

TURING

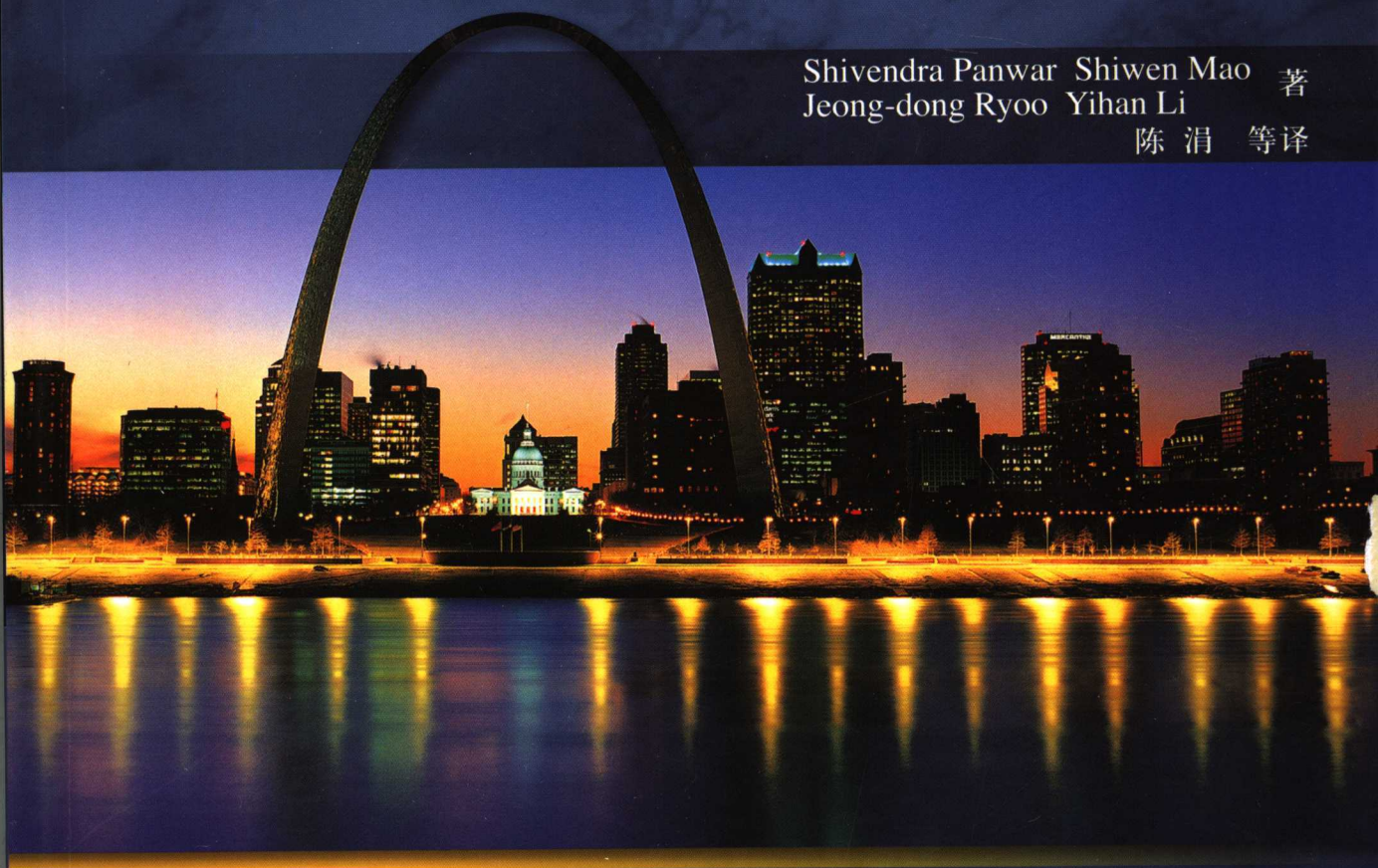
图灵计算机科学丛书

TCP/IP基础教程

基于实验的方法

TCP/IP Essentials: A Lab-Based Approach

Shivendra Panwar Shiwen Mao 著
Jeong-dong Ryoo Yihan Li
陈涓 等译



人民邮电出版社
POSTS & TELECOM PRESS

TCP/IP基础教程

基于实验的方法

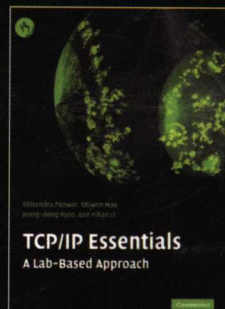
TCP/IP协议族已经成为网络世界事实上的标准和因特网的基础,当前几乎所有的计算机通信系统都使用了TCP/IP协议族。

本书原版由剑桥大学出版社出版。作者独辟蹊径,通过一系列精心设计并经过广泛测试的实验讲述TCP/IP技术,使学生能够直观地地实践中观察各个协议的实现和行为。本书讨论的主题包括网桥、路由器、局域网、静态与动态路由选择、多播与实时业务以及网络管理与安全。书中的实验都是在Linux环境中描述的,同时通过注释说明了在Solaris中实现的情况。本书包含很多练习,同时还为教师提供了许多补充教学材料。

本书适用于计算机相关专业的学生。对工程师来说,本书也是一本理想的TCP/IP指南。

本书具有以下特色

- 理论讲述简明扼要,但不失全面。
- 在理论之后,立即用实验来深化和巩固。
- 书中的实验贴近实战,凝聚了作者多年授课经验。



TCP/IP Essentials A Lab-Based Approach

Shivendra Panwar

是纽约理工大学电气与计算机工程系的教授。他现在是纽约州高级电信技术中心 (Center for Advanced Technology in Telecommunications, CATT) 的主任,发表过80多篇相关论文。

Shiwen Mao

是弗吉尼亚理工学院和州立大学电气与计算机工程系的研究员。

Jeong-dong Ryoo

是韩国大邱电子电信研究院 (ETRI) 的资深研究员。

Yihan Li

是纽约理工大学电气工程系的研究员。

ISBN 7-115-15386-8



9 787115 153869 >

本书相关信息请访问: 图灵网站 <http://www.turingbook.com>
读者/作者热线: (010) 88593802
反馈/投稿/推荐信箱: contact@turingbook.com

上架建议 计算机/网络技术/TCP/IP

人民邮电出版社网址 www.ptpress.com.cn

ISBN7-115-15386-8/TP • 5758

定价: 29.00 元

TURING

图灵计算机科学丛书

TCP/IP基础教程

基于实验的方法

TCP/IP Essentials: A Lab-Based Approach

Shivendra Panwar Shiwen Mao 著
Jeong-dong Ryoo Yihan Li
陈涓 等译



人民邮电出版社
POSTS & TELECOM PRESS

图书在版编目 (CIP) 数据

TCP/IP 基础教程: 基于实验的方法 / (美) 潘沃 (Panwar, S.) 著; 陈涓译.

—北京: 人民邮电出版社, 2006.12

(图灵计算机科学丛书)

ISBN 7-115-15386-8

I. T... II. ①潘...②陈... III. 计算机网络—通信协议—教材 IV. TN915.04

中国版本图书馆 CIP 数据核字 (2006) 第 119845 号

内 容 提 要

TCP/IP 协议实际上已经成为互联网的标准, 因此, 在“计算机网络”课程中, TCP/IP 协议已经是核心内容。在描述 OSI 模型中的各层次时, 也是围绕各层次所对应的协议展开的。本书通过一系列精心设计的实验, 涵盖了 TCP/IP 协议所涉及的各个方面, 包括网桥、路由器、LAN、静态和动态路由、多播和实时服务、网络管理和安全等。作者教授计算机网络课程 8 年多, 书中的内容来源于作者实际的授课经验。所有实验均经过广泛且有效的测试, 其中的实验环境是 Linux 操作系统。本书有配套的教辅材料。

本书适合作为本科电子工程、计算机工程以及计算机专业计算机网络课程的配套教材, 也适合作为专科学校相关专业的教材。

图灵计算机科学丛书

TCP/IP 基础教程: 基于实验的方法

-
- ◆ 著 Shivendra Panwar Shiwen Mao Jeong-dong Ryoo Yihan Li
译 陈 涓 等
责任编辑 杨海玲 陈兴璐
 - ◆ 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>
北京顺义振华印刷厂印刷
新华书店总店北京发行所经销
 - ◆ 开本: 787×1092 1/16
印张: 13.5
字数: 342 千字 2006 年 12 月第 1 版
印数: 1—4 000 册 2006 年 12 月北京第 1 次印刷

著作权合同登记号 图字: 01-2006-1280 号

ISBN 7-115-15386-8/TP · 5758

定价: 29.00 元

读者服务热线: (010) 88593802 印装质量热线: (010) 67129223

版 权 声 明

Original edition, entitled *TCP/IP Essentials: A Lab-Based Approach* by Shivendra Panwar, Shiwen Mao, Jeong-dong Ryoo, Yihan Li, ISBN 052160124X published by Cambridge University Press in 2004.

This translation edition is published with the permission of the Syndicate of the Press of the University of Cambridge, Cambridge, England.

© Cambridge University Press 2004.

本书原版由剑桥大学出版社出版。

本书中文翻译版由剑桥大学出版社授权人民邮电出版社出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

版权所有，侵权必究。

前 言

你能学会用世界上所有的语言来叫一种鸟的名字,但当你知道了所有这些名字的叫法之后,对这种鸟却还是一无所知……所以,让我们来观察这只鸟,看看它在干什么——这才是最重要的。我很早就明白了知道一个事物的名字和了解这个事物是不同的。

理查德·费曼 (Richard Feynman, 美国物理学家, 诺贝尔奖得主)

顾名思义,本书通过实验简明扼要地讲授 TCP/IP。说它简明扼要,是因为它在我们认为足以掌握 TCP/IP 基础知识的深度上,提供了一种很独特的主题选择方式。我们的目的不是写一本参考书。在实验室里,学生可以实际观察 TCP/IP 协议的实现,这是很重要的。通过这些观察并从中得出结论,会为通常有些枯燥的网络协议学习带来活力,并激励学生去了解更多与协议有关的知识。

只有那些负责建立实验室的教师才需要阅读附录 A。我们一直尽力降低费用,所以只有那些非常吝啬的大学管理者才会对本书所需实验室设备的费用皱眉头。(至于实验室的空间,那是另一个问题了!)我们假定学生有连网方面的基本知识,这些知识可能是从前一门课程中学来的,或者学习时先对网络进行了总体的课堂介绍,再回过头来进行本书的实验。第 0 章对 TCP/IP 进行了快速的介绍,目的主要有两个:提供对 TCP/IP 协议栈的概述,并且为本书的其余部分提供一个框架。第 1 章~第 9 章具有相同的结构:每一章都在课堂讲授部分提供了适于讲授的介绍性材料,之后都跟着一个在实验室中进行的实验。实验应该在课堂讲授部分之后进行,课堂讲授部分为学生提供了完成实验以及通过实验课中的观察得到见解所需的基本知识。每个实验都设计为 3 个小时之内完成。

作者在纽约理工大学讲授了 8 年的计算机网络课程,这些实验都是在此基础上发展起来的。开始时我们使用的是基于 Solaris 操作系统的 SUN 工作站,但现在已经转换到基于 Linux 操作系统的计算机上了。本书使用的操作系统主要是 Linux,但是当 Solaris 中的命令与 Linux 的命令不同时,我们也提供了相应的 Solaris 命令。学生们可能对 Linux 操作系统不太熟悉,所以第 1 章对 Linux 进行了介绍。这一章还介绍了后继实验中会用到的一些关键工具,如 tcpdump 和 Ethereal。第 2 章介绍了网络接口、ping 和 IP 地址。第 3 章介绍了网桥(也称作第二层交换机)、网桥/路由器配置和 Cisco IOS。第 4 章的重点是路由选择,研究了 RIP 和 OSPF 路由协议,以及一些有用的 traceroute 实用程序。第 5 章介绍了 UDP 和 FTP。第 6 章介绍了 TCP,包括对 TCP 拥塞控制机制的学习。大多数情况下,这 6 章就足以将 TCP/IP 的基础知识介绍给学生了。但是,对那些希望将 TCP/IP 的基本知识与应用程序结合起来的学生来说,后面 3 章是非常重要的。第 7 章介绍了 IP 多播和一些实时应用程序。第 8 章介绍了一些主要的应用,包括 Web、DHCP、NTP 和 NAT,此外,还对套接字编程进行了简要介绍。按理说,网络管理与安全是学生需要了解的最重要的两个特性,至少要在最基本的层面对其有所了解。第 9 章简要介绍了这方面内容,

这些内容可以很容易地成为一门独立课程的主题。本书的末尾提供了一个关键的 RFC 列表。

可以用多种不同的方式讲授书中的内容。对连网知识的概要了解是学习本书的前提。但是，网络方面的入门课程也可以与前 6 个实验结合起来，在课程的最后回过头来通过实验演示协议栈中最下面 4 层的工作原理。对计算机研究人员来说，在讲授网络课程时，可能更喜欢采用自顶向下的方法。在这种情况下，可以将这些实验重新排列以便将重点放在高层上面。

教师注意

附加的课程资料，包括演讲幻灯片、实验报告实例、家庭作业、测验以及勘误表，都可在网站 www.cambridge.org/052160124X 上获得。

致 谢

作者要对纽约理工大学、美国国家科学基金会、纽约州科学、科技和学术研究办公室 (NYSTAR, New York State Office of Science, Technology and Academic Research) 以及证券业自动化公司 (SIAC, Securities Industries Automation Corporation) 提供的支持表示感谢。尤其是, 我们是在与 SIAC——负责纽约和美国证券交易所的连网及系统需求的公司——的合作中得到了最初的启发。要特别感谢 SIAC 的 Andrew Bach、Joseph Kubat、Michael Lamberg、Darko Mrakovcic 和 Dror Segal 提供的支持。向为实验的最初版本提供帮助的 Nitin Gogate 博士及其所有研究生们表示感谢。我们要感谢 Jeffrey (Zhifeng) Tao、Yanming Shen 和 Pei Liu, 他们对实验进行了校对和测试。我们还要感谢下列教师, 这些年他们也在纽约理工大学教授过这门课程: Malathi Veeraraghavan、John (Zheng-Xue) Zhao 和 Jorg Liebeherr。

通用约定

下列约定的使用贯穿全书。

- 在正文段落中，Linux、UNIX 和 Cisco IOS 命令都以黑体表示，比如 **telnet** 和 **enable**。
- 在带有选项和参数的复合命令中，命令和选项都以黑体表示的，而参数则以斜体表示。例如，在

tcpdump -enx host *ip_addr1* and *ip_addr2*

中，命令 **tcpdump** 使用了选项 **-e**、**-n** 和 **-x**。在后面的过滤器中，**host**、**and**、**not**、**or** 等关键字也都以黑体表示。参数是 *ip_addr1* 和 *ip_addr2*，在练习中可以用相应的 IP 地址来替代。

以下示例命令

/etc/init.d/snmpd start|stop,

中使用了两个选项。可以使用 **start**，也可以使用 **stop**，但两者不能同时使用。

- 在实验报告中要回答每个练习的**实验报告**部分的问题。例如，对第 1 章的练习 2 来说，学生就需要在实验报告中回答下列问题：

实验报告打开一个新的命令窗口时的默认目录是什么？你的工作目录是什么？

- 在本书中，我们的重点是 Linux 操作系统。但是，本书也可用于 Sun 公司的 Solaris 操作系统。在后续内容中，有些是 Linux 专用的，也有些是这两种操作系统都通用的，而 Solaris 专用部分则加方框突出显示出来。

读者意见交流卡

亲爱的读者:

感谢您对我们的支持与爱护。为了今后为您提供更优秀的图书,请您抽出宝贵时间填写本表(或通过我们的网站 www.turingbook.com 填写本表),将您的意见及时告知我们。您将有机会免费获赠我们出版的图书,并能获得最新的出版信息和更多服务,谢谢!

系列书名: 图灵程序设计丛书

本书名: TCP/IP 基础教程: 基于实验的方法

读者资料:

姓名: _____ 性别: ☐男 ☐女 年龄: _____
职业: _____ 文化程度: _____ 通信地址: _____
电话: _____ 传真: _____ 电子信箱(E-mail): _____

1. 您是如何得知本书的:

- ☐别人推荐 ☐书店 ☐出版社图书目录
☐杂志、报纸、网络等的介绍(请指明)
☐其他(请指明) _____

2. 您从何处购得本书:

- ☐新华书店 ☐电脑专业书店 ☐网上书店 ☐其他 _____

3. 影响您购买本书的因素:

- ☐内容和质量 ☐装帧设计 ☐价格
☐内容提要、前言或目录 ☐书评广告
☐出版社名气 ☐作者名气 ☐其他 _____

4. 您对本书封面和封底设计的满意度:

- ☐很满意 ☐比较满意 ☐一般 ☐较不满意 ☐不满意
☐建议 _____

5. 您认为本书:

- 价格: ☐高 ☐合适 ☐低
翻译质量: ☐高 ☐一般 ☐差
图书印刷质量: ☐高 ☐一般 ☐差

6. 您希望本书哪些方面进行改进?

7. 您感兴趣或希望出版的图书有:

请寄: 北京市西四环北路 140 号京鼎原商务楼 405 房间 人民邮电出版社图灵公司 市场部收

邮编: 100097 电话: 010-88593802 传真: 010-88593803

电子信箱(E-mail): contact@turingbook.com 网址: www.turingbook.com

目 录

第0章 TCP/IP 概述..... 1	1.6 与诊断工具有关的练习..... 28
0.1 因特网..... 1	1.7 与端口号有关的练习..... 30
0.2 TCP/IP 协议..... 2	第2章 单段网络..... 31
0.3 网际互连设备..... 4	2.1 目的..... 31
0.4 封装与多路复用..... 5	2.2 局域网..... 31
0.5 命名与编址..... 6	2.2.1 点对点网络..... 31
0.5.1 域名..... 6	2.2.2 以太网式局域网..... 32
0.5.2 端口号..... 8	2.2.3 IEEE 802.11 无线局域网..... 33
0.5.3 IP 地址..... 8	2.2.4 地址解析协议..... 34
0.5.4 IPv6..... 10	2.3 网络接口..... 36
0.5.5 介质访问控制地址..... 11	2.3.1 网络接口的操作..... 36
0.6 多路访问..... 11	2.3.2 网络接口的配置..... 37
0.7 路由选择与转发..... 12	2.4 因特网控制报文协议..... 38
0.8 拥塞控制与流量控制..... 12	2.5 Sock 流量生成程序..... 39
0.9 错误检测及控制..... 13	2.6 网络接口练习..... 39
0.10 协议的首部格式..... 14	2.7 ARP 练习..... 40
0.10.1 以太网帧格式..... 14	2.8 关于 ICMP 和 ping 的练习..... 42
0.10.2 IP 首部格式..... 14	2.9 关于 IP 地址和子网掩码的练习..... 43
0.10.3 UDP 首部格式..... 15	第3章 网桥、LAN 和 Cisco IOS..... 45
0.10.4 TCP 首部格式..... 15	3.1 目的..... 45
0.11 举例: TCP/IP 协议族如何共同工作..... 16	3.2 以太网网桥..... 45
第1章 Linux 与 TCP/IP 网络互联..... 19	3.2.1 网桥的使用..... 45
1.1 目的..... 19	3.2.2 网桥的操作..... 46
1.2 Linux 与 TCP/IP 实现..... 19	3.2.3 生成树算法..... 46
1.2.1 TCP/IP 实现..... 19	3.3 网桥或路由器的配置..... 49
1.2.2 网络守护进程及服务..... 20	3.3.1 Cisco 互联网操作系统..... 49
1.2.3 网络配置文件..... 21	3.3.2 Cisco IOS 配置模式..... 49
1.3 Linux 命令与工具..... 22	3.3.3 网桥/路由器的配置过程..... 51
1.3.1 基本的 Linux 命令..... 22	3.3.4 透明网桥的配置..... 52
1.3.2 文本编辑器..... 23	3.4 与 Cisco IOS 有关的练习..... 53
1.3.3 视窗转储..... 24	3.5 一个简单的网桥实验..... 54
1.3.4 软盘的使用..... 25	3.6 生成树练习..... 55
1.4 诊断工具..... 25	3.7 关于 Cisco IOS 的 Web 浏览器 UI 的练习..... 57
1.4.1 Tcpdump..... 25	第4章 静态和动态路由选择..... 58
1.4.2 Ethereal..... 26	4.1 目的..... 58
1.5 与 Linux 命令有关的练习..... 26	

4.2 静态和动态路由选择	58	6.6 TCP 诊断工具	94
4.2.1 下一跳路由选择	58	6.6.1 分布式基准系统	94
4.2.2 静态路由选择与动态路由选择	60	6.6.2 NIST Net	95
4.2.3 路由选择中 ICMP 报文的使用	60	6.6.3 TCP 分组的 tcpdump 输出	95
4.2.4 动态路由选择	62	6.7 关于 TCP 连接控制的练习	95
4.2.5 多协议标记交换与流量工程	67	6.8 关于 TCP 交互数据流的练习	96
4.3 操纵路由表	67	6.9 关于 TCP 大数据量数据流的练习	97
4.3.1 工作站的路由表	67	6.10 关于 TCP 定时器和重传机制的练习	97
4.3.2 路由器的路由表	67	6.11 其他练习	98
4.4 Traceroute	68	6.12 关于 DBS 和 NIST Net 的练习	98
4.5 一个简单的路由器实验	69	第 7 章 多播和实时业务	102
4.6 RIP 练习	70	7.1 目的	102
4.7 用 ICMP 进行的路由选择实验	72	7.2 IP 多播	102
4.8 OSPF 练习	73	7.2.1 多播寻址	103
4.9 静态路由选择实验	74	7.2.2 多播组管理	105
4.10 Traceroute 实验	75	7.2.3 多播路由选择	106
第 5 章 UDP 及其应用	76	7.2.4 多播骨干网: MBone	108
5.1 目的	76	7.2.5 配置一台多播路由器	108
5.2 用户数据报协议	76	7.3 实时多媒体流	109
5.3 MTU 与 IP 分段	77	7.3.1 实时流	109
5.3.1 IP 分段	77	7.3.2 支持多媒体流业务的协议	111
5.3.2 路径 MTU 发现	77	7.3.3 Java 媒体框架和 JMStudio	114
5.4 客户机-服务器应用程序	78	7.4 简单的多播练习	114
5.4.1 客户机-服务器结构	78	7.5 IGMP 练习	116
5.4.2 TFTP	78	7.6 多播路由选择练习	118
5.4.3 FTP	79	7.7 多播视频流练习	119
5.5 使用 sock 程序	81	第 8 章 Web、DHCP、NTP 和 NAT	121
5.6 UDP 练习	81	8.1 目的	121
5.7 路径 MTU 发现练习	81	8.2 超文本传输协议	121
5.8 FTP 和 TFTP 的练习	83	8.2.1 超文本传输协议与 Web	121
第 6 章 学习 TCP	85	8.2.2 通用网关接口协议	123
6.1 目的	85	8.2.3 Apache Web 服务器	123
6.2 TCP 服务	85	8.3 动态主机配置协议 (DHCP)	124
6.3 管理 TCP 连接	85	8.3.1 DHCP 协议	124
6.3.1 TCP 连接的建立	86	8.3.2 配置 DHCP	127
6.3.2 TCP 连接的终止	86	8.4 网络时间协议 (NTP)	128
6.3.3 TCP 定时器	87	8.4.1 NTP 协议	128
6.4 管理 TCP 数据流	87	8.4.2 配置 NTP	129
6.4.1 TCP 差错控制	87	8.4.3 网络计时工具	129
6.4.2 TCP 交互数据流	90	8.5 IP 网络地址转换设备	130
6.4.3 TCP 大数据量数据流	91	8.5.1 网络地址转换与端口地址转换	130
6.5 调整 TCP/IP 内核	93	8.5.2 配置 NAT 路由器	131

8.6 套接字编程简述.....	132	9.6.2 安全 Apache 服务器.....	152
8.7 HTTP 练习.....	134	9.7 网络层安全.....	153
8.8 DHCP 练习.....	136	9.8 系统安全.....	155
8.9 NTP 练习.....	137	9.8.1 防火墙.....	155
8.10 NAT 练习.....	138	9.8.2 iptables.....	155
8.11 套接字编程练习.....	141	9.8.3 审计和入侵检测.....	156
第 9 章 网络管理与安全	142	9.9 SNMP 练习.....	156
9.1 目的.....	142	9.10 与安全应用程序有关的练习.....	158
9.2 网络管理.....	142	9.11 与安全 Apache 服务器有关的练习.....	158
9.2.1 简单网络管理协议.....	142	9.12 与防火墙和 iptables 有关的练习.....	159
9.2.2 MIB 结构.....	144	9.13 与审计和入侵检测有关的练习.....	160
9.2.3 NET-SNMP.....	144	参考文献和进阶阅读	161
9.3 网络安全综述.....	145	参考文献.....	161
9.4 加密、机密性和认证.....	146	进阶阅读.....	161
9.4.1 数据加密.....	146	附录 A 教师指南	163
9.4.2 公共密钥加密方案.....	148	附录 B 路由器的初始配置	175
9.4.3 散列和报文认证.....	148	附录 C 源代码	177
9.4.4 数字签名与认证.....	149	附录 D 关键的 RFC 列表	191
9.5 应用层安全.....	150	缩略语	194
9.5.1 安全命令解释器协议和 OpenSSH.....	150	索引	197
9.5.2 Kerberos.....	151		
9.6 传输层和 Web 安全.....	151		
9.6.1 安全套接字层协议.....	151		

第 0 章

TCP/IP 概述

因特网的基本结构基于这样的假设：因特网是一个分组交换的通信设施，其中，很多可区分的网络由被称为网关的分组通信处理器连接在一起，网关采用的是存储转发方式的分组转发算法。

——大卫·克拉克（David D. Clark，因特网先驱）

0.1 因特网

因特网是一个由全世界数百万个计算机网络组成的全球信息系统。因特网的用户可以交换电子邮件、访问远程计算机上的资源、浏览网页、体验实时的流式视频或音频，并且可以为其他用户发布信息。随着电子商务（e-commerce）的发展，很多公司都在因特网上提供了各种服务，例如在线银行业务、金融交易、购物以及在线拍卖等。随着所提供业务的扩展，因特网的规模也呈指数级增长。此外，各种各样的电子设备也被连接到因特网上来，比如，移动电话、个人数字助理（personal digital assistant, PDA），甚至电视和冰箱等。

今天的因特网是从 20 世纪 60 年代末期美国高级研究计划署（Advanced Research Projects Agency, ARPA）资助的、只包含四个节点的 ARPANET 发展起来的。1983 年，ARPANET 采纳了最早由 Cerf 和 Kahn 在[1]中提出的传输控制协议/网际协议（TCP/IP）协议族。1984 年，NSF 建立了一个基于 TCP/IP 的骨干网 NSFNET，NSFNET 成为了 ARPANET 的后继者。因特网于 1995 年实现了完全的商业化。现在，“因特网”这个词是指用分组交换技术进行松散互连的、基于 TCP/IP 协议族的全球计算机网络。

1

因特网由很多组织进行管理。这些组织负责控制 TCP/IP 协议，开发并批准新的标准，以及分配因特网地址和其他资源。下面列出了其中的一些组织。

- ❑ 因特网协会（Internet Society, ISOC）。这是一个由因特网专家组成的专业成员组织，这些专家负责对各种政策和实践活动提出意见，并对很多其他处理网络政策问题的委员会和任务组进行监管。
- ❑ 因特网架构委员会（Internet Architecture Board, IAB）。IAB 负责定义因特网的整体架构，为 IETF（参阅下文）提供指导和大的方向。
- ❑ 因特网工程任务组（Internet Engineering Task Force, IETF）。IETF 负责协议的工程化及协议的开发。
- ❑ 因特网研究任务组（Internet Research Task Force, IRTF）。IRTF 负责进行重点的、长期

的研究。

- 互联网名称与数字地址分配机构 (Internet Corporation for Assigned Names and Numbers, ICANN)。ICANN 负责网际协议 (IP) 地址空间的分配、协议标识符的分配、通用和国家代码顶级域名系统的管理以及根服务器系统的管理。这些服务最初是由互联网数字地址分配机构 (Internet Assigned Numbers Authority, IANA) 和其他一些组织来完成的。现在, 则由 ICANN 来实现 IANA 的功能。
- 因特网信息中心 (Internet Network Information Center, InterNIC)。InterNIC 由 ICANN 负责运作, 提供与因特网域名注册业务有关的信息。

因特网标准都以请求评注 (Request For Comments, RFC) 的形式发布, 以强调“基本的原则是: 任何人都可以表达任何看法, 没有什么官方的” [2]。IETF 的网站 <http://www.ietf.org/> 上提供了所有的 RFC。通常, 一项新技术首先会被作为一个因特网草案 (Internet Draft) 提出来, 草案会在六个月后过期。如果这项因特网草案获得了 ISOC 或业界的持续兴趣及支持, 它就会被提升为一个 RFC, 然后成为一个推荐标准 (Proposed Standard), 接着成为一个标准草案 (Draft Standard)。最后, 如果这个提案通过了所有的测试, IAB 就会将其作为一项因特网标准 (Internet Standard) 发布。

0.2 TCP/IP 协议

2 计算机之间的信息交换任务由各种功能组成, 非常复杂。在单个模块中实现所有这些功能, 就算不是不可能的, 起码也是不切实际的。作为一种替代方式, 我们采用了分而治之 (divide-and-conquer) 的方法。将通信任务划分成一些子任务, 并根据它们相互的依赖关系按等级组织起来。具体来说, 就是将那些分别完成一部分通信功能的子任务组织成不同的层。每个较高的层都使用它下面的层提供的服务, 并为它上面的层提供服务。提供给高层的服务是透明的, 低层的异构性和实现细节都对高层隐藏起来。协议用于不同系统中各实体间的通信, 通常用来定义同一层中子任务的操作。

TCP/IP 协议, 或者更正式地称为因特网协议族 (Internet Protocol Suite), 使得互连的异构计算机网络间的通信变得更加容易。它是各种协议的组合, 这些协议通常按照图 0-1 所示分为四层。这里给出了各层中主要的及相关的协议。

- 应用层 (application layer) 由种类繁多的应用程序组成, 下面列出了其中的一些应用。
 - 超文本传输协议 (Hypertext Transfer Protocol, HTTP)。提供万维网 (World Wide Web, WWW) 服务。
 - Telnet。用来对一台计算机进行远程访问。
 - 域名系统 (Domain Name System, DNS)。在域名和 IP 地址间进行转换的分布式服务。
 - 简单网络管理协议 (Simple Network Management Protocol, SNMP)。用于管理本地或远程网络设备的协议。
 - 动态主机配置协议 (Dynamic Host Configuration Protocol, DHCP)。自动配置网络接口的协议。
- 传输层 (transport layer) 为应用层提供数据传输功能, 包括下列协议。

- 传输控制协议 (Transmission Control Protocol, TCP)。在 IP 网络基础上通过面向连接的 (connection-oriented) 的数据传输方式提供可靠的 (reliable) 数据传输。
 - 用户数据报协议 (User Datagram Protocol, UDP)。一个比 TCP 简单, 但不保证可靠性的无连接 (connectionless) 协议。
- 网络层 (network layer) 负责为经过网络的分组进行路由选择, 包括下列协议。
- 网际协议 (Internet Protocol, IP)。TCP/IP 协议栈的“主力”, 提供不可靠 (unreliable) 的、无连接服务。
 - 因特网控制报文协议 (Internet Control Message Protocol, ICMP)。用于差错和控制报文。
 - 因特网组管理协议 (Internet Group Management Protocol, IGMP)。用于多播 (multicast) 成员的管理。
- 链路层 (link layer) 负责处理所有的硬件细节, 以便为网络层提供数据传输服务。有各种不同的链路层技术可以为网络层协议提供支持, 比如下面列举的这些。
- 以太网 (Ethernet)。流行的多点接入 (multiple access) 局域网协议。
 - 无线 LAN。基于 IEEE 802.11 标准的无线多点接入局域网。
 - 点对点协议 (Point to Point Protocol, PPP)。连接成对主机的点对点 (point-to-point) 协议。
 - 地址解析协议 (Address Resolution Protocol, ARP)。负责解析网络层地址。

3

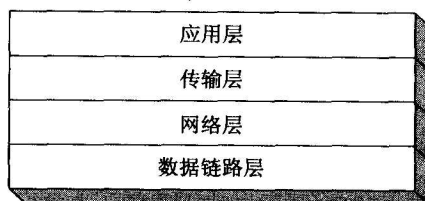


图 0-1 TCP/IP 协议栈

图 0-2 显示的是各层协议间的关系。我们会在后面的章节中对这些协议进行更详细的讨论。

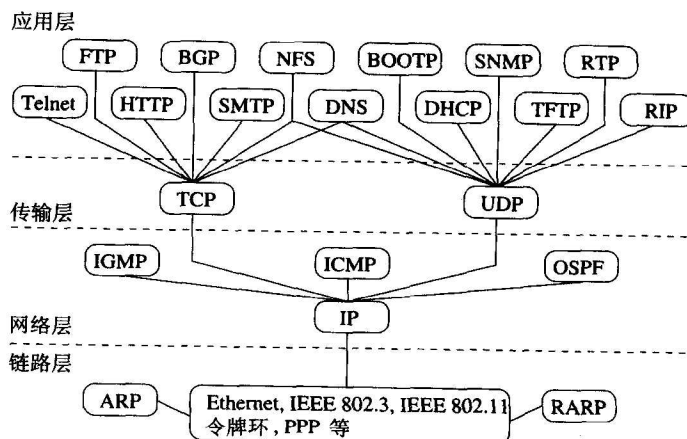


图 0-2 TCP/IP 协议

4

0.3 网际互连设备

因特网是由网际互连设备连接起来的计算机的集合。根据网际互连设备的功能及其操作所处的层次,可以将这些设备分为集线器(hub)、网桥(bridge)、交换机(switch)和路由器(router)。

集线器是物理层设备,用于连接多台主机。集线器只是简单地将从一个端口接收到的数据帧复制到所有其他端口上去,以此来仿真广播介质。网桥,有时又称为第二层交换机(layer two switches)¹,是链路层设备。网桥不对高层信息进行检查,因此可以快速转发通信流。网桥可用来连接远程终端,这样就扩大了网络的有效规模。第3章将对网桥进行进一步的讨论。

路由器,也称为第三层交换机(layer three switches),是包含路由选择功能的网络层设备。每台路由器都维护着一张路由表(routing table),路由表中的每个条目都包含一个目的地址和下一跳地址。任何一个路由器中都没有抵达目的地址的完整路由信息。分组到达时,路由器在其路由表中查找与分组的地址相符的条目,然后根据条目中的下一跳地址将分组转发出去。第4章将进一步讨论路由选择问题。

图0-3显示的是通过一台以太网集线器相连的两台主机之间进行通信时涉及的网络层次。因为它们使用的是相同的链路层协议,所以主机之间可以直接通信。图0-4显示的是两个使用不同链路层技术的网段是如何通过网桥进行互连的,网桥作为不同链路层协议间的接口,负责进行数据帧的转发。图0-5显示了两个网络是如何通过路由器进行互连的,路由器不仅要像图0-4所示的那样执行第二层的功能,还要执行网络层的主要功能:路由选择和分组转发。

如上面的这些例子所示,通过集线器连接形成的是单网段。多个网段间可以通过网桥或交换机互连以形成一个扩展的局域网,这些扩展局域网通常是与公司或其他社会机构的网络相关的。广域网(Wide Area Network, WAN)是用高速的点对点连接方式将不同企业网的路由器连接在一起而构成的。这些连接通常都建立在SDH/SONET电路交换网络基础之上。

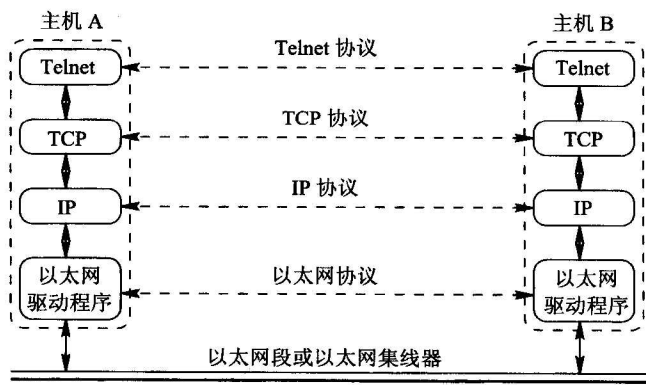


图0-3 两台主机通过同一个以太网段或通过以太网集线器进行通信时所涉及的网络层次说明

1. 工业上也会混用术语智能集线器(smart hub)来表示交换机。