COMPUTER NETWORKS

PROTOCOLS, STANDARDS, AND INTERFACES

Prentice-Hall International Editions

UYLESS BLACK

Computer Networks

Protocols, Standards, and Interfaces

Uyless Black
Center for Advanced Professional Education

Prentice Hall International, Inc.

Library of Congress Cataloging-in-Publication Data

Black, Uyless D.
Computer networks.

Includes index.
1. Computer networks. I. Title.
TK5105.5.B56 1987 004.6 87–2228
ISBN 0-13-166091-8

Editorial/production supervision
and interior design: Evalyn Schoppet
Manufacturing buyer: Gordon Osbourne

Printed and bound in Great Britain by
A. Wheaton & Co. Ltd, Exeter.

10 9 8 7 6 5 4 3

ISBN 0-13-166091-8 025

Prentice-Hall Computer Communications Series
Wushow Chou, Advisor

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Simon & Schuster Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Prentice-Hall, Inc., *Englewood Cliffs, New Jersey*

Preface

In my role as a consultant and lecturer in the field of data communications, I am often asked the following question: "How do computers and terminals actually communicate with each other?" Upon further probing, usually I find the person is actually seeking answers to a more difficult question: "How can I understand the many parts of a communications network and how they fit together?" This is not an easy question to answer, since there are many parts of a communications system. Computer networks can now consist of satellite systems, packet switches, personal computers, private branch exchanges (PBXs), local area networks (LANs), digital systems, and a myriad of other complex technologies.

My goal in writing this book is to provide answers to the questions posed above, and to do so with simple prose rather than with pages upon pages of arcane formulas. Although books based on mathematical concepts and algorithms obviously are quite important, they rarely satisfy the practitioner whose primary concern is with nontheoretical problems, and it is the practitioner for whom this book is intended. Nevertheless, any book on the subject of computer networks must be somewhat technical and detailed if it is to adequately address the subject. Although a background in electronic engineering or a mathematics degree is not required to understand this book, the reader will need to delve into the chapters with some patience and care. If this premise is accepted, then the book should be useful.

This book is written for individuals with varying levels of knowledge and experience, from the beginner who needs an overview to the more advanced data communications professional who needs to fill some information gaps in specific areas. The beginner should read Appendices A and B before moving into Chapter One. After reading Chapter One, Appendix C is available if the reader wishes more information on physical-level interfaces. The more advanced reader can browse through these appendices and then delve into any of the chapters. The book guides the reader to (or around) sections that go into more technical detail. Readers wishing only an overview may choose to skip those sections or chapters.

The chapters have been written to be as self-contained as possible, but each chapter assumes the reader understands the previous chapters and Appendices A, B, and C. Appendices D and E have been included for ease of reference to some of the more important international standards. Each chapter contains suggested readings and, where applicable, notes regarding information sources.

This book is organized around the International Organization for Standardization's Open Systems Interconnection (OSI) layered protocol model, described in Chapter Three. Emphasis is placed on the first four layers of the model, since most of the communications functions reside in those layers. However, information is also provided on some of the more important aspects of the upper three layers.

The book is also structured around the Protocol Classification Tree. The classification tree is not intended to be all-inclusive, but serves as a method to describe some of the more important functions of computer networks. Chapter Two provides a general discussion of the classification tree and also guides the reader to more detailed information in subsequent chapters.

Acknowledgments

It has been said that writing a book is an act of intellectual masochism. This statement has more than an element of truth to it. The author suffers, but derives gratification from the process as well. This is also true of the author's friends and colleagues; in my case, their support and encouragement has been essential during this lengthy project.

My Prentice-Hall editors, Karl Karlstrom and John Wait have been instrumental in seeing this book come to print. Dan Joraanstad from Prentice-Hall has also been very helpful and supportive of my ideas and efforts.

My friend, Phil Dietz, deserves special mention. He provided invaluable support in clarifying and simplifying some highly technical descriptions of electronics in this book, as he did for my previous book.

My four colleagues at the Center for Advanced Professional Education, Ed Sawicki, Ken Sherman, Herb Stern, and Rick Watkins, have played an important role in the production of this book, since they helped keep the bread on my table while I worked on the manuscript.

I also wish to thank the many individuals who have attended my lectures and seminars. I have benefitted enormously from their input to my lectures and to this book. My seminar attendees, not only in the United States but in Europe, Canada, and Asia, have given me immeasurably valuable insights into the data communications industry. I hope this book addresses some of their needs.

The reviews from Prentice-Hall were also very valuable to me. In several instances, they helped smooth out some technical detail, and I owe the reviewers my gratitude.

Holly Velez, Terece Crawford, and Jo Ann Schreiner have displayed uncommon intelligence and patience in dealing with my indecipherable notes and dictation, and with the constant changes. Jean Stum provided very valuable assistance as well.

Finally, the Black family has been the basic foundation supporting my efforts in the past, and they continue to offer this support today. My parents, my brothers, and my son, Tom, have my love and gratitude for their support.

Contents

CHAPTER ONE

Introduction to Computer Networks

The Use of Networks

Due to the tremendous impact of computers and computer networks on society during the past decade, this period in history has come to be called the "information age." The productivity and profitability of both organizations and individuals have been enhanced significantly by these revolutionary tools. Hardly a day goes by without an individual using a computer network to conduct personal and professional business. This trend is accelerating as more businesses and homes discover the power of computers and communications networks. The day-to-day transactions at department stores, banks, reservation counters, and other businesses are all dependent upon computer networks. The information age is equally dependent on the computer *and* the computer network.

What is a computer network? Several definitions are accepted in the industry. Perhaps the simplest is: A number of computers (and usually terminals) interconnected by one or more transmission paths. The transmission path is often the telephone line, due to its convenience and universal presence. The network exists to meet one goal: the transfer and exchange of data between the computers and terminals. This data exchange provides for the many computer-based services we often take for granted in our daily lives, such as bank teller machines, point-of-sale terminals, check-verification devices, and even the guidance of the space shuttle.

Advantages of Networks

Computer networks provide several important advantages to businesses and individuals.

1. Modern organizations today are widely dispersed, with offices located in diverse parts of a country and the world. Many of the computers and

terminals located at the sites need to exchange information and data, often on a daily basis. A network provides the means to exchange data among these computers and to make programs and data available to the people of the organization.

2. The networking of computers permits the sharing of resources of the machines. For instance, if a computer becomes saturated with too much work at one site, the work can be loaded through the network path onto another computer in the network. Such load sharing permits a more even and better utilization of resources.
3. Networking also provides the critical function of back-up. In the event that one computer fails, its counterpart can assume its functions and workload. Back-up capability is especially important in systems such as air traffic control. In the event a computer malfunctions, back-up computers rapidly take over and assume control of operations without endangering air travelers.
4. The use of networking allows a very flexible working environment. Employees can work at home by using terminals tied through networks into the computer at the office. Many employees now carry terminals or portable personal computers on trips and tie into their networks through hotel room telephones. Other employees travel to remote offices and use telephones and networks to transmit and receive critical sales, administrative, and research data from computers at company headquarters.

The information age is aptly named, for our society now relies on information to reduce the costs to produce our goods as well as to improve the overall quality of our lives. Communications systems and computer networks provide for rapid exchange of information residing on computers throughout a country.

Structure of the Communications Network

Before proceeding further, it will be helpful to define some terms. Figure 1-1 illustrates a simple data communications system. The *application process* (AP) is the end-user application. It usually consists of software such as a computer program, or it could be an end-user terminal. Typical examples are an accounts receivable program, a payroll program, an airline reservation system, an inventory control package, or a personnel system.

In Figure 1-1, site A could execute an application process (AP_{A1}) in the form of a software program to access an application process at site B (which is in this case a program [AP_{B1}] and a data base). Figure 1-1 also shows a site B program (AP_{B2}) accessing a file at site A through an application program

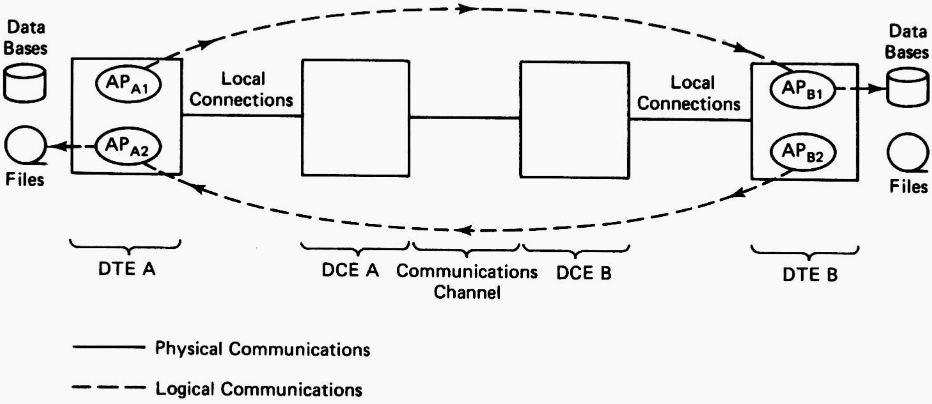

Figure 1-1. A Communications System

(AP_{A2}). (This book uses the term *application process* to describe end-user applications, unless otherwise noted.)

The application resides in the *data terminal equipment,* or DTE. DTE is a generic term used to describe the end-user machine, which is usually a computer or terminal. The DTE could be a large mainframe computer, such as a large IBM or ICL machine, or it could be a smaller machine, such as a terminal or a personal computer. The DTE takes many forms in the industry. Here are several examples:

- a work station for an air traffic controller
- an automated teller machine in a bank
- a point-of-sale terminal in a department store
- a sampling device to measure the quality of air
- a computer used to automate the manufacturing process in a factory
- an electronic mail computer or terminal
- a personal computer in the home or office.

The function of a communications network is to interconnect DTEs together so that they can share resources, exchange data, provide back-up for each other, and allow employees and individuals to perform their work from any location.

Figure 1-1 shows that a network provides *logical* and *physical* communications for the computers and terminals to be connected. The applications and files use the physical channel to effect logical communications. Logical, in this context, means the DTEs are not concerned with the physical aspects of the communications process. Application A1 need only issue a logical *Read* request with an identification of the data. In turn, the communications system is responsible for sending the *Read* request across the physical channels to application B1.

Figure 1-1 also shows the *data circuit-terminating equipment,* or DCE (also called data communications equipment). Its function is to connect the DTEs into the communication line or channel.[1] The DCEs designed in the 1960s and 1970s were strictly communications devices. However, in the last few years the DCEs have incorporated more user functions, and today some DCEs contain a portion of an application process. Notwithstanding, the primary function of the DCE remains to provide an *interface* of the DTE into the communications network. The familiar modem is an example of a DCE.

The interfaces are specified and established through *protocols*. Protocols are agreements on how communications components and DTEs are to communicate with each other. They may include actual regulations which stipulate a required or recommended convention or technique. Typically, several levels of interfaces and protocols are required to support an end-user application.

Today, many organizations are adapting common interfaces and protocols as a result of worldwide efforts to publish recommended *standards* that are vendor and product independent. Our goal is to gain an understanding of these valuable protocols, standards, and interfaces.

Point-to-point and Multidrop Circuits

DCEs and DTEs are connected in one of two ways. As illustrated in Figure 1-1, they are connected in a point-to-point configuration in which only two DTE devices are on the line or channel. Illustrated in Figure 1-2 is another approach called a multidrop configuration. In this configuration, more than two devices are connected to the same channel.

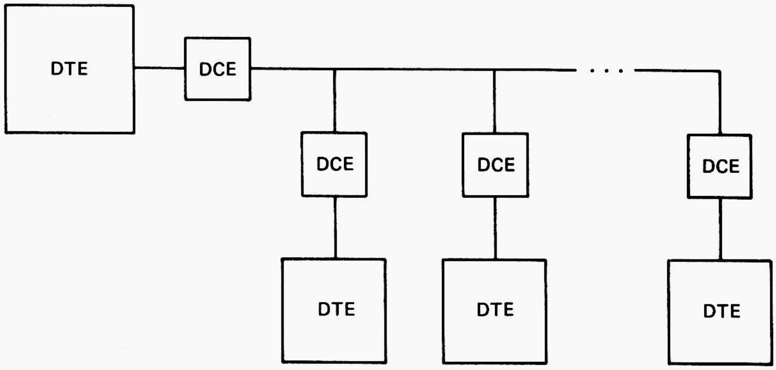

Figure 1-2. Multidrop Circuits

Data Flow and Physical Circuits

The DTEs and DCEs send communications traffic to each other in one of three methods:

Simplex: Transmission in one direction only

Half-Duplex: Transmission in both directions, but only one direction at a time (also called two-way alternate [TWA])

Full-Duplex (or Duplex): Transmission simultaneously in both directions (also called two-way simultaneous [TWS]).

Simplex transmission is common in television and commercial radio. It is not as common in data communications because of the one-way nature of the process, but simplex systems are found in some applications, such as telemetry. Half-duplex transmission is found in many systems, such as inquiry/response applications wherein a DTE sends a query to another DTE and waits for the applications process to access and/or compute the answer and transmit the response back. Terminal-based systems (keyboard terminals and terminals with CRT screens) often use half-duplex techniques. Full-duplex (or simply duplex) provides for simultaneous two-way transmission, without the intervening stop-and-wait aspect of half-duplex. Full-duplex is widely used in applications requiring continuous channel usage, high throughput, and fast response time.

Thus far, the terms half-duplex and full-duplex (duplex) have been used to describe how data moves across the circuit. We have focused on these terms as they are used in the data communications industry. Figure 1-3 shows the physical circuit itself, without regard to how data moves. Some

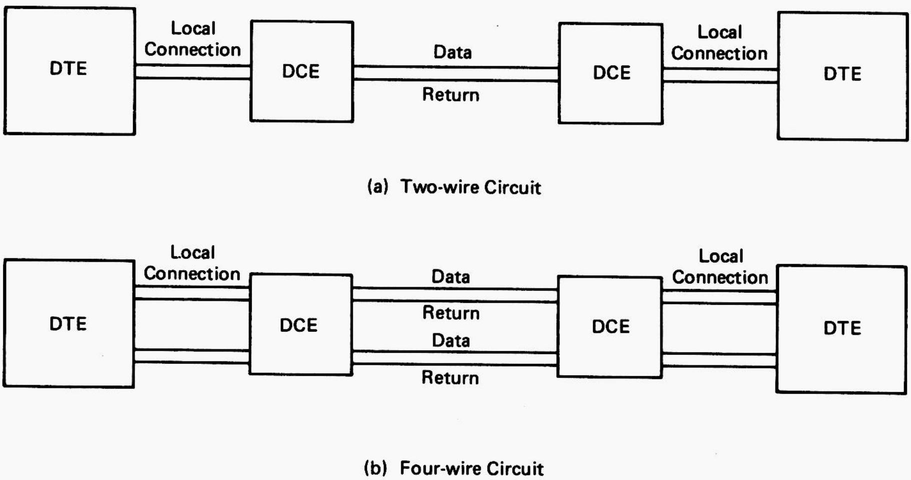

Figure 1-3. Two-Wire and Four-Wire Circuits

people in the industry use half-duplex and full-duplex to describe data flow *and* the physical circuit itself. The following discussion explains the physical circuit.

In telephone communications, the terms *two-wire* and *four-wire* circuits are often used to describe the channel. A two-wire circuit is sometimes called a half-duplex circuit. One wire is for the transmission data and the other is for the return circuit. (See Appendix A for a discussion of return circuits.) A four-wire circuit is sometimes called a full-duplex circuit. Two pairs of two wires exist, two for the data and two for the return circuits. The telephone company usually describes a two-wire circuit as a switched dial-up circuit, and a four-wire circuit as a leased, nonswitched circuit. However, exceptions exist and the reader is encouraged to check with the specific telephone company. The tradeoffs of switched and nonswitched circuits are discussed later in the chapter.

The advantages cited earlier of communications networks cannot be realized without the addition of an important component to the system. This component is the *data switching equipment,* or DSE. Figure 1-4 illustrates the use of the DSE in conjunction with the DTE and DCE. As the name implies, the major function of the DSE is to switch or route traffic (user data) through the network to the final destination. The DSE provides the vital functions of network routing around failed or busy devices and channels. The DSE may also route the data to the final destination through intermediate components, perhaps other switches.

Figure 1-4 illustrates a simple arrangement of the DCE, DTE, and DSE in the network. Later discussions will reveal that the configurations can be considerably more complex.

Network Topologies

Topologies and Design Goals

A network configuration is also called a *network topology.* A network topology is the shape (or the physical connectivity) of the network. The term topology is borrowed from geometry to describe the form of something. The network designer has three major goals when establishing the topology of a network.

- provide maximum possible reliability to assure proper receipt of all traffic (alternate routing)
- route the traffic across the least-cost path within the network between the sending and receiving DTEs (although the least-cost route may not be chosen if other factors, such as reliability, are more important)
- give the end user the best possible response time and throughput.

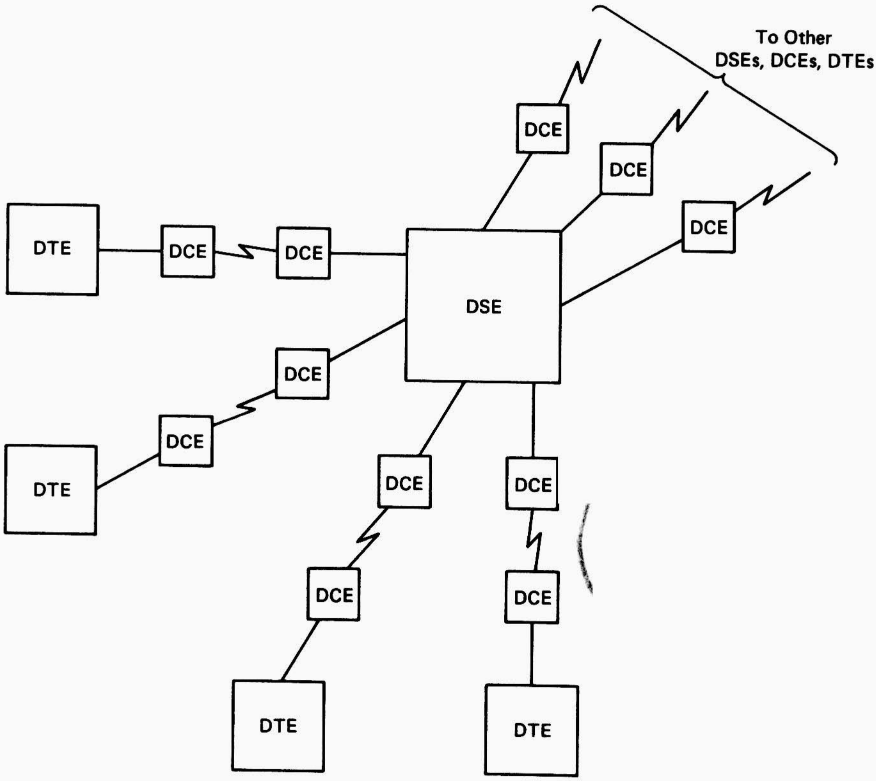

Figure 1-4. Data Switching Equipment (DSE)

Network reliability refers to the ability to deliver user data correctly (without errors) from one DTE to another DTE. It entails the ability to recover from errors or lost data in the network, including channel, DTE, DCE, or DSE failure. Reliability also refers to the maintenance of the system, which includes day-to-day testing; preventive maintenance, such as relieving faulty or failing components of their tasks; and fault isolation in the event of problems. When a component creates problems, the network diagnostic system can pinpoint the error readily, isolate the fault, and perhaps isolate the component from the network.

The second major goal in establishing a topology for the network is to provide the least-cost path between the application processes residing on the DTEs. This involves:

1. minimizing the actual length of the channel between the components, which usually entails routing the traffic through the fewest number of intermediate components;