Addison Wesley

TURING

图灵原版计算机科学系列

TCP/IP详解

卷1：协议

英文版

TCP/IP Illustrated

Volume 1: The Protocols

[美] W. Richard Stevens 著

人民邮电出版社
POSTS & TELECOM PRESS

TURING

图灵原版计算机科学系列

TCP/IP详解

卷1：协议

英文版

人民邮电出版社
北　京

图书在版编目（CIP）数据

TCP/IP详解. 卷1，协议＝TCP/IP Illustrated, Volume 1：The Protocols：英文/（美）史蒂文斯（Stevens, W. R.）著. —北京：人民邮电出版社，2010.4

（图灵原版计算机科学系列）

ISBN 978-7-115-22259-6

I. ①T… II. ①史… III. ①计算机网络－通信协议－英文 IV. ①TN915.04

中国版本图书馆CIP数据核字（2010）第017227号

内 容 提 要

本书是TCP/IP领域的经典之作！书中主要讲述TCP/IP协议，不仅仅讲述RFC的标准协议，而且结合大量实例讲述了TCP/IP协议族的定义原因，以及在各种不同的操作系统中的应用及工作方式，使读者可以轻松掌握TCP/IP的知识。本书内容详尽且具权威性，几乎每章都提供精选的习题，并提供了部分习题的答案。

本书适合任何希望理解TCP/IP协议如何实现的人阅读，更是TCP/IP领域研究人员和开发人员的权威参考书。无论是初学者还是功底深厚的网络领域高手，本书都是案头必备。

图灵原版计算机科学系列

TCP/IP详解 卷1：协议（英文版）

- 著 [美] W. Richard Stevens

责任编辑 杨海玲

- 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

北京艺辉印刷有限公司印刷

- 开本：800×1000 1/16

印张：37.5

字数：715千字 2010年4月第1版

印数：1-2 500册 2010年4月北京第1次印刷

著作权合同登记号 图字：01-2010-0309号

ISBN 978-7-115-22259-6

定价：79.00元

读者服务热线：（010）51095186 印装质量热线：（010）67129223

反盗版热线：（010）67171154

版 权 声 明

献给Brian Kernighan和John Wait，
感谢他们过去5年来给予的鼓励、信任和支持

前　言

概述

本书介绍的是TCP/IP协议族，但是视角却不同于其他TCP/IP教科书。我们将用一种流行的诊断工具来动态地监视协议，而不仅仅是描述协议及其功能。通过观察不同环境下协议的运作情况，可以更好地理解其工作原理和设计方案的由来。此外，本书还分析了协议的实现，读者无须花费精力去阅读数千行的源代码。

在网络协议从20世纪60年代到20世纪80年代的发展过程中，必须要使用昂贵的专用硬件才能监视到分组在线路上的传送情况。要理解由硬件显示的分组，还必须对协议极为熟悉。硬件分析器的功能也受限于硬件设计者所提供的内置功能。

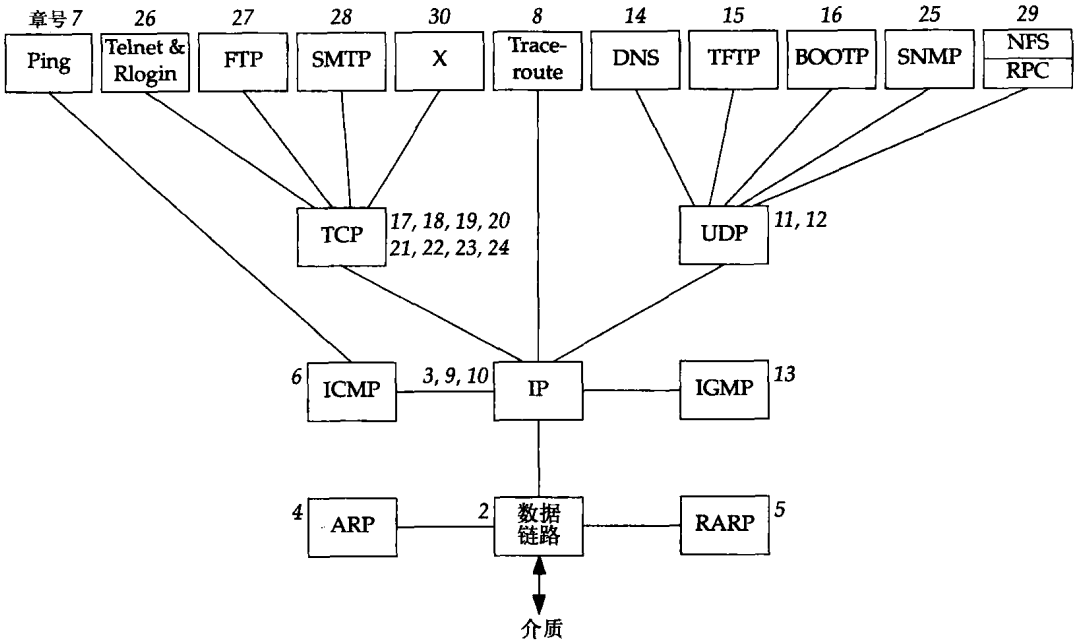
现在的情况有了显著的变化：随处可见的工作站就能监视局域网了[Mogul 1990]。只要在网络上连接一个工作站，然后运行一些公用软件（详见附录A），就能够观察线路上发生的情况。很多人可能会认为这只是一个诊断网络问题的工具，实际上它也非常有助于理解网络协议的工作原理，这正是本书的目标。

本书面向所有希望了解TCP/IP协议运行原理的读者：编写网络应用的程序员、利用

TCP/IP维护计算机系统与网络的系统管理员以及那些需要每天与TCP/IP应用打交道的用户。

本书的结构

下图给出了本书涉及的各种协议和应用，方框上的斜体数字指明了该协议或应用在哪一章讨论。



(图中略去的许多细节将在相应的章节中讨论。例如，DNS和RPC都用到了TCP，但从图中看不出来。)

我们采用一种自底向上的方式来介绍TCP/IP协议族。第1章介绍TCP/IP的基础知识；随后从链路层（第2章）开始向上介绍协议栈。这样做可以为不熟悉TCP/IP或者网络的读者提供阅读后续章节所需的背景知识。

本书在讲述过程中还采用了一种功能性的方法，而没有死板地遵循自底向上的顺序。例如，第3章描述了IP层和IP首部，但IP首部中的大量字段放在具体的应用（这类应用会用到特定的字段或者受特定字段的影响）场景中介绍才是最合适的。例如，分片技术(fragmentation）从UDP（第11章）角度最好理解，因为该协议经常会受其影响。生存时间（time-to-live）字段在第8章研究Traceroute程序时详细描述，因为该字段是该程序运行的基础。类似地，ICMP的很多特性也在后面的章节中讲述，主要考虑协议或应用使用特定ICMP报文的方式。

当然，我们也并不是想把所有的好东西都藏到最后。因此，只要具备了足以理解某种TCP/IP应用的知识基础，我们就会尽快对其进行介绍。在讨论完IP和ICMP之后，我们对Ping和Traceroute进行了讲解。在分析完UDP之后，我们描述了基于UDP的应用（多播、

DNS、TFTP和BOOTP)。不过TCP应用和网络管理必须在彻底讲清楚TCP之后才能加以介绍，因此只能留到最后了。本书侧重于介绍这些应用如何使用TCP/IP协议，而不会提供运行这些应用的所有细节。

致读者

本书自成一体，阅读之前不需要掌握具体的网络或TCP/IP知识。本书还提供了大量的参考文献，对具体领域感兴趣的读者可以进一步阅读这些资料。

本书有多种使用方式：对TCP/IP协议族的所有细节感兴趣的读者可以把本书用作自学参考书，从头看到尾；具有一定TCP/IP背景知识的读者可以跳过前面几章，直接从第7章开始阅读，然后选择感兴趣的章节进行重点研究。每章的末尾都安排了一些习题，附录D给出了大部分习题的答案，这样做是为了本书更适合作为自学参考书。

如果将本书的内容作为一或两个学期的计算机网络课程的一部分，重点应该放在IP(第3章和第9章)、UDP（第11章)、TCP（第17章至第24章）以及一些应用章节上。

全书贯穿了许多交叉引用，并提供了完整的索引，因此读者可以独立阅读各章的内容。书中用到的缩略词及相应的复合术语都详细列在索引之后。

如果你可以访问网络，建议你下载本书中使用的软件（附录F)，并在自己的计算机上进行实验。动手进行协议方面的实验可以大大丰富知识（并且使学习的过程更有乐趣)。

用于测试的系统

本书中的每一个示例都是在实际网络中运行过的，并将输出结果保存在文件中。图1-11（第18页）展示了书中所用的不同主机、路由器以及网络。(为方便读者阅读本书时随时查阅，本书最前面也放上了该图。）这种网络集合非常简单，其拓扑结构不会造成读者对示例的误解；此外，由于使用了4个系统作为路由器，我们可以看到路由器产生的错误报文。

多数系统的名字指明了所用软件的类型，如`bsdi`、`svr4`、`sun`、`solaris`、`aix`、`slip`等。这样，我们只要查看所显示出的系统名就可以确定软件的类型。

书中用到了多种操作系统和TCP/IP实现。

- 主机`bsdi`和`slip`上使用的是Berkeley软件设计公司的BSD/386版本1.0。这个系统源自BSD网络软件的2.0版。(我们在第17页的图1-10中给出了各种BSD版本的演进关系。)
- 主机`svr4`上使用的是U.H.公司的Unix System V/386 Release 4.0的2.0版。这是很普通的SVR4，包含Lachman Associates公司的TCP/IP标准实现(SVR4的大多数版本都采用该标准实现)。
- 主机`sun`上使用的是Sun公司的SunOS 4.1.3。SunOS 4.1.x系统可能是应用最为广泛的TCP/IP实现了，其TCP/IP代码源自4.2BSD和4.3BSD。

- 主机solaris上使用的是Sun公司的Solaris 2.2。Solaris 2.x系统的TCP/IP实现与之前的SunOS 4.1.x系统以及SVR4都不同。(该操作系统实际上是SunOS 5.2，但一般称为Solaris 2.2)
- 主机aix上使用的是IBM的AIX 3.2.2。该TCP/IP实现基于4.3BSD Reno版。
- 主机vangogh.cs.berkeley.edu上使用的是来自加州大学伯克利分校计算机系统研究组的4.4BSD。这个系统拥有伯克利最新的TCP/IP发布版本。(该系统并未在图1-11中给出，但是可以通过因特网获得。)

这些都是Unix系统，但实际上TCP/IP是独立于操作系统的，而且流行的非Unix系统几乎都支持TCP/IP。尽管有些程序（比如Traceroute）不是所有系统都提供的，但本书中的大部分内容都适用于非Unix实现。

排版约定

在展示交互式的输入和输出时，我们用粗体显示键入内容，以等宽正体显示计算机的输出，以斜体显示注释，示例如下：

```
bsdi % telnet svr4 discard      connect to the discard server
Trying 140.252.13.34...         this line and next output by Telnet client
Connected to svr4.
```

另外，我们将系统名（本例中是bsdi）作为shell提示符的一部分，以表明命令正在哪种主机上运行。

整本书中，我们随时会插入缩进的小号字体段落来描述历史问题或实现细节。

我们有时用ifconfig(8)来引用Unix手册中命令的完整描述。这种在命令名后面跟着括起来的数字的表示法是引用Unix命令的标准方式。括号中的数字是Unix手册中命令“手册页”的段号，在那里可以查到更加详尽的信息。但各种UNIX系统的手册的编排方式不尽相同，命令的段号也可能不同。我们采用的是BSD风格的段号（对于衍生自BSD的系统，比如SunOS 4.1.3，这个段号是相同的），但你的手册可能不是这样组织的。

致谢

尽管封面上只出现了作者一人的名字，但高质量教科书的出版需要许多人的共同努力。首先要特别感谢我的家人，他们在我写书的那段时间里饱受煎熬。再次感谢你们：Sally、Bill、Ellen和David。

顾问编辑Brian Kernighan无疑是计算机行业中最优秀的人物。他是最早阅读本书各次草稿的人，用红笔做了数不清的批注。他对细节的关注、对可读性的追求和对全稿的透彻审阅对作者来说是一笔巨大的财富。

技术审稿人从另一个角度给予了支持，他们发现了不少技术性的错误。他们的意见、建议以及（最重要的）批评使本书质量有了极大的提高。感谢Steve Bellovin、Jon Crowcroft、Pete Haverlock和Doug Schmidt为整部书稿所提的意见，也感谢Dave Borman、

Tony DeSimone、Bob Gilligan、Jeff Gitlin、John Gulbenkian、Tom Herbert、Mukesh Kacker、Barry Margolin、Paul Mockapetris、Burr Nelson、Steve Rago、James Risner、Chris Walquist、Phil Winterbottom和Gary Wright等人对书稿各部分提出了同样很有价值的意见。此外，还要特别感谢Dave Borman对所有TCP章节进行了全面审稿，感谢Bob Gilligan与作者合作完成了附录E。

人不可能孤立地工作，所以我想感谢下列曾经给予过我帮助的人（尤其是回答了我在电子邮件里提出的大量问题的人）：Joe Godsil、Jim Hogue、Mike Karels、Paul Lucchina、Craig Partridge、Thomas Skibo和Jerry Toporek。

我曾经被问过很多TCP/IP方面的问题，但却无法立即给人以答复，由此我产生了写这么一本书的念头。我意识到获得答案的最简单的方法就是做一些小实验，强制特定的条件发生，然后直接观察现象。感谢Pete Haverlock提出了有关探测的问题，也感谢Van Jacobson为本书解决问题提供了很多的公用软件。

完成网络方面的书籍需要工作在一个可以接入因特网的真正网络中。感谢美国国家光学天文台，尤其是授权我们接入其网络和主机的Sidney Wolff、Richard Wolff和Steve Grandi。Steve Grandi不仅解答了很多问题，还提供了各种主机上的账号，在此特别致谢。我还要感谢加州大学伯克利分校计算机系统研究组的Keith Bostic和Kirk McKusick授权我们使用最新的4.4BSD系统。

最后，出版社把所有的内容集成起来，并做了大量工作才得以向读者奉献最终版本。这一切都归功于编辑John Wait的组织协调，他真是最优秀的编辑！与John以及Addison-Wesley的其他专业人士合作是一件非常愉快的事情，他们的专业精神和对细节的重视都体现在本书的最终版中。

作者用James Clark编写的Groff软件包制作了本书的最终电子版——Troff硬拷贝。欢迎读者以电子邮件的方式反馈意见、提出建议或订正错误。

W. Richard Stevens

1993年10月于亚利桑那州图森市

Contents