



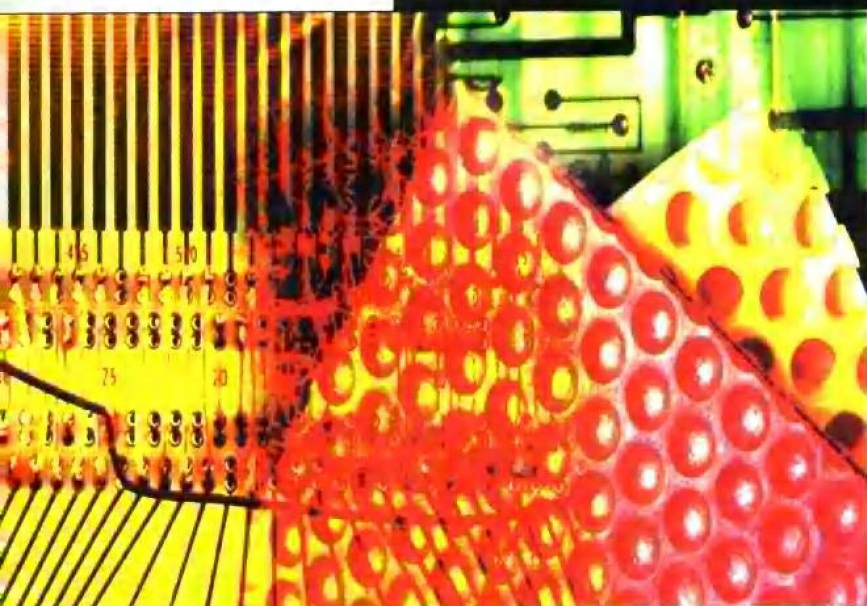
TCP/IP^{and} Related Protocols

(美) *Uyless Black* 著

良友翻译组 译

由经验丰富的专家撰写

TCP/IP 及相关协议



特别适合 TCP/IP 初学者阅读



机械工业出版社



CMP

计算机网络基础与应用系列丛书

TCP/IP 及相关协议

(美) Uyless Black 著

良友翻译组 译

机械工业出版社

本书是一本全面介绍 TCP/IP 及其发展的专著, 是这一领域的最新成果。本书的内容包括: TCP/IP 与互连网络、Internet 中的命名、编址及路由选择、网际协议、网际控制报文协议、TCP 和用户数据报协议、主要应用层协议、TCP/IP 与其他协议的协同操作、网络安全及 IP、帧中继和 ATM 等共 15 章。

本书既可作为一本用户指南, 同时也可以作为了解和学习 TCP/IP 协议的教程。

Uyless Black: TCP/IP and Related Protocols

Authorized translation from the English language edition published by McGraw-Hill

Copyright 1992 by McGraw-Hill

All rights reserved

本书中文简体字版由机械工业出版社出版, 未经出版者书面许可, 本书的任何部分不得以任何方式复制或抄袭。

版权所有, 翻印必究。

本书版权登记号: 图字: 01-98-0619

图书在版编目 (CIP) 数据

TCP/IP 及相关协议 / (美) 伯莱克 (Black, U.) 著; 良友翻译组译. - 北京: 机械工业出版社, 1998

(计算机网络基础与应用系列丛书)

书名原文: TCP/IP and Related Protocols

ISBN 7-111-06308-2

I.T… II.①伯… ②良… III.因特网-通信协议, TCP/IP IV.TP393.4

中国版本图书馆 CIP 数据核字 (98) 第 12002 号

出版人: 马九荣 (北京市百万庄大街 22 号 邮政编码 100037)

责任编辑: 温莉芳 王志刚

中国建筑工业出版社密云印刷厂印刷·新华书店北京发行所发行

1998 年 5 月第 1 版第 1 次印刷

787mm×1092mm $\frac{1}{16}$ ·16 印张

印数: 0 001-5 000 册

定价: 28.00 元

凡购本书, 如有倒页、脱页、缺页, 由本社发行部调换

目 录

第 1 章 TCP/IP 与互连网络	1	2.2.4 选择分组路径	16
1.1 Internet 的组织	2	2.3 帧中继	16
1.2 网络标准请求说明 (RFC)	2	2.4 信元交换 (信元中继)	17
1.3 获取 Internet 信息	2	2.5 局域网 (LAN)	17
1.4 TCP/IP 和 OSI	3	2.5.1 局域网的组成	17
1.5 互连网络系统结构	3	2.5.2 局域网的类型	17
1.5.1 术语和概念	3	2.6 载波侦听/冲突检测	18
1.5.2 无连接和面向连接的协议	4	2.6.1 CSMA/CD 和 IEEE 802.3	18
1.5.3 评价无连接和面向连接的系统	5	2.6.2 令牌环	19
1.6 Internet 层	5	2.7 逻辑链路控制子层 (LLC)	20
1.7 层操作举例	6	2.8 中继器、网桥、路由器和网关	20
1.8 设备驱动程序	7	2.9 小结	23
1.9 近观 TCP/IP 模型	7	第 3 章 Internet 中的命名、编址及路由选择	24
1.10 端口和套接字简介	8	3.1 上层、网络层、数据链路层及物理层的名字和地址	24
1.11 互连网络的挑战	8	3.1.1 物理地址 (物理层地址)	24
1.11.1 不同的网络所使用的 PDU 的长度不同	8	3.1.2 通用的物理地址和协议标识符	25
1.11.2 各个子网的定时器值、超时值和重试值不同	8	3.1.3 CSMA/CD 帧和 MAC 物理地址	26
1.11.3 网络可以使用不同的寻址习惯	8	3.1.4 链路层地址 (LSAP)	27
1.11.4 网络具有不同的性能	9	3.1.5 对 LSAP 信头的扩展 (SNAP)	27
1.11.5 网络可以采取不同的路由选择方法	9	3.2 拨号进入 Internet	28
1.11.6 网络可以要求不同类型的用户接口	9	3.2.1 网络层地址	30
1.11.7 网络可以要求不同的安全级别	9	3.2.2 物理和网络地址的解析	30
1.11.8 差错的定位、诊断及网络维护因网而异	9	3.2.3 是否真的需要两种地址	31
1.12 典型的互连网络拓扑结构	9	3.2.4 上层协议 (ULP) 标识符和名字	31
1.13 小结	11	3.2.5 完整的命名、编址和标识操作	32
第 2 章 网络、网桥、网关、路由器介绍	12	3.3 IP 地址的结构	34
2.1 基本分类	12	3.4 私用 IP 地址的分配	35
2.2 广域网 (WAN)	13	3.5 目的地址和路由选择	36
2.2.1 电路交换	13	3.5.1 直接和间接目的端	36
2.2.2 报文交换	14	3.5.2 IP 路由选择逻辑	37
2.2.3 分组交换	14	3.6 IP 地址和 MAC 地址的关系	38
		3.7 地址解析的问题	39
		3.8 地址解析协议 (ARP)	39
		3.9 代理 ARP	42

3.10 反向地址解析协议 (RARP)	42	5.4.6 路由记录选项	74
3.11 BOOTP	43	5.4.7 时间戳选项	74
3.12 Internet 服务提供者 (ISP)	44	5.4.8 路径跟踪	75
3.13 IP 和 X.121 地址映射	44	5.4.9 分段与重组	76
3.13.1 数据网络的国际编号方案 (X.121)	44	5.5 IP 地址和路由选择表	79
3.13.2 DDN IP 地址	45	5.5.1 地址表	79
3.14 X.121 和 IP 地址的映射	45	5.5.2 路由选择表	80
3.15 子网、子网编址和地址映射	45	5.6 再谈 IP 路由选择和子网划分	81
3.15.1 子网掩码	46	5.7 IP 服务定义和原语	82
3.15.2 子网掩码举例	47	5.7.1 IP/ULP 原语	82
3.15.3 广播地址和其他的唯一性地址	48	5.7.2 IP/SNP 原语	83
3.16 无类域间路由选择 (CIDR)	50	5.7.3 其他的 IP/SNP 服务定义	83
3.17 IPv6 和地址解析	51	5.7.4 网络层/LLC 原语	84
3.18 小结	51	5.8 多址传输	84
第 4 章 域名系统	52	5.9 网际组管理协议 (IGMP)	85
4.1 域名系统结构	52	5.10 IP 和 X.25	86
4.2 域名 (Domain Name)	54	5.11 IP、帧中继和 ATM	86
4.3 顶级域	54	5.12 对 IP 的其他观点	87
4.4 域名解析和域名到地址的映射	55	5.13 无类域间路由选择 (CIDR)	87
4.5 名字服务器操作	56	5.14 无连接方式网络服务 (ISO 8473)	87
4.6 资源记录	57	5.14.1 ISO 8473 PDU	87
4.7 RDATA 字段	58	5.14.2 服务质量功能	88
4.8 对 DNS 类型的解释	59	5.14.3 协议功能	89
4.9 IN-ADDR-ARPA	60	5.14.4 子网间的通信管理	89
4.10 DNS 报文	64	5.14.5 网间路由选择	90
4.11 资源记录压缩	65	5.14.6 CLNP 和 IP	91
4.12 有关名字的更详细的信息	65	5.15 IPv6	91
4.12.1 NSLOOKUP	65	5.15.1 IPv4 的问题	91
4.12.2 谁使用这个名字	66	5.15.2 IPv6 的设计宗旨	91
4.12.3 如何获得一个名字	66	5.15.3 IPv6 地址	92
4.12.4 BIND	66	5.15.4 分层地址	92
4.13 小结	66	5.15.5 特殊地址	92
第 5 章 网际协议 (Internet 协议)	67	5.15.6 IPv6 数据报	93
5.1 IP 协议的主要特征	67	5.15.7 IPv6 与 IPv4 的比较	94
5.2 IP 和网络	67	5.15.8 IPv6 扩展报头	94
5.3 IP 数据报	68	5.16 小结	96
5.4 主要的 IP 服务	72	第 6 章 网际控制信报文协议	97
5.4.1 网际报头检查例程	72	6.1 ICMP 报文格式	97
5.4.2 IP 源路由选择	73	6.2 ICMP 差错和状态报告过程	98
5.4.3 路径记录	73	6.2.1 超时	98
5.4.4 路由选择操作	73	6.2.2 参数不可理解	98
5.4.5 松散的和严格的路由选择	74	6.2.3 报宿不可到达	99

6.2.4 报源抑制	99	8.1.2 基于最少站点(距离向量)的 路由选择	130
6.2.5 回声请求和应答	100	8.1.3 基于链路状态度量的路由选择...	131
6.2.6 重定向	100	8.1.4 核心网关与非核心网关	132
6.2.7 时间戳和时间戳应答	101	8.1.5 外部网关和内部网关	133
6.2.8 信息请求和应答	101	8.1.6 边界路由器和边缘路由器	133
6.2.9 地址掩码请求和应答	101	8.1.7 自治系统之间或区之间如何交换 信息	134
6.2.10 路由器通告和请求	102	8.1.8 谁参与交换路由选择信息和交换 路由选择逻辑	136
6.3 对 ICMP 的其他考虑	102	8.1.9 澄清术语	136
6.4 IPv6 和 ICMP	102	8.2 边缘网关协议	136
6.5 小结	103	8.2.1 BGP 操作	137
第 7 章 传输控制协议和用户数据报 协议	104	8.2.2 BGP 报文	137
7.1 运输层值	104	8.3 内部网关协议	138
7.2 TCP 概述	104	8.4 域内路由选择协议 (IDRP)	138
7.3 TCP 的主要特点	105	8.5 路由选择信息协议 (RIP)	139
7.4 再谈端口和套接字	106	8.5.1 RIP 定时器	139
7.4.1 端口分配和端口捆扎的例子	107	8.5.2 RIP 操作例子	140
7.4.2 使用套接字(socket)支持多路复 用	108	8.5.3 RIP 问题	141
7.5 被动打开和主动打开	108	8.5.4 RIP 报文格式	142
7.6 传输控制块	109	8.5.5 请求与响应	142
7.7 TCP 窗口和流量控制机制	109	8.6 RIP 的修订版 (RIP-2)	142
7.8 重发操作	111	8.6.1 RIP 与 OSPF	143
7.9 估算定时器的超时和重发时间	112	8.6.2 RIP 和 IPX	143
7.10 TCP 和用户接口	113	8.7 Hello 协议	143
7.11 数据片	115	8.7.1 Hello 报文格式	144
7.12 TCP 连接管理操作	117	8.7.2 Gated	144
7.12.1 TCP 打开	117	8.8 距离向量协议总结	144
7.12.2 TCP 数据传输	120	8.9 使用最短路径算法选择最优路径	144
7.12.3 TCP 关闭	121	8.10 开放最短路径优先协议	148
7.12.4 TCP 连接表	122	8.10.1 OSPF 操作	148
7.13 TCP 对大量数据和交互数据的处理...	123	8.10.2 路由器的分类	149
7.14 TCP 和 IP 报头的压缩	124	8.10.3 通告类型	150
7.15 潜在的绕回问题	125	8.10.4 OSPF 最短路径举例	150
7.16 使用 TCP 的其他考虑	126	8.10.5 OSPF 数据结构	152
7.17 用户数据报协议 (UDP)	126	8.10.6 OSPF 分组	155
7.18 IPv6 和 TCP/IP	127	8.10.7 OSPF 与 RIP	161
7.19 小结	128	8.11 小结	162
第 8 章 路径发现协议	129	第 9 章 主要的应用层协议	163
8.1 术语和概念	129	9.1 Telnet 协议	163
8.1.1 距离向量协议和链路状态度 量协议	129	9.1.1 网络虚拟终端	163
		9.1.2 Telnet RFC	164

9.1.3 Telnet 命令	166	10.5.3 高层 MIB 定义	195
9.1.4 Telnet 命令举例	166	10.6 SNMP	196
9.2 Rlogin	167	10.6.1 SNMP 管理关系	196
9.3 平凡文件传输协议 (TFTP)	167	10.6.2 SNMP 操作举例	197
9.3.1 TFTP 和其他协议	168	10.6.3 SNMPv1 和 SNMPv2	198
9.3.2 TFTP 分组	168	10.6.4 SNMP PDUs	198
9.4 文件传输协议 (FTP)	170	10.6.5 SNMP MIB 被管理对象	199
9.4.1 数据类型	170	10.7 CMOT	199
9.4.2 FTP 命令与应答	171	10.7.1 CMOT 层	200
9.4.3 FTP 会话的操作顺序	172	10.7.2 轻量级表示协议 (LLP)	201
9.4.4 FTP 操作举例	173	10.8 小结	201
9.4.5 文件检索举例	174	第 11 章 TCP/IP 与其他协议的协	
9.5 简单邮件传输协议 (SMTP)	174	同操作	202
9.5.1 SMTP 模型	175	11.1 最小 TCP/IP 和局域网协议栈	202
9.5.2 地址字段格式	176	11.2 协议栈对操作系统的依赖性浅析	202
9.5.3 SMTP 操作举例	176	11.3 位于 LLC 之上的 TCP/IP	203
9.5.4 SMTP 和域名系统 (DNS)	177	11.4 使用 UDP 代替 TCP	203
9.6 X Window	177	11.5 在 TCP 或 UDP 之上的 NetBIOS	204
9.6.1 X Window 系统协议	178	11.6 IP 位于 NetBIOS 之上	205
9.6.2 显示连接	179	11.7 XNS 位于 IP 之上	206
9.7 远程过程调用 (RPC)	180	11.8 IP 路由器协议栈	206
9.8 网络文件系统	180	11.9 IP 和 LAN 网桥的关系	207
9.9 远程执行守护进程	181	11.10 IP 和 X.25	208
9.10 PING	181	11.10.1 IP、X.25 和局域网	208
9.11 宿主监督协议	181	11.10.2 IP、X.25 和公共数据网	
9.12 丢弃协议	182	(PDNs)	208
9.13 Finger	182	11.10.3 IP、X.25 和业余分组无线	
9.14 网络时间协议	183	传输	209
9.15 Daytime	184	11.11 在 UDP/IP 网络中使用 IPX	209
9.16 点到点协议	184	11.12 在 IPX 网络中传输 802 LLC 数据	210
9.17 获得 Internet 服务	186	11.13 在 FDDI 网络中传输 IP 数据报	210
9.18 万维网 (WWW)	186	11.14 交换式多兆位数据服务上的 IP	211
9.19 浏览工具包 (Netscape)	187	11.15 OSI 运输协议类 0 位于 TCP 之上	212
9.20 Java	187	11.16 OSI 无连接协议位于 UDP 之上	212
9.21 小结	188	11.17 TCP/IP 位于 ISDN 之上	213
第 10 章 Internet 网络管理协议	189	11.18 TCP/IP 位于帧中继或 ATM 之上	213
10.1 Internet 网络管理标准概论	189	11.19 小结	214
10.2 Internet 网络管理的分层体系结构	190	第 12 章 网络安全	215
10.3 Internet 命名体系结构	190	12.1 私有和公开密钥加密系统	215
10.4 管理信息结构	191	12.1.1 私有密钥加密	215
10.5 管理信息库	192	12.1.2 公开密钥加密	215
10.5.1 对象组概况	192	12.2 PGP 程序	216
10.5.2 对象描述模板	194		

12.3	SNMPv2 的安全性	217	14.1	操作和系统结构	230
12.4	IPv6 和网络安全	218	14.2	IP 和帧中继/ATM 之间的关键区别	230
12.5	MD5	218	14.2.1	帧中继和 ATM 方法	230
12.6	密钥的分发	218	14.2.2	IP 方法	231
12.7	小结	218	14.3	协议栈	231
第 13 章	TCP/IP 和操作系统	219	14.4	其他接口	232
13.1	UNIX 和 TCP/IP	219	14.4.1	接口处所提供的服务	232
13.1.1	面向连接的服务	219	14.4.2	专用网络 - 网络接口 (PNNI)	233
13.1.2	其他输入/输出调用	221	14.5	帧中继和 ATM 网络中的 IP 操作	234
13.1.3	数据报服务	221	14.6	比较 IP、帧中继和 ATM	236
13.1.4	关闭一个连接	222	14.7	IP 与帧中继和 ATM 虚电路的协作	237
13.1.5	其他的系统调用	222	14.7.1	传输	239
13.1.6	程序调用 UNIX TCP/IP 服务的例子	222	14.7.2	接收	240
13.2	PC 接口程序	222	14.8	使用 ARP 实现 IP 地址、帧中继和 ATM 虚电路间的映射	240
13.2.1	通过 SMTP 发送邮件	226	14.9	IP 交换	241
13.2.2	通过 TFTP 发送一个文件	226	14.10	小结	241
13.2.3	通过 FTP 发送文件	227	第 15 章	管理方面的考虑	242
13.2.4	使用 tn 调用 Telnet 登录服务	228	15.1	厂商的 Internet 产品策略	242
13.2.5	检索网络统计信息	228	15.2	对 TCP/IP 和 OSI 协议栈的简单比较	243
13.2.6	使用 PING 以获得回声服务	228	15.2.1	IP 和 CLNP	244
13.2.7	使用 route 命令操作 IP 路由选择表	229	15.2.2	TP4 和 TCP	245
13.3	小结	229	15.3	小结	246
第 14 章	IP、帧中继和 ATM	230			

第 1 章 TCP/IP 与互连网络

数据通信网络的发展，不仅使用户可以相互通信，同时也允许用户共享网络的计算机和信息资源。随着计算机的应用几乎渗透到社会的各个角落，尽管单一网络的作用不可低估，但显然已不能满足企业及个人的信息需求。例如，一个网络的用户常常需要访问和共享属于其他网络的计算机和数据库资源。但把所有的资源放在一个网络中会产生很高的复杂性，价格也惊人的昂贵。

在 60 年代末、70 年代初期，不同网络的用户无法共享资源。网络管理人员出于安全性和成本等方面的考虑，也不愿让用户使用某些资源。结果，用户在网络上将一个信息系统提供其他人使用会很困难。而网络则由于管理方面的问题，互不兼容或不允许相互通信。

那时，许多人已经开始考虑在多个用户间共享资源。要达到这一目的，网络管理人员必须在一系列共同的技术和标准方面达成一致，以便网络能够相互通信。各种应用，如电子邮件和文件传输，也必须进行标准化，以允许在端用户的应用程序间建立连接。

在 70 年代初，国际上有几个工作小组开始谈到网络和应用的兼容性问题。那时，网络互连 (internetworking) 这个术语被创造了出来，它的含义是相互连接的计算机和网络。网络互连的概念是由国际电信联盟 (International Telecommunication Union, ITU) 的标准化部 (ITU-T)、国际标准化组织 (ISO) 和 ARPANET 最初的设计人员所提出来的。ARPA 指的是高级研究计划署，它是美国国防部 (DOD) 的一个组织。

在 ISO 和 ITU-T 对网络互连产生兴趣之前，ARPA 的协议就已经很好地应用了。ARPANET 从 1968 年开始付诸实施，选择 Honeywell 316 作为接口信息处理机 (IMP)，Bolt Berak & Newman (BBN) 进行了最初的 ARPANET 的工作。ARPANET 的节点最初被安装在加利福尼亚大学洛杉矶分校 (UCLA)、加利福尼亚大学圣贝纳迪诺分校、斯坦福研究院 (Stanford Research Institute) 和犹他大学 (University of Utah)。

在这群才华卓著、勤奋投入的工程师们所做的开创性工作之后，这些工作由 ARPANET 网络工作组进行组织。该小组解散于 1971 年。DARPA (国防部高级研究计划署) 承担了早期的组织工作。70 年代初期，DARPA 所做的工作导致了一个早期协议的发展，即网络控制程序，后来成为传输控制协议和网际协议 (TCP/IP)。两年后，Internet 的最重要的部分首先投入运行。同时，DARPA 开始将网上的部分计算机转向 TCP/IP 协议。1983 年 1 月 1 日，DARPA 要求所有与 ARPANET 连接的计算机都要使用 TCP/IP 协议。

TCP/IP 最初用于连接 ARPANET、PRNET (分组广播网) 和 SATNET (分组卫星网)。用户所使用的计算机大多数还是大型机，用户终端通过终端访问服务器与计算机相连。随着 ARPANET 的不断发展，美国国防部决定将它分成两个网络，另一个网络叫 MILNET，专用于军事目的。ARPANET 继续发挥原来的作用，即作为支持科研和发展 (R&D) 的一项应用。80 年代中期，ARPA 互连网络被称为 Internet。到了 1990 年，最初的 ARPANET 节点已不再使用了。

DARPA 决定在 UNIX 系统中实现 TCP/IP 也许对 TCP/IP 的发展最为重要。同样重要

的还有加利福尼亚大学伯克利分校公开了 TCP/IP 代码。一些开发人员认为，公开这些功能丰富、复杂性高的代码，就等于给偷窃者颁发许可证。不管当时人们对此事的看法如何，这一举动在业界产生了重大而积极的影响。因为 TCP/IP 代码并非私有财产，它在大学、私营公司及研究机构中迅速传播。实际上，它已成为基于 UNIX 的计算机用于数据通信的标准协议。

在此期间，由于美国政府及其他科研机构的资助，使用 TCP/IP 的其他网络也诞生了。国家科学基金会建成了大容量的 NSFnet，该网络已使用至今。NSF 不仅在资金方面，而且在策略方面对 Internet 的发展都起到了关键的作用。

MCI、Sprintlink 和 IBM 三家公司共同组成了高级网络服务机构 (ANS)，管理 NSF 的运作。通过网络访问点 (NAP) 连接包括专用网络和公共网络在内的其他网络，从而提供对 Internet 的访问。

1994 年 11 月，由于美国政府不再参与 Internet 公共部分的事务，NSF 通知多所大学和研究机构寻找其他的 Internet 入口。大多数的入口来源于同 NSF 主干网相连的现有网络。NSF 宣布将在几年内提供一定的资金。从 1995 年开始逐步断开网络访问点 (NAP)。

今天，Internet 已经商业化了。Internet 服务提供商 (ISP) 向自己的客户提供对全世界其他 ISP 及许多文件、数据库（许多是免费使用的）的访问并负责收取费用。

1.1 Internet 的组织

随着 Internet 的发展，它的组织和管理工作由 Internet 建议委员会 (IAB) 承担 (见图 1-1)。最初，IAB 包括许多附属机构，它们的主要作用是协助 Internet 任务组的各项工作。1989 年，Internet 任务组划分成 IAB 下的两个工作组：Internet 研究任务工作小组 (IRTF) 和 Internet 工程任务工作小组 (IETF)。IRTF 负责进行 Internet 未来发展方面的研究，而 IETF 主要从技术实现和工程问题方面进行研究。

1.2 网络标准请求说明 (RFC)

在本章的前面曾提过 RFC (Request for Comments)，每个 RFC 是一个对 Internet 的请求的技术说明。它们代表了有关 Internet 的文档。

一些 RFC 最终成为 TCP/IP 的标准，而其他一些作为技术信息发表，还有一些 RFC 是某些研究项目的成果，很有可能成为未来的标准。目前已有超过 1000 个 RFC 请求，尽管其中已有许多规范被淘汰了。

1.3 获取 Internet 信息

本书的目的在于使读者对 TCP/IP 有个基本的了解，但更详细的技术问题则需要参考 RFC 原文档。如果你的计算机连接 Internet 的话，你可以通过匿名 FTP 或 e-mail 从网上得到这些资料。

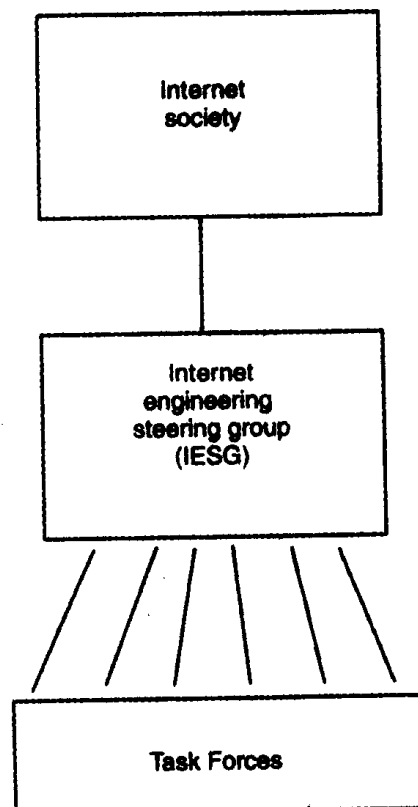


图 1-1 Internet 组织

你也可以给地址 `rfc-info@isi.edu` 发送电子邮件，无需标题 (Subject)，信体部分写明 “help: way-to-get-rfc”。

目前，所有 Internet 网络 IP 地址及域名由名字为 RS.INTERNIC.NET 的计算机维护。另外，RFC1400 提供了如何获取 Internet 信息的详细介绍，推荐的 Web 站点是 `http://isi.edu/rfc-editor`。

1.4 TCP/IP 和 OSI

在讨论网络系统结构之前，由于 TCP/IP 及相关协议使用的越来越多，正给开放系统互连 (OSI) 模型带来一些有趣的问题。由于种种原因，许多人相信 TCP/IP 更为可行。因为，首先，TCP/IP 已经存在并且工作得很好；第二，使用 TCP/IP 协议的产品很丰富；第三，有一个健全、有效的管理机构；第四，TCP/IP 有许多容易获得的文档资料；第五，许多 UNIX 产品都在使用 TCP/IP。

OSI 专门的协议并未广泛应用在端用户 (end-user) 的机器上，主要是因为 TCP/IP 取得了巨大的成功。在高端网络交换机和较大的基于 SONET 的网络上，OSI 是很流行的，特别是基于 OSI 的网络管理协议 CMIP (公共管理信息协议)。

最后一点，OSI 已被证实是一个很好的模型，这也是它的主要作用。许多成功的系统都是基于 OSI 模型的 (如帧中继、ATM、SS7、ISDN 和 FDDI)。

1.5 互连网络系统结构

要掌握 TCP/IP 的操作，首先要了解一些术语和概念。在解释了这些概念之后，我们将更全面地讨论系统结构。

1.5.1 术语和概念

在 Internet 中使用“网关” (gateway) 或“路由器” (router) 来描述在网络间完成中继功能的机器。虽然现在更多地使用路由器这个术语，为了尊重过去的习惯，本书中仍采用这两种术语。图 1-2 显示了网络 A、B、C 中间的网关 (第 2 章中将给出路由器和网关的定义)。

如图 1-2 所示，网络 A、B 和 C 常被称为子网 (Subnetwork)。子网并非意味着它比传统的网络功能要少，相反，这三个网络都是逻辑上完整的网络，并且每个子网都对网际互连发挥着作用。换句话说，许多子网共同构成了互连网络。

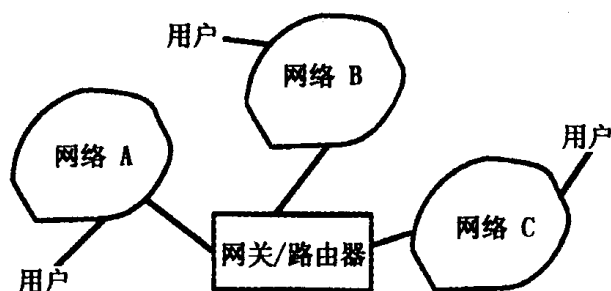


图 1-2 网关和网络（或子网）

互连网络中的网关对端用户的应用来说是透明的。实际上，端用户的应用位于与网络相连的主机上，很少在网关上运行用户应用，主要由于以下原因：第一，网关无需背上应用层协议的包袱。当无需实现应用层协议时，网关可以专注于较少的任务（如，管理网络间的数据传输）。它不必关心诸如数据库访问、电子邮件和文件管理等应用层的功能。第二，网关因此而可以支持任何类型的数据应用。因为网关把应用信息看成是透明的协议数据单元 (PDU)。

除了实现应用层的透明性之外，大多数的网络设计者试图使网关对子网以及子网对网关都具有透明性。也就是说，网关不必关心它所连接的网络的类型。网关的主要任务是接受一个包含足够寻址信息的 PDU，从而使网关为该 PDU 选择到其目的地或到达下一网关的路径。透明性具有吸引力的另一个原因是由于它在一定程度上使网关模块化，这意味着网关可以用在多种不同的网络中。

1.5.2 无连接和面向连接的协议

无连接和面向连接操作的概念对任何通信协议来说都是基础。在 Internet 的许多标准中，两个概念都被用到，因此对它们各自的特点应当有清楚的了解。它们最主要的特征如下：

面向连接操作：在传输数据前，用户和网络间首先建立起一个逻辑连接。然后，通过该用户/网络连接传输数据单元，被传输的数据单元间保持着某种关系。

无连接操作：在数据传输前，用户和网络间不建立逻辑连接，每个数据单元都作为一个独立的单元传输。

面向连接的服务要求在两个端用户及服务提供者（如，网络）之间达成三方协议。它同时也允许通信各方协商某些选项和服务质量（QOS）。连接建立后，各方都存储其他方的信息，如地址和 QOS 特征。一旦数据传输开始，协议数据单元（PDU）无需携带许多协议控制信息（PCI）。因为数据报中不包含进行动态路由选择所需的足够的地址，所以这种方法常用于网络中固定路由选择。对于面向连接的服务，通信一方只需提供一个简短的标识符，其他各方通过查询表格，就可以确定完整的地址及服务质量（QOS）等信息。因为通信参数可以进行协商，所以通信各方无需事先了解其他方的各方面的特性。如果无法按要求提供某种服务，任何一方可以提出一个级别更低的服务进行协商或简单地拒绝连接请求。

以往，面向连接的服务为所有的数据单元（除个别例外）提供确认。如果传输过程中出了问题，面向连接的协议提供了重新传输出错数据单元的机制。除此之外，大多数面向连接的协议还保证了被传输的数据按正确的顺序到达目的地。现在，面向连接的服务可以不再包括一些服务（如顺序控制和通信量的可计算性）。在最近出现的系统（如帧中继）中，并不自动提供通信管理的功能，而是有其他实体（如，位于端用户工作站的用户程序）提供。图 1-3 总结了面向连接的网路的特点。

- 网络中的连接映射
- 缩写的编址
- 网络间常常采用固定的路由选择
- 提供或不提供统计功能

图 1-3 面向连接的网路

无连接服务在管理用户的 PDU 时，把每个 PDU 作为一个独立的实体。连续传输的数据间无需保持关系，网络中用户到用户的通信记录也很少被保持，不进行任选项的协商，也不建立和维护用于数据传输的表格。在少数系统中，通信实体必须就如何通信事先达成一致，并预先确定服务质量（QOS）。而在更多的情况下，为每个被传输的 PDU 单独提供 QOS，而且每个 PDU 中包含一些字段以标识服务的类型和级别。

从理论上讲，无连接的网路能实现数据完整性支持功能，如肯定确认（ACKS）、否定确认（NAKS）和顺序控制。实际上，大多的无连接系统不提供这些功能。从本质上说，无连接的服务能实现如下功能：

■ 对子网中的特定协议的高度独立性

- 子网间的较强的独立性
- 子网对用户特定协议的高度独立性

无连接的网络比面向连接的网络可能更具有健壮性，因为每个 PDU 都是作为一个独立的实体处理的。因此，数据单元可以采取不同的路径以避免网络中的故障节点和拥塞节点。与面向连接的协议相比，无连接协议的协议数据单元（PDU）的头信息较长，对 PDU 中用户数据部分的比例也大，开销花费也较大。无连接协议的特点如图 1-4 所示。

- 不存在端到端的映射
- 数据单元使用完整的地址
- 可选择其它路径
- 不具备或有限的可统计性

图 1-4 无连接服务

在结束无连接和面向连接的协议话题之前，我们应当知道，实际网络中的管理人员应了解一个系统中面向连接和无连接层的各自的优点，最后的选择应由终端用户所需的服务类型及获得该服务的系统开销和资金花费来决定。因为多数厂商提供众多的面向连接和无连接产品，你只需根据自己需要的服务类型进行选择。如果你最关心的是数据的无错传输，你必须保证在系统之中有一个能妥善处理任何通信情况的协议。

1.5.3 评价无连接和面向连接的系统

许多互连网络是无连接的，很少支持或不支持顺序控制和确认功能。大多数局域网属于这一类（如 IP）。而 TCP 却是一个面向连接并提供了多种支持数据完整性功能的协议。我们注意到一些较新的技术，如帧中继（Frame Relay）和异步传输方式（ATM）都正在使用面向连接的技术，主要原因之一是因为面向连接的技术可以在 PDU 中使用较短的信息头，如缩写的标识符（一个虚电路标识），“位”的开销较小，从而提高操作效率。在本书中，我将更详细地将帧中继和 ATM 同 IP 协议作对照。

1.6 Internet 层

在 TCP/IP 网络中，软件和硬件都常常包含了大量的支持通信行为的功能。网络的设计人员要处理这些数量繁多、复杂性高的功能，其任务是艰巨的。为了解决这些问题，网络设计人员通过功能分层（layering the functions）的办法使互连网络结构化。

虽然现代网络建立了七个概念层，互连网络的体系结构却分为四层。图 1-5 描述了互连网络分层体系结构。互连网络模型的底层包含子网，因此被称为子网层（Subnetwork Layer），它包括子网接口。这些子网允许数据在每个网络内传递。子网可以是 X.25、帧中继、ATM 和以太局域网（LAN）。虽然该层包括一个子网，而在实际实现时，所有与子网或网关进行通信的机器中都必须有数据链路层和物理层。图 1-5 只是一个抽象，因为本层必须也包括数据链路层和物理层。从后面的图

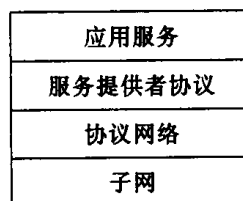


图 1-5 Internet 的各层

子网层的上面为互连网络层，该层提供了将众多的网络和网关连成一体的功能。该层负责将数据从源端传输到目的端。它包含了 IP 协议和 Internet 控制报文协议（ICMP）。稍后会讨论到，其他一些协议（如支持路由选择和地址映射的协议）也同 IP 协议一起在该层中实现。

第三层被称为服务提供者协议层（Service Provider Protocol Layer）。该层负责实现端到端的通信。如果是面向连接的通信，它将提供可靠性措施及处理网间流量控制的机制。该层

包含 TCP 和用户数据报协议 (UDP)。

最上面一层被称为应用服务层 (Application Service Layer)，该层支持与端用户应用的直接接口。互联网络的应用负责诸如文件传输、远程终端访问 (Remote Terminal Access)、远程作业执行 (Remote Job Execution) 和电子邮件等。该层包含了几个广泛应用的协议，如文件传输协议 (FTP)。

1.7 层操作举例

图 1-6 给出了子网和网关与分层协议的关系。这些层与图 1-5 中描述的不同，低层改为数据链路和物理层，在上层中也使用国际更为常用的术语。本图中，假定主机 A 中的用户发送一个应用 PDU 给主机 B 中的应用层协议 (如文件传输系统)。文件传输软件执行各种功能并将文件记录传回主机 A。在许多系统中，主机 B 的操作被称为服务器操作，主机 A 的操作被称为客户操作。

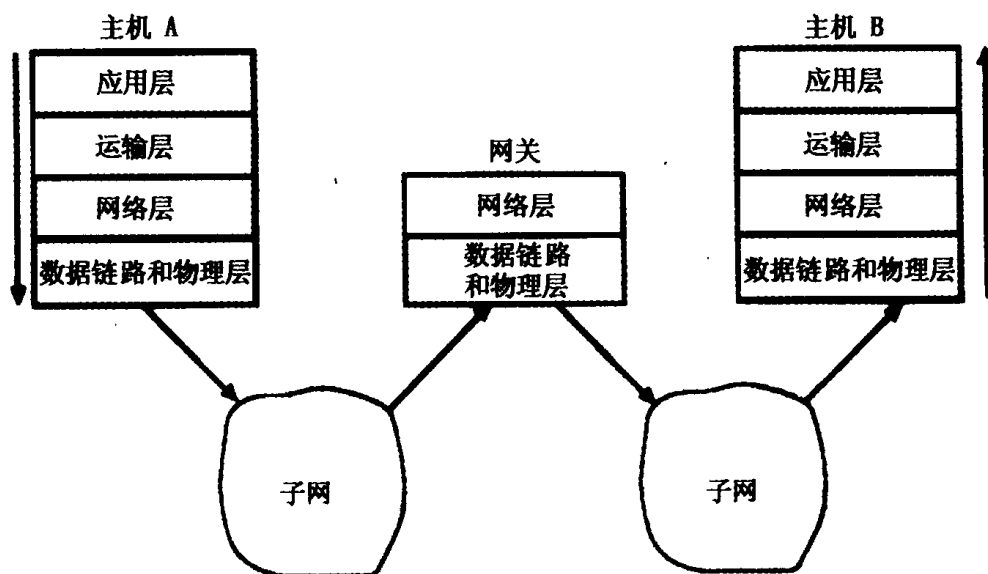


图 1-6 互联网络层操作的例子

在图 1-6 的主机 A 一方，正如协议栈中的向下箭头所示，数据单元被传送给运输层协议。该层执行了许多操作 (在后面章节中讨论)，并给应用层传送来的 PDU 添加一个信息头，此时，该数据单元被称为一个数据片 (Segment)，来自上层的 PDU 被看成运输层数据。然后，运输层将数据传递给网络层 (也称为 IP 层)，该层也执行一定的服务并添加一个信息头。现在，这个数据单元被称为数据报，并传给更低一层。在数据链路层，数据被装配了数据链路层信息头和尾，这时的数据单元被称为帧，并由物理层发送到网络上。当然，对于主机 A 收到主机 B 的数据，其数据在每层处理过程和层间传送的方向刚好与上述过程相反。

互联网络协议不了解网络中发生的一切，网络管理人员也无需对 PDU 进行操作和管理，在一些情况下，当 PDU (包括信息头和数据) 传送经过子网时并未发生任何改动。在图 1-6 中，PDU 出现在网关中并被低层处理，然后传给 IP 层。在 IP 层中，要根据主机提供的目的地址进行路由选择。

选择了路径之后，PDU 被传送给与适当的子网 (包括底层) 相连的通信链路。PDU 被重新封装成数据链路层的帧并传给下一个子网。该数据单元透明地经过子网，最终到达目的端主机。

目的主机（主机 B）通过低层接收到主机 A 发来的数据单元，并执行与主机 A 相反的处理过程。它要在各层中将该层的信息头和数据进行分离。信息头用于决定该层应采取何种处理。因此，信息头主宰了每层所进行的操作。

主机 A 的应用服务层中的文件传输应用所产生的 PDU 被最终传递给了主机 B 中的文件传输应用。如果主机 A 和主机 B 是大型计算机，这两个应用很可能是同一软件。该应用根据收到的头信息来决定要执行的操作。收到的数据也可以再传给主机 B 中的另一个应用，但是，在许多情况下，主机 A 的用户仅希望获得某个服务器协议（如文件传输或电子邮件）所提供的服务。在这种情况下，无需在主机 B 中调用端用户应用进程。

当主机 B 中的服务程序将查询到的数据返回给主机 A 中的客户程序时，处理过程正好相反。在主机 B 中，数据由上层往下传给低层，通过网络、网关和下一个网络，到达主机 A，再由低层逐层向上层传递，最后到达端用户应用。

1.8 设备驱动程序

TCP/IP 并不关心低层（第一层和第二层，特别是第一层）的一个原因是存在一个被称为设备驱动程序的低层软件，其目的在于将上层协议（操作系统）与通信链路中的物理接口隔离开来。每个硬件设备必须同其相应的设备驱动程序协同工作，设备驱动程序负责操作它们所管辖的硬件。

网络设备驱动程序位于介质访问子层（MAC）并管理网络硬件（在第 2 章中讨论），它常被称为介质访问子层驱动程序（MAC Driver）。因此，IP 可在这些驱动程序之上操作而不必关心网络硬件的每个细节。若需了解更详细的信息，您可以研究一下 FTP 软件公司开发的包驱动程序。

1.9 近观 TCP/IP 模型

图 1-7 描述了 TCP/IP 模型及几个主要的相关协议。具体使用何种协议取决于网络用户的需求和网络设计人员的要求。现在你看到某些协议特别是 IP 和 TCP 已经在前面讨论中解释过。位于 TCP（和 UDP）之上的，是一些应用服务层协议的例子，也就是图 1-5 中所示的应用服务。低两层代表数据链路层和物理层。正如图 1-7 所示，具体实现时可供选择的标准和协议很多，该图也在本书中后面的章节中被用来详细论述每层中的操作。

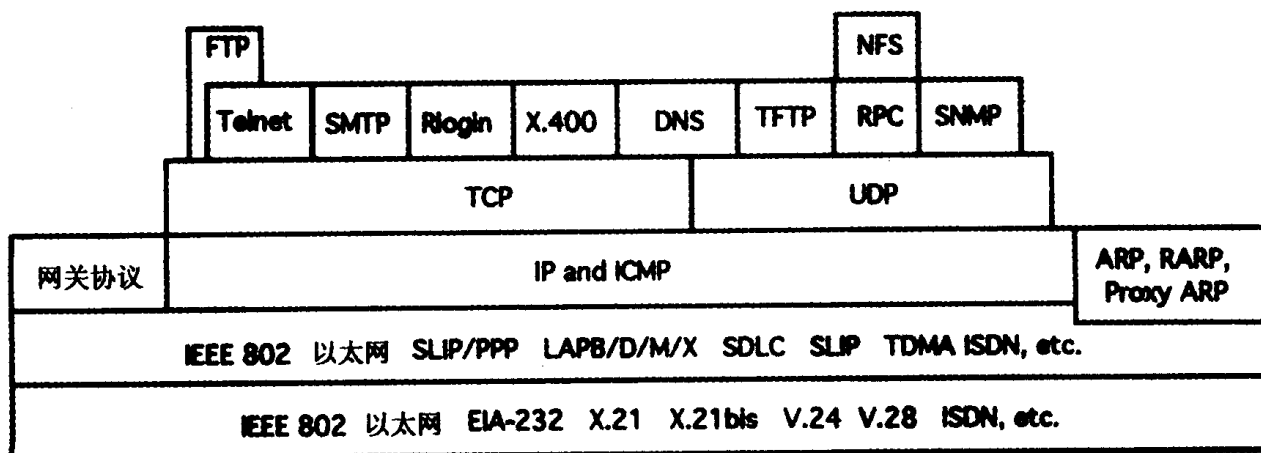


图 1-7 TCP/IP 协议（部分协议）

1.10 端口和套接字简介

使用 TCP/IP 协议的每个应用层进程必须用一个“端口号”来标识自己。两台相互通信的主机使用端口号确定接收到的数据应交给哪个应用程序处理。使用端口号也提供了多路复用的功能，即允许多个用户程序同时与一个应用程序通信。每个端口号指定一个应用程序的入口，这个概念与 OSI 模型中的服务访问点（SAP）极为相似。

除了端口之外，TCP/IP 协议还使用另一个称为套接字（socket）的标识符。套接字来源于 4.3BSD UNIX 系统中的网络输入/输出操作。使用套接字收发网络数据，很类似于 UNIX 系统中使用文件标识符对文件的访问。

在互联网中，一些端口号预先分配给专门的服务程序，这些端口被称为众所周知端口（Well-Known Ports），众所周知的端口所使用的端口范围从 1 到 1023。在以后的章节中我们将更详细地讨论端口和套接字。

1.11 互连网络的挑战

有了以上的背景知识，我们就可以继续讨论网络管理人员在提供互联网络服务时会遇到的各种情况和问题。特别是有关 TCP/IP 的讨论将贯穿本书的始终。现在，我们先讨论一些一般性问题。

1.11.1 不同的网络所使用的 PDU 的长度不同

如果各网络或网关使用的 PDU 长度不同，它们必须提供将数据单元分段的功能，而且必须保持对数据单元的标识。即使数据单元很长，也必须保证数据有序地到达目的地。在第 5 章，我们将了解到 IP 网关具有将数据单元分段的功能，而接收端主机具有将各个分段重组成一个完整的 PDU 的功能。

1.11.2 各个子网的定时器值、超时值和重试值不同

假定网络 A 转发一个数据单元时设定一个等待确认定时器。该定时器保证在限定时间内收到数据单元的接收方发来的端到端的确认。若数据单元传送至网络 B，但是该网络没有端到端的定时器，这就产生了问题。网络 A 是否应该发送一个确认帧，告诉发送方用户已经把数据传送给网络 B 并且假定网络 B 已继续传送数据单元呢？这样作显然并不安全，因为数据单元可能并未到达目的地。TCP/IP 提供了端到端的定时功能，但数据单元经子网传输时并未使用此功能，而是只有主机方才使用。因此，网络可以任意设计所需要的类型的定时器。

1.11.3 网络可以使用不同的寻址习惯

例如，一个网络可能使用逻辑名称寻址，而另一个网络可能使用物理地址寻址。在这种情况下，两个网络的地址解析和映射可能会不同。实际上，大多数网络使用专门的网络地址，例如，SNA 的地址与 DECnet 的地址不同。幸运的是，有一些有关几种类型地址的互连网络标准。另外，一些互连网络系统能够支持网络层地址到物理层地址的映射，其他一些系统可以根据具有用户友好性的名字获得网络地址。

1.11.4 网络具有不同的性能

一个网络同另一个网络相比可能速度较慢，延迟的时间较长，吞吐量较小。互连网络协议并非特别关心这些问题。正如在本章前半部分了解的那样，协议的设计使它可以在不同类型的网络上操作。当然网络管理人员应当关注这些问题，并使用互连网络上的一些工具软件解决这些问题（以后会讨论到 TCP 在这方面很有价值的工具）。

1.11.5 网络可以采取不同的路由选择方法

一个网络可能使用固定的路由选择表，而另一个网络则可能使用具有自适应能力的路由选择表。对于前者，网络的重新排序逻辑很少，对于后者，重新排序的逻辑是很复杂的。TCP/IP 协议族包含许多支持路由选择的协议，尽管已有许多组织在网络中使用互连网络协议，但 TCP/IP 协议族不涉及特定组织的路由选择方法。

1.11.6 网络可以要求不同类型的用户接口

一个网络可能采用面向连接的用户-子网接口，而另一个网络可能使用无连接的数据报协议。接口的类型影响到错误的恢复和流量控制。正如你将在本书的几个章节中看到的那样，互连网络无连接和面向连接协议给网络管理者带来了很具有挑战性的问题（和机遇）。

1.11.7 网络可以要求不同的安全级别

例如，一个网络可能要求对传输的数据进行加密，而另一个网络可能只支持纯文本传输。在 1990 年以前，TCP/IP 除允许用户为传输的数据指定安全级别外别无其他安全性措施。安全性因不同的网络选择方式而异。随着时代的发展，安全性已经成为互连网络中的一个主要问题。

1.11.8 差错的定位、诊断及网络维护因网而异

一个网络中产生的问题可能会波及另一个网络，而受到影响的网络可能缺乏错误分析及纠正的功能。互连网络协议实现了网络管理信息在网间的透明交换。互连网络管理标准的目的在于为网络管理提供有效的支持。

虽然网络互连的任务很复杂，需要进行大量的分析和预测，然而困难并非不可克服，你将会看到提供 TCP/IP 的厂商和标准化组织已经开发并实现了许多有效的网络互连技术。

1.12 典型的互连网络拓扑结构

图 1-8 显示了由一个网关（G）连接的两个网络，网络由网络地址来标识。图左边的网络的地址为 11.4，另一个网络的地址是 128.1。该图中使用了“网络云”（Network Cloud）的术语，因为网络的拓扑结构及操作是未知的。该图中各网络的拓扑结构在下图中描述。

在某些情况下（当无需讨论网络内部的操作时），常常使用网络云进行描述（如图 1-8a 所示）。在另一些情况下涉及到网络内部操作，图 1-8b 描述了一个传统的交换网络（网络 11.4）的拓扑结构及以太网（网络 128.1）的拓扑结构。网络 11.4 中名为 A、B、C 和 D 的方框表示与通信链路相连的分组交换机。它们也可以是同时具有主机和分组交换机功能的计算机。网络 128.1 中的方框代表主机、工作站、路由器或文件/打印服务器。图 1-8b 中的两