McGraw-Hill Series on Computer Communications

# TCP/IP and Related Protocols

## Third Edition

Uyless Black

# TCP/IP 及其相关协议

## 第 3 版

McGraw-Hill Book Co.
世界图书出版公司

# TCP/IP and Related Protocols

**Uyless Black**

Third Edition

世界图书出版公司
McGraw-Hill Book Co

书　　名：TCP/IP and Related Protocols　3rd ed.
作　　者：U. Black
中 译 名：TCP/IP 及其相关协议　第 3 版
出 版 者：世界图书出版公司北京公司
印 刷 者：北京中西印刷厂
发　　行：世界图书出版公司北京公司（北京市朝阳门内大街 137 号　100010）
开　　本：1/32　850 × 1168　　　印　张：13
版　　次：1999 年 3 月第 1 版　　1999 年 3 月第 1 次印刷
书　　号：7-5062-4120-X/ TP·41
版权登记：图字 01-99-0015
定　　价：62.00 元

**Library of Congress Cataloging-in-Publication Data**

Black, Uyless D.
TCP/IP and related protocols
Uyless Black.—3rd ed.
p. cm.—(McGraw-Hill series on computer communications)
Includes index.
ISBN 0-07-913282-0
1. TCP/IP (Computer network protocol) 2. Computer network protocols. I. Title. II. Series.
TK5105.585.B54 1998
004.6'2—dc21 97-33278
CIP

*McGraw-Hill*

*A Division of The McGraw-Hill Companies*



1 2 3 4 5 6 7 8 9 0 DOC/DOC 9 0 2 1 0 9 8 7

ISBN 0-07-913282-0

*The sponsoring editor for this book was Steven M. Elliot, the editing supervisor was Paul R. Sobel, and the production supervisor was Tina Cameron. It was set in Century Schoolbook by Donald A. Feldman of McGraw-Hill's Professional Book Group composition unit.*



IE ISBN: 0-07-116109-0

# Preface

This third edition of *TCP/IP* reflects changes that have occurred in this protocol suite during the past few years, as well as changes that have occurred in other parts of the industry that have a bearing on TCP/IP. The changes deal with (a) the release of a new verison of IP, called IPv6, (b) the growth of Frame Relay networks and the need to run on Frame Relay systems, and (c) the entry of ATM into the data communications area, and the potential need to interwork IP with ATM.

As with the first two editions, this book remains as a user guide, and a general tutorial on the TCP/IP protocol suites.

The title of this book for the first two editions included a subtitle, *Related Protocols*. Although no longer in the book's title, the book maintains the same approach of explaining, not just TCP/IP, but other complementary protocols that make up the TCP/IP protocol suite.

I hope you find this book a useful addition to your library. If you have questions, or would like to communicate with me, I can be reached at 102732.3535@compuserve.com.

Uyless Black

# Acronyms

| | |
|---|---|
| AAL | ATM adaptation layer |
| ABM | asynchronous balanced mode |
| ACK | positive acknowledgment |
| ACSE | association control service element |
| AFI | authority format identifier |
| ARP | address resolution protocol |
| AS | autonomous system |
| ASN.1 | *Abstract Syntax Notation One* |
| ATM | Asynchronous transfer mode |
| AUI | attachment unit interface |
| BECN | backward explicit congestion notification |
| BER | basic encoding rules |
| BGP | border gateway protocol |
| B-ISDN | broadband ISDN |
| BIU | bus interface unit |
| BOOTP | bootstrap protocol |
| CATV | coaxial cable TV |
| CDIR | Classless Inter-Domain Routing |
| CEF | congestion encountered forward |
| CEI | Connection endpoint identifier |
| CLNP | connectionless network protocol |
| CLP | cell loss priority |
| CMIP | common management information protocol |
| CMOT | common management information services and protocol over TCP/IP |
| CRC | cyclic redundancy check |

| | |
|---|---|
| CSMA/CD | carrier sense, multiple access collision detection |
| DA | destination address |
| DE | discard eligibility |
| DIC | data network identification code |
| DLCI | data link connection identifier |
| DNS | domain name system |
| DPA | destination physical address |
| DQDB | distributed-queue-dual-bus |
| DSAP | destination service access point |
| DSP | domain-specific part |
| DSU | digital service unit (also data service unit) |
| DTE | data terminal equipment |
| DTP | data transfer protocol |
| EGP | external gateway protocol |
| EOT | end of transmission |
| FCS | frame-check sequence |
| FDDI | fiber distributed data interface |
| FNC | Federal Networking Council |
| FTAM | file transfer and access management |
| FTP | file transfer protocol |
| GGP | gateway-to-gateway protocol |
| GOES | geosynchronous orbiting environmental satellite |
| GOSIP | U.S. Government OSI Profile |
| HDLC | high-level data link control |
| HMP | host monitoring protocol |
| I | information (field) |
| I/G | individual/group (address bit) |
| IAC | interpret as command |
| ICI | intercarrier interface |
| ICMP | internet control message protocol |
| IDI | initial domain identifier |
| IDP | initial domain part |

| | |
|---|---|
| IEEE | Institute of Electrical & Electronics Engineers |
| IETF | Internet Engineering Task Force |
| IGMP | Internet group management protocol |
| IGP | internal gateway protocol |
| IMIB | Internet Management Information Base |
| IMP | interface message processor |
| INOC | Internet Network Operations Center |
| IP | internet protocol |
| IPDU | internetwork PDU |
| IPX | Internet packet exchange protocol |
| IRTF | Internet Research Task Force |
| ISDN | integrated services digital network |
| ISO | International Standards Organization |
| ISS | initial send sequence |
| ITU-T | International Telecommunication Union—Telecommunication Standardization Sector |
| IVD | integrated voice/data |
| IWU | internetworking unit |
| LAN | local area network |
| LAPB | link access protocol, balanced |
| LCN | Logical channel number |
| LCP | link control protocol |
| LLC | logical link control |
| LPP | lightweight presentation protocol |
| LSAP | link service access point |
| LSDU | link service data unit |
| MAC | media access control |
| MAN | metropolitan area network |
| MAU | medium attachment unit |
| MCF | MAC convergence function |
| MDI | medium dependent interface |
| MIB | Management Information Base |
| MTU | maximum transmission unit |
| MX | mail exchange |
| NAK | negative acknowledgment |

| | |
|---|---|
| NCC | network control center |
| NFS | network file system |
| NIC | Network Information Center |
| NIST | National Institute of Standards and Technology |
| NLPID | network level protocol identifier |
| NN | national number |
| NNI | network-to-network interface |
| NS | name server |
| NSAP | network service access point |
| NTN | network terminal number |
| NTP | network time protocol |
| NVT | network virtual terminal |
| OSI | Open Systems Interconnection |
| OSPF | open shortest path first |
| OUI | organization unique identifier |
| PA | prearbitrated access |
| PAD | packet assembly and disassembly |
| PARC | Palo Alto Research Center |
| PCI | protocol control information |
| PDN | public data networks |
| PDU | protocol data unit |
| PID | protocol identifier |
| PL | physical layer |
| PLP | packet layer procedures |
| PMA | physical medium attachment |
| PNNI | Private Network-Network Interface |
| QA | queued-arbitrated (access) |
| QOS | quality of service |
| RARP | reverse address resolution protocol |
| RD | receive delay |
| RDN | relative distinguished name |
| RER | residual error rate |
| RFCs | Request for Comments |
| RIP | routing informaiton protocol |

| | |
|---|---|
| RN | relative name |
| RNR | receive not ready |
| ROSE | remote operations service element |
| RPC | remote procedure call |
| RR | resource record |
| RRQ | read request |
| RTT | round-trip time |
| SABM | set asynchronous balanced mode |
| SAP | service application point |
| SAR | segmentation and reassembly |
| SCCS | service specific convergence sublayer |
| SDH | synchronous digital hierarchy |
| SD | send delay |
| SDH | Synchronous Digital Hierarchy |
| SDLC | synchronous data link control |
| SEQ | sequence (field) |
| SFD | start frame delimiter |
| SGMP | simple gateway monitoring protocol |
| SIP | simple internet protocol |
| SMDS | Switched Multimegabit Data Service |
| SMI | Structure for Management Information |
| SNAP | subnetwork access protocol |
| SNMP | simple network management protocol |
| SOA | start of zone authority |
| SRI | Stanford Research Institute |
| SRTT | smoothed round trip time |
| SSAP | source service access point |
| STDM | statistical time division multiplexing |
| SVC | switched virtual circuit |
| TCB | transmission control block |
| TCP | transmission control protocol |
| TID | transfer identifier |
| TLV | type, length, value |
| TOS | type-of-service (field) |
| TP4 | transport protocol class 4 |
| TP0 | transport Protocol Class 0 |

| | |
|---|---|
| TTL | time-to-live |
| TUBA | TCP & UDP with bigger addresses |
| U/L | universal/local (bit) |
| UDP | user datagram protocol |
| UI | unnumbered frame |
| ULP | upper layer protocol |
| UNI | user-network-interface |
| UT | Universal Time |
| VCI | virtual circuit identifier |
| VCI | virtual channel identifier |
| VPI | virtual path identifier |
| WAIS | wide area information servers |
| WAN | wide area network |
| WEB | worldwide web |
| XDR | external data representation |
| XID | exchange identification |
| XNS | Xerox network system |

# Contents

Chapter

# 1

# TCP/IP and the Internet

Data communication networks were developed to allow users to share computer and information resources as well as a common communication system. As organizations have brought the computer into almost every facet of business, it has become obvious that a single network, while very useful, is inadequate to meet the information needs of businesses and individuals. A user of one network, for example, often needs to access and share the resources of computers and databases that "belong" to another network. Merging all resources into one network, however, is prohibitively complex and expensive.

In the late 1960s and early 1970s, networks were built so users residing on different networks could not share resources. Network administrators also were reluctant to allow users to tap into resources because of concerns about security as well as excessive usage of network resources. As a result, it was difficult for a user to extend the use of an information system to another user across networks. The networks were either incompatible with each other or were not allowed to communicate because of administrative problems.

During this time, many people began to think about sharing resources among user applications. To do so, however, network administrators had to agree upon a set of common technologies and standards so the networks could communicate with each other. Applications such as electronic mail and file transfer also needed to be standardized to permit interconnections of end-user applications.

In the early 1970s, several groups around the world began to address network and application compatibility. At that time the term *internetworking*, which means interconnecting computers and/or networks, was coined. The concepts of internetworking were pioneered by the International Telecommunication Union, Telecommunications Standardization Sector (ITU-T), the International Standards Organization (ISO), and especially the original designers of the

ARPANET. *ARPA* refers to the Advanced Research Projects Agency, which is a U.S. Department of Defense (DOD) organization.

In fairness to the pioneers of internetworking concepts (and layered protocols, discussed later in this chapter), the ARPA protocols were well in existence before the ISO and ITU-T took an interest in this important subject. The procurement for ARPANET took place in 1968, and the machines selected for this procurement were Honeywell 316 interface message processors (IMPs). The initial effort was contracted through Bolt Bernak & Newman (BBN), and the ARPANET nodes were initially installed at UCLA, University of California at San Bernardino; the Stanford Research Institute (SRI); and the University of Utah. The well-known request for comments (RFCs) resulted from this early work.

After the pioneer work of a group of talented and dedicated engineers, these initial efforts were organized through the ARPANET Network Working Group. The group was disbanded in 1971, and the Defense Advanced Project Research Agency (DARPA) assumed the work of the earlier organization. DARPA's work in the early 1970s led to the development of an earlier protocol, the network control program, later the Transmission Control Protocol and the Internet Protocol (TCP/IP). Two years later, the first significant parts of the Internet were placed into operation. At about this time, DARPA started converting some of its computers to the TCP/IP suite of protocols. DARPA required that all computers connected to ARPANET had to use TCP/IP by January 1, 1983.

TCP/IP was initially used to connect ARPANET, the Packet Radio Net (PRNET), and the Packet Satellite Net (SATNET). Most user computers were large mainframes with terminals attached through terminal access servers. As ARPANET grew, the Department of Defense decided to split it into two networks. The other network was named MILNET and was set up for military purposes. ARPANET continued to be used for its original intent: a network to support R&D (research and development) applications. By the mid-1980s, the "ARPA Internet" was called simply "the Internet." In 1990, the last original ARPANET node was taken out of commission.

Perhaps one of the most significant developments in TCP/IP was DARPA's decision to implement TCP/IP around the UNIX operating system. Of equal importance, the University of California at Berkeley was selected to distribute the TCP/IP code. Some implementors have said that releasing such complex and functionally rich code was a "license to steal." Whatever one's view on the matter, it was a very significant and positive move in the industry. Because the TCP/IP code was nonproprietary, it spread rapidly among universities, private companies, and research centers. Indeed, it has become the standard suite of data communications protocols for UNIX-based computers.

During this period, other networks using TCP/IP were being created, based on funding from the U.S. government and other research agencies. The NSFnet was established as a high-capacity network by the National Science Foundation and is still in existence. NSF has played a key role in the development of the Internet, both in funding and strategic guidance.

The National Science Foundation (NSF) operation was managed by the joint efforts of MCI, Sprintlink, and IBM by forming Advanced Network Services (ANS). Access to the Internet was provided at network access points (NAPs), which connected to other networks, both private and public.

In November 1994, the NSF informed colleges and other institutions to look for another feed into the Internet because the U.S. government was getting out of the public Internet business. Most of the feeds were existing networks that had interconnected into the NSF backbone. NSF announced it would provide some funding for a few more years, and in 1995 it started disconnecting its NAPs.

Today, the Internet is commercial, with the ISPs charging a fee for its customers to use the ISP access to the other ISPs and the many files and data bases available (many of which are free to use) throughout the world.

## Organization of the Internet

As the Internet grew, its organization and management were delegated to the Internet Advisory Board, or IAB (see Figure 1.1). Originally, the IAB consisted of a number of subsidiary organizations, but their main function was to coordinate the Internet task forces. In 1989, the task forces were placed into two major groups within the IAB: Internet Research Task Force (IRTF) and the Internet Engineering Task Force (IETF). The IRTF is responsible for ongoing research activities. The IETF concerns itself with tactical issues, such as implementation and engineering problems.

### Request for comments (RFCs)

The request for comments (RFCs), briefly mentioned earlier in this chapter, are technical notes on an internet protocol. They represent the documentation of the Internet.

Some RFCs are de facto standards for TCP/IP, others are published for informational purposes, and still others are the result of research and might eventually become future standards. Presently, over 1000 RFCs are in existence, although quite a number of these specifications have been superseded.

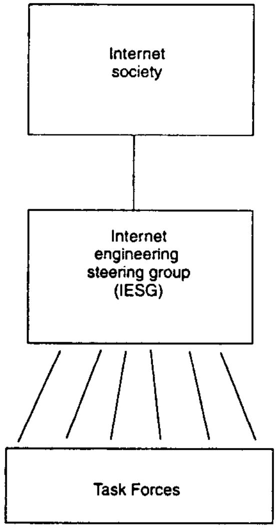


Figure 1.1 The Internet Organization.

## Obtaining Internet Information

The intent of this book is to provide a general understanding of TCP/IP. There is no substitute, however, for the actual source documents (the RFCs). You can obtain them from a wide variety of sources on the Internet if you have electronic access, through anonymous FTP or e-mail (check with your Internet access provider).

You can also send an e-mail message to rfc-info@isi.edu. No subject is required, and the body of the message should be "help: ways_to_get_rfcs."

All Internet IP network numbers and domain names are now maintained at the Internet name RS.INTERNIC.NET. In addition, RFC 1400 provides more details on how to obtain Internet information. As of this writing, the recommended starting point for the Web is http://isi.edu/rfc-editor.

## TCP/IP and the OSI

Before I move into some tutorial discussions of networking architectures, you should know that the use of TCP/IP and related protocols continues to grow, raising some interesting points to the Open

System Interconnection (OSI) Model. Many people believe that TCP/IP is a more viable approach for a number of reasons. First, TCP/IP is here and it works. Second, a wealth of products are available that use the TCP/IP protocol suites. Third, it has a well-founded, functioning administrative structure through the IAB. Fourth, it provides easy access to documentation. Fifth, it is used in many UNIX products.

Without going into a treatise on the subject of OSI, it should be emphasized that OSI-specific protocols have not seen much use in end-user machines, due principally to the success of TCP/IP. In high-end switches and large SONET-based networks, OSI is quite prevalent, especially the OSI-based network management protocol CMIP (Common Management Information Protocol).

The last point is that OSI has proven to be very successful as a model, which is its principle role. Many successful systems are based on the OSI model (Frame Relay, SONET, ATM, SS7, ISDN, and FDDI, to name a few).

## Internetworking Architecture

To grasp the operations of TCP/IP, several terms and concepts must first be understood. Once these concepts are explained, we can discuss the architecture more fully.

### Terms and concepts

The Internet uses the term *gateway* or *router* to describe a machine that performs relaying functions between networks. The preferred term today is *router*, but this book uses both terms in deference to past practice. Figure 1.2 shows a gateway placed between networks A, B, and C. (Routers and gateways are defined further in Chapter 2.)

As shown in Figure 1.2, networks A, B, and C are often called *subnetworks*. The term does not mean that they provide fewer functions than a conventional network. Rather, the three networks consist of a

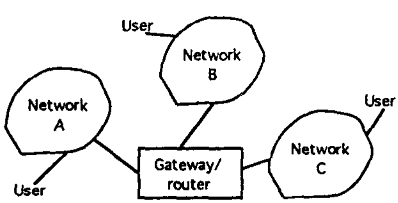


**Figure 1.2** The Gateway and Networks or Subnetworks.