parity bit is set to 1, while space parity means that the parity bit is set to 0.

In the asynchronous communications world, two common sets of parameters are used by most bulletin boards, information utilities and supported by mainframe computers. The first set consists of seven data bits and one stop bit with even parity checking employed, while the second set consists of eight data bits and one stop bit using no parity checking. Table 1.15 compares the communications parameter settings of two popular information utilities.

Table 1.15 Communication parameter settings

Parameter	Information utility	
	CompuServe	Dow Jones
Data bits	7/8	8
Parity	even/none	none
Stop bits	1	1
Duplex	full	full

File transfer problems

Although visual identification of parity errors in an interactive environment is possible, what happens if you transfer a large file over the switched telephone network? During the 1970s a typical call over the switched telephone network resulted in the probability of a random bit error occurring of approximately 1 in 100 000 bits at a data transmission rate of 1200 bps. If you desired to upload or download a 1000-line program containing an average of 40 characters per line, a total of 320 000 data bits would have to be transmitted. During the 4.4 minutes required to transfer this file you could expect 3.2 bit errors to occur, probably resulting in several program lines being received incorrectly if the errors occur randomly. In such situations you would prefer an alternative to visual inspection. Thus, a more efficient error detection and correction method was needed for large data transfers.

Block checking

In this method, data is grouped into blocks for transmission. A checksum character is generated and appended to the transmitted block and the checksum is also calculated at the receiver, using the same algorithm. If the checksums match, the data block is considered to be received correctly. If the checksums do not match, the block is considered to be in error and the receiving station will request the transmitting station to retransmit the block.

One of the most popular asynchronous block checking methods is included in the XMODEM protocol, which was the first method developed to facilitate file transfer and is still extensively used in personal computer communications. Although the operation of this protocol is covered in detail later in this chapter, we will focus our attention at his time on its error detection and correction capability. For information concerning the actual operation of data transfers under the XMODEM protocol the reader is referred to Section 1.15.

Under the XMODEM protocol groups of asynchronous characters are blocked together for transmission and a checksum is computed and appended to the end of the block. The checksum is obtained by first summing the ASCII value of each data character in the block and dividing that sum by 255. Then, the quotient is discarded and the remainder is appended to the block as the checksum. Thus, mathematically the XMODEM checksum can be represented as

$$\text{Checksum} = R\left[\frac{\sum_{1}^{128} \text{ASCII value of characters}}{255}\right]$$

where R is the remainder of the division process.

When data is transmitted using the XMODEM protocol, the receiving device at the other end of the link performs the same operation upon the block being received. This 'internally' generated checksum is compared to the transmitted checksum. If the two checksums match, the block is considered to have been received error-free. If the two checksums do not match, the block is considered to be in error and the receiving device will then request the transmitting device to resend the block.

Start of header	Block number	One's complement block number	128 data characters	Checksum

Figure 1.26 XMODEM protocol block format. The start of header is the ASCII SOH character whose bit composition is 00000001, while the one's complement of the block number is obtained by subtracting the block number from 255. The checksum is formed by first adding the ASCII values of each of the characters in the 128 character block, dividing the sum by 255 and using the remainder

Figure 1.26 illustrates the XMODEM protocol block format. The start of header is the ASCII SOH character whose bit composition is 0 0 0 0 0 0 0 1, while the one's complement of the block number is obtained by subtracting the block number from 255. The block number and its complement are contained at the beginning of each block to reduce the possibility of a line hit at the beginning of the transmission of a block causing the block number to be corrupted but not detected.

The construction of the XMODEM protocol format permits errors to be detected in one of three ways. First, if the start of header is damaged, it will be detected by the receiver and the data block will be negatively acknowledged. Next, if either the block number or the one's complement field is damaged, they will not be the one's complement of each other, resulting in the receiver negatively acknowledging the data block. Finally, if the checksum generated by the receiver does not match the transmitted checksum, the receiver will transmit a negative acknowledgement. For all three situations the negative acknowledgement will serve as a request to the transmitting station to retransmit the previously transmitted block.

Data transparency

Since the XMODEM protocol supports an 8-bit, no parity data format it is transparent to the data content of each byte. This enables the protocol to support ASCII, binary and extended ASCII data transmission, where extended ASCII is the additional 128 graphic characters used by the IBM PC and compatible computers through the employment of an 8-bit ASCII code.

Error detection efficiency

While the employment of a checksum reduces the probability of undetected errors in comparison to parity checking, it is still possible for undetected errors to occur under the XMODEM protocol. This can be visualized by examining the construction of the checksum character and the occurrence of multiple errors when a data block is transmitted.

Assume a 128-character data block of all 1 s is to be transmitted and each data character has the format 0 0 1 1 0 0 0 1, which is an ASCII 49. When the checksum

```
       0                       10
..10011000~~1~~|00110001|001100~~0~~~~1~~|...
```

Figure 1.27 Multiple errors on an XMODEM data block

is computed the ASCII value of each data character is first added, resulting in a sum of 6272 (128×49). Next, the sum is divided by 255, with the remainder used as the checksum, which in this example is 152.

Suppose two transmission impairments occur during the transmission of a data block under the XMODEM protocol, affecting two data characters as illustrated in Figure 1.27.

Here the first transmission impairment converted the ASCII value of the character from 49 to 48, while the second impairment converted the ASCII value of the character from 49 to 50. Assuming no other errors occurred, the receiving device would add the ASCII value of each of the 128 data characters and obtain a sum of 6272. When the receiver divides the sum by 255, it obtains a checksum of 152, which matches the transmitted checksum and the errors remain undetected. Although the preceding illustration was contrived, it illustrates the potential for undetected errors to occur under the XMODEM protocol.

To make the protocol more efficient with respect to undetected errors, several derivatives of the XMODEM protocol have gained popularity and are now commonly available in most communications programs designed for operation on personal computers as well as supported by information utilities and bulletin board systems. These derivatives of the XMODEM protocol include XMODEM/CRC and YMODEM, each of which uses a cyclic redundancy check (CRC) in place of the checksum for error detection. The use of CRC error detection reduces the probability of undetected errors to less than one in a million blocks and is the preferred method for ensuring data integrity. The concept of CRC error detection is explained later in this section under synchronous transmission, as it was first employed with this type of transmission. The operation of the XMODEM, XMODEM/CRC and YMODEM protocols are examined in more detail in the protocol section of this chapter.

Synchronous transmission

The majority of error-detection schemes employed in synchronous transmission involve geometric codes or cyclic code. However, several modifications to the original XMODEM protocol, such as XMODEM-CRC, use a cyclic code to protect asynchronously transmitted data.

Geometric codes attack the deficiency of parity by extending it to two dimensions. This involves forming a parity bit on each individual character as well as on all the characters in the block. Figure 1.28 illustrates the use of block parity checking for a block of 10 data characters. As indicated, this block parity character is also known as the 'longitudinal redundancy check' (LRC) character.

Geometric codes are similar to the XMODEM error-detection technique in the fact that they are also far from foolproof. As an example of this, suppose a 2-bit

			Character parity bit
Character	1	1011011	0
Character	2	0100101	0
	3	0110100	0
	4	1001001	0
	5	0111101	0
	6	1010000	1
	7	0101110	1
	8	0111001	1
	9	1000110	0
	10	0110101	1
Block parity character (LRC)	. .	1110101	1

Figure 1.28 VRC/LRC geometric code (odd parity checking)

duration transmission impairment occurred at bit positions 3 and 4 when characters 7 and 9 in Figure 1.28 were transmitted. Here the two 1 s in those bit positions might be replaced by two 0 s. In this situation each character parity bit as well as the block parity character would fail to detect the errors.

A transmission system using a geometric code for error detection has a slightly better capability to detect errors than the method used in the XMODEM protocol and is hundreds of times better than simple parity checking. While block parity checking substantially reduces the probability of an undetected error in comparison to simple parity checking on a character by character basis, other techniques can be used to further decrease the possibility of undetected errors. Among these techniques is the use of cyclic or polynomial code.

Cyclic codes

When a cyclic or polynomial code error-detection scheme is employed the message block is treated as a data polynomial $D(x)$, which is divided by a predefined generating polynomial $G(x)$, resulting in a quotient polynomial $Q(x)$ and a remainder polynomial $R(x)$, such that

$$D(x)/G(x) = Q(x) + R(x)$$

The remainder of the division process is known as the cyclic redundancy check (CRC) and is normally 16 bits in length or two 8-bit bytes. The CRC checking method is used in synchronous transmission and in asynchronous transmission using derivatives of the XMODEM protocol similar to the manner in which the checksum is employed in the XMODEM protocol previously discussed. That is, the CRC is appended to the block of data to be transmitted. The receiving devices uses the same predefined generating polynomial to generate its own CRC based upon the received message block and then compares the 'internally' generated CRC with the transmitted CRC. If the two match, the receiver transmits a positive acknowledgement (ACK) communications control character to the transmitting device which not only informs the distant device that the data was received

Table 1.16 Common generating polymonials

Standard	Polynomial
CRC-16 (ANSI)	$X^{16} + X^{15} + X^5 + 1$
CRC (ITU)	$X^{16} + X^{12} + X^5 + 1$
CRC-12	$X^{12} + X^{11} + X^3 + 1$
CRC-32	$X^{32} + X^{26} + X^{23} + X^{22} + X^{16} + X^{12} + X^{11} + X^{10} + X^8$ $+ X^7 + X^5 + X^4 + X^2 + X + 1$

correctly but also serves to inform the device that if additional blocks of data remain to be transmitted the next block can be sent. If an error has occurred, the internally generated CRC will not match the transmitted CRC and the receiver will transmit a negative acknowledgement (NAK) communications control character which informs the transmitting device to retransmit the block previously sent.

Table 1.16 lists four generating polynomials in common use today. The CRC-16 is based upon the American National Standards Institute and is commonly used in the United States. The ITU CRC is commonly used in transmissions in Europe while the CRC-12 is used with 6-level transmission codes and has been basically superseded by the 16-bit polynomials. The 32-bit CRC is defined for use in local networks by the Institute of Electrical and Electronic Engineers (IEEE) and the American National Standards Institute (ANSI). For further information concerning the use of the CRC-32 polynomial the reader is referred to the IEEE/ANSI 802 standards publications.

The column labelled polynomial in Table 1.16 actually indicates the set bits of the 32-bit, 16-bit or 12-bit polynomial. Thus, the CRC-16 polynomial has a bit composition of 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1.

Hardware computations

Originally CRC computations were associated with synchronous transmission in which the computations were performed using special hardware known as shift registers. Figure 1.29 illustrates the basic format of a multisection shift register used for the computation of the ITU CRC. The contents of this register are initially set to zero.

As each bit in a data block is output onto the communications line it is also applied to the location marked X in Figure 1.29. This results in that bit and subsequent bits being placed into bit positions 15, 10, and 3. Note that those bit positions reference the positions within the register. Since the least significant bit (LSB) is output from the register first, that bit position corresponds to the X^{16} term of the CRC.

As bits are shifted from left to right through the segments of the shift register they eventually reach exclusive-OR gates located between each section of the register. The input to each of these exclusive-OR gates consists of bits shifted through sections of the registers and new bits output onto the communications line.

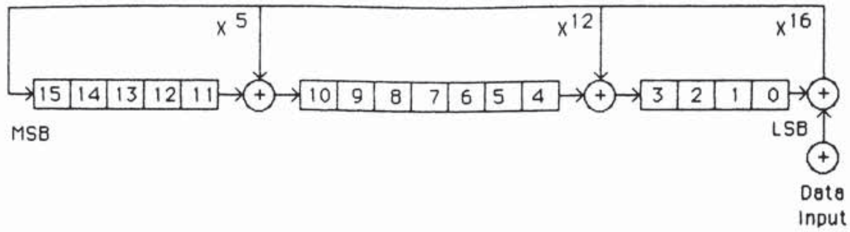


Figure 1.29 Multisection shift used for ITU CRC computation + exclusive-OR gate.

Input	Output
00	0
01	1
10	1
11	0

The operation of an exclusive-OR gate results in the output of a 0 if the inputs are both 0 or both 1. If the inputs differ, the output of the exclusive-OR gate is a 1. Once all of the bits in a data block entered the shift register the transmitter retrieves the contents of the shift register and transmits the 16 bits as a two-byte CRC. This two-byte CRC is commonly referred to as the block check characters (BCC) in communications literature. Even though a CRC-32 uses a four-byte CRC, that CRC is also commonly referred to as a BCC. After transmitting the CRC the transmitter initializes the shift register positions to zero prior to sending bits from the next block into the register. At the opposite end of the transmission link the receiving device applies the incoming bit stream to point X of a similar multisection shift register. Once the receiver's CRC is computed it is then extracted from the shift register and compared to the CRC appended to the transmitted data block. If the two CRCs match the data block is considered to be received without error.

Although hardware-based CRC computations are still normally used with synchronous transmission, almost all such computations are performed by software using the microprocessor of personal computers. To calculate the 16-bit CRC the message bits are considered to be the coefficients of a polynomial. This message polynomial is first multiplied by X^{16} and then divided by the generator polynomial $(X^{16} + X^{12} + X^5 + 1)$ using modulo 2 arithmetic. The remainder after the division is the desired CRC.

International transmission

Due to the growth in international communications, one frequently encountered transmission problem is the employment of dissimilar CRC generating polynomials. This typically occurs when an organization in the United States attempts to communicate with a computer system in Europe or a European organization attempts to transmit to a computer system located in the United States.

When dissimilar CRC generating polynomials are employed, the two-byte block-check character appended to the transmitted data block will never equal the block-check character computed at the receiver. This will result in each transmitted data block being negatively acknowledged, eventually resulting in a threshold of negative acknowledgements being reached. When this threshold is reached the protocol aborts the transmission session, causing the terminal operator to reinitiate the communications procedure required to access the computer system they wish to connect to. Although the solution to this problem requires either the terminal or a port on the computer system to be changed to use the appropriate generating polynomial, the lack of publication of the fact that there are different generating polynomials has caused many organizations to expend a considerable amount of needless effort. One bank which the author is familiar with monitored transmission attempts for almost 3 weeks. During this period, they observed each block being negatively acknowledged and blamed the communications carrier, insisting that the quality of the circuit was the culprit. Only after a consultant was called and spent approximately a week examining the situation was the problem traced to the utilization of dissimilar generating polynomials.

Forward error correcting

During the 1950s and 1960s when mainframe computers used core memory circuits, designers spent a considerable amount of effort developing codes that carried information which enabled errors to be detected and corrected. Such codes are collectively called forward error correction (FEC) and have been employed in Trellis Coded Modulation modems under the ITU-T V.32, V.32 bis, V.33, and V.34 recommendation, and will be described later in this book.

One popular example of a forward error correcting code is the Hamming code. This code can be used to detect one or more bits in error at a receiver as well as to determine which bits are in error. Since a bit can have only one of two values, knowledge that a bit is in error allows the receiver to reverse or reset the bit, correcting its erroneous condition.

The Hamming code uses m parity bits with a message length of n bits, where $n = 2^m - 1$. This permits k information bits where $k = n - m$. The parity bits are then inserted into the message at bit positions 2^{j-1} where $j = 1, 2, \ldots, m$. Table 1.17 illustrates the use of a Hamming code error correction for $m = 3$, $k = 4$, and $n = 7$.

In the Hamming code encoding process the data and parity bits are exclusive-ORed with all possible data values to determine the value of each parity bit. This is illustrated in Figure 1.30 which results in P_1, P_2, and P_3 having values of 0, 1, and 1, respectively.

Table 1.17 Hamming code error correction example

Message length	= 7 bits
Information bits	= 4 bits = 1100
Parity	= 3 bits = P_1, P_2, P_3

$$[P_1P_21P_3100] \begin{bmatrix} 0\,0\,1 \\ 0\,1\,0 \\ 0\,1\,1 \\ 1\,0\,0 \\ 1\,0\,1 \\ 1\,1\,0 \\ 1\,1\,1 \end{bmatrix} = \begin{cases} P_1 + 0 + 1 + 0 + 1 + 0 + 0 = 0 \\ 0 + P_2 + 1 + 0 + 0 + 0 + 0 = 0 \\ 0 + 0 + 0 + P_3 + 1 + 0 + 0 = 0 \end{cases}$$

Figure 1.30 The Hamming code encoding process. Using exclusive-OR arithmetic the three equations yield $P_1 = 0$, $P_2 = 1$, $P_3 = 1$

$$[0\ 1\ 1\ 1\ 0\ 0\ 0] \begin{bmatrix} 0\,0\,1 \\ 0\,1\,0 \\ 0\,1\,1 \\ 1\,0\,0 \\ 1\,0\,1 \\ 1\,1\,0 \\ 1\,1\,1 \end{bmatrix} = \begin{cases} 0\,0\,1\,0\,0\,0\,0 = 1 \\ 0\,1\,1\,0\,0\,0\,0 = 0 \\ 0\,0\,0\,1\,0\,0\,0 = 1 \end{cases}$$

Figure 1.31 The Hamming code decoding process

Using the values obtained for each parity bit results in the transmitted message becoming 0111100. Now assume an error occurs in bit 5, resulting in the received message becoming 0111000. Figure 1.31 illustrates the Hamming code decoding process in which the received message is exclusive-ORed against a matrix of all possible values of the three parity bits, yielding the sum 101_2, which is the binary value for 5. Thus, bit 5 is in error and is corrected by reversing or resetting its value.

1.13 STANDARDS ORGANIZATIONS, ACTIVITIES AND THE OSI REFERENCE MODEL

The importance of standards and the work of standards organizations have proved essential for the growth of worldwide communications. In the United States and many other countries, national standards organizations have defined physical and operational characteristics that enable vendors to manufacture equipment compatible with line facilities provided by communications carriers as well as equipment produced by other vendors. At the international level, standards organizations have promulgated several series of communications related recommendations. These recommendations, while not mandatory, have become highly influential on a worldwide basis for the development of equipment and facilities and have been adopted by hundreds of public companies and communications carriers.

In addition to national and international standards a series of *de facto* standards has evolved through the licensing of technology among companies. Such *de facto* standards, as an example, have facilitated the development of communications software for use on personal computers. Today, consumers can purchase communications software that can control modems manufactured by hundreds of vendors since most modems are now constructed to respond to a uniform set of control codes.

In this section we will first focus our attention upon national and international standards bodies as well as discuss the importance of *de facto* standards. Due to the importance of the Open Systems Interconnection (OSI) Reference Model in data communications we will conclude this section with an examination of that model. Although this examination will only provide the reader with an overview of the seven layers of the OSI model, it will provide a foundation for a detailed investigation of several layers of that model presented in subsequent chapters in this book.

National standards organizations

Table 1.18 lists six representative national standards organizations—four whose activities are primarily focused within the United States and two whose activities are directed within specific foreign countries. Although each of the organizations listed in Table 1.18 is an independent entity, their level of coordination and cooperation with one another and international organizations is very high. In fact, many standards adopted by one organization have been adopted either 'as is' or in modified form by other organizations.

Table 1.18 Representative national standards organizations

United States
American National Standards Institute
Electronic Industries Association
Federal Information Processing Standards
Institute of Electrical and Electronic Engineers
Foreign
British Standards Institution
Canadian Standards Association

ANSI

The American National Standards Institute (ANSI) is the principal standards forming body in the United States. This non-profit, non-governmental organization is supported by approximately 1000 professional societies, companies and trade organizations. ANSI represents the United States in the International Organization for Standardization (ISO) and develops voluntary standards that are designed to benefit both consumers and manufacturers.

Perhaps the most well known part of ANSI is its X3 standards committee. Established in 1960 to investigate computer industry related standards, the X3

Table 1.19 ANSI X3S3 data communications technical committee task groups

Task group	Responsibility
X3S31	Planning
X3S32	Glossary
X3S33	Transmission format
X3S34	Control procedures
X3S35	System performance
X3S36	Signaling speeds
X3S37	Public data networks

standards committee has grown to 25 technical committees to include the X3S3 Data Communications Technical Committee which is composed of seven task groups. Table 1.19 lists the current task groups of the ANSI X3S3 Data Communications Technical Committee and their responsibilities.

Foremost among ANSI standards are X3.4 which defines the ASCII code and X3.1 which defines data signaling rates for synchronous data transmission on the switched telephone network.

Table 1.20 lists representative ANSI data communications standards publications by number and title. Many of these standards have been adopted by the US government as Federal Information Processing Standards (FIPS). The reader is referred to Table 1.23 at the end of this section for a list of the addresses of standards organizations mentioned in this section.

EIA

The Electronic Industries Association (EIA) is a trade organization headquartered in Washington, DC, which represents most major US electronics industry manufacturers. Since its founding in 1924, the EIA Engineering Department has published over 500 documents related to standards.

In the area of communications, the EIA established its Technical Committee TR-30 in 1962. The primary emphasis of this committee is upon the development and maintenance of interface standards governing the attachment of data terminal equipment (DTE), such as terminals and computer ports, to data communications equipment (DCE), such as modems and digital service units.

TR-30 committee standards activities include the development of the ubiquitous RS-232 interface standard which describes the operation of a 25-pin conductor which is the most commonly used physical interface for connecting DTE to DCE.

Two other commonly known EIA standards and one emerging standard are RS-366A, RS-449, and RS-530. RS-366 describes the interface used to connect terminal devices to automatic calling units; RS-449 was originally intended to replace the RS-232 interface due to its ability to extend the cabling distance between devices, while RS-530 may eventually evolve as a replacement for both RS-232 and RS-449 as it eliminates many objections to RS-449 that inhibited its

Table 1.20 Representative ANSI publications

Publication number	Publication title
X3.1	Synchronous Signaling Rates for Data Transmission
X3.4	Code for Information Interchange
X3.15	Bit Sequencing for the American National Standard Code for Information Interchange in Serial-by-Bit Data Transmission
X3.16	Character Structure and Character Parity Sense for Serial-by-Bit Data Communications Information Interchange
X3.24	Signal Quality at Interface Between Data Processing Technical Equipment for Synchronous Data Transmission
X3.25	Character Structure and Character Parity Sense for Parallel-by-Bit Communications in American National Standard Code for Information Interchange
X3.28	Procedures for the Use of the Communications Control Characters for American National Standard Code for Information Interchange interchange in Specific Data Communications Links
X3.36	Synchronous High-Speed Signaling Rates Between Data Terminal Equipment and Data Communication Equipment
X3.41	Code Extension Techniques for Use with 7-Bit Coded Character Set of American National Standard Code for Information Interchange
X3.44	Determination of the Performance of Data Communications Systems
X3.57	Structure for Formatting Message Headings for Information Interchange for Data Communication System Control
X3.66	American National Standard for Advanced Data Communication Control Procedures (ADCCP)
X3.79	Determination of Performance of Data Communications System that Use Bit-Oriented Control Procedures
X3.39	Data Encryption Algorithm

adoption. The reader is referred to Section 1.14 for detailed information covering the previously mentioned EIA standards.

The TR-30 committee works closely with both ANSI Technical Committee X3S3 and with groups within the International Telecommunications Union (ITU) Standardizations Sector (ITU-T) which was formerly known as the Consultative Committee for International Telephone and Telegraph (CCITT). In fact, the ITU V.24 standard is basically identical to the EIA RS-232 standard, resulting in hundreds of communications vendors designing RS-232/V.24 compatible equipment.

As a result of the widespread acceptance of the RS-232/V.24 interface standard, a cable containing up to 25 conductors with a predefined set of connectors can be used to cable most DTEs to DCEs. Even though there are exceptions to this interface standard, this standard has greatly facilitated the manufacture of communications products, such as terminals, computer ports, modems, and digital service units that are physically compatible with one another and which can be easily cabled to one another.

Another important EIA standard resulted from the joint efforts of the EIA and the Telecommunications Industry Association (TIA). Known as EIA/TIA-568, this standard defines structured wiring within a building and is primarily used for supporting local area networking. This standard will be covered in Chapter 3 when we focus our attention upon LANs.

FIPS

As a result of US Public Law 89-306 (the Brooks Act) which directed the Secretary of the Department of Commerce to make recommendations to the President concerning uniform data processing standards, that agency developed a computer standardization program. Since Public Law 89-306 did not cover telecommunications, the enactment of Public Law 99-500, known as the Brooks Act Amendment, expanded the definition of automatic data processing (ADP) to include certain aspects related to telecommunications. FIPS, an acronym for Federal Information Processing Standards, are the indirect result of Public Law 89-306 and is the term applied to standards developed under the US government's computer standardization program.

FIPS specifications are drafted by the National Institute of Standards and Technology (NIST), formerly known as the National Bureau of Standards (NBS). Approximately 100 FIPS have been adopted, ranging in scope from the ASCII code to the Hollerith punched card code and such computer languages as COBOL and FORTRAN. Most FIPS have an ANSI national counterpart. The key difference between FIPS and their ANSI counterparts is that applicable FIPS must be met for the procurement, management and operation of ADP and telecommunications equipment by federal agencies, whereas commercial organizations in the private sector can choose whether or not to obtain equipment that complies with appropriate ANSI standards.

IEEE

The Institute of Electrical and Electronic Engineers (IEEE) is a US-based engineering society that is very active in the development of data communications standards. In fact, the most prominent developer of local area networking standards is the IEEE, whose subcommittee 802 began its work in 1980 prior to the establishment of a viable market for the technology.

The IEEE Project 802 efforts are concentrated upon the physical interface of equipment and the procedures and functions required to establish, maintain and release connections among network devices to include defining data formats, error control procedures and other control activities governing the flow of information. This focus of the IEEE actually represents the lowest two layers of the ISO model, physical and data link, which are discussed later in this section and in more detail in Chapter 3.

BSI

The British Standards Institution (BSI) is the national standards body of the United Kingdom. In addition to drafting and promulgating British National Standards, BSI is responsible for representing the United Kingdom at ISO and other international bodies. BSI responsibilities include ensuring that British standards are in technical agreement with relevant international standards, resulting in, as an example, the widespread use of the V.24/RS-232 physical interface in the United Kingdom.

CSA

The Canadian Standards Association (CSA) is a private, non-profit organization which produces standards and certifies products for compliance with their standards. CSA functions similar to a combined US ANSI and Underwriters Laboratory, the latter also a private organization which is well known for testing electrical equipment ranging from ovens and toasters to modems and computers.

In many instances, CSA standards are the same as international standards, with many ISO and ITU standards adopted as CSA standards. In other instances, CSA standards represent modified international standards.

International standards organizations

Two important international standards organizations are the International Telecommunications Union (ITU) and the International Organization for Standardization (ISO). The ITU can be considered as a governmental body as it functions under the auspices of an agency of the United Nations. Although the ISO is a non-governmental agency, its work in the field of data communications is well recognized.

ITU

The International Telecommunications Union (ITU) is a specialized agency of the United Nations headquartered in Geneva, Switzerland. The ITU is tasked with direct responsibility for developing data communications standards and consists of 15 Study Groups, each tasked with a specific area of responsibility.

The work of the ITU is performed on a four-year cycle which is known as a study period. At the conclusion of each study period, a plenary session occurs. During the plenary session, the work of the ITU during the previous four years is reviewed, proposed recommendations are considered for adoption, and items to be investigated during the next four-year cycle are considered. The ITU's Tenth Plenary Session met in 1992 and its 11th session occurred during 1996. Although approval of recommended standards is not intended to be mandatory, ITU recommendations have the effect of law in some Western European countries and many of its recommendations have been adopted by both communications carriers and vendors in the United States.

Recommendations

Recommendations promulgated by the ITU are designed to serve as a guide for technical, operating and tariff questions related to data and telecommunications.

ITU recommendations are designated according to the letters of the alphabet, from Series A to Series Z, with technical standards included in Series G to Z. In the field of data communications, the most well known ITU recommendations include Series I which pertains to Integrated Services Digital Network (ISDN) transmission, Series Q which describes ISDN switching and signaling systems, Series V which covers facilities and transmission systems used for data transmission over the PSTN and leased telephone circuits, the DTE–DCE interface and modem operations, and Series X which covers data communications networks to include Open Systems Interconnection (OSI).

The ITU V-series

To provide readers with a general indication of the scope of ITU recommendations, Table 1.21 lists those promulgated for the V-series at the time this book was prepared. In examining the entries in Table 1.21, note that the ITU Recommendation V.3 is actually a slightly modified ANSI X3.4 standard, the ASCII code. For international use, the V.3 recommendation specifies national currency symbols in place of the dollar sign ($) as well as a few other minor differences. You should also note that certain ITU recommendations, such as V.21, V.22 and V.23, among others, while similar to AT&T Bell modems, may or may not provide operational compatibility with modems manufactured to Bell specifications. Chapter 5 provides detailed information concerning modem operations and compatibility issues.

ISO

The International Organization for Standardization (ISO) is a non-governmental entity that has consultative status within the UN Economic and Social Council. The goal of the ISO is to 'promote the development of standards in the world with a view to facilitating international exchange of goods and services'.

The membership of the ISO consists of the national standards organizations of most countries, with approximately 100 countries currently participating in its work.

Perhaps the most notable achievement of the ISO in the field of communications is its development of the seven-layer Open Systems Interconnection (OSI) Reference Model. This model is discussed in detail at the end of this section.

De facto standards

Prior to the breakup of AT&T, a process referred to as divestiture, US telephone interface definitions were the exclusive domain of AT&T and its research subsidiary, Bell Laboratories. Other vendors which developed equipment for use on the AT&T network had to construct their equipment to those interface definitions. In addition, since AT&T originally had a monopoly on equipment that could be connected to the switched telephone network, upon liberalization of that

Table 1.21 CCITT V-series recommendations

General

V.1 Equivalence between binary notation symbols and the significant conditions of a two-condition code.

V.2 Power levels for data transmission over telephone lines.

V.3 International Alphabet No. 5.

V.4 General structure of signals of International Alphabet No. 5 code of data transmission over public telephone networks.

V.5 Standardization of data signaling rates for synchronous data transmission in the general switched telephone network.

V.6 Standardization of data signaling rates for synchronous data transmission on leased telephone-type circuits.

V.7 Definitions of terms concerning data communication over the telephone network.

Interface and voice-band modems

V.10 Electrical characteristic for unbalanced double-current interchange circuits for general use with integrated circuit equipment in the field of data communications. Electrically similar to RS-423.

V.11 Electrical characteristics of balanced double-current interchange circuits for general use with integrated circuit equipment in the field of data communications. Electrically similar to RS-422.

V.15 Use of acoustic coupling for data transmission.

V.16 Medical analogue data transmission modems.

V.19 Modems for parallel data transmission using telephone signaling frequencies.

V.20 Parallel data transmission modems standardized for universal use in the general switched telephone network.

V.21 300 bps duplex modem standardized for use in the general switched telephone network. Similar to the Bell 103.

V.22 1200 bps duplex modem standardized for use on the general switched telephone network and on leased circuits. similar to the Bell 212.

V.22 bis 2400 bps full-duplex two-wire modem standard.

V.23 600/1200 baud modem standardized for use in the general switched telephone network. Similar to the Bell 202.

V.24 List of definitions for interchange circuits between data terminal equipment and data circuit-terminating equipment. Similar to and operationally compatible with RS-232.

V.25 Automatic calling and/or answering on the general switched telephone network, including disabling of echo suppressors on manually established calls. RS-366 parallel interface.

V.25 bis Serial RS-232 interface.

V.26 2400 bps modem standardized for use on four-wire leased telephone-type circuits. Similar to the Bell 201 B.

V.26 bis 2400/1200 pbs modem standardized for use in the general switched telephone network. Similar to the Bell 201 C.

V.26 ter 2400 pbs modem that uses echo cancellation techniques suitable for application in the general switched telephone network.

V.27 4800 bps modem with manual equalizer standardized for use on leased telephone-type circuits. Similar to the Bell 208A.

(continued)

Table 1.21 (*cont.*)

V.27 bis	4800/2400 bps modem with automatic equalizer standardized for use on leased telephone-type circuits.
V.27 ter	4800/2400 bps modem standardized for use in general switched telephone network. Similar to the Bell 208B.
V.28	Electrical characteristics for unbalanced double-current interchange circuits (defined by V.24; similar to and operational with RS-232).
V.29	9600 bps modem standardized for use on point-to-point four-wire leased telephone-type circuits. Similar to the Bell 209.
V.31	Electrical characteristics for single-current interchange circuits controlled by contact closure.
V.32	Family of 4800/9600 bps modems operating full-duplex over two-wire facilities.
V.33	14.4 kbps modem standardized for use on point-to-point four-wire leased telephone-type circuits.
V.34	28.8/33.6 kbps modem standardization for operating full duplex over two-wire facilities.
V.35	Data transmission at 48 kbps using 60–108 kHz group band circuits. CCITT balanced interface specification for data transmission at 48 kbps, using 60–108 kHz group band circuits. Usually implemented on a 34-pin M block type connector (M 34) used to interface to a high-speed digital carrier such as DDS.
V.36	Modems for synchronous data transmission using 60–108 kHz group band circuits.
V.37	Synchronous data transmission at a data signalling rate higher than 72 kbps using 60-108 kHz group band circuits.
V.40	Error indication with electromechanical equipment.
V.41	Code-independent error control system.
V.42	Error control for modems.
V.42 bis	Data compression for use in switched network modems.
V.50	Standard limits for transmission quality of data transmission.
V.51	Organization of the maintenance of international telephone-type circuits used for data transmission.
V.52	Characteristic of distortion and error-rate measuring apparatus for data transmission.
V.53	Limits for the maintenance of telephone-type circuits used for data transmission.
V.54	Loop test devices for modems.
V.55	Specification for an impluse noise measuring instrument for telephone-type circuits.
V.56	Comparative tests of modems for use over telephone-type circuit.
V.90	56 kbps modem standardization for operating downstream and 33.6 kbps upstream over two-wire facilities.

policy third-party vendors had to design communications equipment, such as modems, that was compatible with the majority of equipment in use.

As some third-party vendor products gained market acceptance over other products, vendor licensing of technology resulted in the development of *de facto* standards. Another area responsible for the development of a large number of *de facto* standards is the Internet community. In this section we will examine both vendor and Internet *de facto* standards.

AT&T compatibility

Since AT&T originally had a monopoly on equipment connected to its network, when third-party vendors were allowed to manufacture products for use on AT&T facilities they designed most of their products to be compatible with AT&T equipment. This resulted in the operational characteristics of many AT&T products becoming *de facto* standards. In spite of the breakup of AT&T, this vendor still defines format and interface specifications for many facilities that third-party vendors must adhere to for their product to be successfully used with such facilities. AT&T, like standards organizations, publishes a variety of communications reference publications that define the operational characteristics of its facilities and equipment.

AT&T's catalog of technical documents is contained in two publications. AT&T's *Publication 10000* lists over 140 publications and includes a synopsis of their contents, date of publication and cost. Formally known as the *Catalog of Communications Technical Publications, Publication 10000* includes several order forms as well as a toll-free 800 number for persons who wish to call in their order. AT&T's *Publication 10000A*, which was issued as an addendum to *Publication 10000* lists new and revised technical reference releases as well as technical references deleted from *Publication 10000* and the reason for each deletion. In addition, *Publication 10000A* includes a supplementary list of select codes for publications listed in *Publication 10000*. The select code is the document's ordering code, which must be entered on the AT&T order form. Readers should obtain both documents from AT&T to simplify future orders.

Publication 10000 and 10000*A* and the publication listed therein can be obtained by writing or calling AT&T at the address or telephone numbers listed in Table 1.23.

Table 1.22 is an extract of some of the AT&T technical publications listed in *Publications 10000* and *10000A*. As can be seen, a wide diversity of publications can be ordered directly from AT&T.

Table 1.22 Selected AT&T publications

Publication number	Publication title
CB 142	The Extended Superframe Format Interface Specification
CB 143	Digital Access and System Technical Reference and Compatibility Specification
PUB 41449	Integrated Service Digital Network (ISDN) Primary Rate Interface Specification
PUB 41457	SKYNET Digital Service
PUB 52411	ACCUNET T1.5 Service Description and Interface Specification
PUB 54010	X.25 Interface Specification and Packet Switching Capabilities
PUB 54012	X.75 Interface Specification and Packet Switching Capability
PUB 54075	56 kbps Subrate Data Multiplexing

Cross-licensed technology

As the deregulation of the US telephone industry progressed, hundreds of vendors developed products for the resulting market. Many vendors cross-licensed technology, such as the command set which defines the operation of intelligent modems. Due to this, another area of *de facto* standards developed based upon consumer acceptance of commercial products. In certain cases, *de facto* standards have evolved into *de jure* standards with their adoption by one or more standard-making organizations.

Bellcore

A third area of *de facto* standards is Bellcore. Upon divestiture in 1984, AT&T formed Bell Communications Research, Inc. (Bellcore) with its seven regional Bell holding companies. Bellcore provides technical and research support to these holding companies in much the same way that Bell Laboratories supported AT&T. Bellcore maintains common standards for the nation's telephone systems, ensures a smoothly operating telephone network and coordinates telecommunications operations during national emergencies. With approximately 7000 employees, hundreds of research projects and an annual budget of approximately $1 billion, Bellcore is among the largest research and engineering organizations in the United States.

Like AT&T, Bellcore publishes a catalog of technical publications called *Catalog 10000*. The Bellcore catalog list approximately 500 publications that vary in scope from compatibility guides, which list the interface specifications that must

Table 1.23 Communications reference publication sources

ANSI 1430 Broadway New York, NY 10018, USA (212) 354-3300	ITU General Secretariat International Telecommunication Union Place des Nations 1211 Geneva 20, Switzerland
AT&T Customer Information Center Indianapolis, IN 46219, USA (800) 432-6600 (317) 352-8557	IEEE 345 East 47th Street New York, NY 10017, USA (212) 705-7900
Bell Communications Research Information Operations Center 60 New England Avenue Piscataway, NJ 08854, USA (201) 981-5600	US Department of Commerce National Technical Information Service 5285 Port Royal Rd. Springfield, VA 22161, USA (703) 487-4650
EIA Standard Sales 2001 Eye Street NW Washington DC 20006, USA (202) 457-4966	

be adhered to by manufacturers building equipment for connection to telephone company central offices, to a variety of technical references. *Catalog 10000* and the publications listed therein can be ordered directly from Bellcore or from any one of the seven regional Bell operating companies by mail or telephone. The address of Bell Communications Research is listed in Table 1.23.

Internet standards

The Internet can be considered as a collection of interconnected networks that use the Transmission Control Protocol/Internet Protocol (TCP/IP) suite. The Internet has its roots in experimental packet switching work sponsored by the US Department of Defense Advanced Research Projects Agency (ARPA) which resulted in the development of the ARPANet. That network was responsible for the development of file transfer, electronic mail and remote terminal access to computers which became applications incorporated into the TCP/IP protocol suite. The efforts of ARPA during the 1960s and 1970s were taken over by the Internet Activities Board (IAB) whose name was changed to the Internet Architecture Board in 1992. The IAB is responsible for the development of Internet protocols to include deciding if and when a protocol should become an Internet standard.

While the IAB is responsible for setting the general direction concerning the standardization of protocols, the actual effort is carried out by the Internet Engineering Task Force (IETF). The IETF is responsbile for the development of documents called Requests For Comments (RFCs) which are normally issued as a preliminary draft. After a period allowed for comments the RFC will be published as a proposed standard or, if circumstances warrant, it may be dropped from consideration. If favorable comments occur concerning the proposed standard it can be promoted to a draft standard after a minimum period of six months. After a review period of at least four months the draft standard can be recommended for adoption as a standard by the Internet Engineering Steering Group (IESG). The IESG consists of the chairperson of the IETF and other members of that group and performs an oversight and coordinating function for the IETF. Although the IESG must recommend the adoption of an RFC as a standard the IAB is responsible for the final decision concerning its adoption. Figure 1.32 illustrates

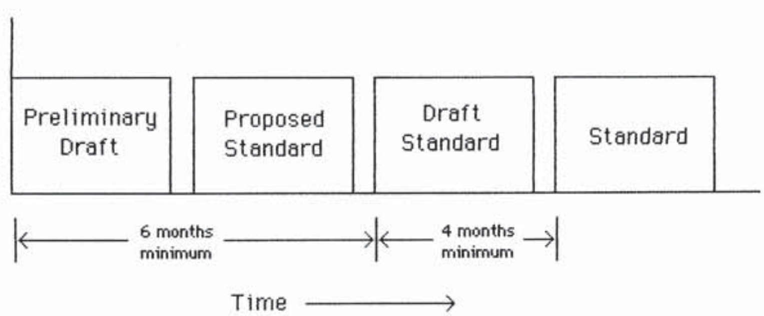


Figure 1.32 Internet standards time track

the time track for the development of an Internet standard. RFCs cover a variety of topics, ranging from TCP/IP applications to the Simple Network Management Protocol (SNMP) and the composition of databases of network management information. Over 1500 RFCs were in existence when this book was published and some of those RFCs will be described in detail later in this book.

The ISO reference model

The International Organization for Standardization (ISO) established a framework for standardizing communications systems called the Open Systems Interconnection (OSI) Reference Model. The OSI architecture defines the communications process as a set of seven layers, with specific functions isolated to and associated with each layer. Each layer, as illustrated in Figure 1.33, covers lower layer processes, effectively isolating them from higher layer functions. In this way, each layer performs a set of functions necessary to provide a set of services to the layer above it.

Layer isolation permits the characteristics of a given layer to change without impacting the remainder of the model, provided that the supporting services remain the same. One major advantage of this layered approach is that users can mix and match OSI conforming communications products to tailor their communications system to satisfy a particular networking requirement.

The OSI Reference Model, while not completely viable with current network architectures, offers the potential to directly interconnect networks based upon the use of different vendor equipment. This interconnectivity potential will be of substantial benefit to both users and vendors. For users, interconnectivity will remove the shackles that in many instances tie them to a particular vendor. For vendors, the ability to easily interconnect their products will provide them with

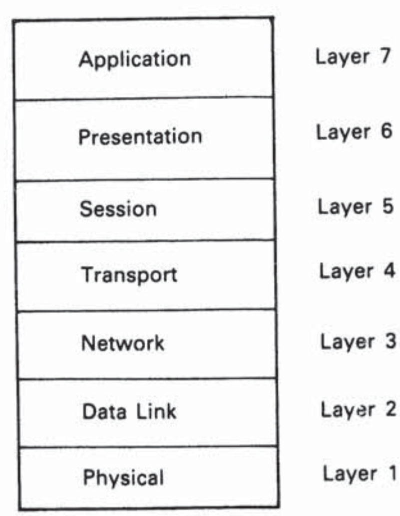


Figure 1.33 ISO reference model

access to a larger market. The importance of the OSI model is such that it has been adopted by the ITU as Recommendation X.200.

Layered architecture

As previously discussed, the OSI reference model is based upon the establishment of a layered, or partitioned, architecture. This partitioning effort can be considered as being derived from the scientific process whereby complex problems are subdivided into functional tasks that are easier to implement on an aggregate individual basis than as a whole.

As a result of the application of a partitioning approach to communications network architecture, the communications process was subdivided into seven distinct partitions, called layers. Each layer consists of a set of functions designed to provide a defined series of services which relate to the mission of that layer. For example, the functions associated with the physical connection of equipment to a network are referred to as the physical layer.

With the exception of layers 1 and 7, each layer is bounded by the layers above and below it. Layer 1, the physical layer, can be considered to be bound below by the interconnecting medium over which transmission flows, while layer 7 is the upper layer and has no upper boundary. Within each layer is a group of functions which can be viewed as providing a set of defined services to the layer which bounds it from above, resulting in layer n using the services of layer $n-1$. Thus, the design of a layered architecture enables the characteristics of a particular layer to change without affecting the rest of the system, assuming the services provided by the layer do not change.

OSI layers

An understanding of the OSI layers is best obtained by first examining a possible network structure that illustrates the components of a typical network. Figure 1.34

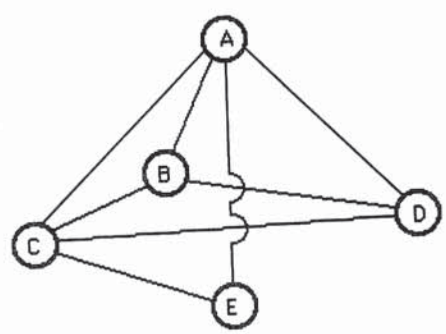


Figure 1.34 Network components. The path between a source and a destination node established on a temporary basis is called a logical connection. (O = Node. Lines represent paths)

illustrates a network structure which is only typical in the sense that it will be used for a discussion of the components upon which networks are constructed.

The circles in Figure 1.34 represents nodes which are points where data enters or exits a network or is switched between two paths. Nodes are connected to other nodes via communications paths within the network where the communications paths can be established on any type of communications media, such as cable, microwave or radio.

From a physical perspective, a node can be based upon the use of one of several types of computers to include a personal computer, minicomputer or mainframe computer or specialized computer, such as front-end processor. Connections to network nodes into a network can occur via the use of terminals directly connected to computers, terminals connected to a node via the use of one or more intermediate communications devices or via paths linking one network to another network.

The routes between two nodes, such as C–E–A, C–D–A, C–A and C–B–A which could be used to route data between nodes A and C are information paths. Due to the variability in the flow of information through a network, the shortest path between nodes may not be available for use or may represent a non-efficient path with respect to other paths constructed through intermediate nodes between a source and destination node. A temporary connection established to link two nodes whose route is based upon such parameters as current network activity is known as a logical connection. This logical connection represents the use of physical facilities to include paths and node switching capability on a temporary basis.

The major functions of each of the seven OSI layers are described in the following seven subsections.

Layer 1—the physical layer

At the lowest or most basic level, the physical layer (level 1) is a set of rules that specifies the electrical and physical connection between devices. This level specifies the cable connections and the electrical rules necessary to transfer data between devices. Typically, the physical link corresponds to established interface standards, such as RS-232. The reader is referred to Section 1.14 for detailed information concerning several physical layer interface standards.

Layer 2—the data link layer

The next layer, which is known as the data link layer (level 2), denotes how a device gains access to the medium specified in the physical layer; it also defines data formats, including the framing of data within transmitted messages, error control procedures, and other link control activities. From defining data formats to include procedures to correct transmission errors, this layer becomes responsible for the reliable delivery of information. Data link control protocols such as binary synchronous communications (BSC) and high-level data link control (HDLC) reside in this layer. The reader is referred to Section 1.15 for detailed information concerning data link control protocols and to Chapter 3 for information concerning the subdivision of that layer for local area network communications.

Layer 3—the network layer

The network layer (level 3) is responsible for arranging a logical connection between a source and destination on the network to include the selection and management of a route for the flow of information between source and destination based upon the available data paths in the network. Service provided by this layer are associated with the movement of data through a network, to include addressing, routing, switching, sequencing and flow control procedures. In a complex network the source and destination may not be directly connected by a single path, but instead require a path to be established that consists of many subpaths. Thus, routing data through the network onto the correct paths is an important feature of this layer.

Several protocols have been defined for layer 3, including the ITU X.25 packet switching protocol and the ITU X.75 gateway protocol. X.25 governs the flow of information through a packet network while X.75 governs the flow of information between packet networks. In the TCP/IP protocol suite the Internet Protocol (IP) represents a network layer protocol. Packet switching networks are described in Chapter 2.

Layer 4—the transport layer

The transport layer (level 4) is responsible for guaranteeing that the transfer of information occurs correctly after a route has been established through the network by the network level protocol. Thus, the primary function of this layer is to control the communications session between network nodes once a path has been established by the network control layer. Error control, sequence checking, and other end-to-end data reliability factors are the primary concern of this layer. Examples of transport layer protocols include the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP).

Layer 5—the session layer

The session layer (level 5) provide a set of rules for establishing and terminating data streams between nodes in a network. The services that this session layer can provide include establishing and terminating node connections, message flow control, dialog control, and end-to-end data control.

Layer 6—the presentation layer

The presentation layer (level 6) services are concerned with data transformation, formatting and syntax. One of the primary functions performed by the presentation layer is the conversion of transmitted data into a display format appropriate for a receiving device. This can include any necessary conversion between different data codes. Data encryption/decryption and data compression and decompression are examples of the data transformation that could be handled by this layer.

Layer 7—the application layer

Finally, the application layer (level 7) acts as a window through which the application gains access to all of the services provided by the model. Examples of functions performed at this level include file transfers, resource sharing and database access. While the first four layers are fairly well defined, the top three layers may vary considerably, depending upon the network used. Figure 1.35 illustrates the OSI model in schematic format, showing the various levels of the model with respect to a terminal accessing an application on a host computer system.

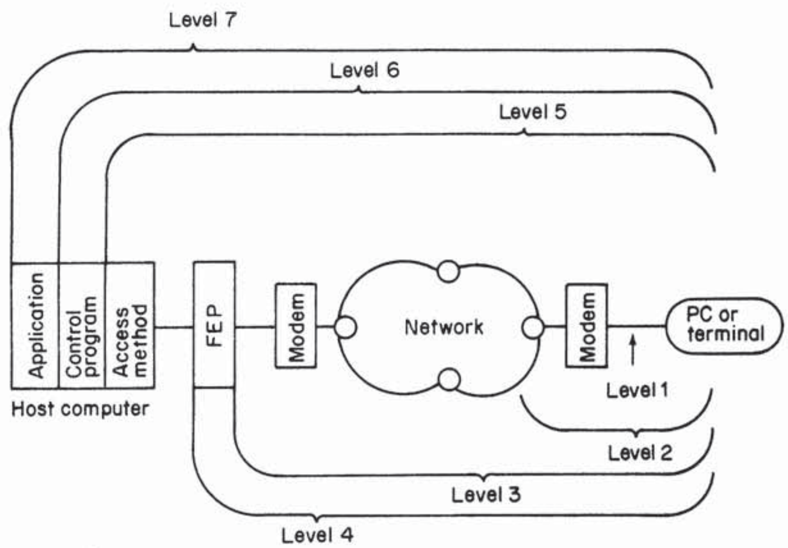


Figure 1.35 OSI model schematic

Data flow

As data flows within an ISO network each layer appends appropriate heading information to frames of information flowing within the network while removing the heading information added by a lower layer. In this manner, layer (n) interacts with layer ($n-1$) as data flows through an ISO network.

Figure 1.36 illustrates the appending and removal of frame header information as data flows through a network constructed according to the ISO reference model. Since each higher level removes the header appended by a lower level, the frame traversing the network arrives in its original form at its destination.

As the reader will surmise from the previous illustrations, the ISO reference model is designed to simplify the construction of data networks. This simplification is due to the eventual standardization of methods and procedures to append appropriate heading information to frames flowing through a network, permitting data to be routed to its appropriate destination following a uniform procedure.

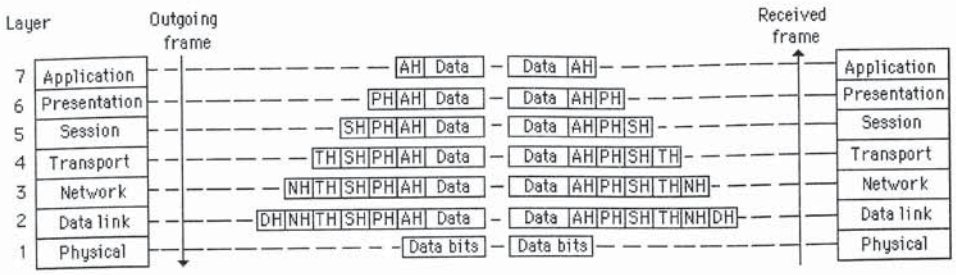


Figure 1.36 Data flow within an ISO reference model network DH, NH, TH, SH, PH and AH are appropriate headers Data Link, Network Header, Transport Header, Session Header, Presentation Header and Application Header added to data as the data flows through an ISO reference model network

1.14 THE PHYSICAL LAYER: CABLES, CONNECTORS, PLUGS AND JACKS

As discussed in Section 1.13, the physical layer is the lowest layer of the ISO reference model. This layer can be considered to represent the specifications required to satisfy the electrical and mechanical interface necessary to establish a communications path. Standards for the physical layer are concerned with connector types, connector pin-outs and electrical signaling, to include bit synchronization and the identification of each signal element as a binary one or binary zero. This results in the physical layer providing those service necessary for establishing, maintaining and disconnecting the physical circuits that form a communications path.

Since one part of communications equipment utilization involves connecting data terminal equipment (DTE) to data communications equipment (DCE), the physical interface is commonly thought of as involving such standards as RS-232, RS-449, ITU V.24 and ITU X.21. Another less recognized aspect of the physical layer is the method whereby communications equipment is attached to communications carrier facilities.

In this section we will first focus attention upon the DTE/DCE interface, examining several popular standards and emerging standards. This examination will include the signal characteristics of several interface standards to include the interchange circuits defined by the standard and their operation and utilization. Since the RS-232/V.24 interface is by far the most popularly employed physical interface, we will examine the cable used for this interface in the second part of this section. This examination will provide the foundation for illustrating the fabrication of several types of null modem cables as well as the presentation of other cabling tricks.

Since communications, in most instances, depend upon the use of facilities provided by a common carrier, in the last of this section we will discuss the interface between customer equipment and carrier facilities. In doing so we will examine the use of plugs and connectors to include the purpose of different types of jacks.

DTE/DCE interfaces

In the world of data communications, equipment that includes terminals and computer ports is referred to as data terminal equipment or DTEs. In comparison, modems and other communications devices are referred to as data communications equipment or DCEs. The physical, electrical and logical rules for the exchange of data between DTEs and DCEs are specified by an interface standard; the most commonly used are the EIA RS-232-C and RS-232-D standards which are very similar to the ITU V.24 standard used in Europe and other locations outside North America.

The EIA refers to the Electronic Industries Association which is a national body that represents a large percentage of the manufacturers in the US electronics industry. The EIA's work in the area of standards has become widely recognized and many of its standards have been adopted by other standards bodies. RS-232-C is a recommended standard (RS) published by the EIA in 1969, with the number 232 referring to the identification number of one particular communications standard and the suffix C designating the revision to that standard.

In the late 1970s it was intended that the RS-232-C standard would be gradually replaced by a set of three standards—RS-449, RS-422 and RS-423. These standards were designed to permit higher data rates than are obtainable under RS-232-C as well as to provide users with added functionality. Although the EIA and several government agencies heavily promoted the RS-449 standard, its adoption by vendors has been limited. Recognizing the fact that the universal adoption of RS-449 and its associated standards was basically impossible, the EIA issued RS-232-D (Revision D) in January 1987 and RS-232-E (Revision E) in July 1991, as well as a new standard known as RS-530.

Four other DTE/DCE interfaces that warrant attention are the EIA RS-366-A and the ITU X.20, X.21, and V.35 standards. The RS-366-A interface governs the attachment of DTEs to a special type of DCE called an automatic calling unit. The ITU X.20 and X.21 standards govern the attachment of DTE to DCE for asynchronous and synchronous operation on public networks, respectively. The ITU V.35 standard governs high-speed data transmission, typically at 48 kbps and above, with a limit occurring at approximately 6 Mbps.

Two emerging standards we will also examine are the High Speed Serial Interface (HSSI) and the High Performance Parallel Interface (HIPPI). HSSI provides support for serial operating rates up to 52 Mbps, while for extremely high bandwidth requirements, such as extending the channel on a supercomputer, HIPPI supports maximum data rates of either 800 Mbps or 1.6 Gbps on a parallel interface.

Figure 1.37 compares the maximum operating rate of the six interfaces we will discuss in this section. Although RS-232 is 'officially' limited to approximately 20 kbps, that limit is for a maximum cable length of 50 feet, which explains why that interface can still be used to connect data terminal equipment devices to include personal computers to modems operating at data rates up to 28.8, 33.6, or even 56 kbps. That is, since pulse distortion is proportional to the cable distance between two devices, shortening the cable enables a higher speed data transfer capability to be obtained between a PC and a modem. However, when the

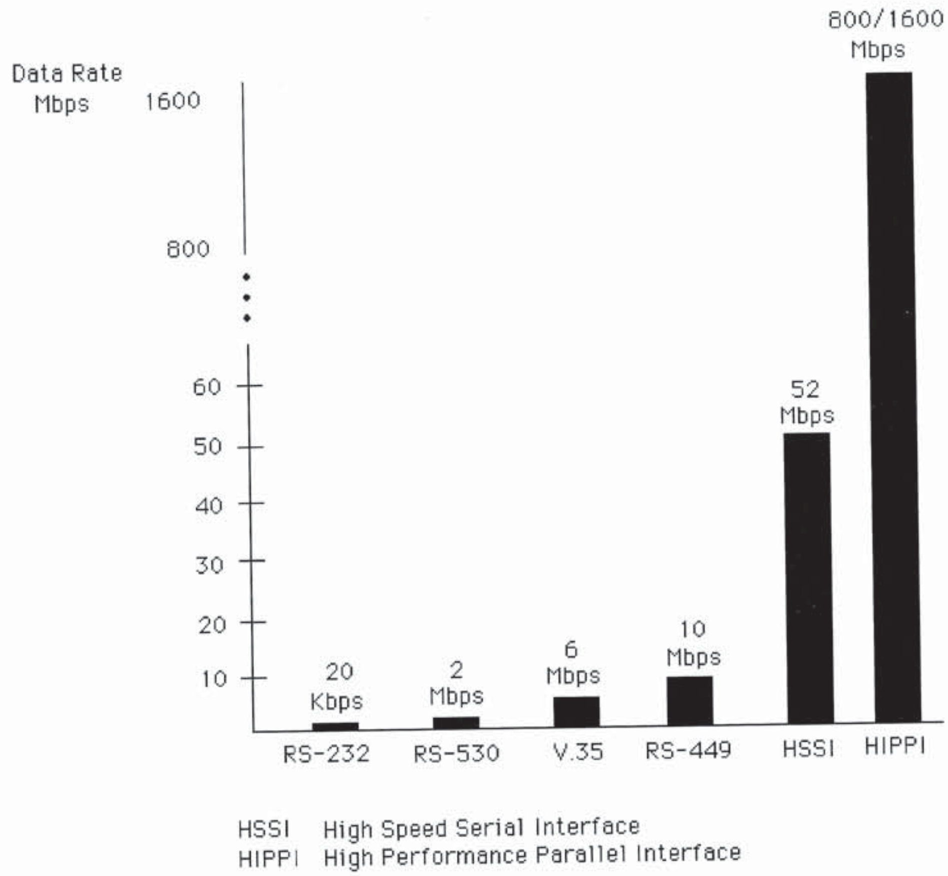


Figure 1.37 Maximum interface data rates

compression feature built into modems is enabled, most vendors recommend you set the interface speed between the DTE and DCE to four times the modem operating rate to permit the modem to perform compression effectively. Unfortunately, even at very short cable lengths the RS-232 interface will lose data at an operating rate above 115.2 kbps. Rather than incorporate a V.35 or another less commonly available interface into their modem, many modem manufacturers currently include a Centronics parallel interface in their products to support a data transfer rate above 115.2 kbps. Since just about every personal computer has a parallel interface, it is readily available or an additional port can be installed at a nominal cost in comparison to the cost associated with obtaining a different interface.

Connector overview

RS-232-D and the ITU V.24 standard as well as RS-530 formally specify the use of a D-shaped 25-pin interface connector similar to the connector illustrated in

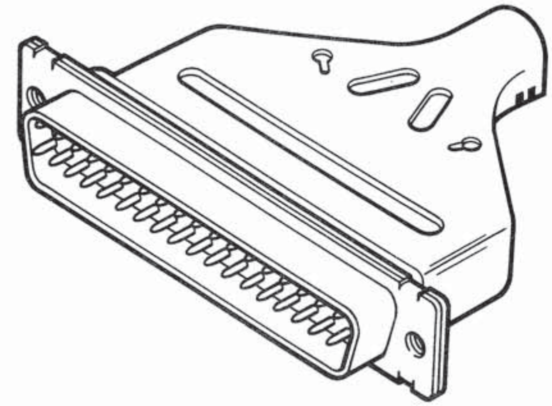


Figure 1.38 The D connector

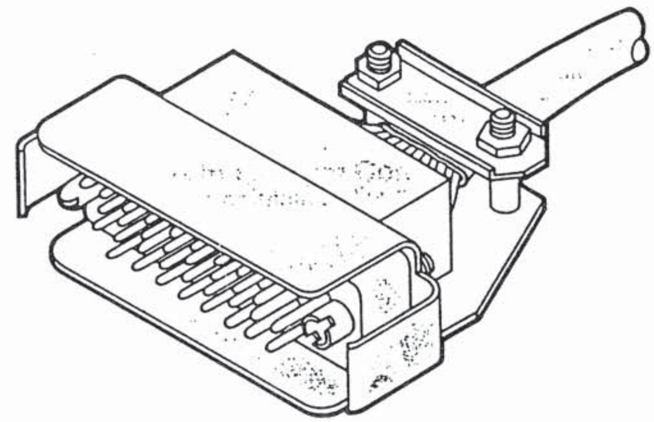


Figure 1.39 The V.35 M series connector

Figure 1.38. A cable containing up to 25 individual conductors is fastened to the narrow part of the connector, while the individual conductors are soldered to predefined pin connections inside the connector.

ITU X.20 and X.21 compatible equipment use a 15-position subminiature D connector that is both smaller and has 10 fewer predefined pins than the 25-position D connector illustrated in Figure 1.38. ITU V.35 compatible equipment uses a 34-position 'M' series connector similar to the connector illustrated in Figure 1.39. RS-449 compatible equipment uses a 37-position D connector and may optionally use a 9-position D connector, while RS-366 compatible equipment uses a 25-position D connector similar to the connector illustrated in Figure 1.38. Although the HSSI connector has 50 pins, it is actually smaller than the 32-position V.35 connector. Another interesting feature of HSSI connectors is their 'genders'. The cable connectors are specified as male, while both DTE and DCE connectors are specified as receptacles. This specification minimizes the need for male–male and female–female adapters commonly required to mate equipment and cables based upon the use of other interface standards.

In comparison to RS-232-D and RS-530, the RS-232-C standard only referenced the connector in an appendix and stated that it was not part of the standard. In spite of this omission, the use of a 25-pin D-shaped connector with RS-232-C is considered as a *de facto* standard. Although a *de facto* standard, many RS-232-C devices, in fact, use other types of connectors. Perhaps the most common exception to the 25-pin connector resulted from the manufacture of a serial/parallel adapter card for use in the IBM PC AT and compatible personal computers, such as the Compaq Deskpro and AST Bravo. The serial RS-232-C port on that card uses a 9-pin connector, which resulted in the development of a viable market for 9-pin to 25-pin converters consisting of a 9-pin and 25-pin connector on opposite ends of a short cable whose interchange circuits are routed in a specific manner to provide a required level of compatibility.

The major differences between RS-232-D and RS-232-C are that the new revision supports the testing of both local and remote communications equipment by the addition of signals to support this function and modified the use of the protective ground conductor to provide a shielding capability.

A more recent revision, RS-232-E, added a specification for a smaller alternative 26-pin connector and slightly modified the functionality of a few of the interface pins. Since RS-232-C and RS-232-D are by far the most commonly supported serial interfaces, we will first focus our attention upon those interfaces. Once this is accomplished we will discuss the newest revision to this popular interface, RS-232-E.

RS-232-C/D

Since the use of RS-232-C is basically universal since its publication by the EIA in 1969, we will examine both revisions C and D in this section, denoting the differences between the revisions when appropriate. When both revisions are similar, we will refer to them as RS-232. In general, devices built to either standard as well as the equivalent ITU V.24 recommendation are compatible with one another. There are, however, some slight differences that can occur due to the addition of signals to support modem testing under RS-232-D.

Since the RS-232-C/D standards define the most popular method of interfacing between DTEs and DCEs in the United States, they govern, as an example, the interconnection of most terminal devices to stand-alone modems. The RS-232-C/D standards apply to serial data transfers between a DTE and DCE in the range from 0 to 19 200 bps. Although the standards also limit the cable length between the DTE and DCE to 50 feet, since the pulse width of digital data is inversely proportional to the data rate, you can normally exceed this 50-foot limitation at lower data rates as wider pulses are less susceptible to distortion than narrower pulses. When a cable length in excess of 50 feet is required, it is highly recommended that low capacitance shielded cable be used and tested prior to going on-line, to ensure that the signal quality is acceptable. This type of cable is discussed later in this section.

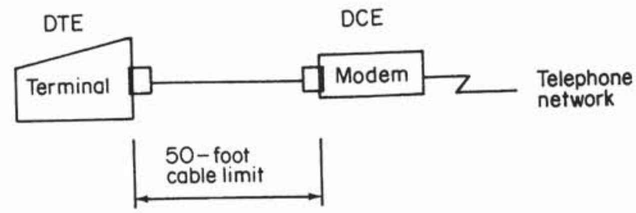


Figure 1.40 The RS-232 physical interface standard cables are typically 6, 10, or 12 feet in length with 'male' connectors on each end

Another part of the RS-232-D standard specifies the cable heads that serve as connectors to the DTEs and DCEs. Here the connector is known as a DB-25 connector and each end of the cable is equipped with this 'male' connector that is designed to be inserted into the DB-25 'female' connectors normally built into modems. Figure 1.40 illustrates the RS-232 interface between a terminal and a stand-alone modem.

Signal characteristics

The RS-232 interface specifies 25 interchange circuits or conductors that govern the data flow between the DTE and DCE. Although you can purchase a 25-conductor cable, normally fewer conductors are required. For asynchronous transmission, normally 9 to 12 conductors are required, while synchronous transmission typically requires 12 to 16 conductors, with the number of conductors required a function of the operational characteristics of the devices to be connected to one another.

The signal on each of the interchange circuits is based upon a predefined voltage transition occurring as illustrated in Figure 1.41. A signal is considered to be ON when the voltage (V) on the interchange circuit is between $+3$ V and $+15$ V. In comparison, a voltage between -3 V and -15 V causes the interchange circuit to be placed in the OFF condition. The voltage range from $+3$ V to -3 V is a transition region that has no effect upon the condition of the circuit.

Although the RS-232 and V.24 standards are similar to one another, the latter differs with respect to the actual electrical specification of the interface. The ITU V.24 recommendation is primarily concerned with how the interchange circuits operate. Thus, another recommendation, known as V.28, actually defines the electrical specifications of the interface that can be used with the V.24 standard.

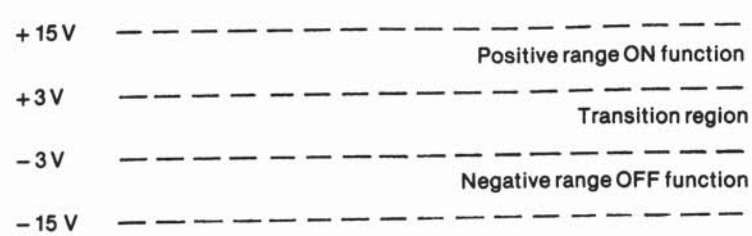


Figure 1.41 Interchange circuit voltage ranges

Table 1.24 Interchange circuit comparison

	Interchange circuit voltage	
	Negative	Positive
Binary state	1	0
Signal condition	Mark	Space
Function	OFF	ON

Table 1.24 provides a comparison between the interchange circuit voltage, its binary state, signal condition and function. As a binary 1 is normally represented by a positive voltage with a terminal device, this means that data signals are effectively inverted for transmission.

Since the physical implementation of the RS-232 standard is based upon the conductors used to interface a DTE to a DCE, we will examine the functions of each of the interchange circuits. Prior to discussing these circuits, an explanation of RS-232 terminology is warranted since there are three ways you can refer to the circuits in this interface.

Circuit/conductor reference

The most commonly used method to refer to the RS-232 circuits is by specifying the number of the pin in the connector which the circuit uses. A second method used to refer to the RS-232 circuits is by the two- or three-letter designation used by the standards to label the circuits. The first letter in the designator is used to group the circuits into one of six circuit categories as indicated by the second column labeled 'interchange circuit' in Figure 1.42. As an example of the use of this method, the two ground circuits have the letter A as the first letter in the circuit designator the signal ground circuit is called 'AB', since it is the second circuit in the 'A' ground category. Since these designators are rather cryptic, they are not commonly used.

A third method used is to describe the circuits by their functions. Thus, pin 2, which is the transmit data circuit, can be easily referenced as transmit data. Many persons have created acronyms for the descriptions which are easier to remember than the RS-232 pin number or interchange circuit designator. For example, transmit data is referred to as 'TD', which is easier to remember than any of the RS-232 designators previously discussed.

Although the list of circuits in Figure 1.42 may appear overwhelming at first glance, in most instances only a subset of the 25 conductors are employed. To better understand this interface standard, we will first examine those interchange circuits required to connect an asynchronously operated terminal device to an asynchronous modem. Then we can expand upon our knowledge of these interchange circuits by examining the functions of the remaining circuits, to include those additional circuits that would be used to connect a synchronously operated terminal to a synchronous modem.

PIN Number	Interchange circuit	ITU equivalent	Description	Gnd	Data		Control		Timing		Testing	
					From DCE	To DCE	From DCE	To DCE	From DCE	To DCE	From DCE	To DCE
1	AA	101	Protective Ground (Shield)	X								
7	AB	102	Signal Ground/Common Return	X								
2	BA	103	Transmitted Data			X						
3	BB	104	Received Data		X							
4	CA	105	Request to Send					X				
5	CB	106	Clear to Send				X					
6	CC	107	Data Set Ready (DCE Ready)				X					
20	CD	108.2	Data Terminal Ready (DTE Ready)					X				
22	CE	125	Ring Indicator				X					
8	CF	109	Received Line Signal Detector				X					
21	(RL)/ CG	110	(Remote Loopback)/Signal Quality Detector				X					
23	CH	111	Data Signal Rate Selector (DTE)					X				
23	CI	112	Data Signal Rate Selector (DCE)				X					
24	DA	113	Transmitter Signal Element Timing (DTE)							X		
15	DB	114	Transmitter Signal Element Timing (DCE)						X			
17	DD	115	Receiver Signal Element Timing (DCE)						X			
14	SBA	118	Secondary Transmitted Data			X						
16	SBB	119	Secondary Received Data		X							
19	SCA	120	Secondary Request to Send					X				
13	SCB	121	Secondary Clear to Send				X					
12	SCF	122	Secondary Received Line Signal Detector				X					
8	—	—	Reserved for Testing									X
9	—	—	Reserved for Testing								X	
18	(LL)		(Local Loopback)									X
25	(TM)		(Test Mode)								X	

Figure 1.42 RS-232-C/D and ITU V.24 interchange circuit by category, RS-232-D additions/changes to RS-232-C are indicated in parentheses

Asynchronous operations

Figure 1.43 illustrates the general signals that are required to connect an asynchronous terminal device to a asynchronous modem. Note that although a 25-conductor cable can be used to cable the terminal to the modem, only 10 conductors are actually required.

By reading the modem vendor's specification sheet you can easily determine the number of conductors required to cable DTEs to DCEs. Although most cables have straightthrough conductors, in certain instances the conductor pins at one end of a cable may require reversal or two conductors may be connected onto a common pin, a process called strapping. In fact, many times only one conductor will be used for both protective ground and signal ground, with the common conductor cabled to pins 1 and 7 at both ends of the cable. In such instances a 9-conductor cable could be employed to satisfy the cabling requirement illustrated in Figure 1.43. With this in mind, let us review the functions of the 10 circuits illustrated in Figure 1.43.

Protective ground (GND, Pin 1)

This interchange circuit is normally electrically bonded to the equipment's frame. In some instances, it can be further connected to external grounds as required by

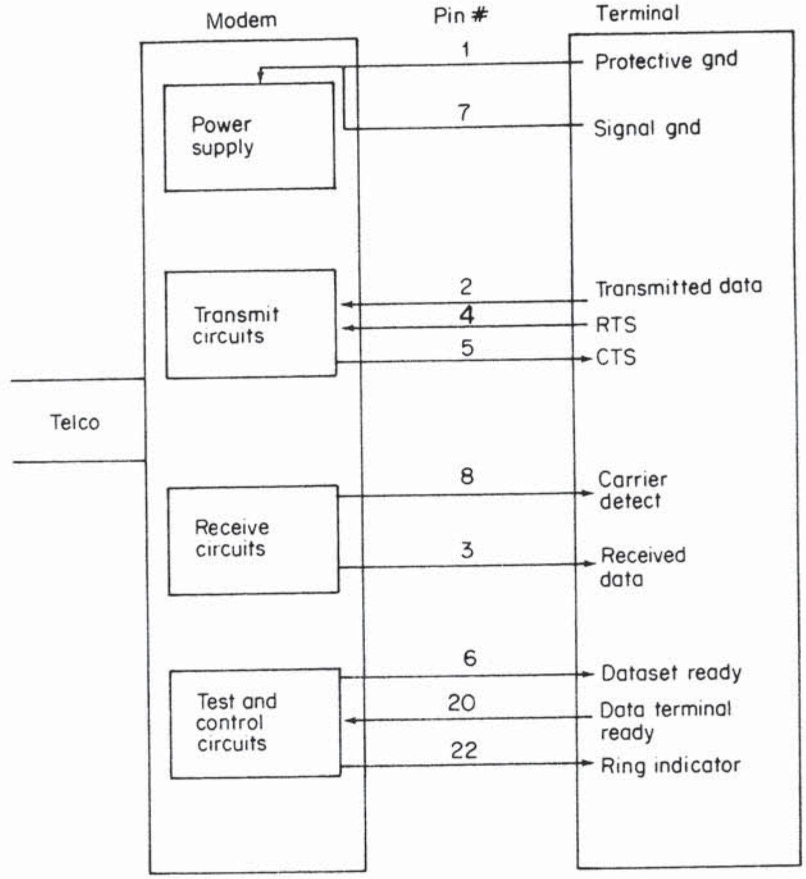


Figure 1.43 DTE-DCE interface example. In this example, pin 7 can be tied to pin 1, resulting in the use of a common 9-conductor cable

applicable regulations. Under RS-232-D this conductor use is modified to provide shielding for protection against electromagnetic or other interference occurring in high-noise environments.

Signal ground (SG, Pin 7)

This circuit must be included in all RS-232 interfaces as it establishes a ground reference for all other lines. The voltage on this circuit is set to zero to provide a reference for all other signals. Although the conductors for pins 1 and 7 can be independent of one another, typical practice is to 'strap' pin 7 to pin 1 at the modem. This is known as tying signal ground to frame ground. Since RS-232 uses a single ground reference circuit the standard results in what is known as an electrically unbalanced interface. In comparison, RS-422 uses differential signaling in which information is conveyed by the relative voltage levels in two conductors, enhancing the data rate and distance for that standard.

Transmitted data (TD, Pin 2)

The signals on this circuit are transmitted from data terminal equipment to data communications equipment or, as illustrated in Figure 1.43, a terminal device to the modem. When no data is being transmitted the terminal maintains this circuit in a marking or logical 1 condition. This is the circuit over which the actual serial bit stream of data flows from the terminal device to the modem where it is modulated for transmission.

Receive data (RD, Pin 3)

The receive data circuit is used by the DCE to transfer data to the DTE. Thus, after data is demodulated by a modem, it is transferred to the attached terminal over this interchange circuit. When the modem is not sending data to the terminal, this circuit is held in the marking condition.

Request to send (RTS, Pin 4)

The signal on this circuit is sent from the DTE (terminal) to the DCE (modem) to prepare the modem for transmission. Prior to actually sending data, the terminal must receive a clear to send signal from the modem on pin 5.

Clear to send (CTS, Pin 5)

This interchange circuit is used by the DCE (modem) to send a signal to the attached DTE (terminal); indicating that the modem is ready to transmit. By turning this circuit OFF, the modem informs the terminal that it is not ready to receive data. The modem raises the CTS signal after the terminal initiates a request to send (RTS) signal.

Carrier detect (CD, Pin 8)

Commonly referred to as received line signal detector (RLSD), a signal on this circuit is used to indicate to the DTE (terminal) that the DCE (modem) is receiving a carrier signal from a remote modem. The presence of this signal is also used to illuminate the carrier detect light-emiting diode (LED) indicator on modems equipped with that display indicator. If this light indicator should go out during a communications session, it indicates that the session has terminated owing to a loss of carrier, and software that samples for this condition will display the message 'carrier lost' or a similar message to indicate this condition has occurred.

Data set ready (DSR, Pin 6)

Signals on this interchange circuit flow from the DCE to the DTE and are used to indicate the status of the data set connected to the terminal. When this circuit is in the ON (logic 0 as in 1, 2, 3) condition, it serves as a signal to the terminal that the modem is connected to the telephone line and is ready to transmit data. Since the RS-232 standard specifies that the DSR signal is ON when the modem is connected to the communications channel and not in any test condition, a modem using a self-testing feature or automatic dialing capability would pass this signal to the terminal after the self-test is completed or after the telephone number of a remote location was successfully dialed. Under RS-232-D this signal was renamed DCE ready.

Data terminal ready (DTR, Pin 20)

The signal on this circuit flow from the DTE to the DCE and is used to control the modem's connection to the telephone line. An ON condition on this circuit prepares the modem to be connected to the telephone line, after which the connection can be established by manual or automatic dialing. If the signal on this circuit is placed in an OFF condition, it causes the modem to drop any telephone connection in progress, providing a mechanism for the terminal device to control the line connection. Under RS-232-D this signal was renamed DTE ready.

Ring indicator (RI, Pin 22)

The signal on this interchange circuit flows from the DCE to the DCE and indicates to the terminal device that a ringing signal is being received on the communications channel. This circuit is used by an auto-answer modem to 'wake-up' the attached terminal device. Since a telephone rings for 1 s and then pauses for 4 s prior to ringing again, this line becomes active for 1 s every 5 s when an incoming call occurs.

The ring indicator circuit is turned ON by the DCE when it detects the ON phase of a ring cycle. Depending upon how the DCE is optioned, it may either keep the RI signal high until the DTE turns DTR low or the DCE may turn the RI signal ON and OFF to correspond to the telephone ring sequence.

If the DTE is ready to accept the call its DTR lead will either by high, which is known as a hot-DTR state, or be placed into on ON condition in response to the RI signal turning ON. Once the RI and DTR circuits are both ON, the DCE will actually answer the incoming call and place a carrier tone onto the line.

If a computer port connected to a modem is not in a hot-DTR state the first ring causes the modem to turn ON its RI circuit momentarily, alerting the computer port to the incoming call. As the computer port turns on its DTR circuit the modem's RI circuit will be turned off as it cycles in tandem with the telephone company ringing signal. Thus, the DCE must wait for the next ON phase of the ring cycle to answer the call, explaining why many modems may require two rings to answer a call.

Control signal timing relationship

The actual relationship of RS-232 control signals varies by time based upon the operational characteristics of devices connected as well as the strapping option

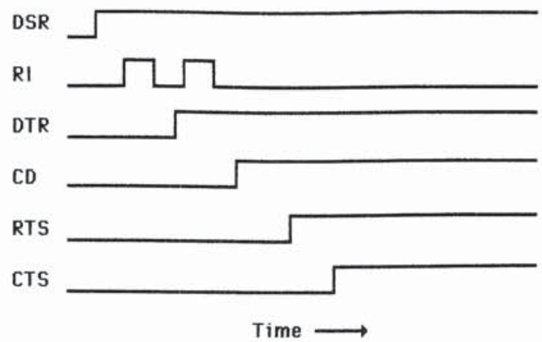


Figure 1.44 Control signal timing relationship. The state of the control signals varies by time based upon the operational characteristics of devices connected as well as the strapping option settings of those devices

settings of those devices. Figure 1.44 illustrates a common timing relationship of control signals between a computer port and a modem.

At the top of Figure 1.44 it is assumed that the data set is powered on, resulting in the data set ready (DSR) control signal being high or in the ON state. Next, two ring indicator (RI) signals are passed to the computer port, resulting in the computer responding by raising its data terminal ready (DTR) signal. The DTR signal in conjunction with the second ring indicator (RI) signal results in the modem answering the call, presenting the carrier detect (CD) signal to the computer port. Assuming the computer is programmed to transmit a sign-on message, it will raise its request to send (RTS) signal. The modem will respond by raising its clear to send (CTS) signal if it is ready to transmit, which enables the computer port to begin the actual transmission of data.

Synchronous operations

One major difference between asynchronous and synchronous modems is the timing signals required for synchronous transmission.

Timing signals

When a synchronous modem is used, it puts out a square wave on pin 15 at a frequency equal to the modem's bit rate. This timing signal serves as a clock from which the terminal would synchronize its transmission of data onto pin 2 to the modem. Thus, pin 15 is referred to as transmit clock as well as its formal designator of transmission signal element timing (DCE), with DCE referencing the fact that the communications device supplies the timing.

Whenever a synchronous modem receives a signal from the telephone line it puts out a square wave on pin 17 to the terminal at a frequency equal to the modem's bit rate, while the actual data is passed to the terminal on pin 3. Since pin 17 provides receiver clocking, it is known as 'receive clock' as well as its more formal designator of receiver signal element timing.

In certain cases a terminal device such as a computer port can provide timing signals to the DCE. In such situations the DTE will provide a clocking signal to the DCE on pin 24 while the formal designator of transmitter signal element timing (DTE) is used to reference this signal.

The process whereby a synchronous modem or any other type of synchronous device generates timing is known as internal timing. If a synchronous modem or any other type of synchronous DCE is configured to receive timing from an attached DTE, such as a computer port or terminal, the DCE must be set to external timing when the DTE is set to internal timing.

Intelligent operations

There are three interchange circuits that can be employed to change the operation of the attached communications device. One circuit can be used to first determine that a deterioration in the quality of a circuit has occurred, while the other two circuits can be employed to change the transmission rate to reflect the circuit quality.

Signal quality detector (CG, Pin 21)

Signals on this circuit are transmitted from the DCE (modem) to the attached DTE (terminal) whenever there is a high probability of an error in the received data due to the quality of the circuit falling below a predefined level. This circuit is maintained in an ON condition when the signal quality is acceptable and turned to an OFF condition when there is a high probability of an error. Under RS-232-D this circuit can also be used to indicate that a remote loopback is in effect.

Data signal rate selector (CH/C1, Pin 23)

When an intelligent terminal device such as a computer port receives an OFF condition on pin 21 for a predefined period of time, it may be programmed to change the data rate of the attached modem, assuming that the modem is capable of operating at dual data rates. This can be accomplished by the terminal device providing an ON condition on pin 23 to select the higher data signaling rate or range of rates while an OFF condition would select the lower data signaling rate or range of rates. When the data terminal equipment selects the operating rate the signal on pin 23 flows from the DTE to the DCE and the circuit is known as circuit CH. If the data communications equipment is used to select the data rate of the terminal device, the signal on pin 23 flows from the DCE to the DTE and the circuit is known as circuit CI.

Secondary circuits

In certain instances a synchronous modem will be designed with the capability to transmit data on a secondary channel simultaneously with transmission occurring on the primary channel. In such cases the data rate of the secondary channel is normally a fraction of the data rate of the primary channel.

To control the data flow on the secondary channel RS-232 standards employ five interchange circuits. Pins 14 and 16 are equivalent to the circuits on pins 2 and 3, except that they are used to transmit and receive data on the secondary channel. Similarly, pins 19, 13 and 12 perform the same functions as pins 4, 5 and 8 used for controlling the flow of information on the primary data channel.

In comparing the interchange circuits previously described to the connector illustrated in Figure 1.38, the reader should note that the location of each interchange circuit is explicitly defined by the pin number assigned to the circuit. In fact, the RS-232-D connector is designed with two rows of pins, with the top row containing 13 while the bottom row contains 12. Each pin has an explicit signal designation that corresponds to a numbering assignment that goes from left to right across the top row and then left to right across the bottom row of the connector. For ease of illustration the assignment of the interchange circuits to each of the pins in the D connector is presented in Figure 1.45 by rotating the connector 90° clockwise. In this illustration, RS-232-D conductor changes from RS-232-C are denoted in parentheses.

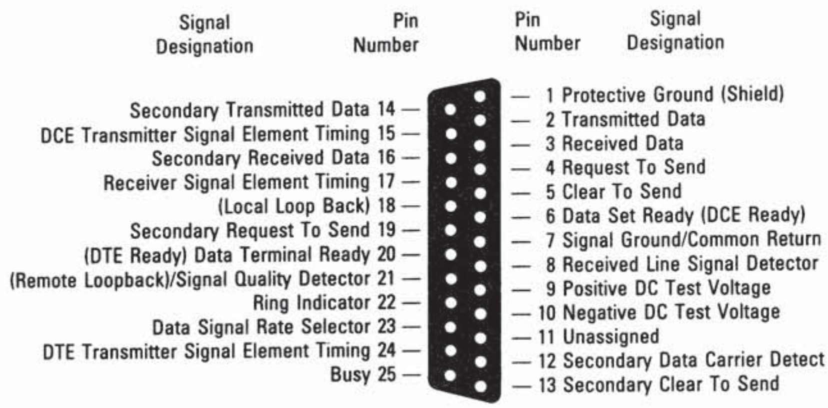


Figure 1.45 RS-232 interface on D connector

Test circuits

RS-232-D adds three circuits for testing that were not part of the earlier RS-232-C standard. The DTE can request the DCE to enter remote loopback (RL, pin 21) or local loopback (LL, pin 18) mode by placing either circuit in the ON condition. The DCE, if built to comply with RS-232-D, will respond by turning the test mode (TM, pin 25) circuit ON and performing the appropriate test.

Connector conversion

Table 1.25 contains a list of the corresponding pins between a DB-9 connector used on an IBM PC AT Compaq Deskpro and other personal computer serial ports and a standard RS-232 DB-25 connector. Data in this table can be used to develop an appropriate DB-9 to DB-25 converter. As an example of the use of

Table 1.25 DB-9 to DB-25 pin correspondence

DB-9		DB-25
1	Carrier detect	8
2	Receive data	3
3	Transmitted data	2
4	Data terminal ready	20
5	Signal ground	7
6	Data set ready	6
7	Request to send	4
8	Clear to send	5
9	Ring indicator	22

Table 1.25, the conductor for carrier detect would be wired to connect pin 1 at the DB-9 connector to pin 8 at the DB-25 connector.

RS-232-E

A more recent revision to RS-232, RS-232-E (Revision E), resulted in several minor changes to the operation of some interface circuits and the specification of an alternate interface connector (Alt A). Although none of the changes were designed to create compatibility problems with prior versions of RS-232, the use of the alternative physical interface can only be accomplished if a Revision E device is cabled via an adapter to an earlier version of that interface or if a dual connector cable is used. The 26-pin Alt A connector is about half the size of the 25-pin version and was designed to support hardware where connector space is at a premium, such as laptop and notebook computers. Pin 26 is only contained on the Alt A connector and presently is functionless.

In addition to specifying an alternative interface connector RS-232-E slightly modified the functionality of certain pins or interchange circuits. First, pin 4 (Request to Send) is defined as Ready for Receiving when the DTE enables that circuit. Next, pin 18 which was used for Local Loopback will now generate a 'Busy Out' when enabled. A third modification is the use of Clear to Send for hardware flow control, a function used by countless vendors over the past 10 years but only now formally recognized by the standard. The term 'flow control' represents the orderly control of the flow of data. By toggling the state of the Clear to Send signal, DCE equipment can regulate the flow of information from DTE equipment, a topic we will discuss in detail later in this book.

RS-232/V.24 limitations

There are several limitations associated with the RS-232 standard and the V.28 recommendation which defines the electrical specification of the interface that can be used with the V.24 standard. Foremost among these limitations are data rate and cabling distance.

RS-232/V.24 is limited to a maximum data rate of 19.2 kbps at a cabling distance of 50 feet. In actuality, speeds below 19.2 kbps allow greater transmission distances while for cable lengths of only a few feet a data rate over 100 kbps becomes possible.

Differential signaling

Over long cabling distances the cumulative cable capacitance and resistance combine to cause a significant amount of signal distortion. At some cabling distance, which decreases in an inverse relationship to the data rate, the signal cannot be recognized. Thus, to overcome the cabling distance and speed limitations associated with RS-232 a different method of signaling was devised. This signaling technique, known as differential signaling, results in information being conveyed by the relative voltage levels in two wires. Instead of using one driver to produce a large voltage swing as RS-232 does, differential signaling uses two drivers to split a signal into two parts.

Figure 1.46 illustrates differential signaling as specified by the RS-422 interface standard. To transmit a logical '1' the A driver output is driven more positive while the B driver output is more negative. Similarly, to transmit a logical '0' the A driver output is driven more negative while the B driver output is driven more positive. At the receiver a comparator is used to examine the relative voltage levels on the two signal wires.

With the use of two wires, RS-449 specifies a mark or space by the difference between the voltages on the two wires. This difference is only 0.4 V under RS-422, whereas it is 6 V (+3 V and −3 V) under RS-232/V.24. Thus, if the difference signal between the two wires is positive and greater than 0.2 V, the receiver will read a mark. Similarly, if the difference signal is negative and more negative than −0.2 V, the receiver will read a space. In addition to requiring a lower voltage

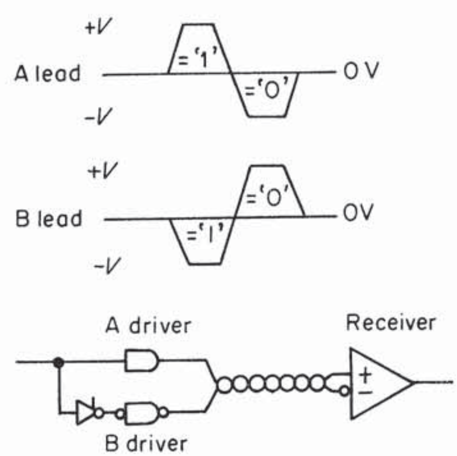


Figure 1.46 Differential signaling. The RS-422 specifies balanced differential signaling since the sum of the currents in the differential signaling wires is zero

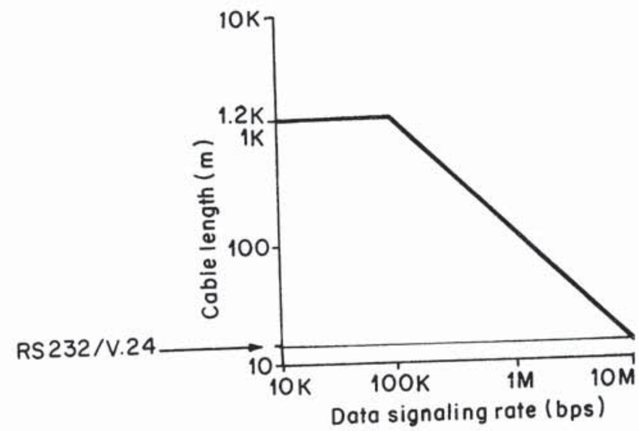


Figure 1.47 RS-422 cable distance plotted against signaling rate

shift, source and load impedance of RS-232, which is approximately 5 kΩ, is reduced to 100 Ω by the use of differential signaling. Another benefit of this signaling method is the effect of noise on signal distortion. Since the presence of noise results in the same voltage being imposed on both wires, there is no change in the voltage difference between the signal wires. Thus, the combination of lower signaling levels and impedances coupled with voltage difference immunity to noise permits differential signaling to drive longer cable distances at higher speeds. In Figure 1.47 you will find a plot of cable distances against signaling rate for the RS-422 standard.

For comparison purposes, the RS-232/V.24 cable distance is plotted against the signaling rate in Figure 1.47. As indicated, RS-422 offers significant advantages over RS-232 with respect to both cabling distance and data signaling rate.

RS-449

RS-449 was introduced in 1977 as an eventual replacement for RS-232-C. This interface specification calls for the use of a 37-pin connector as well as an optional 9-pin connector for devices using a secondary channel. Unlike RS-232, RS-449 does not specify voltage levels. Two additional specifications known as RS-442-A and RS-423 cover voltage levels for a specific range of data speeds. RS-442-A and its counterpart, ITU X.27 (V.11), define the voltage levels for data rates from 20 kbps to 10 Mbps, while RS-423-A and its ITU counterpart, X.26 (V.10), define the voltage levels for data rates between 0 and 20 kbps.

As previously mentioned, RS-442 (as well as its ITU counterpart) defines the use of differential balanced signaling. RS-422 is designed for twisted-pair telephone wire transmission ranging from 10 Mbps at distances up to 40 feet to 100 kbps at distances up to 4000 feet. RS-423 defines the use of unbalanced signaling similar to RS-232. This standard supports data rates ranging from 100 kbps at distances up to 40 feet to 10 kbps at distances up to 200 feet.

RS-232 Designation	Circuit mnemonic	Circuit name	Circuit direction	Circuit type	
Signal Ground	SG	Signal Ground	—	Common	
—	SC	Send Common	to DCE		
—	RC	Receive Common	from DCE		
—	IS	Terminal in Service	to DCE	Control	
Ring Indicator	IC	Incoming Call	from DCE		
Data Terminal Ready	TR	Terminal Ready	to DCE		
Data Set Ready	DM	Data Mode	from DCE		
Transmit Data	SD	Send Data	to DCE	Data	Primary channel
Receive Data	RD	Receive Data	from DCE		
Transmit Timing (DTE)	TT	Terminal Timing	to DCE	Timing	
Transmit Timing (DCE)	ST	Send Timing	from DCE		
Receive Timing	RT	Receive Timing	from DCE		
Request to Send	RS	Request to Send	to DCE	Control	
Clear to Send	CS	Clear to Send	from DCE		
Receive Signal Detector	RR	Receiver Ready	from DCE		
Signal Quality Detector	SO	Signal Quality	from DCE		
—	NS	New Signal	to DCE		
—	SF	Select Frequency	to DCE		
Data Rate Selector (DTE)	SR	Signaling Rate Selector	to DCE		
Data Rate Selector (DCE)	SI	Signaling Rate Indicator	from DCE		
Secondary Transmit Data	SSD	Secondary Send Data	to DCE	Data	Secondary channel
Secondary Receive Data	SRD	Secondary Receive Data	from DCE		
Secondary Request to Send	SRS	Secondary Request to Send	to DCE	Control	
Secondary Clear to Send	SCS	Secondary Clear to Send	from DCE		
Secondary Receiver Signal Detector	SRR	Secondary Receiver Ready	from DCE		
Local Loopback (D/E)	LL	Local Loopback	to DCE	Control	
—	RL	Remote Loopback	to DCE		
—	TM	Test Mode	from DCE		
—	SS	Select Standby	to DCE	Control	
—	SB	Standby Indicator	from DCE		

Figure 1.48 RS-449 interchange circuits

The use of RS-422, RS-423, and RS-449 permits the cable distance between DTEs and DCEs to be extended to 4000 feet in comparison to RS-232's 50-foot limitation. Figure 1.48 indicates the RS-449 interchange circuits. In comparing RS-449 to RS-232, you will note the addition of 10 circuits which are either new control or status indicators and the deletion of three functions formerly provided by RS-232. The most significant functions added by RS-449 are local and remote loopback signals. These circuits enable the operation of diagnostic features built into communications equipment via DTE control, permitting, as an example, the loopback of the device to the DTE and its placement into a test mode operation. With the introduction of RS-232-D a local loopback function was supported. Thus, the column labeled RS-232 Designation with the row entry 'Local

Loopback' indicates that that circuit is only applicable to revisions D and E of that standard by the entries D/E in parentheses after the circuit name.

Although a considerable number of articles have been written describing the use of RS-449, its complexity has served as a constraint in implementing this standard by communications equipment vendors. Other constraints limiting its acceptance include the cost and size of the 37-pin connector arrangement and the necessity of using another connector for secondary operations. By mid-1998, less than a few percent of all communications devices were designed to operate with this interface. Due to the failure of RS-449 to obtain commercial acceptance the EIA issued RS-530 in March 1987. This new standard, which is described later in this section, is intended to gradually replace RS-449.

V.35

The V.35 standard was developed to support high-speed transmission, typically 48, 56 and 64 kbps. Originally the V.35 interface was designed into adapter boards inserted into mainframe computers to support 48 kbps transmission on analog wideband facilities. Today, the V.35 standard is the prevalent interface to 56 kbps common carrier digital transmission facilities in the United States and 48 kbps common carrier digital transmission facilities in the UK. In addition, the V.35 standard can support data transfer operations at operating rates up to approximately 6 Mbps. This has resulted in the V.35 interface being commonly employed in videoconferencing equipment, routers and other popularly used communications devices.

The V.35 electrical signaling characteristics are a combination of an unbalanced voltage and a balanced current. Although control signals are electrically unbalanced and compatible with RS-232 and ITU V.28, data and clock interchange circuits are driven by balanced drivers using differential signaling similar to RS-422 and ITU V.11. V.35 uses a 34-pin connector specified in ISO 2593 similar to the connector illustrated in Figure 1.39.

Figure 1.49 illustrates the correspondence between RS-232 and V.35. Note that the V.35 pin pairs tied together by a brace are differential signaling circuits that use a wire pair.

RS-366-A

The RS-366-A interface is employed to connect terminal devices to automatic calling units. This interface standard uses the same type 25-pin connector as RS-232; however, the pin assignments are different. A similar interface to RS-366-A is the ITU V.25 recommendation, which is also designed for use with automatic calling units. Figures 1.50 and 1.51 illustrate the RS-366-A and ITU V.25 interfaces. Note that for both interfaces each actual digit to be dialed is transmitted as parallel binary information over circuits 14 to 17. The pulse on pin 14 represents the value 2^0 while the pulses on pins 15 to 17 represent the values 2^1, 2^2 and 2^3, respectively. Thus, to indicate to the automatic calling unit that it should dial the digit 9, circuits 14 and 17 would become active.

V.35 Pin	V.35 Name	Direction DCE DTE	Function	RS-232 Pin	RS-232 Name
A	FG		Frame GND	1	AA
B	SG		Signal GND	7	AB
C	RTS	→	Request to Send	4	CA
D	CTS	←	Clear to Send	5	CB
E	DSR	→	Data Set Ready	6	CC
F	RLSD	→	Received Line Signal	8	CF
H	DTR	←	Data Terminal Ready	20	CD
J	RI	→	Ring Indicator	22	CE
R, T	RD	→	Receive Data	3	BB
U, X	SGR	→	Receive Clock	17	DD
P, S	SD	←	Send Data	2	BA
U, W	SCTE	←	Send Clock (EXT)	24	DA
Y, a	SCT	→	Send Clock	15	DB
m	TST	→	Reserved for Test (D/E)	25	TM

Figure 1.49 V.35/RS-232 signal correspondence. Illustrates the correspondence between RS-232 and V.35. Note that the V.35 pin pairs tied together by a brace are differential signaling circuits that use a wire pair

Originally, automatic calling units provided the only mechanism to automate communications dialing over the PSTN. This enabled the use of RS-366 automatic dialing equipment under computer control to re-establish communications via the PSTN if a leased line became inoperative, a process called dial-backup. Another common use of automatic calling units is to poll remote terminals from a centrally located computer in the evening when rates are lower. Under software control the central computer would dial each remote terminal and request the transmission of the day's transactions for processing. Due to the development and wide acceptance of the use of intelligent modems with automatic dialing capability, the use of automatic calling units has greatly diminished.

Until the mid-1980s only intelligent asynchronous modems had an automatic dialing capability, restricting the use of automatic calling units to mainframe computers that required a method to originate synchronous data transfers over the PSTN. In such situations a special adapter needed to be installed in the communications controller of the mainframe, which controlled the operation of the automatic calling unit. The introduction of synchronous modems with automatic dialing capability significantly diminished the requirement for automatic calling units since their use eliminates the requirement to install an expensive adapter in the communications controller as well as the cost of the automatic calling unit. The operation and utilization of intelligent modems is discussed in Chapter 5.

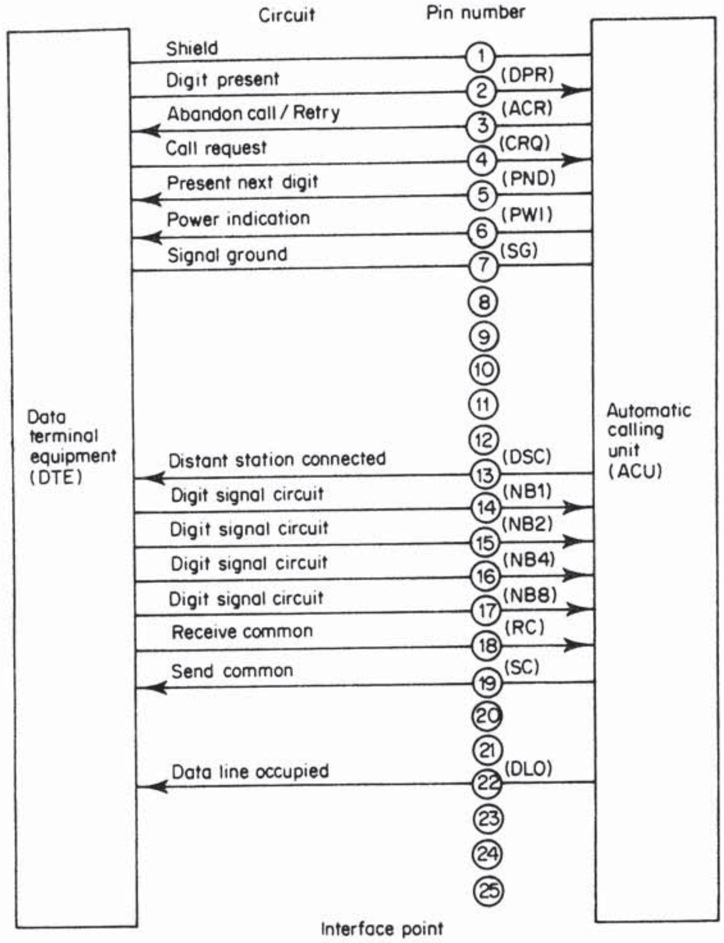


Figure 1.50 RS-336-A interface

X.21 and X.20

Interface standards X.21 and X.20 were developed to accommodate the growing use of public data networks. X.21 is designed to allow synchronous devices to access a public data network while X.20 provides a similar capability for asynchronous devices.

Instead of assigning functions to specific pins on a connector like RS-232, RS-449 and V.35, X.21 assigns coded character strings to each function for establishing connections through a public data network. For example, a dial tone is presented to a computer as a continuous sequence of ASCII '+' characters on the X.21 receive circuit. The computer can then dial a stored number by transmitting it as a series of ASCII characters on the X.21 transmit circuit. Once the call dialing process is completed, the computer will receive call progress signals from the modem on the receive circuit indicating such conditions as number busy and call in progress.

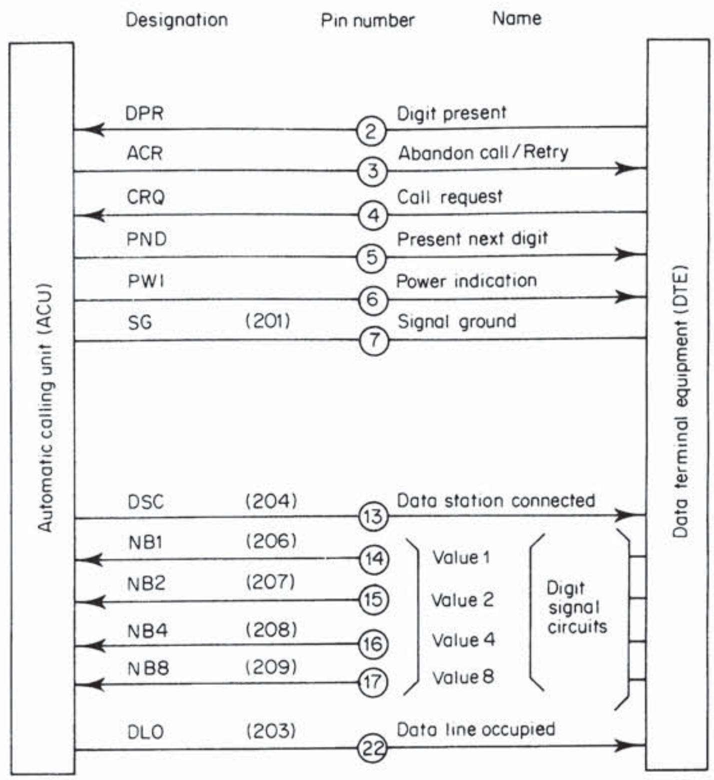


Figure 1.51 V.25 interface

The X.21 interface specifies the use of the balanced signaling characteristics of ITU X.27 (RS-422) for the network side of the interface and either X.27 or X.26 (RS-423) for the terminal equipment side. This specification enables terminal equipment to be designed for several applications. Unlike RS-232 and V.24, the X.21 standard specifies the use of a 15-pin connector. The X.20 interface uses the same 15-pin connector as X.21; however, since it supports asynchronous transmission it only needs to transmit data, and to receive data and ground signals.

Figure 1.52 illustrates the X.21 interchange circuits by circuit type. As indicated, X.21 specifies four categories of interchange circuits—ground, data transfer, control and timing. The operation of the circuits in each category is described below.

Ground signals

Circuit G, signal ground or common return, is used to connect the zero volt reference of the transmitter and receiver ends of the circuit. If X.26 differential signaling is used the G circuit is split into two. The Ga circuit is used as the DTE common return and is connected to ground at the DTE. The G circuit becomes Gb and is used as the DCE common return and is connected to ground at the DCE.

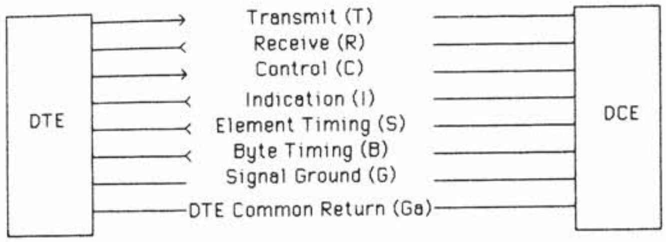


Figure 1.52 CCITT X.21 interchange circuits. Illustrates the X.21 interchange circuits by circuit type

Data transfer circuits

Circuit T is the transmit circuit used by the DTE to transmit data to the DCE. Circuit R is the receive circuit which is used by the DCE to transmit data to the DTE.

Control circuits

The X.21 specification has two control circuits—control and indication. Circuit C is the control circuit used by the DTE to indicate to the DCE the state of the interface. During the data transfer phase in which coding flows over the transmit circuit, circuit C remains in the ON condition.

Circuit I, which is the indication circuit, is used by the DCE to indicate the state of the interface to the DTE. When circuit I is ON the representation of the signal occurs in coded form over the receive circuit. Thus, during the data transfer phase circuit I is always ON.

Timing circuits

There are two timing circuits specified by X.21—signal element timing and byte timing.

Circuit S, which is signal element timing, is generated by the DCE and controls the time of data on the transmit and receive circuits. In providing a clocking signal, circuit S turns ON and OFF for nominally equal periods of time. The second timing circuit, circuit B, which is byte timing, provides the DTE with 8-bit timing information for synchronous transmission generated by the DCE. Circuit B turns OFF whenever circuit S is ON, indicating the last bit of an 8-bit byte. At other times within the period of the 8-bit byte circuit B remains ON. This circuit is not mandatory in X.21 and is only used occasionally.

Limitations

The X.21 standard has not gained wide acceptance for several reasons to include the popularity of the RS-232/V.24 standard and the cost of implementing X.21. With respect to cost, X.21 transmits and interprets coded character strings. This requires more intelligence to be built into the interface, adding to the cost of the interface. Due to the preceding limitations, the ITU defined an alternate interface for public data network access known as X.21 bis, where bis is the Latin term for secondary.

X.21 bis

The X.21 bis recommendation is both physically and functionally equivalent to the V.24 standard, which is compatible with RS-232. The X.21 bis recommendation is designed as an interim interface for X.25 network access and will gradually be replaced by the X.21 standard as more equipment is manufactured to meet the X.21 specification. The X.21 bis connector is the common DB25 connector used by RS-232 and V.24. The connector pins for X.21 bis are defined in an ISO specification called DIS 2110.

RS-530

Like RS-232, RS-530 uses the near-universal 25-pin D-shaped interface connector. Although this standard is intended to replace RS-449, both RS-422 and RS-423 standards specify the electrical characteristics of the interface and will continue in existence. These standards are referenced by the RS-530 standard.

Similar to RS-449, RS-530 provides equipment meeting this specification with the ability to transmit at data rates above the RS-232 limit of 19.2 kbps. This is accomplished by the standard originally specifying the utilization of balanced signals in place of several secondary signals and the Ring Indicator signal included in RS-232. As previously mentioned in our discussion of differential signaling, this balanced signaling technique is accomplished by using two wires with opposite polarities for each signal to minimize distortion.

The RS-530 specification was first outlined in March 1987 and was officially released in April 1988. A revision known as RS-530-A was approved in May 1992. By supporting data rates up to 2 Mbps and using the standard 'D' type 25-pin connector, RS-530 offers the potential to achieve a high level of adoption during the 1990s. One significant change resulting from Revision A is the specification of an alternative 26-position interface connector (Alt A) which is the same optional connector as specified in Revision E to RS-232. Another significant change resulting from Revision A was the addition of support for Ring Indicator which enables the interface to support switched network applications.

Figure 1.53 summarizes the RS-530 interchange circuits and compares those circuits to both RS-232 and RS-449. Note that RS-530 has maintained the standard RS-232 circuit structure for data, clock and control, all of which are balanced signals based upon the RS-442 standard. RS-530 has also adopted the three test circuits specified in RS-232-D: local loop, remote loop and test mode. Like RS-232-D, these three circuits are single ended. RS-530 has maintained pin 1 as frame ground or shield and pin 7 as signal ground.

The original RS-530 interface specified balanced circuits for DCE Ready and DTE ready, using pins 22 and 23 for one pair of each signal, respectively. Under

Designation	RS-530	RS-232	RS-449
Shield	1	1	1
Transmitted data	BA (A) 2 BA (B) 14	2 —	4 Send data 22
Received data	BB (A) 3 BB 16	3 —	6 Received data 24
Request to send	CA (A) 4 CA (B) 19	4 —	7 Request to send 25
Clear to send	CB (A) 5 CB (B) 13	5 —	9 Clear to send 27
DCE ready	CC (A) 6 CC (B) 22	6 —	11 Data mode 29
DTE ready	CD (A) 20 CD (B) 23	20 —	12 Terminal ready 30
Signal ground	AB 7	7	19 Signal ground
Received line signal detector	CF (A) 8 CF (B) 10	8 —	13 Receiver ready 31
Transmit signal element timing (DCE source)	DB (A) 15 DB (B) 12	15 —	5 Send timing 23
Receive signal element timing (DCE source)	DD (A) 17 DD (B) 9	17 —	8 Receive timing 26
Local loopback	LL 18	—	10 Local loopback
Remote loopback	RL 21	—	14
Transmit signal element timing (DTE source)	DA (A) 24 DA (B) 11	24 —	17 Terminal timing 35
Test mode	TM 25	—	18 Test mode
Ring indicator (Revision A)	CD 22	22	15 Incoming call
Signal common (Revision A)	AC 23	—	20 Receive common

Figure 1.53 Pin comparison—RS-530 to RS-232 and RS-449

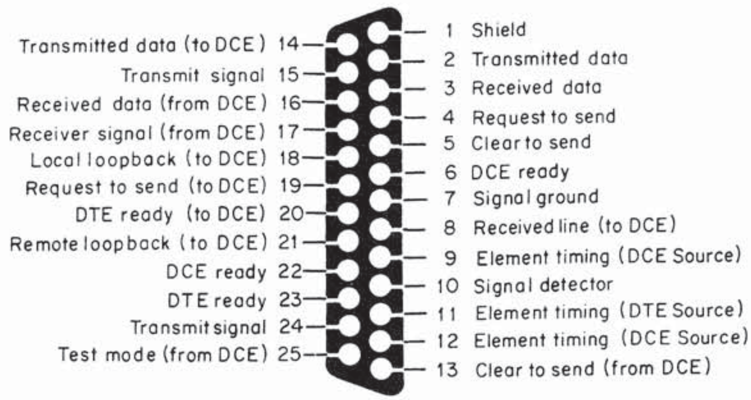


Figure 1.54 RS-530 interface on D connector

Revision A, Ring Indicator support was added through the use of pin 22 while Signal Common support was added through the use of pin 23.

Based upon the RS-530 pin assignments contained in Figure 1.53, the interchange circuits for the D connector specified by the standard are illustrated in Figure 1.54. You can compare Figure 1.54 with Figure 1.45 to see the similarity between RS-530 and RS-232 D-connector interfaces.

High Speed Serial Interface

The High Speed Serial Interface (HSSI) was jointly developed by T3Plus Networking, Inc. and Cisco Systems, Inc. as a mechanism to satisfy the growing demands of high-speed data transmission applications. Although the development of HSSI dates from 1989, its developers were forward looking, recognizing that the practical use of T3 and Synchronous Optical Network (SONET) terminating products would require equipment to transfer information well beyond the capability of the V.35 and RS 422/449 interfaces. The result of their efforts was HSSI, which is a full-duplex synchronous serial interface capable of transmitting and receiving information at data rates up to 52 Mbps between a DTE and a DCE. This standard was ratified by the American National Standards Institute (ANSI) in July 1992 and was being reviewed by the International Organization for Standardization (ISO) and the ITU-T for standardization when this book was revised. ANSI document SP2795 defines the electrical specifications for the interface at data rates up to 52 Mbps while document SP2796 specifies the operation of the DTE–DCE interface circuits.

Rationale for development

The rationale for the development of HSSI was not only the operating limit of 6 Mbps for V.35 and 10 Mbps for RS-422/449 but, in addition, the problems that occur when those standards are extended to higher operating rates. Several manufacturers developed proprietary methods to increase the data transfer rate of those standards; however, doing so resulted in an increase in radio frequency interference (RFI) which in some instances resulted in the disruption of the operation of other nearby equipment and cable connections.

HSSI eliminates potential RFI problems while obtaining a high speed data transfer capability through the use of emitter-coupled logic (ECL). ECL is faster than complementary metal-oxide semiconductor (CMOS) or transistor-to-transistor logic (TTL) commonly used in other interfaces, while generating a lower amount of noise. To accomplish this, ECL has a voltage swing of 0.8 between defining 0 and 1 bits, which is considerably smaller than the voltage swings in CMOS and TTL logic.

Signal definitions

HSSI can be viewed as a simple V.35 type interface based upon the use of emitter-coupled logic for transmission levels, with 12 signals currently defined. Figure 1.55 illustrates the 12 HSSI currently defined interchange circuits, with the normal dataflow indicated by arrowheads on each circuit.

Under HSSI signaling the DCE manages clocking, similar to the V.35 and RS-499 standards. The DCE generates Receive Timing (RT) and Send Timing (ST) signals from the network clock. In comparisons, the DTE returns the clocking signal as Terminal Timing (TT) with data on circuit SD (Send Data) to ensure data is in phase with timing.

HSSI signaling was designed to support continuous as well as gapped, or discontinuous clocking. The latter is associated with the DS3 signal used for T3 transmission at 44.736 Mbps. Under DS3 signaling every 85th frame is a control frame, requiring the DCE clock to run for 84 contiguous pulses and then miss one

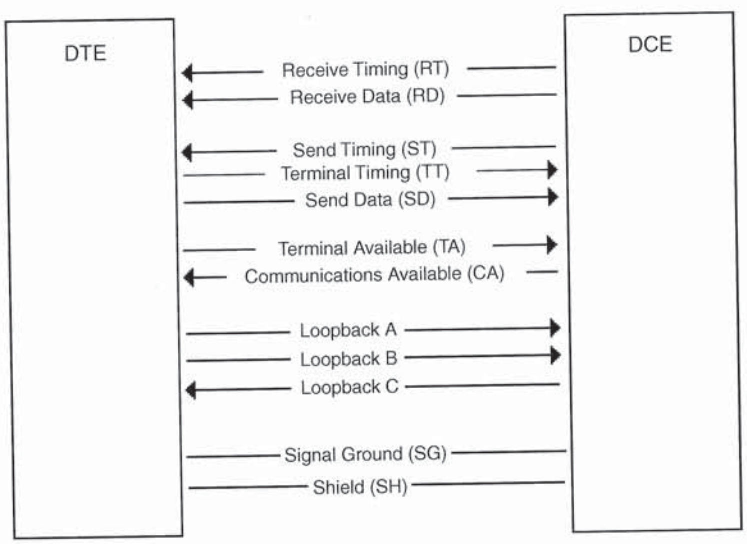


Figure 1.55 HSSI signaling between DTE and DCE

pulse or gap over it to correctly achieve the DS3 frequency. We can obtain an appreciation of the operation of HSSI by examining the operation and functionality of each of the 12 signals currently supported by the interface.

Receive Timing (RT)

The Receive Timing circuit presents the DCE clocking obtained from the network to the attached terminal device. As previously discussed, RT is a gapped clock and has a maximum bit rate of 52 Mbps. The clocking signal on RT provides the timing information necessary for the DTE to receive data on circuit RD.

Receive Data (RD)

Data received by the DCE from an attached communications circuit are transferred on the RD circuit to the DTE.

Send Timing (ST)

The Send Timing circuit transports a gapped clocking signal with a maximum bit rate of 52 Mbps from the DCE to the DTE. This clock provides transmit signal element timing information to the DTE which is returned via the Terminal Timing (TT) circuit.

Terminal Timing (TT)

The Terminal Timing circuit provides the path for the echo of the Send Timing clocking signal from the DTE to the DCE. The clocking signal on this circuit provides transmit signal element timing information to the DCE which is used for sampling data forwarded to the DCE on circuit SD.

Send Data (SD)

The flow of data from the DTE to the DCE occurs on circuit SD. As previously mentioned, clocking on circuit TT provides the DCE with the timing signals to correctly sample the SD circuit.

Terminal Available (TA)

Terminal Available can be considered as the functional equivalent of Request to Send (RTS); however, unlike TRS, TA is asserted by the DTE independently of DCE when the DTE is ready to both send and receive data. Actual data transmission will not occur until the DCE has asserted a Communications Available (CA) signal.

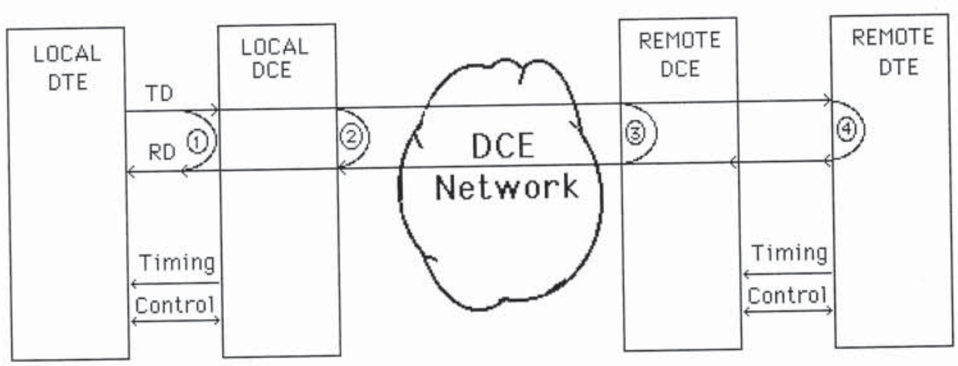
Communications Available (CA)

The Communications Available (CA) signal can be considered as functionally similar to the Clear to Send (CTS) signal. However, the CA signal is asserted by the DCE independently of the TA signal whenever the DCE is prepared to both transmit and receive data to and from the DCE. The assertion of voltage on circuit CA indicates that the DCE has a functional data communications channel; however, transmission will not occur until the TA signal is asserted by the DTE.

Through the elimination of Data Set Ready (DSR) and Data Terminal Ready (DTR) signals commonly used in other interfaces, the HSSI interface becomes relatively simple to implement. This in turn simplifies the DTE-to-DCE data exchange by eliminating the complex handshaking procedures required when using other interfaces.

Loopback circuits

Through the use of three loopback circuits HSSI provides expanded diagnostic testing capability that can be extremely valuable when attempting to isolate transmission problems. Circuits LA and LB are asserted by the DTE to inform the near or far end DCE to initiate one of three diagnostic loopback modes—loopback at the remote DCE line, loopback at the local DCE line, or loopback at the remote DTE. The third loopback circuit, LC, is optional and is used to request the local DTE to provide a loopback path to the DCE. When the LC circuit is asserted the DTE would set TT = RT and SD = RD, enabling testing of the DCE to DTE interface independent of the DTE. The ST circuit would not be used as it cannot be relied upon as a valid clocking source.



1. Local digital loopback
2. Local 'analog' loopback
3. Remote digital loopback
4. Remote 'analog' loopback

Figure 1.56 HSSI supports four loopbacks

Figure 1.56 illustrates the four possible HSSI loopbacks with respect to the local DTE. Although data flows end-to-end in digital form, the term 'analog' used to reference two loopbacks is a carry-over from modem loopback terminology and indicates that data is converted from unipolar to bipolar in the same manner as certain modem loopbacks convert digital to analog to test the modulator of a modem.

Signal Ground (SG)

Signal Ground is used to ensure that transmit signal levels remain within the common input range of receivers. The SG circuit is grounded at both ends.

Shield (SH)

The shield is used to limit electromagnetic interference. To accomplish this the shield encapsulates the HSSI cable.

Pin Assignments

HSSI employs a 50-pin plug connector and receptacle. The connectors are mated to a cable consisting of 25 twisted pairs of 28 AWG cable and are limited to a 50 foot length. The 25 twisted pairs are encapsulated in a polyvinyl chloride (PVC) jacket. Table 1.26 lists the pin assignments. Note that the signal direction is indicated with respect to the DCE.

Table 1.26 HSSI pin assignments

Signal name	Signal direction	Pin no. (+side)	Pin no. (−side)
SG Signal Ground	N/A	1	26
RT Receive Timing	←	2	27
CA DCE Available	←	3	28
RD Receive Data	←	4	29
LC Loopback circuit C (optional)	←	5	30
ST Send Timing	←	6	31
SG Signal Ground	N/A	7	32
TA DTE Available	→	8	33
TT Terminal Timing	→	9	34
LA Loopback circuit A	→	10	35
SD Send Data	→	11	36
LB Loopback circuit B	→	12	37
SG Signal Ground	N/A	13	38
Reserved for future use	→	14–18	39–43
SG Signal Ground	N/A	19	44
Reserved for future use	←	20–24	45–49
SG Signal Ground	N/A	25	50

Applications

Since its initial development in 1989 HSSI has been incorporated into a large number of products designed to support high speed serial communications. In addition, its relatively simple interface has resulted in its use at data rates that would normally require the use of a V.35 or RS-499 interface. For example, many routers, multiplexers, inverse multiplexers and Channel Service Units operating at 1.544 Mbps can now be obtained with HSSI as well as products designed to operate at T3 (44.736 Mbps) and SONET Synchronous Transmission Service Level 1 (STS-1) (51.84 Mbps).

High Performance Parallel Interface

The High Performance Parallel Interface (HIPPI) represents an ANSI switched network standard which was originally developed to support direct communications between mainframes, supercomputers and directly attached storage devices. A series of ANSI standards currently define the physical layer operation of HIPPI as well as data framing, disk and tape connections and link encapsulation. Table 1.27 lists ANSI HIPPI related standards and their areas of standardization.

Table 1.27 ANSI HIPPI-related standards

ANSI standard	Area covered
X3.183-1991	Physical layer
X3.222-1993	Switch control
X3.218-1993	Link encapsulation
X3.210-1992	Framing protocol
ANSI/ISO 9318-3	Disk connections
ANSI/ISO 9318-4	Tape connection

Transmission distance

Although its name implies the use of a parallel interface, a number of extensions to HIPPI resulted in its ability to support a 300 meter serial interface over multimode copper as well as parallel transmission via a 50-pair shielded twisted-pair wiring group for relatively short distances. In its basic mode of operation a HIPPI based network consists of two computers connected via a pair of 50-pair copper cables to HIPPI channels on each device. Each 50-pair cable supports transmission in one direction, resulting in the pair of 50-pair cables providing a full-duplex transmission facility. That transmission facility can extend up to 25 meters and operate at either 800 Mbps or 1.6 Gbps, the latter accomplished by doubling the data path.

Through the use of one or more HIPPI switches you can develop an extended HIPPI network. That network can use copper cable between switches which permits cabling runs up to 200 meters in length, or a fiber extender can be used to support extending the distance between switches up to 10 km. The fiber extender

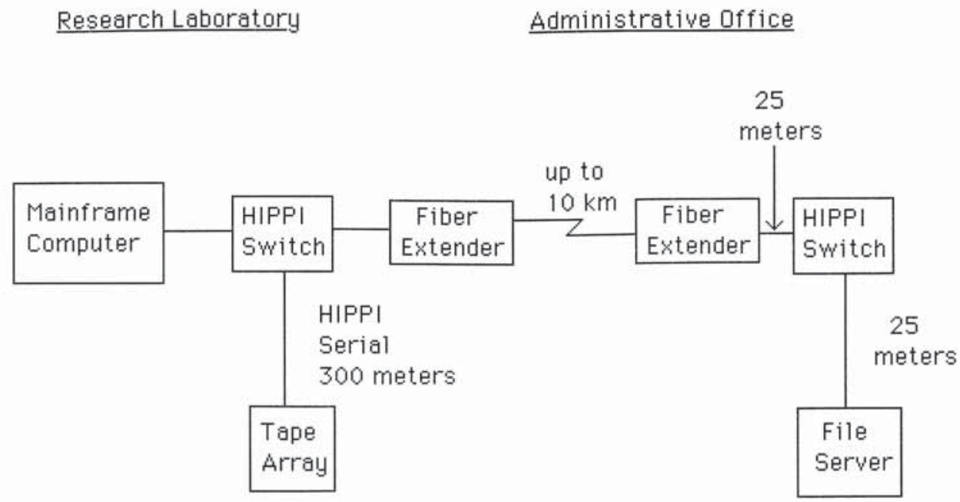


Figure 1.57 Creating a HIPPI-based network

functions as a parallel to serial converter as well as an electrical to optical converter to support serial light transmission between switches.

Figure 1.57 illustrates the creation of HIPPI based network on a college campus to connect a research laboratory to an administrative file server. HIPPI interfaces are now available for a wide range of products to include personal computers, routers and gateways as well as mainframes and supercomputers.

Operation

HIPPI operates by framing data to be transmitted as well as by using messages to control data transfer operations. A HIPPI connection is set up through the use of three messages. A Request message is used by the data originator to request the establishment of a connection. A Connect message is returned by the desired destination to inform the requestor that a connection was established. The third message is Ready, which is issued by the destination to inform the originator that it is ready to accept data.

Cables and connectors

Numerous types of cables and connectors can be employed in data transmission systems. To familiarize readers with cabling options that can be considered, we will now focus our attention upon several types of cables and their connectors, as well as several cabling tricks based upon our previous examination of the operation of RS-232/V.24 interchange circuits.

Twisted-pair cable

The most commonly employed data communications cable is the twisted-pair cable. This cable can usually be obtained with 4, 7, 9, 12, 16 or 25 conductors, where each conductor is insulated from another by a PVC shield.

For EIA RS-232 and ITU V.24 applications, those standards specify a maximum cabling distance of 50 feet between DTE and DCE equipment for data rates ranging from 0 to 19 200 bps; and, normal industry practice is to use male connectors at the cable ends which mate with female connectors normally built into such devices as terminals and modems. Figure 1.58 illustrates the typical cabling practice employed to connect a DTE to a DCE.

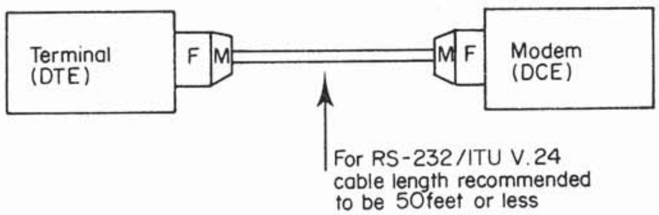


Figure 1.58 DTE to DCE cabling

Low-capacitance shielded cable

In certain environments where electromagnetic interference and radio frequency emissions could be harmful to data transmission, you should consider the utilization of low-capacitance shielded cable in place of conventional twisted-pair cable. Low-capacitance shielded cable includes a thin wrapper of lead foil that is wrapped around the twisted-pair conductors contained in the cable, thereby providing a degree of immunity to electrical interference that can be caused by machinery, fluorescent ballasts and other devices.

Ribbon cable

Since an outer layer of PVC houses the individual conductors in a twisted-pair cable, the cable is rigid with respect to its ability to be easily bent. Ribbon or flat cable consists of individually insulated conductors that are insulated and positioned in a precise geometric arrangement that results in a rectangular rather than a round cross-section. Since ribbon cable can be easily bent and folded, it is practical for those situations where you must install a cable that must follow the contour of a particular surface.

The RS-232 null modem

No discussion of cabling would be complete without a description of a null modem, which is also referred to as a modem eliminator. A null modem is special cable that is designed to eliminate the requirement for modems when interconnecting two collocated data terminal equipment devices. One example of this would be a requirement to transfer data between two collocated personal computers that do not have modems and use different types of diskettes, such as an IBM PC which uses a $5\frac{1}{4}$-inch diskette and an IBM PS/2 which uses a $3\frac{1}{2}$-inch diskette. In this situation, the interconnection of the two computers via a null modem cable would permit programs and data to be transferred between each personal computer in spite of the media incompatibility of the two computers. Since DTEs transmit data on pin 2 and receive data on pin 3, you could never connect two such devices together with a conventional cable as the data transmitted from one device would never be received by the other. In order for two DTEs to communicate with one another, a connector on pin 2 of one device must be wired to connector pin 3 on the other device. Figure 1.59 illustrates an example of the wiring diagram of a null modem cable used to connect two DB-25 connectors together, showing how pins 2 and 3 are cross-connected as well as the configuration of the control circuit pins on this type of cable.

Since a terminal will raise or apply a positive voltage in the 9 V to 12 V range to turn on a control signal, you can safely divide this voltage to provide up to three different signals without going below the signal threshold of 3 V previously illustrated in Figure 1.41. In examining Figure 1.59, we should note the following control signal interactions are caused by the pin cabling:

(1) Data terminal ready (DTR, pin 20) raises data set ready (DSR, pin 6) at the other end of the cable. This makes the remote DTE think a modem is connected to the other end and powered ON.
(2) Request to send (RTS, pin 4) raises data carrier detect (CD, pin 8) on the other end and signals clear to send (CTS, pin 5) at the original end of the cable. This makes the DTE believe that an attached modem received a carrier signal and is ready to modulate data.
(3) Once the handshaking of control signals is completed, we can transmit data onto one end of the cable (TD, pin 2) which becomes receive data (RD, pin 3) at the other end.

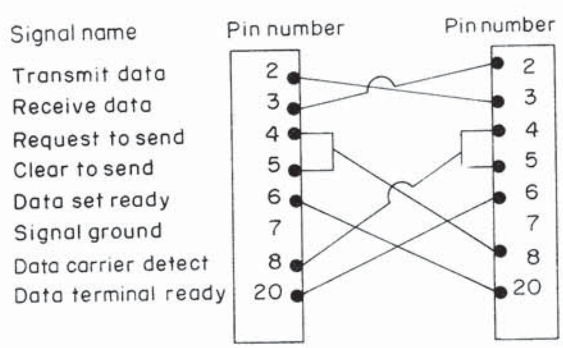


Figure 1.59 DB-25 to DB-25 null modem cable

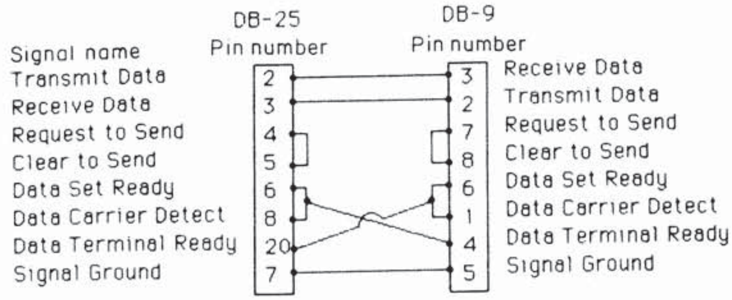


Figure 1.60 DB-25 to DB-9 null modem cable

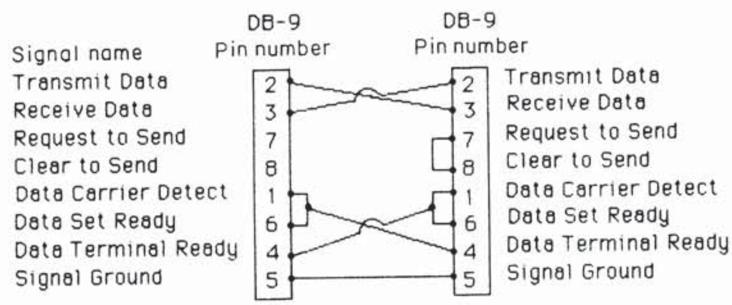


Figure 1.61 DB-9 to DB-9 null modem cable

You can also use a null modem cable to connect a DB-25 connector port to a DB-9 connector port as well as two DB-9 connector ports to each other. Figure 1.60 illustrates the connector wiring for a DB-25 to DB-9 null modem cable, while Figure 1.61 illustrates the conductor wiring for a null modem cable used to connect two DB-9 connectors.

Since a large number of personal computers use DB-9 connectors, the null modem cables illustrated in Figures 1.60 and 1.61 are popularly employed to directly cable two personal computers to one another as well as PCs to other types of terminal devices, including mainframe computer ports and the ports of a data PBX.

In comparing the wiring of the conductors in Figure 1.59 to the wiring of conductors illustrated in Figures 1.60 and 1.61, you will note the similarity between each type of null modem cable. That is, transmit data is always routed to receive data at the opposite end of the cable, RTS and CTS control signals are always tied together, and the tying of DSR to DCD is routed to the DTR signal at the other end of the cable. You will also note that, although the routing of conductors is consistent for all three types of null modem cables, the actual routing of conductors to specific pins will vary due to the difference in the assignment of conductors to pins on the DB-25 and DB-9 conductors.

The cable configurations illustrated in Figures 1.59 and 1.61 will work for most data terminal equipment interconnections; however, there are a few exceptions. The most common exception is when a terminal device is to be cabled to a port on a mainframe computer that operates as a 'ring-start' port. This means that the computer port must obtain a ring indicator (RI, pin 22) signal. In this situation, each null modem must be modified so that data set ready (DSR, pin 6) on a DB-25 connector is jumpered to ring indicator (RI, pin 22) at the other end of the DB-25 cable to initiate a connect sequence to a 'ring-start' system.

Owing to the omission of transmit and receive clocks, the previously described null modem cables can only be used for asynchronous transmission. For synchronous transmission you must either drive a clocking device at one end of the cable or employ another technique. Here you would use a modem eliminator which differs from a null modem by providing a clocking signal to the interface. If a clocking source is to be used, DTE timing (pin 24) on a DB-25 connector is normally selected to develop a synchronous null modem cable. In developing this cable, pin 24 is strapped to pins 15 and 17 at each end of the cable as illustrated in Figure 1.62. Then, DTE timing provides transmit and receive clocking signals at both ends of the cable.

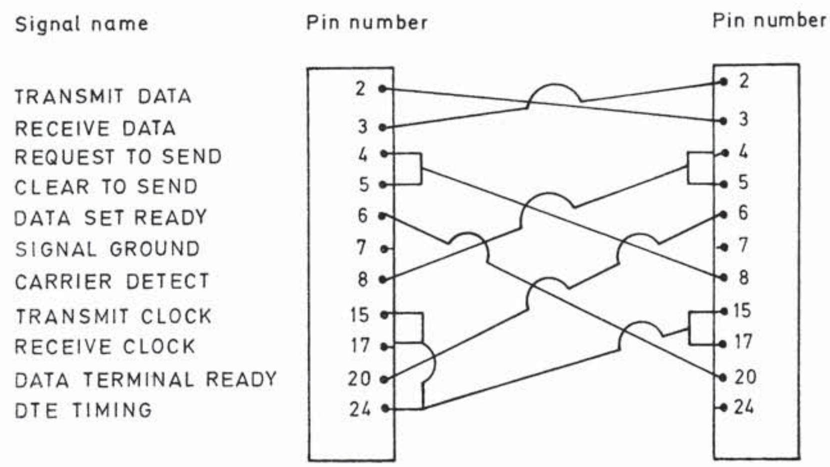


Figure 1.62 Synchronous null modem cable

RS-232 cabling tricks

A general purpose 3-conductor cable can be used when there is no requirement for hardware flow control and a modem will not be controlled. Here the term flow control refers to the process that causes a delay in the flow of data between DTE and DCE, DCE and DTE or two DCEs or two DTEs resulting from the changing of control circuit states. Figure 1.63 illustrates the use of a 3-conductor cable for DTE–DCE and DTE–DTE or DCE–DCE connections. When this situation occurs it becomes possible to use a 9-conductor cable with three D-shaped connectors at each end, with each connected to three conductors on the cable connector which eliminates the necessity of installing three separate cables.

Figure 1.64 illustrates a 5-conductor cable that can be installed between a DTE and DCE (modem) when asynchronous control signals are required. Similar to the

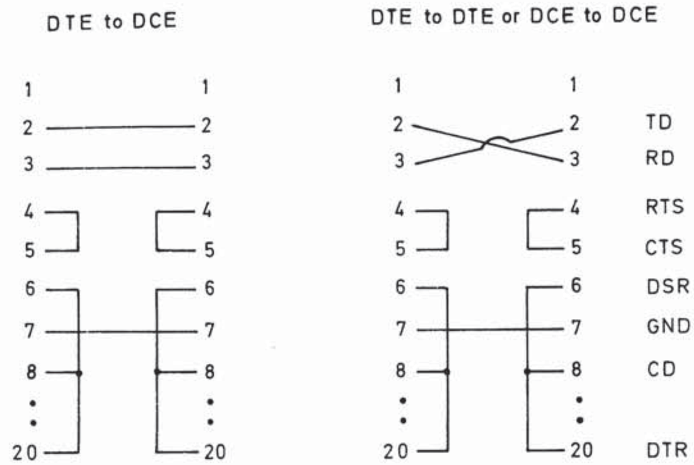


Figure 1.63 General purpose 3-conductor cable

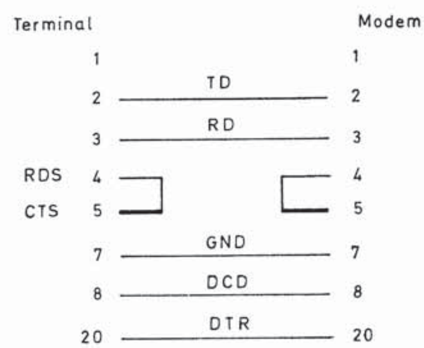


Figure 1.64 General purpose 5-conductor cable

use of 9-conductor cable to derive three 3-conductor connections, standard 12-conductor cable can be used to derive two 5-conductor connections.

Plugs and jacks

Modern data communications equipment is connected to telephone company facilities by a plug and jack arrangement as illustrated in Figure 1.65. Although the connection appears to be, and in fact is, simplistic, the number of connection arrangements and differences in the types of jacks offered by telephone companies usually ensures that the specification of an appropriate jack can be a complex task. Fortunately, most modems and other communications devices include explicit instructions covering the type of jack the equipment must be connected to as well as providing the purchaser with information that must be furnished to the telephone company in order to legally connect the device to the telephone company line.

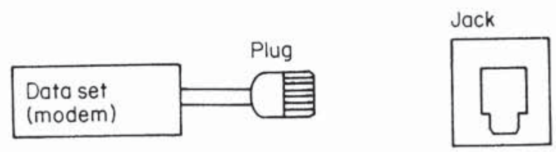


Figure 1.65 Connection to telephone company facilities. Data communications equipment can be connected to telephone company facilities by plugging the device into a telephone company jack

Most communications devices designed for operation on the PSTN interface the telephone company network via the use of an RJ11C permissive or an RJ45S programmable jack.

Figure 1.66 illustrates the conductors in the RJ11 and RJ45 modular plugs. The RJ11 plug is primarily used on two-wire dial lines. This plug is used in both the home and office for connecting a single instrument telephone to the PSTN. In addition, the RJ11 also serves as an optional connector for four-wire private lines. Although the RJ11 connector is fastened to a cable containing four or six stranded-copper conductors, only two wires in the cable are used for switched network applications. When connected to a four-wire leased line, four conductors are used.

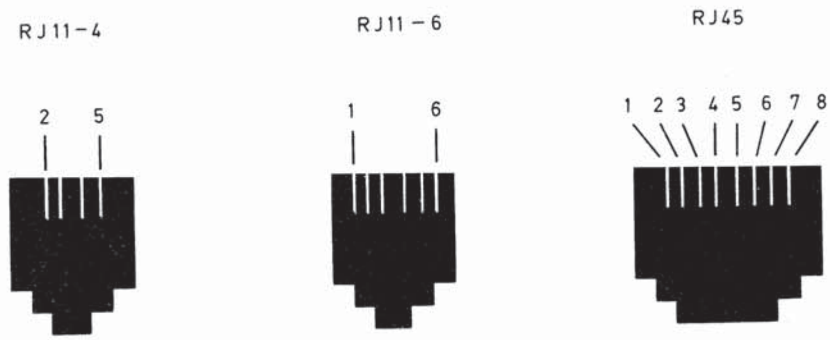


Figure 1.66 RJ11 and RJ45 modular plugs

The development of the RJ11 connector can be traced to the evolution of the switchboard. The plugs used with switchboards had a point known as the 'tip' which was colored red, while the adjacent sleeve known as the ring was colored green.

The original color coding used with switchboard plugs was carried over to telephone wiring. If you examine a four-wire (two-pair) telephone cable, you will note that the wires are coloured yellow, green, red and black. The green wire is the tip of the circuit while the red wire is the ring. The yellow and black wires can be used to supply power to the light in a telephone or used to control a secondary telephone using the same four-wire conductor cable.

Table 1.28 Color identification of telephone cables

Four-wire		Six-wire	
Pair	Color	Pair	Color
1	Yellow Green	1	Blue Yellow
2	Red Black	2	Green Red
		3	Black White

The most common types of telephone cable used for telephone installation are four-wire and six-wire conductors. Normally, a four-wire conductor is used in a residence that requires one telephone line. A six-wire conductor is used in either a residential or business location that requires two telephone lines and can also be used to provide three telephone lines from one jack. Table 1.28 compares the color identification of the conductors in four-wire and six-wire telephone cable.

During the late 1970s, telephone companies replaced the use of multiprong plugs by the introduction of modular plugs which in turn are connected to modular jacks.

The RJ11C plug was designed for use with any type of telephone equipment that requires a single telephone line. Thus, regardless of the use of either four-wire or six-wire cable only two wires in the cable need be connected to an RJ11C jacks. The RJ11 plug can also be used to service an instrument that supports two or three telephone lines; however, RJ14C and RJ25C jacks must then be used to provide that service. These two jacks are only used for voice. For data transmission both four- and six-conductor plugs are available for use, with conductors 1, 2, 5 and 6 in the jack normally reserved for use by the telephone company. Then, conductor 4 functions as the ring circuit while conductor 5 functions as the tip to the telephone company network.

The RJ45 plug is also designed to support a single line although it contains eight positions. In this plug, positions 4 and 5 are used for ring and tip and a programmable resistor on position 8 in the jack is used to control the transmit level of the device connected to the switched network.

The RJ45 plug and jack connectors are also used in some communications products to provide an RS-232 DTE–DCE interface via twisted-pair telephone wire. In certain cases an RJ45 to DB25 adapter may be needed. This adapter will, as an example, permit the cabling of a cable terminated with an RJ45 plug to DB25 connector or a DB25 connector cable end to a RJ45 socket. RJ45 connectors typically support the transmitted data, received data, data terminal ready (DTE ready), data set ready (DCE ready), data detect (received line signal detector), request to send, clear to send and signal ground circuits.

The physical size of the plugs used to wire equipment to each jack as well as the size of each of the previously discussed jacks are the same. The only difference between jacks is in the number of wires cabled to the jack and the number of contacts in the jack which are used to pass telephone wire signals.

Connecting arrangements

There are three connecting arrangements that can be used to connect data communications equipment to telephone facilities. The object of these arrangements is to ensure that the signal received at the telephone company central office in the United States does not exceed −12 dBm.

Permissive arrangement

The permissive arrangement is used when you desire to connect a modem to your organization's switchboard, such as a private branch exchange (PBX). When a permissive arrangement is employed, the output signal from the modem is fixed at a maximum of −9 dBm and the plug that is attached to the data set cable can be connected to three types of telephone company jacks as illustrated in Figure 1.67. The RJ11 jack can be obtained as a surface mounting (RJ11C) for desk sets or as a wall-mounted (RJ11W) unit; however the RJ41S and RJ45S are available only for surface mounting.

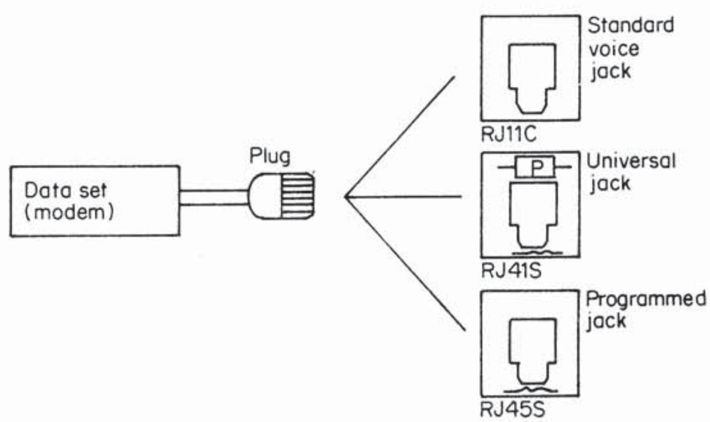


Figure 1.67 Permissive arrangement jack options

Since permissive jacks use the same six-pin capacity miniature jack used for standard voice telephone installations, this arrangement provides for good mobility of terminals and modems.

Fixed loss loop arrangement

Under the fixed loss loop arrangement the output signal from the modem is fixed at a maximum of −4 dBm and the line between the subscriber's location and the telephone company central office is set to 8 dBm of attenuation by a pad located within the telephone company provided jack. As illustrated in Figure 1.68, the

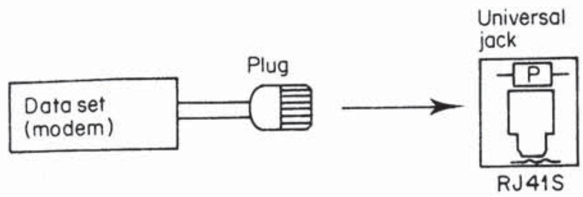


Figure 1.68 Fixed loss loop arrangement

only jack that can be used under the fixed loss loop arrangement is the RJ41S. This jack has a switch FLL-PROG, which must be placed in the FLL position under this arrangement. Since the modem output is limited to −4 dBm, the 8 dB attenuation of the pad ensures that the transmitted signal reaches the telephone company office at −12 dBm. As the pad in the jack reduces the receiver signal-to-noise ratio by 8 dB, this type of arrangement is more susceptible to impulse noise and should only be used if one cannot use either of the two other arrangements.

Programmable arrangement

Under the programmable arrangement configuration a level setting resistor inside the standard jack provided by the telephone company is used to set the transmit level within a range between 0 and −12 dBm. Since the line from the user is directly routed to the local telephone company central office at installation time, the telephone company will measure the loop loss and set the value of the resistor based upon the loss measurement. As the resistor automatically adjusts the transmitted output of the modem so the signal reaches the telephone company office at −12 dBm, the modem will always transmit at its maximum allowable level. As this is a different line interface in comparison to permissive or fixed loss data sets, the data set must be designed to operate with the programmability feature of the jack.

Either the RJ41S universal jack or the RJ45S programmed jack can be used with the programmed arrangement as illustrated in Figure 1.69. The RJ41S jack is installed by the telephone company with both the resistor and pad for programmed and fixed loss loop arrangements. By setting the switch to PROG, the programmed arrangement will be set. Since the RJ45S jack can operate in either the permissive or programmed arrangement without a switch, it is usually preferred as it eliminates the possibility of an inadvertent switch reset.

Telephone options

Prior to the use of modular jacks, telephones were hardwired to the switched telephone network. Even with the growth in the use of modular connecting arrangements, there are still a few locations where telephones are connected the 'old-fashioned way'. Those telephone sets require the selection of specific options to be used with communications equipment. As part of the ordering procedure you must specify a series of specific options that are listed in Table 1.29.