

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

TCP/IP 网络编程

TCP/IP Programming

任泰明 编著

- 原理知识与实用技能融为一体
- 全面介绍Winsock网络程序设计技术
- 避免使用晦涩难懂专业术语，语言浅显易懂



高校系列

21世纪高等学校计算机规划教材

21st Century University Planned Textbooks of Computer Science

TCP/IP 网络编程

TCP/IP Programming

任泰明 编著



高校系列

人民邮电出版社

北京

图书在版编目 (CIP) 数据

TCP/IP网络编程 / 任泰明编著. —北京: 人民邮电出版社, 2009. 8
21世纪高等学校计算机规划教材
ISBN 978-7-115-21054-8

I. T… II. 任… III. ①计算机网络—通信协议—高等学校—教材②计算机网络—程序设计—高等学校—教材
IV. TN915.04 TP393.09

中国版本图书馆CIP数据核字 (2009) 第125551号

内 容 提 要

TCP/IP 是 Internet 和 Intranet 中计算机或相关设备之间进行“交流”的协议, Winsock 是应用最为广泛的, 基于 TCP/IP 的网络程序的编程接口。

本书以通俗易懂的语言详细介绍了 TCP/IP 及其工作原理, 以简单明了的编程实例全面介绍了基于 Winsock 的网络程序设计技术。具体内容主要包含 TCP/IP 的结构与工作原理、网络接口层基本知识、网络层技术及其协议、传输层技术及其协议、应用层常用协议、TCP/IP 的实现技术、网络程序设计基本知识、TCP 编程、UDP 编程、网络综合程序设计、Winsock 常用函数介绍及其应用等内容。

本书可作为大专院校“网络协议及其网络编程”课程的教材使用, 也可作为网络相关工作技术人员的参考书或作为 IT 培训机构网络技术方面的培训教程使用。

21 世纪高等学校计算机规划教材

TCP/IP 网络编程

-
- ◆ 编 著 任泰明
责任编辑 刘 博
 - ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京楠萍印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 20.5
字数: 534 千字
印数: 1—3 000 册
 - 2009 年 8 月第 1 版
2009 年 8 月北京第 1 次印刷

ISBN 978-7-115-21054-8/TP

定价: 35.00 元

读者服务热线: (010)67170985 印装质量热线: (010)67129223
反盗版热线: (010)67171154

出版者的话

现今社会对人才的基本要求之一就是应用计算机的能力。在高等学校,培养学生应用计算机的能力,主要是通过计算机课程的体制改革,即计算机教学分层、分类规划与实施;密切联系实际,恰当体现与各专业其他课程配合;教学必须以市场需求为导向,目的是培养高素质创新型人才。

人民邮电出版社经过对教学改革新形势充分的调查研究,依据目前比较成熟的教学大纲,组织国内优秀的有丰富教学经验的教师编写一套体现教学改革最新形势的“高校系列计算机教材”。在本套教材的出版过程中,我社多次召开教材研讨会,广泛听取了一线教师的意见,也邀请众多专家对大纲和书稿做了认真的审读与研讨。本套教材具有以下特点。

1. 覆盖面广, 突出教改特色

本套教材主要面向普通高等学校(包括计算机专业和非计算机专业),是在经过大量充分的调研基础上开发的计算机系列教材,涉及计算机教育领域中的所有课程(包括专业核心骨干课程与选修课程),适应了目前经济、社会对计算机教育的新要求、新动向,尤其适合于各专业计算机教学改革的特点特色。

2. 注重整体性、系统性

针对各专业的特点,同一门课程规划了组织结构与内容不同的几本教材,以适应不同教学需求,即分别满足不同层次计算机专业与非计算机专业(如工、理、管、文等)的课程安排。同时本套教材注重整体性的策划,在教材内容的选择上避免重叠与交叉,内容系统完善。学校可根据教学计划从中选择教材的各种组合,使其适合本校的教学特点。

3. 掌握基础知识, 侧重培养应用能力

目前社会对人才的需要更侧重于其应用能力。培养应用能力,须具备计算机基础理论、良好的综合素质和实践能力。理论知识作为基础必须掌握,本套教材通过实践教学与实例教学培养解决实际问题的能力和知识综合运用能力。

4. 教学经验丰富的作者队伍

高等学校在计算机教学和教材改革上已经做了大量的工作,很多教师在计算机教育与科研方面积累了相当多的宝贵经验。本套教材均由有丰富教学经验的教师编写,并将这些宝贵经验渗透到教材中,使教材独具特色。

5. 配套资源完善

所有教材均配有 PPT 电子教案,部分教材配有实践教程、题库、教师手册、学习指南、习题解答、程序源代码、演示软件、素材、图书出版后要更新的内容等,以方便教与学。

我社致力于优秀教材的出版,恳请大家在使用的过程中,将发现的问题与提出的意见反馈给我们,以便再版时修改。

人民邮电出版社

前言

在 Internet 普及的今天,作为 Internet 工作基础的 TCP/IP 及其编程已经成为 IT 从业人员所要具备的基本知识与技能。打开国内外各大知名网站的招聘页面,都可以看到类似于“熟悉 TCP/IP、掌握 socket 通信开发”的要求。本书就是为了满足读者这方面知识的需求而编写的一本 TCP/IP 与基于 TCP/IP 编程方面的书籍。

本书有以下几个方面的特点。

(1) 内容的组织上按照协议原理与协议编程分为上、下两篇。上篇主要介绍 TCP/IP 簇中的常用协议,下篇专门介绍网络编程知识与技能。

(2) 在编写每一节的内容时将原理知识与实用技能融为一体,方便读者学习。

(3) 考虑到 TCP/IP 比较抽象,学习起来有一定的难度,所以全书尽量避免使用晦涩难懂的专业术语,而用浅显易懂的语言说明问题,努力将本书打造成一本人人都能读懂的书籍。

(4) 初学网络程序设计的人员,往往感到网络程序设计内容多,学习起来比较复杂。针对这一问题,本书在讲解网络程序设计时,根据网络程序固有的特点,先总结了网络程序设计的通用模式,然后再举例说明网络程序的设计,使读者易于入手。

(5) Winsock 函数内容多,使用起来比较复杂,针对这一问题,笔者在写作时将常用的 Winsock 函数分散到各种实例中去介绍,在最后一章将所有常用的 Winsock 函数一一作了较为详细说明,并在每个函数后面加入了其应用实例或使用说明。

本书分为上、下两篇,上篇包含 6 章,各章主要内容如下。

第 1 章:介绍了 TCP/IP 的产生、结构和工作原理,同时还简要介绍了 ISO/OSI RM 知识。

第 2 章:介绍 TCP/IP 层次结构中网络接口层包含的内容,主要有物理层和数据链路层的相关知识。

第 3 章:介绍 TCP/IP 层次结构中网络层及其相关知识。主要内容有 IP 数据报格式、IP 层的功能、IP 地址、ICMP 协议、地址转换协议以及 IP 的最新版本 IPv6 等。

第 4 章:介绍 TCP/IP 层次结构中传输层及其相关知识。主要内容有端口的概念, TCP 和 UDP 的协议数据格式、协议原理, TCP 与 UDP 的比较等内容。

第 5 章:介绍 TCP/IP 层次结构中应用层及其相关知识。主要内容有应用层常用协议 DNS、FTP、Telnet、HTTP、POP 和 SMTP 的格式、工作原理、协议实例等。

第 6 章:简要介绍了 TCP/IP 在 Windows 和 Linux 操作系统下的实现原理以及 TCP/IP 的二进制代码。

下篇包含以下 5 章内容。

第 7 章:介绍了网络程序设计有关的基础知识和一个网络程序入门实例,以及 Winsock 中编写网络程序常用的建立连接、传输数据、关闭连接等有关的函数。

第 8 章: 介绍了 TCP 程序设计流程、基于 C/C++ 的 TCP 程序设计实例和基于 Java 技术的 TCP 程序设计实例。

第 9 章: 介绍了 UDP 程序设计流程、基于 C/C++ 的 UDP 程序设计实例和基于 Java 技术的 UDP 程序设计实例。

第 10 章: 介绍了使用 MFC 中提供的有关类进行网络程序设计的知识。

第 11 章: 介绍了 Winsock API 中各种函数的功能, 并举例说明了一些函数的使用方法。

由于作者水平有限, 错漏之处在所难免, 欢迎广大读者提出宝贵意见。

任泰明

2009 年 4 月

目 录

上篇 TCP/IP

第 1 章 TCP/IP 概述.....2

1.1 TCP/IP 的产生与发展.....2

1.1.1 TCP/IP 的产生.....2

1.1.2 TCP/IP 的发展与应用.....4

1.2 TCP/IP 的体系结构.....4

1.2.1 TCP/IP 的分层体系结构与 协议栈的概念.....4

1.2.2 网络接口层.....5

1.2.3 网际层.....6

1.2.4 传输层.....6

1.2.5 应用层.....7

1.3 TCP/IP 的工作原理.....7

1.3.1 使用 TCP/IP 的 Internet 网络结构.....7

1.3.2 TCP/IP 通信模型.....8

1.3.3 TCP/IP 中数据的封装与解封过程.....10

1.3.4 TCP/IP 的操作系统边界与地址边界.....12

1.4 ISO/OSI 与 TCP/IP 的关系.....13

1.4.1 ISO/OSI 网络体系结构概述.....13

1.4.2 ISO/OSI 与 TCP/IP 的对应关系 分析.....15

1.5 与 TCP/IP 有关的组织与管理机构.....16

1.5.1 网络协议的国际标准化组织.....17

1.5.2 Internet 的管理机构.....17

1.5.3 RFC 文档.....19

小结.....20

习题.....20

第 2 章 网络接口层.....22

2.1 物理层.....22

2.1.1 数据传输介质.....22

2.1.2 数据传输方式.....24

2.1.3 数据编码技术.....28

2.1.4 信道复用技术.....30

2.2 数据链路层.....32

2.2.1 数据链路层的概念.....32

2.2.2 数据链路的流量控制机制.....33

2.2.3 差错控制.....33

2.2.4 点到点协议.....35

2.3 局域网技术.....36

2.3.1 局域网的组成.....36

2.3.2 IEEE 802 局域网技术.....37

2.3.3 以太网技术.....38

2.3.4 VLAN.....39

小结.....40

习题.....40

第 3 章 互联网络层.....41

3.1 网络层的互联技术.....41

3.1.1 网络互联概述.....41

3.1.2 路由器.....42

3.1.3 面向连接的互联技术.....43

3.1.4 无连接的互联技术.....43

3.1.5 IP 互联原理.....43

3.2 IP.....44

3.2.1 IP 概述.....44

3.2.2 IP 数据报格式.....45

3.2.3 IP 数据报的分片与重组.....48

3.2.4 IP 数据报选项.....50

3.3 IP 地址.....52

3.3.1 IP 地址的表示格式	53	4.3.2 TCP 连接的建立与关闭	85
3.3.2 IP 地址的分类	53	4.3.3 TCP 的流量控制和拥塞控制机制	86
3.3.3 子网的划分	54	4.3.4 TCP 的超时重发机制	89
3.3.4 特殊 IP 地址	56	4.4 TCP 与 UDP 的比较	89
3.4 Internet 控制报文协议	57	小结	90
3.4.1 ICMP 报文的封装与格式	57	习题	90
3.4.2 ICMP 差错报文	59		
3.4.3 ICMP 控制报文	60		
3.4.4 ICMP 请求与应答报文	61		
3.5 IP 数据报路由选择与路由协议	63		
3.5.1 路由表	63		
3.5.2 路由选择算法	64		
3.5.3 路由表的建立与刷新	65		
3.5.4 路由选择协议	66		
3.6 地址转换协议	67		
3.6.1 ARP	68		
3.6.2 RARP	69		
3.7 IPv6	70		
3.7.1 IPv6 产生的背景	70		
3.7.2 IPv6 的新特征	71		
3.7.3 IPv6 数据报格式	72		
3.7.4 IPv6 地址	72		
小结	74		
习题	74		
第 4 章 传输层	76	第 5 章 应用层	92
4.1 传输层的通信与端口	76	5.1 网络工作模式	92
4.1.1 点到点通信与端到端通信	76	5.1.1 C/S 模式	92
4.1.2 端口的概念与功能	77	5.1.2 B/S 模式	93
4.1.3 常用的端口	78	5.1.3 B/S 模式与 C/S 模式的比较	95
4.2 用户数据报协议	79	5.2 域名系统	96
4.2.1 UDP 概述	79	5.2.1 域名和域名系统	97
4.2.2 伪首部与校验和	80	5.2.2 域名服务器	99
4.3 传输控制协议	81	5.2.3 域名解析实例	99
4.3.1 TCP 报文段格式	81	5.3 远程登录协议	100
		5.3.1 远程登录概述	100
		5.3.2 Telnet 的工作原理	101
		5.3.3 网络虚拟终端的概念	102
		5.3.4 Telnet 选项协商	104
		5.3.5 Telnet 协议选项协商实例	105
		5.4 电子邮件原理及其协议	108
		5.4.1 TCP/IP 下 E-mail 的工作原理	108
		5.4.2 SMTP 原理与工作实例	110
		5.4.3 POP3 协议原理与工作实例	112
		5.4.4 电子邮件报文的格式与实例	115
		5.4.5 多用途因特网邮件扩展	115
		5.5 超文本传输协议	117
		5.5.1 万维网的工作过程	117
		5.5.2 超文本传输协议与应用实例	118
		5.6 文件传输协议	121
		5.6.1 FTP 简介	121
		5.6.2 FTP 支持的文件类型和文件结构	122

5.6.3 FTP 的工作原理	123	6.1.4 网际层	133
5.6.4 FTP 命令和应答	124	6.1.5 传输控制协议	136
5.6.5 FTP 工作实例	127	6.1.6 用户数据报协议	137
小结	128	6.2 UNIX/Linux 的 TCP/IP 实现	137
习题	128	6.2.1 Linux TCP/IP 网络协议栈	137
第 6 章 TCP/IP 的实现	130	6.2.2 Linux 网络数据处理流程	138
6.1 Windows 的 TCP/IP 实现	130	6.2.3 Linux 的 IP 路由	139
6.1.1 Windows TCP/IP 提供的 功能与服务	130	6.3 TCP/IP 协议代码	140
6.1.2 Windows TCP/IP 体系结构	131	6.3.1 TCP/IP 报文时序模型	140
6.1.3 网络接口层	132	6.3.2 报文代码及其分析	142
		小结	146
		习题	146

下篇 网络编程技术

第 7 章 网络程序设计入门	148	WSASocket)	163
7.1 有关网络程序设计的一些概念	148	7.4.3 地址绑定 (bind)	165
7.1.1 网络应用软件	148	7.4.4 服务器端监听连接 (listen)	167
7.1.2 网络中两个通信程序如何 相互识别对方	148	7.4.5 客户端请求连接 (connect 或 WSAConnect)	168
7.1.3 Windows Sockets 介绍	149	7.4.6 服务器端接受连接 (accept 或 WSAAccept)	171
7.1.4 套接口的概念	150	7.5 Winsock 中的数据传输函数及其应用	172
7.1.5 套接口编程原理	151	7.5.1 有连接的数据发送 (send 或 WSASend)	173
7.1.6 网络字节顺序	153	7.5.2 有连接的数据接收 (recv 或 WSARecv)	175
7.2 网络程序工作模型	153	7.5.3 无连接的数据接收 (recvfrom 或 WSARecvfrom)	177
7.2.1 网络程序要考虑的几个问题	153	7.5.4 无连接数据发送 (sendto 或 WSASendTo)	179
7.2.2 网络程序工作模型	155	7.6 Winsock 中网络连接与套接口的 关闭函数及其应用	181
7.3 一个简单的客户机/服务器程序	156	7.6.1 关闭读写通道	181
7.3.1 程序要求	156	7.6.2 关闭套接口	182
7.3.2 服务器程序	156	7.6.3 终止使用 Winsock	182
7.3.3 客户端程序	159	小结	183
7.3.4 实例运行结果	160		
7.4 Winsock 中建立连接的函数及其应用	161		
7.4.1 加载协议栈 (WSAStartup)	161		
7.4.2 创建套接口 (socket 或			

习题	183	10.1 使用 MFC 类库设计网络程序	216
第 8 章 TCP 程序设计	184	10.1.1 MFC 类库简介	216
8.1 TCP 网络程序设计模型	184	10.1.2 MFC 中与网络程序设计 有关的类	217
8.1.1 服务器端工作流程	184	10.1.3 MFC 中的 Windows 套接口类	218
8.1.2 客户端工作流程	185	10.1.4 客户端程序实例	224
8.1.3 服务器与客户机通信过程	186	10.1.5 服务器端程序实例	228
8.1.4 基于 TCP 的网络程序框架	186	10.2 使用 WinInet 设计客户端程序	230
8.2 TCP 实例程序 1	187	10.2.1 WinInet 程序设计知识	230
8.2.1 实例程序说明	187	10.2.2 实例程序说明	249
8.2.2 服务器程序	188	10.2.3 创建应用程序操作界面	249
8.2.3 客户端程序	190	10.2.4 应用程序代码及其说明	251
8.3 TCP 实例程序 2	191	10.3 使用原始套接口设计网络程序	258
8.3.1 程序要求	191	10.3.1 原始套接口程序设计简介	259
8.3.2 服务器程序	192	10.3.2 实例说明	259
8.3.3 客户端程序	195	小结	266
小结	198	习题	266
习题	198	第 11 章 Winsock 主要函数及其 用法	268
第 9 章 UDP 程序设计	200	11.1 数据格式转换函数	268
9.1 UDP 网络程序设计模型	200	11.1.1 htonl()和 WSAHtonl()	268
9.1.1 UDP 程序工作流程	200	11.1.2 htons()和 WSAHtons()	269
9.1.2 程序框架	201	11.1.3 ntohl()和 WSANtohl()	270
9.2 UDP 实例程序 1	202	11.1.4 ntohs()和 WSANtohs()	270
9.2.1 实例程序说明	202	11.2 IP 地址转换函数及其应用	271
9.2.2 服务器程序	203	11.2.1 inet_addr()	271
9.2.3 客户端程序	205	11.2.2 inet_ntoa()	272
9.3 UDP 实例程序 2	207	11.3 网络信息获取函数及其应用	272
9.3.1 程序要求	208	11.3.1 获得主机名——gethostname()	272
9.3.2 服务器程序	208	11.3.2 获得与套接口相连的远程协议 地址——getpeername()	273
9.3.3 客户端程序	212	11.3.3 获得套接口本地协议地址—— getsockname()	274
小结	215	11.3.4 根据主机名取得主机信息—— gethostbyname()和	
习题	215		
第 10 章 Winsock 程序综合实例	216		

WSAAsyncGetHostByName()	274	WSACreateEvent()	303
11.3.5 根据主机地址取得主机信息—— gethostbyaddr()和 WSAAsyncGetHostByAddr()	277	11.6.2 网络事件注册—— WSAEventSelect()	303
11.3.6 根据协议名取得主机协议信息—— getprotobyname()和 WSAAsyncGetProtoByName()	277	11.6.3 事件对象状态复位—— WSAResetEvent()	304
11.3.7 根据协议号取得主机协议信息—— getprotobynumber()和 WSAAsyncGetProtoByNumber()	278	11.6.4 事件对象状态置位—— WSASetEvent()	304
11.3.8 根据服务名取得相关服务信息—— getservbyname()和 WSAAsyncGetServByName()	279	11.6.5 关闭事件对象—— WSACloseEvent()	304
11.3.9 根据端口号取得相关服务信息—— getservbyport()和 WSAAsyncGetServByPort()	280	11.6.6 等待事件对象—— WSAWaitForMultipleEvents()	305
11.3.10 网络信息获取函数应用实例	280	11.6.7 网络事件查询—— WSAEnumNetworkEvents()	306
11.4 套接口选项函数及其应用	282	11.6.8 事件对象 I/O 管理程序实例	307
11.4.1 套接口选项函数说明	282	11.7 错误处理函数	308
11.4.2 SOL_SOCKET 选项级别	284	11.7.1 获得错误操作代码—— WSAGetLastError()	308
11.4.3 IPPROTO_IP 选项级别	288	11.7.2 设置错误操作代码—— WSASetLastError()	309
11.4.4 IPPROTO_TCP 选项级别	290	11.8 Winsock 2 支持的其他函数	309
11.4.5 套接口属性设置和获取实例	291	11.8.1 共享套接口—— WSADuplicateSocket()	309
11.5 套接口 I/O 处理函数及其应用	293	11.8.2 获取传送协议信息—— WSAEnumProtocols()	310
11.5.1 阻塞与非阻塞通信方式	293	11.8.3 初始化服务质量—— WSAGetQOSByName()	311
11.5.2 设置套接口的工作方式—— ioctlsocket()和 WSAIoctl()	294	11.8.4 返回重叠操作结果—— WSAGetOverlappedResult()	312
11.5.3 套接口 I/O 状态查询——select()	296	11.8.5 叶节点加入多点会话—— WSAJoinLeaf()	313
11.5.4 异步事件通知—— WSAAsyncSelect()	298	11.8.6 终止套接口上的数据接收—— WSARecvDisconnect()	314
11.5.5 取消正在执行的阻塞调用—— WSACancelBlockingCall()	301	11.8.7 终止套接口上的数据发送—— WSASendDisconnect()	314
11.5.6 判断是否有阻塞调用—— WSAIsBlocking()	302	小结	315
11.5.7 取消未完成的一个异步操作—— WSACancelAsyncRequest()	302	习题	315
11.6 事件对象 I/O 管理及其应用	302	参考文献	316
11.6.1 创建事件对象——			

TCP/IP

第 1 章

TCP/IP 概述

1.1 TCP/IP 的产生与发展

人们常说“治学先治史”。为了使读者对 TCP/IP 有一个比较深刻的认识，在本节中先介绍一下 TCP/IP 的产生原因与发展应用情况。

1.1.1 TCP/IP 的产生

人类社会产生以来，人们在不断地探索方便快捷的通信方式。从古代的驿站、烽火台发展到近代的电报、电话等，特别是电话的使用和普及，使相距千里的人们不再遥远，这些都只是部分地解决了“人脑”之间的通信问题。然而，科学的发展是永无止境的，20 世纪 40 年代人们研制出的电子计算机能代替部分人脑的工作，将人们从“繁重的科学计算中解放出来”，因此被人们称之为“电脑”，同样“电脑”之间也需要进行通信。古代人们点燃狼烟是战争中通信的需要，现代电脑之间的通信最初也是由于战争的需要。

在 20 世纪 50 年代初，为了提高防空能力，美国军方将计算机与通信技术相结合，开发出了半自动地面防空系统，起名字叫做“智者”（SAGE）。SAGE 系统将远程雷达与其他测量设施连接起来，将测得的防空信息通过总长为 2.41×10^6 km 的通信线路传输到一台 IBM 大型计算机上，集中进行防空信息处理与控制，即形成了一种以单机为中心的集中式网络通信系统。这种集中式网络通信系统在当时看起来是很先进的，但其有致命的弱点，即一旦网络控制中心受到敌方的攻击，将会使整个网络系统瘫痪，造成非常严重的后果。为了改变这一现状，1958 年，由当时的美国总统艾森豪威尔正式向国会提出了要建立“国防部高级研究计划署”的问题，这一组织的英文缩写为“DARPA”（Defense Advanced Research Project Agency），也常被人们简称为“ARPA”。当初，成立“DARPA”的目的非常明确，就是要“保持美国在技术上的领先地位，防止潜在的对手不可预见的技术进步”。

在 DARPA 的所有项目中，对当今世界影响最大，与普通人关系最密切，改变了人们日常交往和通信方式的是 1968 年 6 月提出的“资源共享的计算机网络”（Resource Sharing Computer Networks）研究计划。该计划的最初目的是让 DARPA 的所有电脑都能互联起来，使大家可以共享人们的研究成果，由于当时该研究是在美国国防部高级研究计划署的组织下进行的，因此人们就把这个网络叫做“ARPANET”，在国内有些资料音译为“阿帕网”。

在 ARPANET 中，通信的计算机需要相互理解对方的“意图”，如用什么信号来打开通信信

道,通信完成后用什么信号关闭信道,这就要求通信双方要达成某种方式的“共识”,这种“共识”就是现在所说的计算机通信“协议”。早在1965年,Larry Roberts和Thomas Merrill第一次将两种不同的电脑连接在一起,之后Merrill就把这种传送文件的方式称作“协议”(Protocol),他的一个同事好奇地问他为什么用“协议”这个词,他说“我觉得这像是在搞外交”。从此,人们就将计算机之间通信要共同遵守的原则称之为通信协议。

在20世纪70年代初,美国加州大学洛杉矶分校的网络工作小组(NWG)制定出“网络控制协议”(Network Control Protocol, NCP)。“网络控制协议”是一台主机直接对另一台主机的通信协议,实质上是一个设备驱动程序。要想真正做到将许多安装了不同操作系统的电脑连接起来,“网络控制协议”就有些力不从心了。而且,“网络控制协议”也没有纠错功能,如果在传输中出现了错误,就只能停止网络中传输的数据,那么这次传输也就失败了。因此,“网络控制协议”有其先天的不足。

随着计算机应用领域的不断扩展,接入ARPANET的电脑数量不断增加,就需要有一种新的协议来管理计算机网络中的通信。最重要的问题是给每一台电脑分配一个唯一的名字,这样别的电脑才能把这台电脑“认”出来。用专业一点的术语来说,就是要给网络中的每一台电脑一个唯一的“地址”。这就需要人们设计出一种新的协议,以解决不同电脑或不同网络之间的互联问题,当时人们为这种新的协议确定了如下一些基本设计原则。

(1) 每一个独立的网络必须按自己的标准建立起来,当这个网络和互联网连接的时候,不需要对其内部做任何改动。

(2) 网络应该在最佳的状态下完成通信。

(3) 如果一个“信包”没有到达目的地,最初发出“信包”的节点将很快重发该“信包”。

(4) 网络之间由叫做“黑匣子”的设备进行互相连接,这里所谓的“黑匣子”就是后来被人们称为类似于“网关”和“路由器”的设备。

(5) 整个互联网不需要在操作层面上进行任何总体控制。

按照这些设计原则,在1973年由研究操作系统通信原理的Kahn和参与过网络通信协议NCP设计的Vinton Cerf两人一起为ARPANET开发具有以上特点的网络互联协议。1974年5月,Vinton Cerf和Kahn两人合作,在IEEE(电气和电子工程师协会)刊物上发表了一篇题为“分组网络互联的一个协议”的论文;同年12月,他们正式发表了第一份TCP的详细说明,该协议负责在互联网上传输和转发信包。由于在后来的实验中证实有时这种TCP在信包丢失时不能得到有效地纠正,因此,他们认识到应该把这一协议分成两个不同的协议:一个是用来检测网络传输中差错的“传输控制协议”(TCP),当检测到传输中有差错时,它就能发出重发信号,源端收到该信号后就重新发送发生差错的信包,通过这种差错重传机制保证数据能够正确传输到目的地;另一个是专门负责对不同网络进行互联的“互联网协议”(IP),为了实现不同类型的局域网互联,它在各种局域网地址标准之上,为互联网络中的所有主机设定了统一的互联网地址(IP地址),以保证不同网络中的主机(当然也可以是同一个网络中的主机)只要接入互联网,它们之间就可以相互识别,进行通信。这两个协议就是我们现在所说的“TCP/IP”。

使用TCP/IP的网络,其最大的优点是可以使不同的网络互联,并且网络中的任何节点或某部分被破坏都不会造成整个网络的瘫痪,这正好符合美国军方的需要。于是,在1982年做出了在ARPANET上使用TCP/IP以代替原来使用的网络控制协议(NCP)的决定。1983年1月1日,ARPANET停止了NCP的使用,从此互联网上的主机都使用TCP/IP,使TCP/IP成了计算机网络领域事实上的工业标准协议。

那么,什么是 TCP/IP 呢?对初学者来说可以简单地这样理解,TCP 是用来检测数据传输中的差错并进行流量控制的协议,TCP 如果发现问题,就要发出重新传输的信号,直到所有数据安全传输到目的地,因此它起到一个信息传输过程中的控制作用。IP 的最主要作用就是给网络中的每一台电脑分配一个唯一地址,否则别的电脑就找不到它,这就像日常生活中寄信要有地址,打电话要有电话号码一样。现在人们将 Internet 中使用的所有协议统称为 TCP/IP。

1.1.2 TCP/IP 的发展与应用

TCP/IP 产生于 20 世纪 70 年代,在 1983 年,著名的 ARPANET 全面使用 TCP/IP 后,大大加快了 TCP/IP 的推广速度。后来人们以 ARPANET 为骨干网建立了 Internet。在 1985 年,美国国家科学基金会(NSF)开始了 TCP/IP 的研究工作,并在 1986 年资助建立了 NSFNET。NSFNET 使用 TCP/IP,作为骨干网连接了全美主要科研机构,后来 NSFNET 代替 ARPANET 成为 Internet 的骨干网,这样,TCP/IP 顺理成章地成为了 Internet 中的“世界语”。

Internet 在全球的应用非常广泛,Internet 的快速发展得益于 TCP/IP 的提出和使用,反过来,Internet 又促进了 TCP/IP 的普及和应用。

这里还要说明一点,在计算机系统的操作系统领域,1983 年,美国加州大学伯克利分校网络开发者发布的最新版的 UNIX 操作系统——4.2BSD (Berkeley Software Distribution) 实现了 TCP/IP,这也大大地促进了 TCP/IP 的推广与应用。

TCP/IP 从产生到现在虽然共出现了 6 个版本,但前 3 个版本基本没有投入实际使用,第 5 版是基于 ISO/OSI 中的 7 层模型提出来的,由于层次结构变化大,实现成本较高,因此没有投入实际使用。现在 Internet 网络中广泛使用的是 TCP/IP 的第 4 版,一般简称为 IPv4。在 IPv4 的使用过程中,人们发现其 32 位的地址范围有限并存在安全等问题,于是提出 IPv6,该协议主要将 IP 进行了修改,使其在地址空间、数据的完整性和保密性、语音和视频数据传输等方面有了很大的改进,能更好地满足当前网络实际应用的需求。本书主要讨论 IPv4 协议,同时对 IPv6 协议的关键内容也进行一些介绍。

1.2 TCP/IP 的体系结构

尽管 TCP/IP 比较复杂,但它有比较清晰的体系结构。在学习 TCP/IP 前,从总体上把握 TCP/IP 的体系结构非常重要。本节介绍 TCP/IP 的分层结构和各层的大概功能。

1.2.1 TCP/IP 的分层体系结构与协议栈的概念

网络系统由硬件系统(如网卡、集线器、路由器等)和软件系统(如操作系统、通信协议等)组成。在进行 TCP/IP 的研究与进行网络软件开发时,一般人们只关心网络系统的软件系统,而软件系统的核心内容就是通信时所使用的网络协议。因此本小节讨论的 TCP/IP 结构是从软件角度来说的。

在网络协议发展的初期,人们并没有想到要将网络协议进行分层设计,只是在后来的研究与应用中才逐步认识到网络协议的复杂性,于是才想到了人类在解决复杂问题时所广泛使用的“分层管理”的思想。如在人类社会,中央政府管理省级政府,省级政府管理地区级政府,地区级政府管理县级政府等,这样,中央政府的管理指令先达到省级政府,省级政府再下达到地区级

政府，地区级政府再下达到县级政府。同样，县级政府有问题先要向上一级地区政府汇报，然后由地区政府向上一级政府汇报。将这一思想应用到网络软件的设计上，人们就把网络系统的软件按功能划分为若干个层次（一般少则分成 4 层，多则可达 7 层），每层负责不同级别的通信功能。

按照上面介绍的思想，人们将 TCP/IP 分为 4 层，这 4 层由高到低分别是应用层、传输层、网络层和网络接口层，如图 1-1 所示。

各层常用的协议	
第 4 层 应用层	Telnet、FTP、HTTP、DNS、SNMP、SMTP 等
第 3 层 传输层	TCP 和 UDP
第 2 层 网际层	IP、ICMP 和 IGMP
第 1 层 网络接口层	设备驱动程序和网络接口卡等

图 1-1 TCP/IP 的层次结构

其中每一层完成不同的通信功能，一个层就好像一个“黑匣子”一样。每个“黑匣子”的内部是如何实现的，其他层并不关心，其他层只关心这个“黑匣子”对外部提供了什么样的服务。在这样一种体系结构中，每一层都向它的上层提供一定的服务，同时每一层可以使用它的下层所提供的功能。这样在相邻层之间就有一个接口，通过这个接口将相邻层之间联系起来。这种网络协议的层次结构主要有如下优点。

(1) 各层之间相互独立。只要保持相邻层之间的接口不变，一个层内部可以用不同的方式来实现。各层独立的另外一个好处是一个层内部实现升级或改变时，不影响与该层相邻的上层和下层。

(2) 易于设计与实现。各层都可以使用最合适的技术来实现，并且不同层都可以由不同的团队来设计。

(3) 易于维护。由于将一个复杂的系统分解为若干个相对简单的层次，并且各层完成的功能不同，所以当系统出现问题时易于分析与维护。

一般情况下，网络的层次结构和每层所使用协议的集合称为网络体系结构（Network Architecture）。一个具体的网络协议其所包含的层数和每层所使用的协议是确定的，如 TCP/IP 分为 4 层，每层所使用的协议如图 1-1 所示。

如上所述，网络协议是分层的，在这种层次结构中各层有明确的分工，不同层的协议从上到下形成了一个栈结构的依赖关系，通常将其形象地称为协议栈（Protocol Stack）。图 1-1 所示是 TCP/IP 的栈结构，从该结构中可以看出，在 TCP/IP 的协议栈中包括很多协议，但 TCP 和 IP 是该协议栈中两个最重要的协议，所以人们常常将该协议栈简称为 TCP/IP。

下面介绍一下 TCP/IP 协议栈中的 4 个层及其主要功能。

1.2.2 网络接口层

网络接口层处于 TCP/IP 协议栈的最低层，它负责将其之上的网络层要发送出去的数据（即 IP 数据报）发送到其下面的物理网络，或接收由物理网络发送到该目标机的数据帧，并抽出 IP 数据报交给网络层。要注意，这里所说的物理网络是指各种实际传输数据的局域网或广域网等。

在 TCP/IP 协议栈中并没有具体定义网络接口层的内容，一般情况，人们认为只要是在其上能进行 IP 数据报传输的物理网络（如常用的局域网中的以太网、令牌环网、光纤分布式数据接口、IEEE 802.3 及 RS-232 串行线路等），都可以当成 TCP/IP 协议栈的网络接口层。

为什么在 TCP/IP 协议栈中没有定义网络接口层呢？这是因为不定义网络接口层的具体内容如下两点好处。

(1) 便于实现不同网络之间的互联。实现不同网络的互联是 TCP/IP 要解决的最主要问题。不同的网络尽管其数据传输介质、数据传输速率等有很大的差异，但都可以实现网络内数据的传输，当然也就可以进行 TCP/IP 协议栈中网络层 IP 数据报的传输。这样 TCP/IP 就可以将重点放在网络之间的互联上，而不用去纠缠各种物理网络的具体实现细节，这样就非常巧妙地解决了不同类型物理网络的互联问题。这也是 TCP/IP 得以广泛应用的一个重要原因。

(2) 为将来物理网络的发展留下了广阔的空间。物理网络与计算机硬件技术和通信技术的发展紧密相连，如物理网络中数据传输率在不断提高，网络中设备的性能也在不断提高，所有这些都不会影响 TCP/IP 的使用。

1.2.3 网际层

网际层 (Internet Layer) 在 TCP/IP 协议栈的第 2 层，也称为互连网络层 (互联层)，因该层的主要协议是 IP，所以也可简称为 IP 层。它是 TCP/IP 协议栈中最重要的一层，主要功能是把源主机上的分组根据需要发送到互联网中的任何一台目标主机上。当然在一般情况下，发送信息的源主机要知道接收信息的目标主机的“地址” (该内容在第 3 章中介绍)。通信时，源主机与目标主机可以在同一个网络中，也可以在不同的网络中。

在网际层传输的数据单位叫 IP 数据报，也称为 IP 分组。

在一个由很多网络组成的互联网中，一台主机 (即源主机) 与不在同一个网络中的另一台主机 (目标主机) 通信时，可能有多条通路相连，网际层的一个重要功能就是要在这些通路中做出选择，这就是所谓的路由选择功能。路由选择是网际层一个非常重要的功能。

网际层的本质是在各种物理网络基础上，使用 IP 将它们互联，组成了一个传输 IP 数据报的虚拟网络，所以网际层实现了不同网络的互联功能。

总之，在 IP 层提供的是一种“尽力而为”的数据报传输服务，它不能保证数据总是可靠地从源主机传输到目标主机。

在 TCP/IP 协议簇中，网际层协议包括 IP (网际协议)、Internet 互联网控制报文协议 (Internet Control Message Protocol, ICMP)、Internet 组管理协议 (Internet Group Management Protocol, IGMP) 等。

1.2.4 传输层

在 TCP/IP 协议簇中，传输层 (Transport Layer) 处于第 3 层。传输层完成通常所说的两台主机之间的通信，其实质是两台主机上对应的应用进程之间的通信。传输层提供的应用进程之间的通信，也叫端到端 (End to End) 的通信。端到端的通信是在传输层两个通信的实体之间进行的，这就好像是在两个通信的实体之间建立了一条逻辑通路一样，它屏蔽掉了 IP 层的路由选择和物理网络等细节。

在实际的通信过程中，传输层应用进程之间对通信质量的要求是不一样的。为了满足不同的需要，在 TCP/IP 协议簇中传输层定义了两个不同的传输协议：一个是 TCP (传输控制协议)；另一个是 UDP (用户数据报协议)。

TCP 为两台主机提供高可靠性的数据通信服务，它可以将源主机的数据流无差错地传输到目标主机。当有数据要发送时，它对应用进程送来的数据进行分片，以适合于在网络层中传输；当