

HZ BOOKS
华章教育

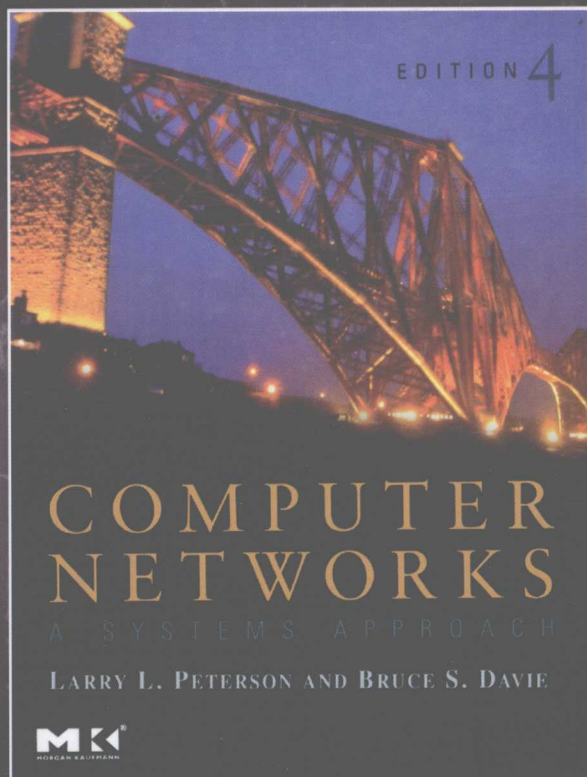


计 算 机 科 学 丛 书

原书第4版

计算机网络 系统方法

(美) Larry L. Peterson Bruce S. Davie 著 薛静锋 张龙飞 胡晶晶 王勇 译 曹元大 审校



Computer Networks A Systems Approach

(Fourth Edition)



机械工业出版社
China Machine Press

计 算 机

TP393/340-1

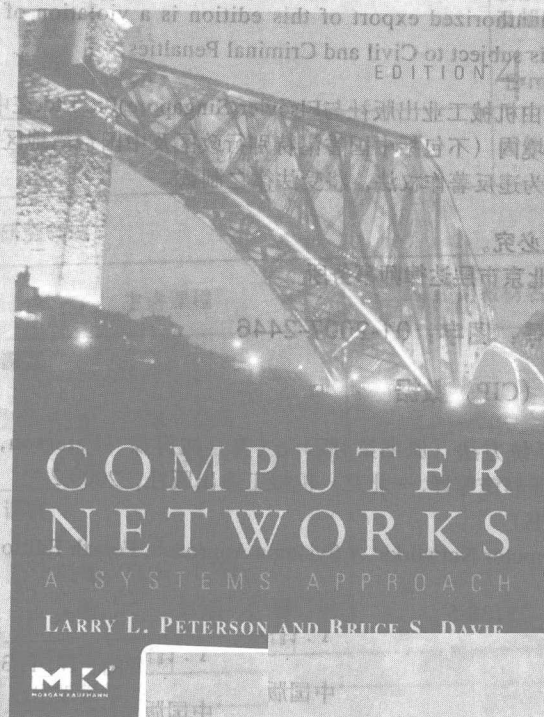
2009

书

原书第4版

计算机网络 系统方法

(美) Larry L. Peterson Bruce S. Davie 著 薛静锋 张龙飞 胡晶晶 王勇 译 曹元大 审校



Computer Networks

A Systems Approach

Fourth Edition



机械工业出版社
China Machine Press

本书介绍计算机网络技术的基本概念和应用,内容翔实、论述严谨。本书采用“系统方法”来分析计算机网络,把网络看作一个由相互关联的构造模块组成的系统,而不是严格地进行分层,介绍了很多网络中的新技术,包括对等网、IPv6、覆盖网络、内容分发网络、MPLS与交换、无线与移动技术等,涉及大量的实际应用。本书还引入丰富的因特网实例,说明实际网络的设计,更便于读者理解。每章后面的习题有助于读者掌握和复习知识要点。

本书适合作为高等院校计算机及相关专业的高年级本科生和研究生的教材,也可供网络专业人员参考。

Larry L. Peterson and Bruce S. Davie: Computer Networks: A Systems Approach, Fourth Edition (ISBN-13: 978-0-12-370548-8 ISBN-10: 0-12-370548-7).

Copyright © 2007 by Elsevier Inc. All rights reserved.

Authorized Simplified Chinese translation edition published by China Machine Press

ISBN: 978-981-259-918-6

Copyright © 2009 by Elsevier (Singapore) Pte Ltd. All rights reserved.

Printed in China by China Machine Press under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由机械工业出版社与Elsevier(Singapore)Pte Ltd.在中国大陆境内合作出版。本版仅限在中国境内(不包括中国香港特别行政区及中国台湾地区)出版及标价销售。未经许可之出口,视为违反著作权法,将受法律之制裁。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

本书版权登记号:图字:01-2007-2446

图书在版编目(CIP)数据

计算机网络:系统方法(原书第4版)/(美)彼得森(Peterson, L. L.)(美)戴维(Davie, B. S.)著;薛静锋等译.一北京:机械工业出版社,2009.2

(计算机科学丛书)

书名原文:Computer Networks: A Systems Approach, Fourth Edition

ISBN 978-7-111-25138-5

I. 计… II. ①彼… ②戴… ③薛… III. 计算机网络 IV. TP393

中国版本图书馆CIP数据核字(2009)第002485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:王 玉 盛思源

北京诚信伟业印刷有限公司印刷

2009年2月第1版1次印刷

184mm×260mm • 30印张

标准书号:ISBN 978-7-111-25138-5

定价:65.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线:(010) 68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域中取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章分社较早意识到“出版要为教育服务”。自1998年开始，华章分社就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专程为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章分社欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



译者序

自20世纪90年代以来,计算机网络技术与应用在我国迅猛发展,网络技术已成为广大计算机用户应用的主流。网络从业人员越来越多,计算机网络方面的专业技术人员和在校学生迫切需要一本面向用户、面向应用和理论联系实际的,介绍新技术、新成果和新趋势的,以及难易程度适当的计算机网络书籍。

为了使广大从事网络应用系统开发和应用的人员尽快和全面地掌握网络的基本理论,熟悉网络环境以及各种网络实用技术,我们将Larry L. Peterson 和 Bruce S. Davie的畅销经典教科书《计算机网络:系统方法》第4版推荐给广大读者。这一版对第3版做了重大的提升和改进,扩充了网络安全、XML、覆盖网络(包括对等网和内容分发网络)、Web服务(包括SOAP和REST体系结构)、无线技术(包括Wi-Fi、WiMAX标准和以3G标准为代表的蜂窝无线技术)、域间路由、多媒体应用的服务质量(包括VoIP和视频流)、拥塞控制机制等方面的最新材料,调整和增加了一些章的习题,并给出了部分习题解答。

本书作者强调网络现有工作方式的成因,采用“系统方法”,把网络看作一个由相互关联的构造模块组成的系统(反对严格地分层),并引入丰富的因特网实例,说明实际网络的设计。书中给出的程序代码不再基于某个特定的操作系统,而是重新改编为适用于通用的环境,说明网络软件是如何实现的,借此让读者了解所有建造网络的基础构件是如何结合在一起的。

本书是确保成功的课堂教学和高效网络运行的重要资源。与通常的网络教材不同,本书用系统化的观点探讨计算机网络。它通过给出构建网络的基础构件,讲述这些基础构件如何架构一个完整的网络,以及网络为什么这样设计。它是一本最新版的计算机网络的优秀教材,为学生和专业人士理解现行的网络技术以及即将出现的新技术奠定了良好的理论基础。

本书由北京理工大学薛静锋副教授主持翻译,参加本书翻译工作的有薛静锋(第5、6章、前言、选题解答、术语)、张龙飞(第3、4、7章)、胡晶晶(第1、2章)和王勇(第8、9章)。全书由北京理工大学曹元大教授审校。

由于译者水平有限,书中可能有错误或不甚完善之处,欢迎读者批评指正。读者如果对翻译方面有意见,可以通过E-mail与译者联系,地址为:xuejf@bit.edu.cn。

译者

2008年2月

序 言

这本经典著作问世已经10余年了，在这些年中发生的许多事情是令人惊奇的。Web已经从实验阶段变革到了万维网时期，基于IP的语音和P2P内容共享已经出现了，百兆技术已经普及，宽带网也已经进入了家庭，僵尸网络和其他可怕的安全问题越来越多。很多事情都发生了变化，新技术不断出现，但因特网的基础依旧存在。

10年中，本书也进行了多次修改，发展到了第4版。但本书的基本创作理念与第1版是相同的，本书给出了你所需要的事实，并将这些事实置于更广泛的背景下，这样，即使细节发生了改变，你所掌握的知识仍然是有用的。本书不仅告诉你网络的今天，而且使你能够为网络的明天做准备。本书增加了新的特征，即通过页边插文来描述书中所提概念的原因，为什么一种观点失败了？为什么一种观点成功了？

本书做了哪些修改呢？有些技术已经从人们的视线中读出，本版中对这些内容做了删减，例如FDDI和ATM局域网。有些技术已经逐步成熟起来并以新的形式出现了，远程过程调用不再是一种基于局域网的低层调用机制，而是因特网范围Web服务的基础。我们欢迎吉比特以太网，对有关无线技术的章节进行了更新和扩展，在路由实现上也进行了改进。关于TCP的内容也做了更新，讨论了新的确认机制和对高速的扩展。

本书在安全方面进行了更多考虑，该章内容进行了全面修订，强调了网络安全的方法，讨论了安全威胁及防范方法。最后一章帮助你把所学知识综合起来，在应用层使用案例研究（VoIP、多媒体和P2P）来讲述如何把前面各章的所有概念组合起来，构造支持这些应用的系统。

网络的发展不会减慢，不久我们就会谈论基于IP的电视效果、因特网和传感器网络的冲突，以及其他许多很新和很令人兴奋的思想。请放松下来，今天你读了这本书，你就会有明天所需要的视野。

David D. Clark

麻省理工学院

第1版序言

“长龙式编码” (spaghetti code) 一词普遍被认为是一种轻蔑的称谓。所有优秀的计算机科学家都推崇模块化，这是因为它能带来许多好处，其中包括在解决问题的同时不必了解问题的所有环节。因此，在一本书中以及在编写代码时，模块化扮演着重要的角色。如果一本书的材料以模块化的方式有效地组织起来，那么读者就会很乐意从头读到尾。

在网络协议领域，采用国际标准的形式，即ISO的7层网络协议参考模型，给出“真正的”模块化，这或许是独一无二的。这种模型反映模块化的一种层次方法，无论设计是符合还是偏离这种模型，这种模型都广泛地用做协议组织讨论的起点。

看上去围绕这种分层的模型来组织一本网络书籍是显而易见的。但事实上这样做是有风险的，因为在组织网络的核心概念时，OSI的模型其实是不成功的。一些基本的需求，如可靠性、流量控制或安全性问题并不能完全纳入OSI的分层中。因此，导致对理解参考模型的巨大混乱，有时甚至产生怀疑。实际上，如果严格地按照层次模型来组织一本书，那么它就具有某些长龙式编码的属性。

本书作者遵循传统的分层模型，但并没有拘泥于完全利用这种模型去帮助我们理解网络中的重大问题。相反，作者采用独立于层次模型的方法来组织基本概念的讨论。因此，在阅读本书后，读者将会理解流量控制、拥塞控制、可靠性增强、数据的表示以及同步等问题。同时，读者将会分别了解这些问题是如何涉及传统分层模型的这一层或那一层的。

这是一本适时的书。本书考察当今使用的主要协议，尤其是有关因特网的协议。作者有丰富的从事因特网工作的经验，因此能够不仅从理论上而且从实践上揭示协议设计的问题。本书介绍许多最新的协议，读者可以从中获得最新的观点。更重要的是，基本问题的讨论是基于问题的本质属性，不受分层参考模型或当前的协议细节的限制。在这一点上，本书既体现了适时性，又不受时间的限制。本书的独特之处在于它将有关的实际问题、现实的事例和基本概念的解释有机地结合在一起。

David D. Clark

麻省理工学院

前言

当本书第1版在1996年出版时，在因特网上购物还是很新奇的事情，如果一个公司用它的域名做广告被认为是超前的。而当今社会，因特网商务已进入日常生活中，“.com”股票已经历了一个完整的兴衰循环。从光交换机到无线网络，一大批新兴技术正在成为主流。似乎关于因特网唯一可以预见的东西就是它会不断地变化。

尽管有这么大的变化，我们在第1版中提出的问题对于今天来说仍然是有效的：使因特网得以运行的基本概念和技术是什么？回答是，TCP/IP体系结构的大部分功能对于今天仍然适用，这一点正像30年前它的创立者预见的那样。这并不是说因特网的体系结构没什么新鲜的，而是正好相反。一个体系结构30年来不仅幸存下来，而且促进因特网这样快速增长和变化，了解其中的设计原理正是我们的出发点。正像前几版一样，第4版把因特网的体系结构“何以如此”作为它的基础。

读者对象

我们的目的是把本书做为内容广泛的网络课程教材，供研究生或高年级本科生使用。我们相信，这本书的核心概念不但对正在进行再培训以便完成网络相关任务的专业人员有吸引力，而且可以帮助网络从业人员理解每天都要接触的网络协议背后的“为什么”，并且明白网络的整体概念。

根据我们的经验，第一次学习网络的学生和专业人员通常会把网络协议理解成一种从高层传到低层的命令，认为多学一些术语缩写词就可以了。事实上，协议是从工程设计原理的应用中开发出来的复杂系统的构件。不仅如此，协议总是根据现实世界的经验不断地被精练、扩展和替换。因此，本书的目标并不单纯地介绍当今使用的协议，更侧重于解释合理的网络设计的基本原理。我们认为把握这些基本原理是应对当今网络领域中瞬息万变的最好办法。

第4版中的变化

尽管我们关注网络的基本原则，但我们还是要使用正在运行的因特网中的例子来阐述这些原则。因此，我们补充了相当多的新材料，跟踪近期网络技术的重要进展。我们同时对已有材料做了删除、重新组织和改变，以反映过去10年发生的变化。

也许自从编写第1版以来，我们所察觉到的最重要的变化就是现在几乎每一位读者都对诸如万维网和电子邮件这样的网络化应用有一定的了解。因此，我们从第1章开始就加大了对应用的侧重。我们把应用当做学习网络技术的动机，并得出一组需求，有用的网络只有满足这些需求才能在全球范围内支持当前的和未来的各种应用。然而，我们保留了先前版本解决问题的方法，即从主机的互联问题开始，逐层向上讨论，最后对应用层的问题进行详细的考察。我们认为从各种应用及其需求起步对于在本书所覆盖的各主题间建立联系是很重要的。同时，我们感到对于诸如应用层协议和传输层协议这样的高层协议的问题，只有在讲明白主机连接和分组交换这样的基本问题之后才能很好地理解。

就像我们在第2版和第3版中所做的一样，我们附加或增大了对重要的新主题的覆盖面，并使其他主题紧跟潮流。这一版中主要的新主题或有实质性改动的主题包括：

- 全面修订并更新了网络安全的内容，致力于构建安全系统，而不仅仅是密码算法；
- 扩展并更新了XML（可扩展标记语言）方面的内容；
- 更新了覆盖网络的内容，包括对等网和内容分发网络；
- 增加了Web服务的内容，包括SOAP和REST（代表性状态传输）体系结构；
- 更新了无线技术的内容，包括802.11（Wi-Fi）和802.16（WiMAX）标准，以及以3G标准为代表的蜂窝无线技术；
- 扩展了域间路由的内容；
- 扩展了用于多媒体应用的协议和服务质量的内容，例如，基于IP的语音（VoIP）和视频流；
- 更新了拥塞控制机制（特别是在高带宽网络环境下）方面的内容。

此外，在这一版中增加了新的特征：“它们现在在哪里”。这个简单的讨论集中于协议在现实世界中的成功和失败。有时它们描述了某个大多数人已经注销了但却比较成功的协议，也有可能是跟踪某个运行了很长时间但却失败了协议的命运。我们的目的是通过这些材料告诉大家，在竞争激烈的网络世界里，这些技术的实际应用。

方法

对于像计算机网络这样动态的和不断变化的领域来说，一本教材能提供的最重要的东西是洞察能力，以便能够分清什么是重要的，什么是不重要的，什么是长久的，什么是表面上的。根据我们致力于网络新技术研究的20多年的经验，和对本科生与研究生讲授网络最新趋势的课堂反馈，以及把先进的网络产品投放市场的经验，我们提炼出了自己的观点，称为系统方法，它是本书的精髓。这种系统方法有以下含义：

- 与其接受现成的网络产品作为准则，不如从最基本的原理开始，让你了解当今网络技术的发展过程。这能让我们解释网络为什么像现在这样设计。根据我们的经验，一旦理解了基本概念，对于遇到的任何新协议，消化和吸收都将变得相对容易。
- 虽然材料是围绕传统的网络层次松散地组织起来，从底层开始沿协议栈向上展开的，但是我们并不采用严格的分层方法。许多主题涉及多层，例如拥塞控制 and 安全性就是这样，所以我们在传统的分层模型之外讨论它们。简言之，我们相信可以很好地使用分层，但是不必受它的限制。采用端到端的观点常常是更有用的。
- 与其抽象地解释协议如何工作，不如使用当今最重要的协议具体地说明网络是如何工作的，许多协议都是源自TCP/IP因特网的。这就允许我们在讨论中借鉴实际经验。
- 虽然在最底层可以用从计算机销售商那里购买的硬件建立网络，并且通信服务可从电话公司租用，但是只有软件才可以使网络提供新的服务，并且迅速地适应新的需求。这就是我们为什么强调网络软件是如何实现的理由，而不是只停留在描述所涉及的抽象算法上。我们还会给出一些来自运行的协议栈的代码段，以说明如何实现某些协议和算法。
- 网络是由许多组件构成的，虽然在解决一个具体问题时忽略一些不重要的元素是必要的，但也必须要理解如何将组件组合在一起，以构成一个具有一定功能的网络。所以我们花大量的时间解释网络总体的端到端行为，而不只是个别的组成部分，以便能够理解一个完整的网络是如何运行的，包括从应用到硬件的所有方面。
- 这种系统方法包含要进行实验性的性能研究，然后使用从定量分析各种设计选项和指导优化实现这两个方面收集的数据。这种强调经验分析的方法贯穿全书。
- 网络很像其他计算机系统，例如操作系统、处理器体系结构、分布式和并行系统等等。它们都很大并很复杂。为了处理这种复杂性，系统设计者常常提出一组设计原则。我们

重点介绍这些贯穿全书的设计原则，并用计算机网络中的例子加以说明。

教学法和特点

第4版保留了以前版本中的一些特点，并增加了一个特点，我们鼓励读者利用这些特点：

- 问题描述。在每一章的开头，我们描述在网络设计中必须解决的一组问题，由它引出本章探讨的一些主题。
- 相关主题。在本书中，相关主题详细说明要探讨的题目或介绍相关的高级主题。在许多情况下，这些主题与实际中的网络技术有关。
- “它们现在在哪里？”。这些新的内容跟踪现实情况下协议的成功与失败。
- 突出的段落。这些段落归纳了在讨论中得出的重要结论，例如广泛使用的系统设计原则。
- 实际的协议。虽然本书着重核心概念而不是现成的协议说明，但实际的协议常用来说明大部分重要的思想。因此本书可以用做许多协议的参考源。为了帮助你找到这些协议的描述，每节标题中用括号括起来的是协议名称，指明在那一节定义的协议。例如，5.2节描述可靠的端到端协议的原则，它提供对TCP的详细描述，TCP是这个协议的典型例子。
- 开放问题。每章的叙述以对一个重要问题的讨论结尾，这个问题是研究领域、业界或整个社会正在探讨的课题。我们发现这些讨论能使读者更关心所讨论的网络课题并对其产生浓厚的兴趣。
- 补充读物。在每一章结尾列有精选的参考书目。这些书目一般包含刚讨论的有关题目的创新性论文。我们竭力推荐高级读者(如研究生)学习这个书目中的文章，以便补充各章所讲的材料。

本书结构和课程使用

本书按以下方式组织：

- 第1章介绍全书使用的核心概念。涉及各种应用，讨论了网络体系结构，介绍了协议实现，并定义了通常促使人们进行网络设计的定量性能标准。
- 第2章综述广泛的低层网络技术，从以太网到令牌环再到无线网络，还描述了所有链路协议必须解决的许多问题，包括编码、组帧和差错检测。
- 第3章讲述交换网（数据报网与虚电路网）的基本模型，并详细地介绍两种流行的交换技术——交换式以太网和ATM。同时也讨论基于硬件的交换机设计问题。
- 第4章讲述网络互联，并且描述网际协议(IP)的基本原理。这一章讨论的一个中心问题是像因特网这样规模的网络如何对分组进行路由选择。单播、多播和域间路由问题也在讨论之列。
- 第5章讲述传输层，详细地描述因特网用于建立客户机/服务器应用的传输控制协议(TCP)和远程过程调用(RPC)，同时介绍了支持多媒体应用的实时传输协议(RTP)。
- 第6章讨论拥塞控制和资源分配。这一章的问题贯穿网络层（第3、4章）和传输层（第5章）。特别注意，这一章会描述拥塞控制如何在TCP上工作，并且介绍在IP中为提供服务质量所使用的机制。
- 第7章考虑通过网络发送的数据，这涉及表示格式和数据压缩方面的问题，也包括XML。压缩的讨论包括解释MPEG视频压缩和MP3音频压缩是如何工作的。
- 第8章讨论网络安全，首先介绍加密工具、密钥分配问题，讨论几种使用公钥和私钥的鉴别技术。这一章的主要内容是使用诸如极好隐私(PGP)、安全外壳(SSH)和IP安全

体系结构 (IPSEC) 建立安全系统。这一章也讨论防火墙。

- 第9章描述网络应用的典型实例和它们使用的协议, 包括像电子邮件和万维网这样的传统应用, 和像IP电话和视频流这样的多媒体应用, 以及像对等文件共享和内容分发网络这样的覆盖网络。这一章也介绍用于开发新应用协议的Web服务体系结构。

对本科生的课程, 可能需要追加课时帮助学生理解第1章的导论材料, 而放弃第6~8章的高级主题。然后, 在第9章转到网络应用的通常主题上。相反, 研究生的指导教师可用一、两次课讲完第1章的内容, 让学生自己更仔细地研究材料, 以腾出更多的时间深入讲授最后四章的内容。研究生和本科生都要完成中间四章 (第2~5章) 的核心材料, 但本科生可以有选择地跳过那些更深入的章节 (如2.2节, 4.4节和4.5节)。

对于自学本书的读者, 我们相信所选的主题涵盖了计算机网络的核心内容, 因此建议从前到后顺序阅读。另外我们提供了详细的参考文献目录, 帮助读者进一步确定感兴趣领域的补充材料。我们还提供了选题解答。

本书采取独特的方法来讨论拥塞控制, 即把有关拥塞控制和资源分配的所有专题集中到第6章。这样做是因为拥塞控制问题不能在任何一层单独解决, 并且我们希望读者同时能够考虑各种设计选择 (这和我们观点是一致的, 即严格的层次性常常模糊了重要的设计权衡)。然而, 对拥塞控制的更传统的处理方式是可能的, 即在学习第3章时参考6.2节的内容, 以及在学习第5章时参考6.4节的内容。

习题

在每一个新版本中都对习题进行了重大改进, 在第2版中基于课堂练习增加了习题的数量, 改进了习题的质量, 在第3版中对习题做了两方面的重大修改, 包括:

- 对那些我们认为很有挑战性或需要本书以外知识 (例如, 概率知识) 的习题, 我们加上★标记以表明它们具有更高层次的难度。
- 在每一章我们都附加了一些范例习题并在书后“选题解答”中给出答案。这些习题用✓标记, 目的是为解决书中的其他习题提供一些帮助。

在新版本中增加了新的习题以反映更新的内容, 现有的习题分为如下几类:

- 分析性的习题, 要求学生做简单的代数计算, 展示他们对基本关系的理解。
- 设计问题, 要求学生提出和评价各种情况下的协议。
- 动手的习题, 要求学生写少量代码去测试一个想法或使用现成的网络工具进行实验。
- 文献研究问题, 能够让学生更深入地了解某个特别的问题。

此外, 正如下面所详细描述的那样, 在线提供基于Socket的编程任务以及模拟实验。

补充材料和在线资源

为便于教学, 我们准备了教学手册, 其中包含一些习题的解答。可以向出版商索取该手册。

其他辅助材料, 包括课件、教材图片、基于Socket的编程练习、测验习题以及编程习题, 可以在Morgan Kaufmann出版公司的网站<http://www.elsevierdirect.com/index.jsp>上找到。建议读者经常访问关于本书的网页, 我们正在逐步增加辅助材料并创建与网络技术相关站点的链接。

最后, 与第3版一样, 本书配有一套实验。这些实验是由马萨诸塞大学达特茅斯分校的Emad Aboelela教授开发的, 通过模拟实验来探讨本书中协议的行为、可扩展性和性能。

致谢

如果没有许多朋友的帮助,本书是不可能问世的。我们非常感谢所有为改进本书做出贡献的人。然而,在致谢之前要提到的是,我们已经尽力改正审阅人指出的错误以及尽量准确地描述同事们给我们解释的协议和机制。如果还有什么错误,那就是我们的责任。如果你发现任何错误,请发电子邮件给我们的出版商Morgan Kaufmann,地址是netbugsPD4e@elsevier.com,我们将在本书再次印刷时进行改正。

首先,我们衷心感谢审阅过全部或部分手稿的人。除了审阅过以前版本的人,我们感谢David Maltz, Bobby Bhattacharjee和Sarvesh Kulkarni对全书的审阅。还要感谢Ric Pruss和Mike Takefman对各章节的审阅。我们也要感谢所有提供反馈意见和信息来帮助我们决定如何写该版的人,他们是Tim Batten, Julio Pontes和Kevin Mills。

普林斯顿网络系统组的几位成员贡献了他们的想法,例如,提供了修订、数据以及本书早期版本的代码段。我们特别感谢Andy Bavier, Tammo Spalink, Mike Wawrzoniak, Zuki Gottlieb, George Tzanetakis和Chad Mynhier。KyoungSoo Park在系统解答、教师手册和课件方面提供了有益的帮助。正如以前一样,我们感谢国防部高级研究计划署、国家科学基金、Intel公司和Cisco公司在过去几年对我们网络研究课题的支持。同时也感谢Cisco公司为我们提供了进行本书工作的机会。

没有Mark Abbott实质性的帮助,我们就不可能完成本书的修订,他为本书提供了大量新材料,我们表示衷心地感谢。

最后,我们衷心地感谢丛书编辑David Clark以及Morgan Kaufmann出版公司在本书编写期间帮助过我们的所有人。还要特别感谢我们原来的责任编辑Jennifer Young,本版的编辑Rick Adams,我们的制作编辑Rachel Roumeliotis和先前版本的助理编辑Karyn Johnson。在本书的写作过程中,与MKP出版公司全体人员合作的过程很令人愉快。

目 录

出版者的话	2.4.1 二维奇偶校验	54
译者序	2.4.2 因特网校验和算法	55
序言	2.4.3 循环冗余校验	56
第1版序言	2.5 可靠传输	59
前言	2.5.1 停止和等待	59
	2.5.2 滑动窗口	61
第1章 基础	2.5.3 并发逻辑信道	67
1.1 应用	2.6 以太网 (802.3)	67
1.2 需求	2.6.1 物理特性	68
1.2.1 连通性	2.6.2 访问协议	69
1.2.2 成本—效益合算的资源共享	2.6.3 以太网的经验	72
1.2.3 支持公共服务	2.7 环网 (802.5, FDDI, RPR)	72
1.3 网络体系结构	2.7.1 令牌环介质访问控制	74
1.3.1 分层和协议	2.7.2 令牌环维护	75
1.3.2 OSI体系结构	2.7.3 FDDI	76
1.3.3 因特网体系结构	2.7.4 弹性分组环 (802.17)	77
1.4 实现网络软件	2.8 无线网络	78
1.4.1 应用编程接口 (套接字)	2.8.1 蓝牙 (802.15.1)	80
1.4.2 应用实例	2.8.2 Wi-Fi (802.11)	81
1.4.3 协议实现的问题	2.8.3 WiMAX (802.16)	84
1.5 性能	2.8.4 蜂窝电话技术	85
1.5.1 带宽与时延	2.9 小结	86
1.5.2 延迟和带宽的乘积	第3章 分组交换	96
1.5.3 高速网络	3.1 交换和转发	96
1.5.4 应用程序性能需求	3.1.1 数据报	98
1.6 小结	3.1.2 虚电路交换	99
第2章 直接连接的网络	3.1.3 源路由选择	104
2.1 网络构件	3.2 网桥和局域网交换机	106
2.1.1 节点	3.2.1 学习型网桥	106
2.1.2 链路	3.2.2 生成树算法	108
2.2 编码 (NRZ、NRZI、Manchester、4B/5B)	3.2.3 广播和多播	111
2.3 组帧	3.2.4 网桥的局限性	111
2.3.1 面向字节的协议 (PPP)	3.3 信元交换 (ATM)	113
2.3.2 面向比特的协议 (HDLC)	3.3.1 信元	113
2.3.3 基于时钟的组帧 (SONET)	3.3.2 分段和重组	116
2.4 差错检测	3.3.3 虚路径	119
	3.3.4 ATM的物理层	120

3.4 实现和性能	121	5.2.5 触发传输	235
3.4.1 端口	123	5.2.6 自适应重传	237
3.4.2 网状结构	125	5.2.7 记录边界	239
3.5 小结	127	5.2.8 TCP扩展	239
第4章 网络互联	135	5.2.9 其他设计选择	240
4.1 简单的网络互联 (IP)	135	5.3 远程过程调用	241
4.1.1 什么是互联网	136	5.3.1 RPC基础	242
4.1.2 服务模型	137	5.3.2 RPC的实现 (SunRPC和DCE)	246
4.1.3 全局地址	144	5.4 实时传输协议 (RTP)	250
4.1.4 IP中的数据报转发	146	5.4.1 需求	251
4.1.5 地址转换 (ARP)	148	5.4.2 RTP细节	252
4.1.6 主机配置 (DHCP)	151	5.4.3 控制协议	254
4.1.7 差错报告 (ICMP)	153	5.5 性能	256
4.1.8 虚拟网络和隧道	153	5.6 小结	258
4.2 路由	155	第6章 拥塞控制和资源分配	266
4.2.1 用图表示网络	156	6.1 资源分配中的问题	266
4.2.2 距离向量 (RIP)	157	6.1.1 网络模型	267
4.2.3 链路状态 (OSPF)	162	6.1.2 分类法	269
4.2.4 度量标准	167	6.1.3 评价标准	270
4.2.5 移动主机的路由	169	6.2 排队规则	272
4.2.6 路由器实现	172	6.2.1 FIFO	272
4.3 全球因特网	174	6.2.2 公平排队	274
4.3.1 划分子网	175	6.3 TCP拥塞控制	276
4.3.2 无类路由 (CIDR)	178	6.3.1 加性增/乘性减	277
4.3.3 域间路由 (BGP)	179	6.3.2 慢启动	278
4.3.4 路由区	185	6.3.3 快速重传和快速恢复	282
4.3.5 IP版本6 (IPv6)	187	6.4 拥塞避免机制	283
4.4 多播	194	6.4.1 DEChit	283
4.4.1 多播地址	194	6.4.2 随机早期检测 (RED)	284
4.4.2 多播路由 (DVMRP, PIM, MSDP)	195	6.4.3 基于源的拥塞避免	287
4.5 多协议标记交换	202	6.5 服务质量	291
4.5.1 基于目的地的转发	203	6.5.1 应用需求	291
4.5.2 显式路由	207	6.5.2 综合服务 (RSVP)	295
4.5.3 虚拟专用网和隧道	208	6.5.3 区分服务 (EF和AF)	301
4.6 小结	210	6.5.4 基于等式的拥塞控制	305
第5章 端到端协议	223	6.6 小结	306
5.1 简单的多路分解协议 (UDP)	224	第7章 端到端的数据	315
5.2 可靠的字节流 (TCP)	225	7.1 表示格式化	316
5.2.1 端到端的问题	226	7.1.1 分类方法	317
5.2.2 报文段格式	227	7.1.2 例子 (XDR、ASN.1、NDR)	319
5.2.3 连接的建立与终止	229	7.1.3 标记语言 (XML)	322
5.2.4 滑动窗口再讨论	232	7.2 数据压缩	324

7.2.1 无损压缩算法	325
7.2.2 图像压缩 (JPEG)	327
7.2.3 视频压缩 (MPEG)	329
7.2.4 在网上传输MPEG	332
7.2.5 音频压缩 (MP3)	335
7.3 小结	336
第8章 网络安全	341
8.1 密码工具	342
8.1.1 密码原理	342
8.1.2 对称密钥密码	344
8.1.3 公钥密码	345
8.1.4 认证码	346
8.2 密钥预分发	348
8.2.1 公钥的预分配	349
8.2.2 对称密钥的预分发	351
8.3 认证协议	352
8.3.1 原始性和实效性技术	352
8.3.2 公钥认证协议	353
8.3.3 对称密钥认证协议	354
8.3.4 Diffie-Hellman密钥协商	356
8.4 安全系统	358
8.4.1 极好隐私 (PGP)	358
8.4.2 安全外壳 (SSH)	359
8.4.3 传输层安全 (TLS, SSL, HTTPS)	361
8.4.4 IP安全 (IPsec)	363

8.4.5 无线安全 (802.11i)	365
8.5 防火墙	366
8.6 小结	369
第9章 应用	374
9.1 传统应用	374
9.1.1 电子邮件 (SMTP、MIME、IMAP)	375
9.1.2 万维网 (HTTP)	380
9.1.3 域名服务 (DNS)	384
9.1.4 网络管理 (SNMP)	390
9.2 Web服务	391
9.2.1 定制应用协议 (WSDL, SOAP)	392
9.2.2 一个通用的应用协议 (REST)	396
9.3 多媒体应用	397
9.3.1 会话控制和呼叫控制 (SDP, SIP, H.323)	397
9.3.2 多媒体应用的资源分配	403
9.4 覆盖网络	406
9.4.1 路由覆盖	408
9.4.2 对等网 (Gnutella, BitTorrent)	412
9.4.3 内容分发网络	419
9.5 小结	423
术语	428
习题选答	443
参考文献	452

第1章 基础

我必须创造一个体系，否则我即将沦为别人体系的附庸；我不要推理和比较，我的工作
是创造。

——威廉·布莱克

问题：建造一个网络

假设我们要建造一个计算机网络，它有发展到全球性规模的潜力，并且能够支持各种各样的应用，如远程会议、视频点播、电子商务、分布式计算和数字化图书馆等。那么要采用什么样的技术作为基础构件，以及使用何种软件体系结构才能把这些构件集成为一个有效的通信服务？本书最主要的目标就是回答这个问题，描述可用的构件，以及说明如何自下而上用它们来建造一个网络。

在我们了解如何设计计算机网络之前，首先应在什么是计算机网络这一问题上达成共识。曾经有一段时期，网络（network）一词是指用于将单一功能终端连接到大型计算机所用线路的集合。一些人认为，“网络”这个词是指语音电话网络；而另一些人认为，该词仅指用于传播视频信号的电缆网络。这些网络的主要共同点是专门处理某种特定类型的数据（按键、音频或视频），并且通常连接到特殊用途的设备（终端、手持接收器和电视机）。

计算机网络与其他类型网络有什么区别？通用性也许是计算机网络的最主要特征。计算机网络主要由通用可编程硬件来构建，并且不会为诸如打电话或传送电视信号那样的特定应用做任何优化。相反，计算机网络能够传输多种不同类型的数据，并且支持广泛的不断增长的应用。本章考察计算机网络的一些典型应用，然后讨论网络设计者为支持这些应用必须了解的内容。

一旦我们弄清楚这些需求，接下来该怎么做呢？幸运的是，我们并不是在建造第一个网络。其他一些人，其中最著名的是因特网的研究人员，他们已经先于我们完成了这项任务。在因特网构建中得到的丰富经验可以用来指导我们的设计。这些经验体现在网络体系结构（network architecture）中，网络体系结构指明可用的软硬件构件，并且说明如何将它们组织起来构成一个完整的网络系统。

为了引导我们如何实现如何建造一个网络，本章将阐述四项内容。首先，揭示不同的应用和不同的群体（如网络用户和网络操作者）对网络的需求。第二，引入网络体系结构的概念，它是本书后续部分的基础。第三，介绍计算机网络实现的几个关键因素。最后，介绍用来衡量计算机网络性能的关键度量标准。

1.1 应用

许多人是通过因特网的各种应用（如万维网、电子邮件、音频和视频流、聊天室和共享音乐文件）来了解因特网的。举例来说，万维网提供了一个直观而且简单的界面。用户浏览有很多文本和图形对象的网页时，点击他们想进一步了解的對象，就会弹出相应的新网页。

大多数人也明白在这个应用的背后，是网页上的每个可选对象都绑定着一个指向下一个页面的标识符。这个标识符被称为统一资源定位器（uniform resource locator, URL），它惟一标识你通过浏览器所能浏览到的每一个网页。举个例子：

`http://www.cs.princeton.edu/~llp/index.html`

这是提供本书的一名作者信息页面的URL：字符串http表明要下载页面应使用超文本传输协议（Hyper Text Transfer Protocol, HTTP），www.mkp.com是提供网页的计算机的名字，而~llp/index.html pd3e惟一标识Larry在该网站的主页。

但是，点击这样一个URL后，大多数用户并不清楚在因特网上可能需要交换多达17条消息才能得到网页，并且网页本身要小到可以存放在一条消息中。这些消息中有6条用来把服务器名（cs.princeton.edu）翻译成它所对应的因特网地址（128.112.136.35），3条消息用来建立从浏览器到服务器之间的传输控制协议（Transmission Control Protocol, TCP）的连接，4条消息用来让浏览器发送HTTP的“get”请求，并让服务器回送被请求的页面（以及双方对收到消息的确认），还有4条消息用来关闭TCP连接。当然，这不包含因特网节点一天交换的数以百万计的消息，这些消息只是让节点相互知道自己的存在，并准备提供网页服务，把机器名翻译成网络地址，并将各种消息向它们的最终目的地转发。

尽管现在还不像网上冲浪那样普及，但是音频和视频的流式播放作为因特网的一项应用正在兴起。虽然可以把整个视频文件从远程主机下载到本地播放，就像下载和显示一个网页的过程一样，但你必须等到视频文件的传送进行到最后一秒才能开始观看。以流的方式播放视频意味着发送方和接收方分别充当视频流的信源（source）和信宿（sink）。也就是说，信源创建一个视频流（也许是使用视频采集卡），以消息的形式在因特网上发送，而信宿在收到消息后将其显示出来。

视频应用有各种不同的类型。视频点播是对视频应用一个例子，它先从硬盘上读取一个已有的电影，然后把它传送到网络上去。另一种应用是视频会议，因为它有很严格的时间约束，所以更具挑战性。就像使用电话一样，参与者之间的交互必须是及时的。当一端的用户做一个动作，这个动作必须尽快被显示在另一端。太长的延迟会造成系统无法使用。相比而言，如果从用户打开视频到第一幅图像被显示出来用了几秒钟，那么这个服务仍可被接受。此外，交互视频意味着有双向流动的视频数据，而视频点播应用大多只向一个方向发送视频数据。

20世纪90年代中、早期UNIX应用程序vic是一个流行的视频会议工具。图1-1显示一次vic会话的控制面板。实际上，vic是Lawrence Berkeley 实验室和加州大学Berkeley分校共同开发的一套视频会议工具之一。其他例子包括白板应用程序（wb）（允许用户互相发送草图和幻灯片），一种可视化音频工具vat，和用来创建和发布视频会议的会话目录（sdr）。所有这些在UNIX上运行的工具（它们的名字采用英文小写）都可以在因特网上免费获得。同时还可以获得用于其他操作系统的类似工具。值得注意的是，在本书出版的2006年，尽管在因特网上的视频尚不成熟，但基于IP支持视频的工具已经存在十余年了。

尽管这只是两个例子，但从网上下载网页和参加视频会议已经能够表明建立在因特网上的应用的多样性，并暗示因特网设计的复杂性。本书余下的部分将从头开始，一次讲述一个问题，解释如何建立一个通用性较强的网络。第9章作为全书的结束部分将重新讨论这两种特定的应用以及其他几种在当今因特网上流行的应用。