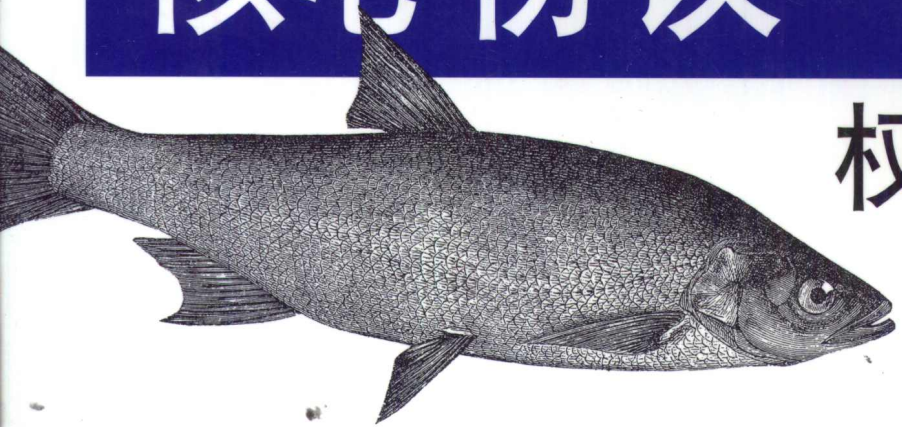
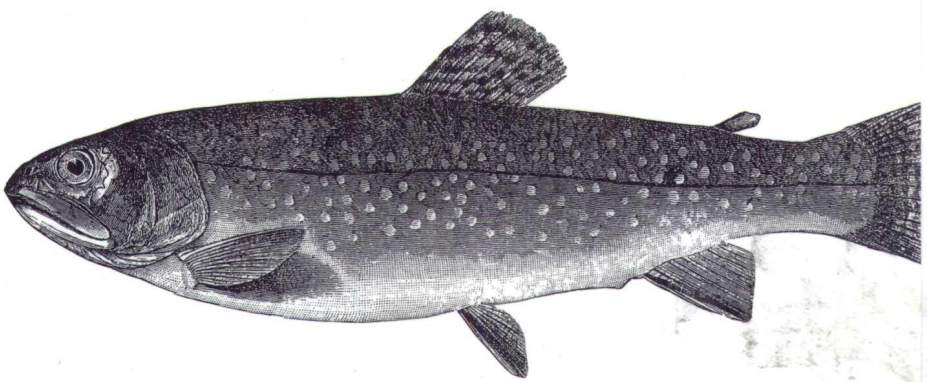
Internet Core Protocols: The Definitive Guide

含光盘

Internet 核心协议

权威指南





O'REILLY®

中国电力出版社

Eric A. Hall 著

张金辉 译

Internet 核心协议权威指南

Eric A. Hall 著

张金辉 译

O'REILLY®

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

O'Reilly & Associates, Inc. 授权中国电力出版社出版

中国电力出版社

图书在版编目（CIP）数据

Internet 核心协议权威指南 /（美）霍尔（Hall, E.）著；张金辉译．-北京：中国电力出版社，2002

书名原文：Internet Core Protoclos: The Definitive Guide

ISBN 7-5083-0941-3

Ⅰ.I... Ⅱ.①霍... ②张... Ⅲ.计算机网络－通信协议 Ⅳ.TP915.04

中国版本图书馆 CIP 数据核字（2002）第 007282 号

北京市版权局著作权合同登记

图字：01-2002-1212 号

书　　名 / Internet 核心协议权威指南

书　　号 / ISBN 7-5083-0941-3

责任编辑 / 陈维宁

封面设计 / Edie Freedman，张健

出版发行 / 中国电力出版社（www.infopower.com.cn）

地　　址 / 北京三里河路 6 号（邮政编码 100044）

经　　销 / 全国新华书店

印　　刷 / 北京市地矿印刷厂

开　　本 / 787 毫米 × 1092 毫米　16 开本　29 印张　424 千字

版　　次 / 2002 年 3 月第一版　2002 年 3 月第一次印刷

印　　数 / 0001-5000 册

定　　价 / 59.00 元（册）

O'Reilly & Associates公司介绍

为了满足读者对网络和软件技术知识的迫切需求，世界著名计算机图书出版机构O'Reilly & Associates公司授权中国电力出版社，翻译出版一批该公司久负盛名的英文经典技术专著。

O'Reilly & Associates公司是世界上在UNIX、X、Internet和其他开放系统图书领域具有领导地位的出版公司，同时是联机出版的先锋。

从最畅销的《The Whole Internet User's Guide & Catalog》（被纽约公共图书馆评为二十世纪最重要的50本书之一）到GNN（最早的Internet门户和商业网站），再到WebSite（第一个桌面PC的Web服务器软件），O'Reilly & Associates一直处于Internet发展的最前沿。

许多书店的反馈表明，O'Reilly & Associates是最稳定的计算机图书出版商——每一本书都一版再版。与大多数计算机图书出版商相比，O'Reilly & Associates公司具有深厚的计算机专业背景，这使得O'Reilly & Associates形成了一个非常不同于其他出版商的出版方针。O'Reilly & Associates所有的编辑人员以前都是程序员，或者是顶尖级的技术专家。O'Reilly & Associates还有许多固定的作者群体——他们本身是相关领域的技术专家、咨询专家，而现在编写著作，O'Reilly & Associates依靠他们及时地推出图书。因为O'Reilly & Associates紧密地与计算机业界联系着，所以O'Reilly & Associates知道市场上真正需要什么图书。

译者序

非常有幸能够翻译这么重要的一本书，因为这毕竟是一个新兴的领域，年轻人一般都会不可避免地陷入其中不能自拔，我也毫不例外。今年3月我试着翻译了本书的其中一章，真实地感受到了其中的奥妙以及翻译带来的快乐。

因为中美之间的文化差异，一些看上去很幽默的句子却只能机械地译成呆板的技术术语。不过这么翻译的结果确实让本书更像是一本技术指导书。在看过本书以后，我就发觉我对网络的认识已经深入了一层，并抹去了对网络的最后一丝神秘感。我的愿望是让本书的所有读者都从中得到我已经得到的知识。

我不得不说的是，本书是一本指导书，它能带给我们的不是对网络的全面认识，而是对网络的基础认识，但它非常透彻地分析了网络基础知识，例如我们常见的ping程序的原理、过程、输出结果和解释都可以在本书找到准确答案。

有时候，我们会抱怨市场上有关网络的书籍花样繁多，但找不到非常权威的书籍。说实话，我以前也有这种感受，但现在我的感受已经变了。我意识到了这本书的重要性及权威性。我不敢说它是最权威的书籍，但它所介绍的知识却绝对具有权威性，因为本书的作者对他所示出的例子都作了试验，从本书中，你可以感觉到他所付出的巨大努力。我所做的只不过是将他的成果转化成我们中国人容易理解的方式而已。

再次谢谢帮助我的所有人，和给我最大精神鼓励的新婚妻子，我在此一并表示深深谢意。

至于本书中一些比较偏僻的词汇和少见的专业术语，我已经列在词汇表中，如果对本书或词汇翻译有什么意见的话，可以给我发信联系，我的电子邮件地址是*zhanggolden@eyou.com*。最后希望大家学有所成！

张金辉

目录

序 1

前言 5

第一章 TCP/IP 概述 15

Internet 简史 15

TCP/IP 结构 20

深层的 TCP/IP 协议和服务 25

应用协议如何在 IP 上通信 39

第二章 IP 43

IP 标准 44

IP 首部 68

使用中的 IP 99

IP 的故障排除 107

第三章 ARP 109
ARP 标准 110
ARP 包 125
使用中的 ARP 138
调试 ARP 问题 144

第四章 多播和 IGMP 147
IP 多播和 IGMP 规范 149
IGMP 消息 161
使用中的多播和 IGMP 177
多播和 IGMP 的故障排除 180

第五章 ICMP 185
ICMP 规范 186
ICMP 消息 209
使用中的 ICMP 245
ICMP 的故障排除 261

第六章 UDP 266
UDP 标准 267
UDP 首部 273
UDP 的故障排除 279

第七章 TCP 284
TCP 标准 285
TCP 首部 340
使用中的 TCP 372
TCP 的故障排除 397

附录一 Internet 标准化进程 411

附录二 IP 地址基础 419

附录三 使用光盘 429

参考文献 437

词汇表 443

序

最初，Internet是关于不同的包交换（packet-switched）网络之间连接的研究成果，它的工作方式决定了连接到每一个包网络上的计算机不需要知道任何其他网络的特性及其存在，只需知道与主机直接连接的网络即可。随之出现的是网络层化设计，这种设计采用封装技术通过中间网络和网关来向终端主机传输端对端Internet包，第一个Internet包含三个大型或中型网络，ARPnet、Atlantic Packet Satellite网络（SATNET）和地面移动Packet Radio Network（PRNET）。最后它把1973年在Xerox PARC开发的第一个3MB的Ethernet也包容了进来。

现在，Internet经过了约25年的发展，已经拥有几十万个网络，并为大约45000000台计算机和150000000个用户提供服务。而且，成员网络中主干线路的速度已由最初的每秒几千字节上升到每秒几十亿字节，而实验室演示用的传输速度已经达到每秒几万亿字节。随着Internet的增长，它的复杂性和用户都有了极大的增长。但是相对于用户甚至是操作员的总人数来说，详细了解Internet运行所用的协议和系统的人数比例却正在下降。

更为糟糕的是，正在使用的协议和服务的数量也从少量几种上升到几百种。过去，只要一个高级管理员就可以管理路由器、域名服务器、邮件服务器和网络中的其他资源，而现在，过度的专业化使得一个人管理所有的事物变得不可能了。在许多大公司里，一些部门专门管理网络路由，而其他部门管理拨号服务，还有其他部门管理网络和邮件系统、域名系统和新闻组。

一个严重的问题是，大公司能够雇用那些了解整个网络各部分的专家，但大部分公司雇用不起这么多专家，所以不得不雇用少数几位，这些专家必须知道所有必需的东西。并且，调试和分析Internet问题跟专业化是相悖的。网络上的问题往往是因网络中不同部分之间的相互作用引起的。如果电子邮件没有发送，问题出现在邮件服务器自身吗？是路由、域名系统或者Ethernet地址到Internet地址映射的低层协议出了问题了吗？想让一个人在这么多领域内（加起来有一打多）诊断问题是不可能的，但是很多网络操作员每天都面临着这个问题。

出现问题时，管理员用很多工具来调试。这些工具包括包分析器，它可以显示网络通信量的内部核心，但是它不能说明通信量的含义。厂家自己的文档是另外一系列工具，但是有时厂家的说明文档也可能会跟有问题的软件一样误解了规范。最后一个选择是让管理员遍览协议的技术专业知识来找到问题的根源。但是如果在凌晨4点，芝加哥的Web服务器与亚特兰大的数据库服务器断开了连接，那么这些专业知识的作用也是有限的。这些文档大部分内容是行为的严格定义，通常不会描述协议是怎么出错的。

这就是出版此系列书的原因。通过本系列书，Eric Hall会向我们揭示IP网络上所使用的协议背后的功能和原理，并提供隐藏在这些协议工作机制背后的理论的完全检验。并且，他还使用从真正的监听工具中获取的包来支持他的指南式讨论。当需要知道包中的某个特定字段的情况时，他也提供了必不可少的参考。另外，Hall还介绍了出现问题时的症状，提供了最常见的互用性问题的详细线索和论述。

指南、参考书和调试指导三方面的结合使得这些书成为包罗万象的基于IP网络的用户手册。对于任何与Internet技术打交道的网络管理员来说，这一系列书都是很具有吸引力的，特别是现在Internet正在经历由近乎指数增长所带来的越来越多的痛苦，该系列书就更为重要。尽管现在网络已经连接了至少44000000个设备，但是所有的信息表明，到2006年将会有10亿个设备连接到网络中，包括使用IP的传感器、车库开门器、录像机、IP电话和所有其他的办公和家用设备。当然，这些设备中的很多需要新协议，网络也会变得更加复杂。

现在网络研究已经转向适用于星际距离的网络了（这时就要考虑速度问题了）。目前已经计划要建立一个启用Internet的火星基站，通过一系列的星际网关连接到陆地。NASA的火星任务开始于1998年，并会在这一千年继续对它进行研究。这些探索的

其中一部分就是Internet网络的组成部分，也许有一天那些对邻近行星、月亮和太阳系外的大行星卫星的探索者和殖民者会将此作为重要的通信工具。

但是反回来，随着Internet的无节制的增长，我们还需要在现有的Internet上付出大量的努力。我们还需要诸如这一系列书的帮助，来帮助分析我们现有Internet和计划中的未来Internet上所会出现的问题。

-- Vint Cerf

前言

某星期五的下午4：45，当你想早一点离开公司回家时，电话铃响了，是你的一个用户打来的，他不能连接邮件服务器，怎么试都不行。糟糕的是，他还要在下班以前给他的老板发送一个报告，这就意味着你必须在回家以前先解决这个问题。

但在修复之前，你必须先确切地知道问题是什么。是用户提供的用户名和密码有误吗？是用户使用了旧的电子邮件客户机程序，而这个客户机程序和新服务器的一些新特性不兼容？或许用户的邮箱被别的操作给锁住了？或者是一些基本的网络连接性问题使得计算机不能够通信？

不幸的是，Internet协议和应用程序从未有过的成功和其大范围适用性使得它产生的复杂性也同样史无前例的多。尽管关于如何安装某一特定厂家的产品有一大堆文档和说明，但其中很少能够详细说明这些产品的基础协议是如何实现的。关于灵巧的电子邮件过滤器已经有了大量的文献，但支持POP3和SMTP命令的资料却没有。这至少使故障的排除变得困难了。更糟的是，当厂家在互相指责的时候，你却在自己的身上找原因。

问题是，为了能够有效地设计、实现、管理和支持基于Internet的、使用核心标准的协议和服务的不同实现，你必须知道它们是如何工作的。所有的一切最终都归结于协议（包括失败的命令和它们产生的错误）。最快的解决途径还是需要理解协议层的工作过程。

这时就需要你能够捕获网络中的通信量，（更重要的是）能够理解正在观察的包。本书的目的在于揭示现有以Internet为核心的网络中最常见协议的细节。通过阅读本书，我们将了解到在TCP/IP网络中使用的每一种核心协议的设计的背景知识，还可以了解到它们所具有的选项和参数的详细参考信息。此系列书的其他书籍将会用同样的方式阐述应用层协议。如果需要知道一些事物不能正确工作的原因，本书和协议分析器的结合将帮助我们解决问题。

读者对象

本书主要是为那些设计、建立、管理和支持使用基于Internet协议和服务的计算机网络的用户而写的。尽管本书对于那些高级用户和程序员可能也有用，但它主要可用作那些与TCP/IP息息相关的人们的参考书。

本书对于那些对计算机网络及其工作机理已经有了初步了解的人，以及那些可能已经知道了一些关于TCP/IP工作机制的知识、但还想对TCP/IP了解更多的人来说，是一本非常好的书。如果你不知道如何给计算机指定IP地址，本书可帮不了你。你应该看Craig Hunt著的《TCP/IP Network Administration》或Craig Hunt和Robert Bruce Thompson合著的《Windows NT TCP/IP Network Administration》（这两本书都是由O'Reilly公司出版的）（译注1）。但是，如果你想知道更多关于IP的生命期或服务类型参数和它们对网络的影响的话，本书就很适合。

应该注意的是，本书并不是针对所有的特定实现和应用程序的参考书。我可能会提到某个特定的实现，这只是为了举例说明，并不能用来代替官方的产品说明。

组织

本书阐述的是基本协议，这些协议提供了所有TCP/IP应用程序和服务所使用的网络和传输服务。包括有关IP、UDP、TCP和通用的支持协议，如ICMP、IGMP和ARP的章节。本书最后附录了与这些协议工作原理间接相关的资料。

译注1：这两本书的中文版《TCP/IP网络管理》和《Windows NT TCP/IP网络管理》已由中国电力出版社引进出版。详情请访问*http://www.infopower.com.cn*。

下面分章详细阐述本书的组织:

- 第一章，TCP/IP概述，从总体上介绍了TCP/IP的历史、设计目标和不同协议之间的内部关系。
- 第二章，IP，详细讨论了网际协议，包括基本环节如IP地址、包的转发、服务提供的有限可靠性、分段存储和优先存储。
- 第三章，ARP，阐述了IP设备如何在网络中互相定位，还阐述了用于不同类型任务中的ARP的变化。
- 第四章，多播和IGMP，描述了网络中多播是如何工作的，以及设备是如何用多播路由来注册并加入分布式多播输入的。
- 第五章，ICMP，讨论了IP使用的错误报告服务，如何实现不同的ICMP消息，并且解释了如何使用ICMP上提供的交互服务来诊断网络。
- 第六章，UDP，阐述了那些不需要TCP可靠性服务的应用程序所用的轻便的、易于出错的传输协议。
- 第七章，TCP，介绍了极其复杂的传输协议的各个主要方面，包括流控制、可靠性和几乎所有基于Internet的应用程序都在使用的网络和应用程序管理服务。
- 附录一，Internet标准化进程，介绍了Internet开发人员写出提案并最终成为标准的进程，也介绍了监视这些过程的权威机构。
- 附录二，IP寻址基础，详细描述了IP地址和它们的格式化规则。
- 附录三，使用光盘，介绍了如何安装Shomiti Surveyor Lite，Shmiti Surveyor Lite是随书所附光盘上的网络分析工具。这张光盘也包含所有现在已公布的RFC（包括本书已用到的）。

每章都大致分为三部分：协议简介、协议语法的细节，以及一些实时用法和故障排除注意事项。你可以根据自己的实际情况来安排如何读这本书。

初学者

如果你对TCP/IP网络很陌生，并且只是想学一些协议（或者大体上说是Internet）的主要概念和结构，那么你可以看第一章，然后看第二章和第七章介绍性部分。读了这些部分以后，你就会对TCP/IP是如何工作的有一个深刻的理解。

工作中的管理员

如果你负责管理网络，并希望对核心协议有一个深入的了解，那么你可以阅读第二章开始部分的介绍性材料、第七章和第五章。事实上，在碰到问题之前，你就应该尽快把这些部分看完。然后当问题突然出现时，可以回过头来看一看参考材料和故障排除提示。

如果你已经碰到了关于某一协议或服务的问题，那么你可能需要开始捕获包，并且详细研究那个出了问题的协议的参考部分。研究包捕获，努力找到问题的出处。然后观察出了问题的包，在相应章节的参考部分找到相应的部分，看看是否能够明白问题之所在。

最后，光盘中包括Shomiti Surveyor Lite，这是一个分析网络通信量的全能工具。（想获得更多的信息，可以访问*www.shomiti.com*。）它也包含了所有RFC的全部文本，这是网络管理人员的另外一个必备工具。归根结底是RFC（而不是本书）规定了网络该如何工作。网络上提供了所有的RFC，但如果你的网络不能工作的话，你就得不到这些东西。有了本书、网络分析器和RFC，你就拥有了解决问题所需要的东西。

怎样阅读本书

本书没有使用任何代码例子，并且在例子中也几乎不用程序输出。如果用了程序输出，往往会提供应用程序的屏幕显示，程序输出并没有在本书的文本中逐行显示。

术语

大部分网络管理人员用“包”或者“数据报”来表示那些通过网络传输的数据。但是，随着TCP/IP的演变，用来描述特定协议传输的数据单元的词语有了很多种。RFC1122把所有这些词语收集在一起，并根据特定的协议定义了每个词语的用法。这些词语在本书中的使用方式与在RFC1122中是一致的。

帧

帧（*frame*）是指用适用于某网络的链接层协议通过该网络传送的数据单元。包括链接层封装技术，如Ethernet II 帧、802.3 Ethernet帧和Token Ring帧。

IP 数据报

IP 数据报（*datagram*）是 IP 协议管理的数据单元，包括传输的所有数据，还包括与这个数据相关的 IP 首部。实际上，IP 数据报就是 IP 工作的直接对象。

IP 包

IP 包（*packet*）是 IP 数据报的另外一个名称，但是这个名称往往指的是帧的数据报部分，而不是数据报本身。例如，收发系统能把一个 IP 数据报看作一个实体，但是为了在一系列中间网络上传输，这个数据报可能需要分为多个 IP 包。一般来说，主机处理 IP 数据报，而路由器处理 IP 包。

消息

消息（*message*）是指从一个高层协议（诸如 UDP 或者 TCP）发送过来的数据单元，它包括传输的数据和与这个数据相关的特定传输首部。尽管大多数情况下消息数据是由应用程序特定的协议产生的，但是 ICMP 和 IGMP 也可以直接和 IP 通信，因此也会产生消息数据。最终消息成为 IP 数据报的数据部分。

TCP 段

尽管由 TCP 产生的首部和数据被看作是消息，但 TCP 消息能够通过多重消息来传播，在这种情况下，一般将消息看作为段（*segment*）。

图像

本书使用不同的图像表示不同类型的网络设备，包括主机、路由器、调制解调器和其他的基本构造设备。为了尽量减少混淆，我们将给出在本前言部分的三个图中所出现的记号在各个章节中的通用解释。

例如，图 P-1 示出了 Token Ring 和 Ethernet 网中常见的应用程序客户机、服务器和网络路由器的符号。

图 P-2 画出了广域网（WAN）中常见的符号，包括调制解调器、卫星、微波无线电接收装置和一般的 WAN（例如帧中继或者租用线路网络）。可以看到网络路由器、应用程序客户机和服务器与基于 LAN 的拓扑结构中的图形是一样的。

请注意，有时某个普通主机用“应用客户机”符号来表示，这意味着这个设备正在向另外一台网络设备发送或者接收数据，另外那台设备可以是客户机或者服务器，这就是说这些设备所扮演的角色不在现在的讨论之列。