

IP

技术基础

编址和路由

- ▶ IP Fundamentals: What Everyone Needs to
- ▶ Know About Addressing & Routing
- ▶ (美) Thomas A. Maufer 著
- ▶ 赵军锁 李志 等译



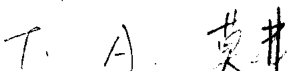
机械工业出版社
China Machine Press

Prentice Hall

IP网络技术丛书

IP技术基础

编址和路由


(美) Thomas A. Maufer 著

赵军锁 李 志 等译



机械工业出版社
China Machine Press

IP协议已经成为当今占统治地位的连网协议，其应用仍在迅速增长。本书全面介绍了IP协议的基础知识，并在兼顾全面、专业、易于理解的基础上重点介绍IP编址和IP路由。全书共分三部分：第一部分介绍IP编址，并透彻地讲解了操作IP地址和子网掩码；第二部分主要讨论IP包如何在最为常见的LAN和WAN子网技术上进行传送；第三部分介绍基于标准的IP路由协议。

本书叙述清晰，实用性强，是一本基础性的、同时也是非常专业的书籍。本书对IP的讲解透彻、全面，是从事计算机网络开发、应用、研究等工作的人员及大专院校师生不可多得的实用的参考书。

Thomas A. Maufer: IP Fundamentals: What Everyone Needs to Know About Addressing & Routing.

Authorized translation from the English language edition published by Prentice Hall PTR.

Copyright ©1999 by Prentice-Hall.

All rights reserved.

Chinese simplified language edition published by China Machine Press.

Copyright © 2000 by China Machine Press.

本书中文简体字版由美国Prentice Hall公司授权机械工业出版社独家出版，未经出版者书面许可，不得以任何方式复制或抄袭本书的任何部分。

版权所有，翻印必究。

本书版权登记号：图字：01-1999-2349

图书在版编目(CIP)数据

IP技术基础：编址和路由/(美)莫弗(Maufer, T. A.)著；赵军锁等译。—北京：机械工业出版社，2000.1

(IP网络技术丛书)

书名原文：IP Fundamentals ~~What Everyone Needs to Know About Addressing & Routing~~
ISBN 7-111-07434-3

I. I… II. ①莫… ②赵… III 计算机网络—传输控制协议，IP IV. TP393

中国版本图书馆CIP数据核字(1999)第64288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：刘立卿

北京牛山世兴印刷厂印刷·新华书店北京发行所发行

2000年1月第1版第1次印刷

787mm×1092mm 1/16·18.5印张

印数：0 001-6 000册

定价：38.00元

凡购本书，如有倒页、脱页、缺页、由本社发行部调换

“计算机实用软件丛书”编委会

高级顾问 张效祥 胡启恒

主 任 牛田佳

副 主 任 李树岭 罗晓沛

特约编委 谭浩强 陈树楷

编 委 (按姓氏笔画排序)

毛 波 方 裕 史美林 孙中臣

孙家骥 刘炳文 刘德贵 吴文虎

张国锋 周山芙 周堤基 钟玉琢

柳克俊 侯炳辉 赵桂珍 聂元铭

徐国平 徐修存 寇国华 戴国忠

丛书前言

随着计算机、通信和信息技术的迅速发展与广泛应用,人类正在进入信息化社会。计算机技术的应用与推广,将直接推动社会信息化的发展;而计算机技术的应用与推广,实质上取决于计算机软件的应用和推广,可以说,没有软件,就没有计算机的应用;学习、使用计算机,从根本上讲就是学习和掌握软件的使用。

为了适应当前计算机技术发展的需要,满足读者学习、使用计算机软件的需求,人民邮电出版社约请有关专家编写出版了这套“计算机实用软件丛书”。

这套丛书的特点是:普及兼顾提高,应用兼顾开发,各书独立成册形成系列,并注重其相关性,使丛书成为广大计算机应用和开发人员学习使用计算机的必备用书。

这套丛书的内容包括:程序设计语言、操作系统技术、数据库技术、软件开发技术及工具、网络技术、多媒体技术等。

在计算机技术飞速发展的今天,软件产品更新快,经常有新产品或新版本问世,因此我们不但介绍当前流行和优秀的软件,而且力求尽快把国内外最新的软件产品也介绍给读者。

我们将全心全意为读者服务,也热切期待广大读者对丛书提出宝贵意见,以进一步提高丛书的质量。让我们共同努力,为提高我国的计算机开发、应用水平做出贡献。

“计算机实用软件丛书”编委会

译者序

IP协议(Internet Protocol, 网际协议)已经逐渐成为一种通用的组网协议, 如在Internet和Intranet中的应用。不仅如此, IP协议的使用仍在迅速增长。也许在某一天, 人们会打出这样一个口号: “网络就是IP, IP就是网络”。这句话也许并不确切, 也不一定会有这样一个事实, 但可以表明我的意思: 网络无处不在, IP无处不在, 有网络的地方就会有IP。

但是人们对于IP的知识和认识呢? 大多数人对IP究竟能做什么及其各部分之间如何相互作用却不甚明了。关于Internet的书籍很多, 但大多数谈论的是关于聊天室、如何创建时髦的Web站点以及其他常用的应用程序。本书所讨论的是Internet的基础设施内幕——IP协议, 而不是用户所看到的表面部分。在本书中你将学到IP究竟如何工作的方方面面。尽管本书的重点是讨论基础设施的本质, 但并非是一本有关通信原理的书; 本书注重实用性、面向解决方案, 尤其适合于那些需要快速掌握IP如何工作的人们。

当我们拿到这本书的时候, 并没有马上就动手翻译, 而是首先全面浏览了一下本书的内容, 带着欣赏的眼光“把玩”了一番。为什么呢? 因为它的独到之处——对某个主题的详尽、全面、透彻的讲解, 当然还有作者的文采和严谨给我们以深深的吸引力。现在, 我们很荣幸能够有机会承担本书的翻译工作, 并抱着认真的态度将这本书的中文版奉献给您, 希望您能够从本书中有所收获, 这是作者的初衷, 也是我们良好的愿望!

本书由赵军锁、李志组织翻译, 参加本书翻译、录排、校对工作的人员有: 赵军锁、李志、龚波、田丽楹、李林、张巧莉、陈曙晖、邓波、邓涛、李卓林、聂宛析、田敏、金光、小光、龚露娜、马军、马丽、田军、田洗县、王小将等。本书的出版是集体劳动的结晶, 在此特别感谢前导工作室的全体工作人员。

由于时间仓促, 且译者经验和水平有限, 译文难免有不妥之处, 恳请读者批评指正!

译者

1999年8月

序 言

1998年10月16日，Jon Postel离开人间。他对Internet或者开放标准的目标做出了无可比拟的贡献。他的思想为Internet工程任务组（IETF, Internet Engineering Task Force）和Internet体系结构研究会（IAB, Internet Architecture Board）奠定了基础。除此之外，在IETF标准的制定上也无人可以和他媲美。没有人写出如此之多的请求注释（RFC, Request for Comment），由他写成的许多RFC对今天的Internet仍然是相当重要的。事实上，将RFC作为统一文档系列的整个思想都是他提出来的。他写出的RFC有网际协议（IP）、Internet控制信息协议（ICMP, Internet Control Message Protocol）、传输控制协议（TCP, Transmission Control Protocol）、用户数据报协议（UDP, User Datagram Protocol）、简单邮件传输协议（SMTP, Simple Mail Transport Protocol）、Telnet、文件传输协议（FTP, File Transfer Protocol）及其他协议。

除了他实际写的协议之外，他作为一个IAB成员、Internet编号管理局（IANA, Internet Assigned Numbers Authority）和RFC编辑，做了大量的幕后工作，对许多并没有署有其名的标准的制订做出了贡献。大多数使用Internet的人们都对Jon是谁和他有多么的重要没有了解。据我所知，他对得到人们的承认一点也不关心——他关心的只是能把事情做好。

IP协议本身是由Jon在1981年说明的。Jon从一开始便参加了ARPANET的建设，他协调IP在ARPANET上的部署，帮助在早期的Internet上部署域名系统（DNS, Domain Name System）。DNS是推动Internet早期爆炸性增长的主要原因。本书的大部分内容都是关于IP本身的理解，IP改变了我们的生活，IP是Internet的基础。在理解IP的过程中，我们将逐渐理解Jon的某些工作，但本书并不十分周详。

我本人并不认识Jon，但他去世的消息让我感到震惊。现在我可以想像出以前当美国人民听到富兰克林或者托马斯·杰斐逊去世时有多么的悲哀了。Jon和其他一些杰出的人物，如Vint Cerf、Steve Crocker、Bob Kahn等，可以说是Internet的奠基人。Internet已经使得数据通信世界发生了革命性的变化，并且正在使话音通信产生革命，但是Internet的影响却远远超出了技术领域。它正在改变着老式的关于国界和政府的概念。

Jon的努力帮助Internet取得了令人难以置信的成功。一项纯粹的研究，在开始时只是研究一种进行无连接分组交换的途径，竟然深入到数以亿计的人们的生活之中，真让人感到惊讶万分。我相信历史将判定Internet是20世纪最为关键的技术成就，甚至比微处理器还要重要。从这方面来看，Jon肯定是20世纪最有影响的人物之一。当回首这个喧嚣纷杂的世纪时，我希望人们能够记住在Internet的诞生和发展上Jon起了多么关键的作用。他将永垂不朽！

本书原书书号：ISBN 0-13-975483-0

原出版社网址：<http://www.phptr.com>

前言

IP协议已经成为当今占统治地位的连网协议，其使用仍在迅速增长。我们应该感谢WWW(World-Wide Web)这样的工具为Internet提供了友好的界面，使得用户在使用Internet时几乎不用经历什么痛苦的摸索和熟悉过程。站在应用的角度看，WWW引起了Internet巨大的改变，因为在此之前的Internet还不能真正称得上是一种面向用户的服务。但是，在促进Internet发展的问题上，WWW却几乎什么也没有做；如果非要说WWW做了一些什么的话，那么可以说它加重了基础设施的负荷（由于它而大大增多了用户的数量，并且Web数据又具有突出的突发性）。

现在的子网定义和路由器配置等等都变得比以前复杂了。虽然现在有一些工具可以帮助进行网络管理，但是网络管理员却仍有可能不具备正确设计其网络的基本知识。假定新网络在连续不断地连接到Internet上，那么管理这些新网络需要的人数肯定也会显著地增长。为了把工作干好，网络管理员需要比一般的用户更为深刻地理解IP。本书就是写给那些需要在更深层次上理解IP连网有关问题，以便使其网络能够正确设计和运行的人们。一般来说，用户大多只具有使用经验，而不知道在表面之下到底都有些什么，是如何运转的。

第一部分研究IP编址。在一个子网中应该具有多少台主机？该子网的广播地址是什么？一个网络的地址边界应该如何定义？为什么每个子网的IP前缀必须惟一（即不重叠）？网络管理员必须能够设计编址方案，并在他们的路由器拓扑中加以实施，还要能够配置路由协议，理解整个系统正常工作所需要的方方面面，以便在出现问题之后可以马上进行修复。同时，值得指出的是，一个设计良好的网络并非只是比草率设计（或者根本没有设计）的网络运行更为稳定。操作IP地址是一项关键的基本技能，否则IP路由协议和转发便没有什么意义。

在操作IP地址和子网掩码方面有许多至关重要的基本技能，这在许多书中或者RFC中都没有详细地讲述，但在这里你可以看到透彻的解释。我不赞成简单地陈述规则的做法。所以提供了一些精心设计的例子，并且每章的后面都包含一些练习用于检验你对所讲概念的理解程度。所有的练习都提供了答案，并且还包含一些理论解释该答案是如何得出的。有经验的网络管理员是在实践和反复试验中学到这些知识的。本书的目标之一便是帮助新的网络管理员尽快能够应付基本的编址和路由问题。掌握了第一部分内容之后便可以保证一个网络管理员不会被与IP地址有关的问题所难倒。

第二部分主要讨论两个相关的问题。首先，需要理解每个路由器用来处理包的转发决策。其次，需要弄明白IP包实际上是如何在路由器间的子网“跳”中传递的。经过30多年的发展之后^[1]，几乎定义了IP如何在任何一种子网技术上运行的标准。幸运地是，我们并不需要理解全部的标准。因为只有几个标准占据了统治地位，并且从这些技术中学到的知识可以推而广之到其他技术上。所以，第二部分的内容主要是讨论IP包如何在最为常见的LAN^[2]和WAN^[3]子网技术上进行传送。

第二部分首先从IP转发决策过程开始讨论。IP包头中包含一个完整的目的地址，路由器使用路由协议找出在去往目的地方向上最佳的外出接口（有望使该包距最终的目的地更近一

步)。每一个中间的路由器使用的都是同一个决策过程。但是在每个中间子网“步跳”中，具体的特定于子网的细节将有所不同。

IP层是一种抽象，为传输层协议提供了子网无关性。即便IP使这些子网在高层协议看起来十分相似，但是各个子网技术彼此之间其实却存在很大的差异。

IP可以运行在很多种物理介质之上，从ArcNet到X.25，数目众多。在本书中，我们将研究最为常见的LAN（以太网、令牌环网、FDDI）和WAN（PPP和帧中继）介质。在理解了这些之后，推而广之，IP在其他介质上的运行也就很容易理解了。在这里没有提到的子网技术都是一些有被“束之高阁”倾向的协议，不会引起公众的兴趣^[4]。

不但与IP地址有关的问题只有少部分IP用户理解，而且地址和路由的相互作用也同样没多少人能很好地理解。一个具有良好编址方案的网络，可能会大大缩小位于网络核心的路由表的尺寸。第三部分概述路由技术，详细讨论了目前最为流行的两个基于标准的路由协议RIP和OSPF的运行过程。本书的这一部分，尤其是关于路由域互连的一章，可以使人们真正理解编址和路由之间的相互关系。

一个设计良好的编址方案可以使网络的故障排除变得容易，因为在这种网络中使用的地址更有实际意义或者可能与子网编址方案有关（例如，在一个帧中继WAN中）。在这样的网络设计中，“相距较近”的地址将具有共同的前缀或者前导比特，而“相距较远”的前缀编号差别很大，这样便可以通过编号知道它指的是网络的哪一部分。部署新网络的人清楚编址和路由问题是十分重要的。人们有时倾向于更为注重网络应用程序，而忘记网络本身，但是，只有设计良好的网络才能为所有的应用程序提供平稳的运行环境。

本书的最后是一系列的附录。附录A讲述ping和tracert，这是两个最为常用的IP故障排除工具。附录B讨论动态主机配置协议（DHCP，Dynamic Host Configuration Protocol），这是当前经常使用的一个协议，用于在一个网络的定义好的子网前缀内减轻IP地址的管理难度。附录C是对新的IEEE 802.1Q和802.1p标准较为全面的综述，它是为VLAN、在MAC层进行多点传送过滤、服务类别（CoS，Class-of-Service）优先化以及其他功能的标准化而提供的。这些新的特征可能会对将来部署以太网的方式产生较大的影响。因为它们仍然相当新，将它们包含到以太网的主章节中好像还不太合适，故而放在附录C讲述。

注释

[1] IP最早在Internet上部署是在1981年。

[2] 以太网、令牌环网和FDDI。

[3] PPP和帧中继。

[4] 注意，本书中没有详细讨论IP over ATM主要是基于以下考虑：首先，当在PVC模式中使用时，ATM和帧中继有些相似。第二，IP over ATM是一个广泛并且复杂的主题，需要专门用一本书进行介绍。第三，就IP来说，IP over ATM LAN仿真和IP over Ethernet或IP over Token Ring几乎相同。

目 录

译者序
序言
前言

第一部分 IP体系结构、编址和路由

第1章 IP导论	1	3.3 明确的子网划分	33
1.1 何谓IP	1	3.4 层次信息的隐藏	33
1.2 LAN和WAN上的通信	4	3.5 扩展网络前缀	35
1.3 IP体系结构概述	5	3.6 子网编址方案的设计	36
1.3.1 入站包的处理	6	3.6.1 定义子网掩码/扩展前缀长度	37
1.3.2 TCP详解	8	3.6.2 定义每一个子网号	37
1.4 IP包头概述	8	3.6.3 全0或者全1子网	37
1.4.1 IP包的分段	9	3.6.4 确定每个子网中的主机地址	38
1.4.2 其他重要的IP“辅助性”协议	10	3.6.5 定义子网掩码/扩展网络前缀的 长度	39
1.5 何谓路由选择	12	3.6.6 定义子网掩码/扩展网络前缀的 长度	40
1.6 Internet为何如此有用	14	3.6.7 定义各个子网号	40
1.7 Internet应用	15	3.6.8 确定每个子网中的主机地址	41
参考文献	16	3.6.9 确定每个子网的广播地址	42
注释	17	3.7 子网的附加练习	46
第2章 Internet 编址约定和扩展问题	18	3.7.1 子网划分练习1	46
2.1 引言	18	3.7.2 子网划分练习2	47
2.1.1 Internet的扩展问题	18	3.7.3 子网划分练习1答案	47
2.1.2 有类别IP编址	20	3.7.4 子网划分练习2答案	48
2.1.3 有类别编址的局限性	24	3.7.5 思考题	49
2.2 有类别IP编址的补充练习	24	3.7.6 思考题答案	49
2.2.1 练习题	24	注释	50
2.2.2 练习题答案	25	第4章 一般的子网划分	52
参考文献	25	4.1 可变长子网掩码和超级网络介绍	52
注释	25	4.2 VLSM	52
第3章 典型的子网划分	27	4.2.1 固定长度(有类别)子网掩码	52
3.1 引言	27	4.2.2 有效使用IP地址空间	54
3.2 早期的子网划分替代方案	27	4.2.3 路由聚合	55
3.2.1 隐含子网划分	27	4.3 VLSM的设计思路	56
3.2.2 代理ARP	30	4.4 分层编址举例	57
		4.5 VLSM举例	59
		4.5.1 创建192.168.0.0/16的16个子网	59
		4.5.2 确定子网#3 ₄ (192.168.48.0/20)的	

主机地址	60	6.2 在OSI的第一层进行互连	103
4.5.3 确定子网#14 ₄ (192.168.224.0/20) 的 16个子子网	61	6.2.1 10BASE-T的配线	104
4.5.4 确定子网#14 ₄ -3 ₄ (192.168.227.0/24) 的主机地址	61	6.2.2 第一层: 中继器	105
4.5.5 为子网#14 ₄ -14 ₄ (192.168.238.0/24) 定义8个子子网	62	6.3 在OSI的第二层进行互连	106
4.5.6 为子网#14 ₄ -14 ₄ -2 ₄ (192.168.238.64/ 27) 确定主机地址	63	6.4 在OSI的第三层进行互连	111
4.6 VLSM练习	64	参考文献	112
4.6.1 VLSM练习题	64	注释	112
4.6.2 VLSM练习题答案	66	第7章 以太网技术	114
4.7 子网和超级网络的操作	67	7.1 引言	114
4.7.1 分解	70	7.2 介质访问控制: CSMA/CD	115
4.7.2 合并	71	7.2.1 二进制指数回退重传算法	116
注释	75	7.2.2 没有冲突的以太网	117
第5章 可升级的Internet地址管理	76	7.3 以太网的帧格式	117
5.1 Internet 中的CIDR与Intranet 中的 VLSM	76	7.4 以太网介质	120
5.1.1 利用CIDR更高效地分配IPv4的 地址空间	77	7.4.1 粗缆以太网: 10BASE5	120
5.1.2 控制Internet路由表的增长	79	7.4.2 细缆以太网: 10BASE2	122
5.1.3 CIDR分配举例之一	81	7.4.3 运行在双绞线上的以太网: 10BASE-T	123
5.1.4 CIDR分配举例之二	82	7.4.4 运行在光纤上的以太网: 10BASE-F	125
5.2 CIDR练习题	84	7.4.5 运行在CATV上的以太网: 10BROAD36	126
5.3 CIDR练习题答案	85	7.4.6 高于10Mbps的以太网	126
5.4 CIDR部署中端站的隐含行为	87	7.4.7 高于100Mbps的以太网	128
5.5 升级Internet地址空间的新方案	89	7.4.8 以太网还能更快吗?	129
5.5.1 呼吁归还未使用的IP网络前缀	90	7.5 以太网的故障排除	130
5.5.2 专用Internet地址分配	90	参考文献	131
5.5.3 从保留的A类地址空间进行地址 分配	90	注释	132
5.5.4 地址分配策略的隐含行为	91	第8章 IP over Ethernet	133
参考文献	92	8.1 协议多路分用	133
注释	92	8.2 获取相邻节点的地址	134
第二部分 IP下的LAN和WAN子网		8.3 在IEEE 802.3上运行IP	136
第6章 LAN互连	103	8.4 IP的转发决策	137
6.1 引言	103	8.5 较大的交换LAN上的IP升级问题	141
		参考文献	144
		注释	145
		第9章 令牌环和FDDI	146
		9.1 令牌传递环技术简介	146
		9.2 令牌环的细节	147

9.2.1 令牌环的功能描述	147
9.2.2 令牌环的帧格式	148
9.2.3 令牌环LAN互连	155
9.3 FDDI和令牌环	155
9.4 在令牌环和FDDI上如何运行IP	156
参考文献	158
注释	158
第10章 点到点协议	159
10.1 串行线路的分类	159
10.1.1 全双工、半双工和单工	159
10.1.2 带宽对称的电路和带宽不对称 的电路	160
10.2 拨号连接和租用线路	161
10.3 PPP概述	161
10.4 链路控制协议	162
10.4.1 LCP: 建立链路	162
10.4.2 LCP: 维持链路	165
10.4.3 LCP: 撤消链路	166
10.5 鉴别	166
10.5.1 PAP: 口令鉴别协议	166
10.5.2 CHAP: 质询握手鉴别协议	167
10.5.3 EAP: 可扩展的鉴别协议	168
10.6 网络层控制协议	169
10.6.1 IP控制协议 (IPCP) 的 功能	169
10.6.2 无编号链路的问题	172
10.7 帧格式和IP封装	173
参考文献	175
注释	177
第11章 帧中继	179
11.1 广域“云”技术的发展	179
11.2 TDM WAN的可伸缩性问题	182
11.3 帧中继特性	186
11.4 帧格式	186
11.5 本地管理接口协议	189
11.6 IP封装	190
11.7 帧中继网络编址需要考虑的事项	191
11.8 InARP协议描述	194
参考文献	195
注释	196

第三部分 基于标准的IP路由协议

第12章 路由知识导论	197
12.1 编址和路由之间的关系	199
12.2 路由协议的作用和类型	201
12.2.1 距离-向量路由协议	203
12.2.2 链接-状态路由协议	206
12.2.3 静态路由	208
参考文献	210
注释	210
第13章 RIP和OSPF	212
13.1 路由信息协议	212
13.1.1 RIPv1	213
13.1.2 RIPv1小结	217
13.1.3 RIPv2的增强	217
13.1.4 RIPv2小结	220
13.2 开放最短路径优先	220
13.2.1 呼叫协议	221
13.2.2 构成邻接关系	224
13.2.3 DR与BDR选举	226
13.2.4 链接状态数据库	226
13.2.5 OSPF区域	229
13.2.6 OSPF小结	232
13.3 OSPF常识	233
13.3.1 对聚合的控制	233
13.3.2 限制专用OSPF路由器的工作量	234
13.3.3 OSPF的比例关系: 每个ABR中 的区域数	235
13.3.4 OSPF的比例关系: 每个路由器 接口上的邻居数目	235
13.3.5 OSPF的比例关系: 每个区域中 的路由器数目	235
13.4 OSPF配置小结	236
13.4.1 OSPF辅助故障排除	236
13.4.2 从RIP迁移到OSPF	237
13.5 RIP和OSPF小结	238
参考文献	238
注释	241
第14章 异种单点传送路由域互连	243

14.1 互连技术介绍	244
14.2 聚合	245
14.2.1 聚合引起的路由循环	247
14.2.2 小结	248
14.3 显式路由导入/导出	249
14.4 交换局域网、静态路由、 默认路由	253
14.4.1 静态和默认路由	253
14.4.2 默认路由的使用限制	258

14.4.3 动态路由	259
参考文献	260
注释	260

第四部分 附 录

附录A 特殊IP故障排除工具的操作 与使用	261
附录B IEEE 802.1Q和802.1P	275
附录C 动态主机配置协议	280

第一部分 IP体系结构、 编址和路由

第1章 IP 导论

1.1 何谓IP

尽管IP已经逐渐成为一种通用的组网协议，（从Internet和Intranet的爆炸性增长可以证明这一点），但是大多数人们对IP究竟能做什么及其各部分之间如何相互作用却不甚明了。关于Internet的书籍很多，但大多数谈论的是关于聊天室、如何创建时髦的Web站点以及其他常用的应用程序。本书所讨论的是Internet的基础设施内幕，而不是用户所看到的表面部分。在本书中你将学到IP究竟如何工作的方方面面。尽管本书的重点是讨论基础设施的本质，但是本书并非是一本有关通信原理的书；本书注重实用性，面向解决方案，尤其适合于那些需要快速掌握IP如何工作的人们。那么，什么是IP呢？IP其实就是多个计算机用于交换信息的基本包（Packet）格式。

什么是包呢？包是一种事先定义好的格式，通常由一个包头（header）后跟一些数据组成，在这里数据可以是：1）部分文件；2）虚拟终端应用程序中的击键和字符回波；3）一段E-mail消息。在Internet上传输的所有信息将被划分为独立的包。包交换（有时也称为分组交换）是在60年代开发出来的，当时主要用于在军事应用程序环境中提供健壮的通信基础设施。由于将信息流分为了包，各个包可以沿自己的路径通过网络，所以便可以避免网络的某些部分负荷太重，换句话说，即无法正常运转。

本书讨论的既不是Web浏览器，也不是如何使用XML设计Web页面，更不是如何在Internet上查找信息。在本书中，你将学到：如何操作IP地址以及如何为一个网络设计编址方案；IP在最为常见的LAN和WAN介质是如何运行的，包括常见连接问题的诊断；几种常见非专有路由协议的功能特性，包括如何在网络上应用这些协议的具体例子。在可能情况下，将结合具体的例子和说明性的练习对所讨论的技术信息进行举例说明。

在任何“分组交换”协议中，进行通信的两台计算机都是把它们的数据划分为分组（也就是包），然后将这些分组通过“分组交换网络”进行传输。分组交换协议有很多，但是在本书中我们将只讨论IP协议。IP因作为Internet的构件块而得到广泛的流行，并且这一趋势仍在日渐加强。其他常见的分组交换协议还有ITU-T^[1]的X.25——第一个（国际性的）基于标准的分组交换技术、IBM的系统网络体系结构（SNA, Systems Network Architecture）、Novell的国际数据包交换协议（IPX, Internetwork Packet Exchange）。另外还有一些其他的分组交换协议，包括AppleTalk Phases I和AppleTalk Phases II、Banyan Vines、ChaosNet，数字设备公司的DECnet Phase IV和V、IPv6、开放系统互连（OSI, Open System Interconnection）、Unix-to-Unix CoPy（UUCP）、施乐公司的Xerox Network Services（XNS）等等。

IP允许数据流过内部部署了IP的计算机网络，例如Internet和众多公司内部Intranet。由

IP包所携带的数据既可以是传统的计算机数据，也可以是新近才在IP上使用的数字化音频和视频流。话音和视频被数字化之后其实就是数据，但是，与传统的数据传输（例如文件传输）不同，它们在传输上具有特殊的需求。话音和视频是时间敏感的，并且对延迟的容忍十分有限，或者说延迟可变性较小。很奇怪，话音和音频流可以容忍丢失一些数据而不会产生听觉或者视觉上的失真，但数据业务却必须保证传输的正确性，也就是说需要花费时间来重传丢失或者损坏的数据包以保证整个传输完好无损。

在IP世界的早期，分组交换设备通常称为“网关”，具体原因可能是因为它们常常作为本地校园网和广域网（例如，ARPANET）之间的网关的缘故。现在，这些设备在更多的情况下被称为“路由器”^[2]。

图1-1所示为一个连接了五个子网的路由器图标。路由器可以在其任何接口上接收数据包，然后决定如何将之转发到其目的地，一般来说都会通过某个不同的接口从路由器将数据包转发出去。图中的双向箭头用以指示数据包既可以从该接口进入路由器，也可以从该接口离开路由器。

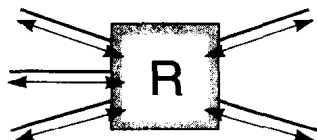


图1-1 一个路由器的表示

路由器是基于IP的网络（包括Internet）的基本构件。

为什么要进行分组交换？

为什么计算机在通信时需要进行“分组交换”呢？为什么不在计算机之间建立一条临时的“电话呼叫”呢？这种差别主要来源于电话网的构建方式和计算机通信的基本特性。大多数计算机通信常常十分短暂——一般在秒的量级，但是却十分强烈。用“突发性”这个词来描述这种情形最为合适。我们不妨将之和话音通信做一下比较，话音通信通常会持续很长一段时间，但是在整个通话期间所使用的“带宽”数量却是可以预测出来的。

一般的电话网都是为进行话音呼叫而设计的，而话音通信常常会持续一段较长（相对于其建立呼叫的时间）的时间。拨号、等待对方拿起电话的时间量级通常都在5秒左右。从电话网络的角度来看，为呼叫建立起资源的所有较为困难的工作都是在前面这几秒钟的时间内完成的，在呼叫建立之后，后面的通话将一直使用这些分配给它的资源。由于在建立呼叫上的工作和维持一条已经建立好的呼叫相比前者更为困难，所以有史以来，电话公司对第一分钟的收费往往比后面的时间要贵一些。

另外，可以确定计算机相互之间的呼叫在什么时候能够结束。而在话音呼叫中，通话的终止是由双方手工完成的。此外，当很快还会向对方发送数据时，计算机可能希望将某个打开的信道保持一段时间。在专线网络（比如说，话音电话网）中维持大量打开的空闲连接会浪费网络资源。并且，假设一个计算机想同时维持10个活动的呼叫，那么将需要10个单独的物理线路，不管用与不用，每个月都要为此支付一定的费用。

除了数据的突发性与话音的非突发性之间存在很大差异之外，话音呼叫的建立速度还受到了人们使用手指进行拨号的限制，因为我们不可能拨得太快而使电话网的控制或者“信令”信道负担过重。而计算机可能需要在很短的时间之内向多个对等体（层）发送一段较短的数据比特，强行在每秒内建立多个连接。

分组交换具有两个主要的操作模式，即“虚电路”和“数据报”，前者是“面向连接”的，后者是“无连接”的。是什么把一条连接变为“虚”连接的呢？从本质上讲，

多个虚连接可以共享某个单独的物理设备，比如说一个单独拨号连接或者到广域网(WAN)交换机的一条专用串行链路。

虽然让人同时进行多个会话较为困难，但是计算机却没有这方面的限制，它可以很容易地同时和多个对等计算机之间维持“会话”。如果每个连接都需要惟一的物理设备，例如它们自己的电话线，那么到10个对等体(层)的10个连接将暗示着呼叫方计算机需要10条物理的电话线路(或者其他任何合适的专用物理介质)。

但是，当把计算机之间的通信划分为分组之后，各个分组便可以共享一条单独线路，其中某个分组可能是发往对等体#1的，而另三个分组可能是发往对等体#7的等等。共享单个物理线路的所有“虚”连接便构成了一个分组交换网络。每个分组的头将告诉网络该分组应该去往何方，从而使得发送方计算机不再需要到每个对等体都需要一条单独物理连接。并且，“分组交换”比“电路交换”更为高效。

分组交换协议有两种基本的类型，即“虚电路”和“数据报”。在虚电路类型中，在每个虚呼叫的开头仍然需要建立呼叫，但是所有先前建立起来的虚连接将共享同一条到分组交换网的物理访问链路。这便消除了“电路交换”存在的“问题”——即多个连接需要有多条物理线路。对于分组交换，所有的虚连接将共享同一条到分组交换机的“物理线路”，而每个分组都具有其自己的目标地址，该目标地址取决于在呼叫建立阶段所分配的虚电路ID。虚电路ID是一个较短的“标识”或者“句柄”，分组交换机可以很容易地据此通过简单地查表将该分组转发到目的地。

在这里挂断就更不成问题了，因为一个建立好的连接只不过是两台计算机之间的分组交换机内的一个表项而已。但是，假如分组交换机只具有有限的内存，最好是拆掉那些不再需要的虚电路，以便为其他虚电路让出一些空间。建立虚电路所消耗的内存比起空闲电话呼叫所占用的带宽而言，代价要小的多。这与物理连接的情形形成鲜明的对比，在物理连接中，规模具有边界限制，也就是说，如果一个计算机有16个调制解调器，而有16个以上的计算机想要向该计算机发送它们的数据，那么，将超出该系统的物理能力。

“无连接”是分组交换的另一种主要形式，IP和许多其他的网络层协议都属于这一范畴。在无连接模式中，仍然保持了虚电路方案中的“单一访问链路”方面，通过到分组交换网络的同一条连接可以到达多个目的地。但是，在无连接网络中，并非在任何数据发送之前都明确地进行呼叫建立，然后使用所分配的特定于该连接的短句柄作为每个分组的目的地地址，而是让每个分组都携带完整的目的地址。即，只管将分组发送到网络中，至于如何将分组转发到它所指明的目的地最为合适则完全依赖于中间的每个分组交换机(即IP路由器)。

分组交换和ARPANET

通常，转发分组的网络元素被称为“分组交换机”。分组交换是由Paul Baran和其他人一起在60年代初期发明出来的。到1966年，人们开始酝酿ARPANET的建设，并且在1969年最终建成了它^[3]。在ARPANET中，分组交换机被称为“接口报文处理器(IMP, Interface Message Processor)”。当时的IMP使用的是内存仅有12KB的Honeywell 516小型机，在上面运行特定于ARPANET的分组交换软件。IMP软件是由Bolt Beranek

and Newman公司（BBN）开发研制的，该公司当前已经成为GTE的一部分。

从1970年起，ARPANET开始使用网络控制协议（NCP，Network Control Protocol）包格式，直到ARPANET转换为在IP（TCP/IP）协议栈上运行传输控制协议（TCP，Transmission Control Protocol）为止（时间为1983年1月1日）。ARPANET的研究人员设计了IP，该协议是在ARPANET提供的真实试验台上进行多年分组交换实践的结晶。ARPANET逐渐成长并发展为今天的Internet。从70年代到80年代，ARPANET取得了显著的增长，并且为深入研究计算机网络提供了必要的支持。在80年代后期，ARPANET逐渐演化为构成Internet的上千个IP网络中的一个成员。在进行了20多年的忠实服务之后，ARPANET在1990年被关闭。

1.2 LAN和WAN上的通信

当一个路由器和另一个路由器或者端站（通常位于局域网或者LAN之上）进行交谈时，需要以某种方式向其相邻节点发送数据包。一个包仅使用其IP地址是无法发送给其相邻节点的；IP地址是高层地址。实际的操作是把IP包用特定于子网类型的帧格式封装起来。在该帧中包含有目标节点的地址，并且通常还包含有路由器的子网层源地址。图1-2所示为IP协议栈中各层之间的关系以及封装的概念。在该图中，可以用用户数据报协议(UDP)代替TCP；在当前的分层讨论中它们是相同的。

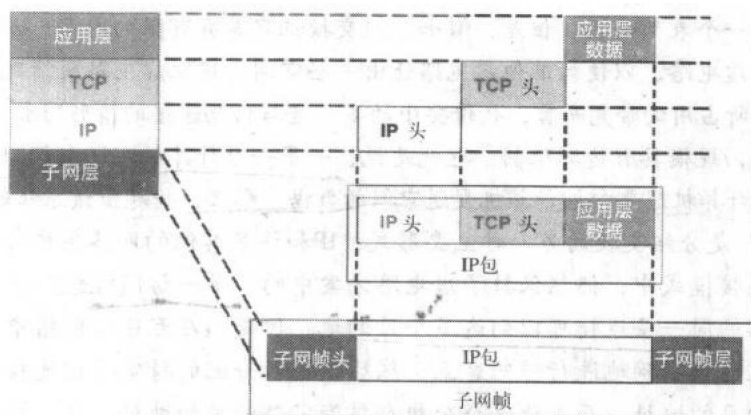


图1-2 分层和封装

LAN和WAN“子网”具有各自的编址方案，路由器在相互通信时必须使用相应的地址。通常除了需要知道其相邻节点的IP地址外，路由器还需要知道其相邻节点的子网层地址。针对每种子网介质，IP将使用不同的技术来获取其相邻节点的子网地址是什么。

在日常生活中，具有多层地址的系统随处可见。我们不妨考虑一下两个位于不同办公楼内的人之间进行通信的情形。其中，一个位于Highgate Tower，2301房间，另一个位于Trumpet Place，3107房间。现在，为了能够进行通信，他们需要知道对方的街道地址。在现实生活中，他们有可能也知道对方所在办公楼的街道地址，但是这并非问题的关键。问题的关键是我们的确每天都在和各种格式的具有多层地址的系统打交道。