计 算 机 科 学 丛 书

TCP/IP详解

卷1：协议

TCP/IP Illustrated

Volume 1: The Protocols

(美) W. Richard Stevens 著

范建华 胥光辉 张涛 等译

谢希仁 校

机械工业出版社 China Machine Press

Addison-Wesley

计算机科学丛书

TCP/IP详解
卷1：协议

(美) W. Richard Stevens 著
范建华 胥光辉 张 涛 等译
谢希仁 校

机械工业出版社
China Machine Press

《TCP/IP详解，卷1：协议》是一本完整而详细的TCP/IP协议指南。描述了属于每一层的各个协议以及它们如何在不同操作系统中运行。作者用Lawrence Berkeley实验室的tcpdump程序来捕获不同操作系统和TCP/IP实现之间传输的不同分组。对tcpdump输出的研究可以帮助理解不同协议如何工作。

本书适合作为计算机专业学生学习网络的教材和教师参考书。也适用于研究网络的技术人员。

本书版权登记号：图字：01-1999-2855

图书在版编目(CIP)数据

TCP/IP详解，卷 1：协议 /(美) 史蒂文斯(W. Richard Stevens)著；范建华等译. -北京：机械工业出版社，2000.4

(计算机科学丛书)

书名原文：TCP/IP Illustrated, Volume 1:The Protocols

ISBN 7-111-07566-8

Ⅰ. T… Ⅱ. ①史… ②范… Ⅲ. 计算机网络-传输控制协议 Ⅳ. TP915.04

中国版本图书馆CIP数据核字(2000)第15348号

机械工业出版社(北京市西城区百万庄大街22号 邮政编码 100037)

责任编辑：陈贤舜

北京牛山世兴印刷厂印刷 · 新华书店北京发行所发行

2000 年4月第1版 · 2001 年3月第4次印刷

787mm × 1092mm 1/16 · 27.5印张

印数：20 001-28 000册

定价：45.00元

译 者 序

我们愿意向广大的读者推荐W. Richard Stevens关于TCP/IP的经典著作(共3卷)的中译本。本书是其中的第1卷:《TCP/IP详解，卷1: 协议》。

大家知道，TCP/IP已成为计算机网络的事实上的标准。在关于TCP/IP的论著中，最有影响的就是两部著作。一部是Douglas E. Comer的《用TCP/IP进行网际互连》，一套共3卷（中译本已由电子工业出版社于1998年出版）。而另一部就是Stevens写的这3卷书。这两套巨著都很有名，各有其特点。无论是从事计算机网络的教师还是进行科研的技术人员，这两套书都应当是必读的。

本书的特点是内容丰富，概念清楚、准确，讲解详细，例子很多。作者在书中举出的所有例子均在作者安装的计算机网络上通过实际验证。各章都留有一定数量的习题。在附录D中，作者对部分习题给出了解答。在本书的最后，作者给出了许多经典的参考文献，并一一写出评注。

参加本书翻译的有: 范建华(序言和第1章~第11章)、杨明(第12章~第18章)、张涛(第19章~第24章)、徐哲(第25章~第26章)、蒋慧(第27章~第28章）和胥光辉(第29章~第30章以及全部附录和各参考文献中作者写的评注)。全书由谢希仁教授校阅。

限于水平，翻译中不妥或错误之处在所难免，敬请广大读者批评指正。

译 者

于解放军理工大学，南京

1999年12月

主要译、校者介绍

谢希仁，中国人民解放军理工大学(南京)计算机系教授，全军网络技术研究中心主任，博士研究生导师，1952年毕业于清华大学电机系电信专业。所编写的《计算机网络》于1992年获全国优秀教材奖。1999年再版的《计算机网络》第二版为普通高等教育“九五”国家级重点教材。近来还主持翻译了Comer写的《TCP/IP网际互联》计算机网络经典教材一套三卷本(电子工业出版社1998年出版)，Harnedy写的《简单网络管理协议教程》(电子工业出版社1999年出版)。

范建华，1971年1月出生，1992年毕业于中国人民解放军通信工程学院(南京)计算机应用专业，获学士学位，1996年获硕士学位，1999年获博士学位。现在总参第六十三研究所工作，讲师职称。曾获军队科技进步三等奖1项，在国内外期刊及会议上发表论文10余篇，翻译专业著作5本。目前从事通信装备模块化技术、数据开采和知识发现以及计算机网络应用等方向的研究工作。

胥光辉，1970年1月出生，1992年毕业于南京大学计算机系软件专业，获学士学位。1995年获南京大学计算机软件专业硕士学位。同年分配到解放军通信工程学院(南京)计算机系任教。现为解放军理工大学通信与信息系统专业博士研究生，讲师职称。主要研究方向：计算机网络和网络管理。曾参加国家自然科学基金资助项目、国家“863”项目以及其他有关网络项目的研究，发表学术论文8篇，译著2本。

《TCP/IP详解，卷1：协议》评语

“这本书肯定能成为TCP/IP开发人员和用户的权威著作。在我随手拿起书来翻阅的短短几分钟内，我就看到了过去曾经困扰我和我的同事们的好几个问题。Stevens解开了许多曾经只有网络专家才能领会的奥秘。作为一名从事多年TCP/IP开发工作的技术人员，我认为这是一本到目前为止最好的书。”

——3Com公司Synernetics部网络工程师　Robert A. Ciampa

“Stevens所有的书可读性都很强，而且有很高的技巧，这本书尤其如此。尽管已经有很多描述TCP/IP协议的书问世，但是Stevens在他的书中却提供了别人没有的深度和实现细节。它用可视化的方法使读者深入TCP/IP内部了解该协议是如何工作的。”

——《Unix评论》网络专栏　Steven Baker

“《TCP/IP详解，卷1：协议》对于开发人员、网络管理员以及其他需要理解TCP/IP技术的人来说，都是一本优秀的参考书。《TCP/IP详解》覆盖面广泛，涉及TCP/IP各个方面的主题，对专家来说它可以提供足够的细节，而对于初学者来说它则给出了详细的背景和注释。”

——NetManage公司市场部副总监　Bob Williams

“……区别就在于Stevens不仅希望给出协议的描述，而且要对它们进行解释。他主要的教学方法是直接阐述，并在每章后面附加习题，同时逐个字节地给出首部中的各个字段以及类似的内容，另外还给出真实的通信例子。”

——《Unix世界》　Walter Zintz

“比纯理论要好得多……W. Richard Stevens把基于多主机的配置作为解释TCP/IP的例子。《TCP/IP详解，卷1：协议》根据实例来阐述理论——使得该书与其他相关主题的书有所区别，不但可读性强，而且信息量大。”

——IBM TCP/IP开发部顾问　Elizabeth Zinkann

“他所使用的图表都很好，写和风格清晰易懂。总的来说，Stevens使一个复杂的问题变得容易理解。这本书值得每个人注意。请一定要阅读它并保留在您的书架上。”

——系统管理员　Elizabeth Zinkann

“W. Richard Stevens写了一本优秀的教科书和参考书。这本书组织结构良好，文字非常清新，提供了许多很好的例子来详细解释IP、TCP以及支撑协议和应用程序的内部逻辑和操作细节。”

——哈佛大学OIT/NSD顾问　Scott Bradner

前　言

概述

本书采用了一种不同于其他教科书的方式描述了TCP/IP协议族。用一个流行的诊断工具来观察这些协议的运作过程，而不是简单地描述这些协议以及它们做些什么。通过观察这些协议在不同环境下的操作过程，我们可以更好地理解它们是如何工作的，以及为什么要那样设计。另外，本书还对协议的实现进行了概述，当然并不要读者去阅读数千行的源代码。

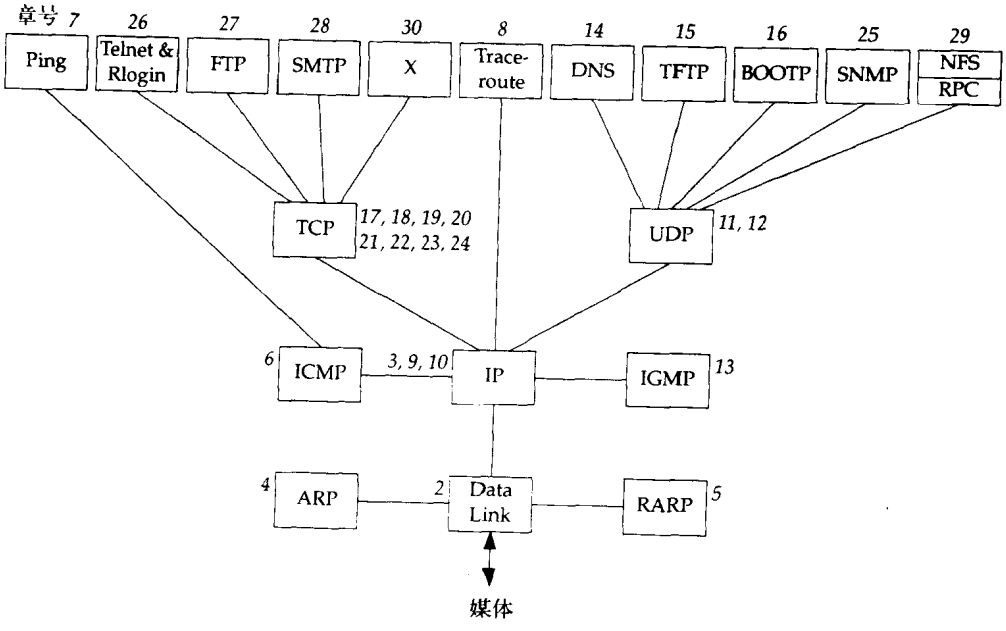
在20世纪60年代到80年代期间开发网络协议时，需要昂贵的专用硬件来观察分组“通过线路”的情况。同时，要理解硬件所显示的分组信息还需要对协议本身极其熟悉。硬件分析人员的作用受限于硬件本身的设计。

现在，对局域网进行监测的工作站随处可见，情况发生了激动人心的变化[Mogul 1990]。只需要在你的网络上连上一个工作站，并运行一些可公开获得的软件（参见附录A），那么你就可以对线路上的流通情况进行观察。许多人认为它只是一个诊断网络故障的工具，但是，它也是一个理解网络协议运作的强有力的工具，这才是本书的目标。

本书适用于那些希望理解TCP/IP协议如何操作的人：编写网络应用程序的程序员，负责维护基于TCP/IP协议的计算机系统和网络的系统管理员，以及经常与TCP/IP应用程序打交道的用户。

本书的组织

本书所涉及到的所有协议和应用程序如下图所示。每个方框旁边的斜体数字表示本书描述该协议或应用程序的章节。



本书各章中所讨论的细节要点并没有在上图中完全画出来。例如，采用TCP协议的DNS和RPC在这里就没有画出来。

我们采用自底向上的方法来讨论TCP/IP协议族。在第1章对TCP/IP协议进行简单介绍之后，我们将在第2章开始从链路层逐个向上讨论整个协议栈中的协议。这样，为那些在总体上不熟悉TCP/IP或网络的读者阅读后面的章节提供必需的背景知识。

本书还采用了功能方法来讨论各个协议，而并不遵循严格的自底向上次序。例如，第3章描述了IP层和IP首部。但是，IP首部中的许多字段最好是放到使用或受其影响的应用程序中进行解释。例如，分片最好通过UDP协议来理解，因为该协议经常受其影响。当我们在第8章讨论Traceroute程序时就完整地描述了生存时间字段，因为这个字段是该程序进行操作的基础。类似地，ICMP协议的许多特性留在后面的章节中才被讨论，以显示某个特定的ICMP报文如何被某个协议或应用程序所使用。

当然，我们并不想把所有的好戏都留在最后。对于TCP/IP应用程序来说，我们是一旦有了理解基础就对它们进行描述。Ping和Traceroute程序是在讨论完IP和ICMP协议之后就进行描述的。基于UDP协议的应用程序（如多播传送、DNS、TFTP以及BOOTP等）是在讨论完UDP以后才进行描述的。但是，对于TCP应用程序和网络管理来说，则必须留到最后，即完全讨论了TCP协议之后才对它们进行描述。本书着重讨论这些应用程序如何使用TCP/IP协议，而不提供它们运行的所有细节。

读者

本书的内容是独立的，不需要特定的网络或TCP/IP知识。如果读者对特定主题的某些细节感兴趣，本书提供了大量的文献可供参考。

本书可用于多个途径。它可用作自学参考书，适用于那些对TCP/IP协议族所有细节都感兴趣的读者。对于具有一定TCP/IP背景知识的读者可以跳过前面部分章节而从第7章开始阅读，然后着重阅读感兴趣的某些章节。在每一章的最后都附有习题，大多数习题的答案参见附录D。这使得本书可以最大限度地用作自学参考书。

如果作为计算机网络专业的教材，学习一个或两个学期，那么重点应放在IP（第3章和第9章），UDP（第11章），TCP（第17章~第24章），以及一些讨论应用程序的章节。

本书提供了许多相关的参考文献，各章的学习均可以单独进行。全书使用的缩写词以及组合词组都列在本书的最后。

如果你能够上网，那么应该通过网络获得本书中所使用的软件（附录F），并在你的机器上进行实验。对协议进行实际的实验可以获得最好的知识（而且更加有趣）。

用作测试的系统

本书中的每个例子都在一个实际的网络上运行过，输出结果都存在一个文件中。所使用的主机、路由器和网络如图1-11所示（这个图还复制在扉页前的插页中，以方便读者阅读本书时进行参考）。这个网络组合非常简单，其拓扑结构使得例子非常清楚。其中四个系统用作路由器，我们可以观察到路由器所产生的错误。

大多数的系统都有一个名字，以表示所使用的软件类型：`bsdi`、`svr4`、`sun`、`solaris`、`aix`、`slip`等等。这样，我们可以通过打印输出中的系统名字来识别软件的类型。

实际网络中使用了多种不同的操作系统和TCP/IP实现：

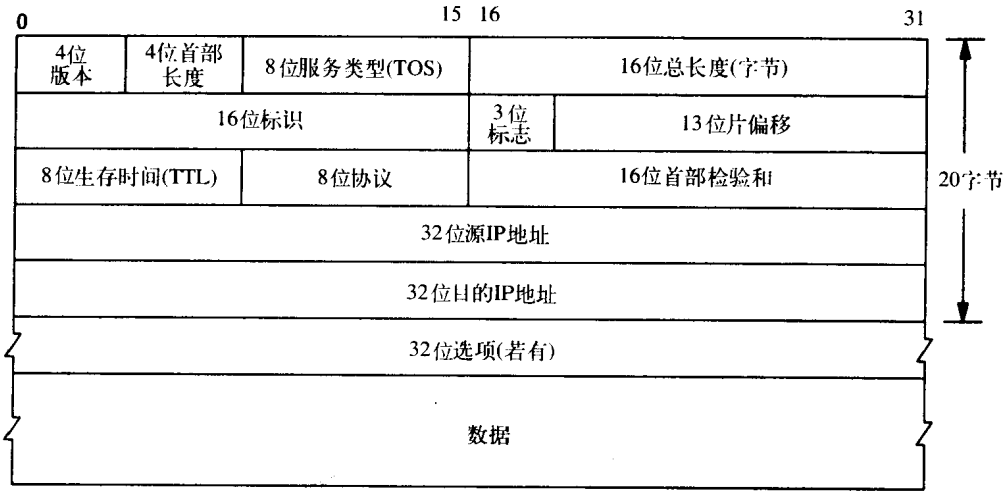
- 主机bsdi和slip上运行的操作系统是伯克利软件设计公司的BSD/386 1.0版。这个系统是从BSD网络软件2.0版派生出来的（不同版本的BSD系统的派生关系如图1-10所示）。
- 主机svr4上运行的操作系统是U. H.公司的UNIX V/386 4.0版。这个系统包含了Lachman协会大多数SVR4版本都使用的实现内容。
- 主机sun上运行的操作系统是Sun Microsystems公司的SunOS 4.1.3版。SunOS 4.1.x系统很可能是使用最广泛的TCP/IP实现。TCP/IP代码是从4.2BSD和4.3BSD派生而来的。
- 主机solaris上运行的操作系统是Sun Microsystems公司的Solaris 2.2。Solaris 2.x系统是一个与SunOS 4.1.x和SVR4都不相同的TCP/IP实现（操作系统实际上是SunOS 5.2，但通常称作Solaris 2.2）。
- 主机aix上运行的操作系统是IBM公司的AIX 3.2.2，其TCP/IP实现是基于4.3BSD Reno发行版。
- 主机vangogh.cs.berkeley.edu上运行的是伯克利加利福尼亚大学计算机系统研究小组开发的4.4BSD系统。该系统是伯克利发行的TCP/IP最新版本（这个系统没有画在扉页前图中，但是可以通过Internet访问它）。

尽管这些都是Unix系统，但是TCP/IP是独立于操作系统的，几乎在每个非Unix系统中都可以实现。本书中的大部分内容也可以应用于非Unix系统，但有些程序（如Traceroute）可能不是所有系统都能提供。

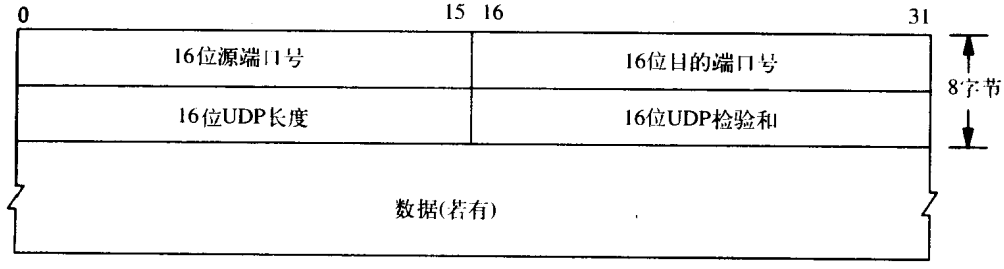
有时我们会参考Unix手册中关于某个命令的完整描述，如ifconfig(8)。命令名后面跟一个带括号的数字是参考Unix命令的一般方法。括号中的数字是该命令位于Unix手册中的小节号，在那里可以找到关于该命令的其他信息。不幸的是，并不是所有的Unix系统手册组织结构都一致，即通过小节号来区分不同的命令组。这里，我们采用的是BSD风格（BSD派生系统都是一样的，如SunOS 4.1.3），但是你的手册可能与此并不相同。

W. Richard Stevens
rstevens@noao.edu
Tucson, Arizona
1993年10月

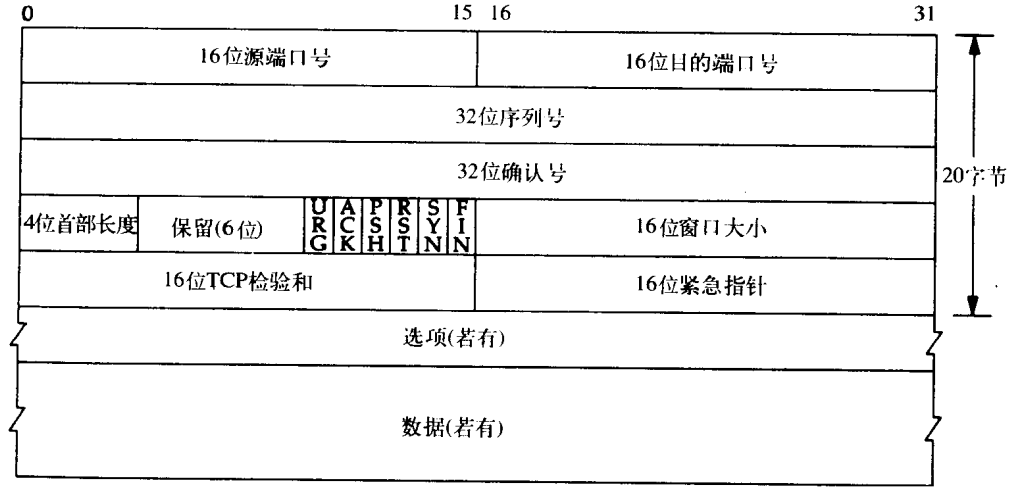
IP首部

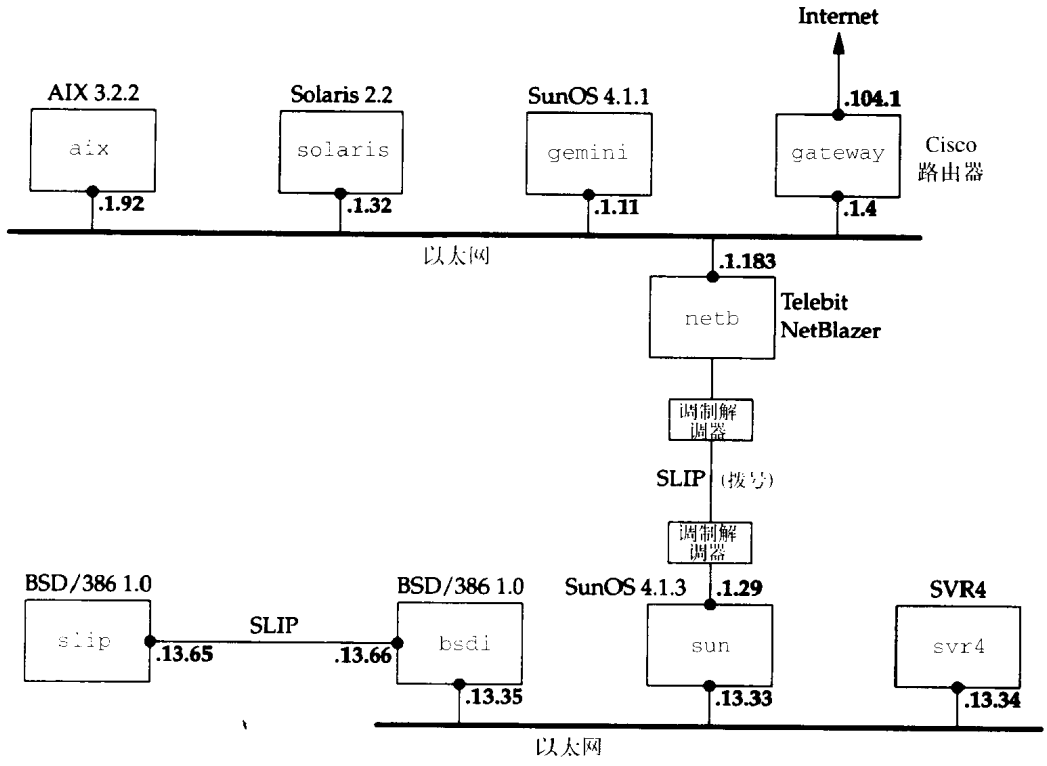


UDP首部



TCP首部





本书所有的例子使用的B类网络**140.252**中的部分地址。
所有的主机都在`tuc.noao.edu`域中。

目　录