



“十二五”普通高等教育本科国家级规划教材

计算机网络

(第7版)

谢希仁 编著



中国工信出版集团



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.phei.com.cn>



“十二五”普通高等教育本科国家级规划教材

计 算 机 网 络

(第 7 版)

谢希仁 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书自 1989 年首次出版以来,曾于 1994 年、1999 年、2003 年、2008 年和 2013 年分别出了修订版。在 2006 年本书通过了教育部的评审,被纳入普通高等教育“十一五”国家级规划教材;2008 年出版的第 5 版获得了教育部 2009 年精品教材称号。2013 年出版的第 6 版是“十二五”普通高等教育本科国家级规划教材。现在的第 7 版又在第 6 版的基础上进行了一些修订。

全书分为 9 章,比较全面系统地介绍了计算机网络的发展和原理体系结构、物理层、数据链路层(包括局域网)、网络层、运输层、应用层、网络安全、互联网上的音频/视频服务,以及无线网络和移动网络等内容。各章均附有习题(附录 A 给出了部分习题的答案和提示)。全书课件(PowerPoint 文件)放在电子工业出版社悦学多媒体课程资源平台上(http://yx.51zhy.cn/mtres/phei_cnetwork.jsp),供读者下载参考。

本书的特点是概念准确、论述严谨、内容新颖、图文并茂,突出基本原理和基本概念的阐述,同时力图反映计算机网络的一些最新发展。本书可供电气信息类和计算机类专业的大学本科生和研究生使用,对从事计算机网络工作的工程技术人员也有参考价值。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

计算机网络 / 谢希仁编著. —7 版. —北京:电子工业出版社,2017.1

“十二五”普通高等教育本科国家级规划教材

ISBN 978-7-121-30295-4

I. ①计… II. ①谢… III. ①计算机网络—高等学校—教材 IV. ①TP393

中国版本图书馆 CIP 数据核字(2016)第 269601 号

策划编辑:郝志恒

责任编辑:牛晓丽

印 刷:三河市华成印务有限公司

装 订:三河市华成印务有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱

邮编:100036

开 本:787×1092 1/16

印张:29

字数:742.4 千字

版 次:1999 年 4 月第 2 版

2017 年 1 月第 7 版

印 次:2017 年 1 月第 1 次印刷

定 价:45.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010) 88254888, 88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:QQ 9616328。

前 言

本教材第6版被纳入“十二五”普通高等教育本科国家级教材规划。由于本教材所讲授的是计算机网络最基本的原理，而这些基本原理是比较成熟和稳定的，因此，介绍基本原理的部分相对稳定，不会有很大的变动。

第7版的教材有以下一些改动。

互联网的发展非常快，编者水平很有限，只能把最重要的一些新内容增加到新版教材中。所有各章应参考的 RFC 文档和参考文献也都尽可能进行了更新。对于重点内容适当地增加了一些习题。

删除了第10章。由于原第10章中的 IPv6 和 MPLS 已属于比较成熟的技术，现在改放在第4章中介绍。原第10章中的 P2P 应用，现在放在第6章中介绍，并增加了一些新的内容。

另一个改动是 Internet 的译名不再使用推荐译名“因特网”，而改为目前大家已普遍使用的“互联网”。虽然推荐译名“因特网”有其权威性，但实践证明，各界人士大都不愿意用。考虑到我国的实际情况，编者决定从第7版起改用事实上的标准译名“互联网”。

此外，比特这个单位以前用英文字母 b 来表示，从第7版起改为 bit。这样可能更加清楚明确些。

本教材的参考学时数为70学时左右。在课程学时数较少的情况下可以只学习前6章，这样仍可获得有关互联网的最基本的知识。

书后共有三个附录，附录A是部分习题解答（而不是详细解题步骤）、附录B是英文缩写词，附录C是参考文献与网址。

考虑到现在从网上下载已非常方便，因此原附在书中的 CD-ROM 已非必要。读者可以访问电子工业出版社悦学多媒体课程资源平台(http://yx.51zhy.cn/mtrcsRes/phei_cnetwork.jsp)下载有关的参考内容。

为满足不同院校、不同学时、不同专业的教学需要，第7版同时推出了精简版，也就是由电子工业出版社出版的《计算机网络简明教程》第3版。

林波博士对第7章网络安全提出了许多修改意见。吴自珠副教授一直对本教材的出版给予全力支持。对此，编者均表示诚挚的谢意。由于编者水平所限，书中难免还存在一些缺点和错误，殷切希望广大读者批评指正。

谢希仁

2016年8月

于解放军理工大学，南京

编者的电子邮件地址：xiexiren@tsinghua.org.cn

欢迎指出书中内容的不足和错误，但作者无法满足一些深入探讨和科研项目咨询的需求，请予谅解。若需索取解题的详细步骤，请参考作者编著的《计算机网络释疑与习题解答》。

目 录

第 1 章 概述	1
1.1 计算机网络在信息时代中的作用	1
1.2 互联网概述	4
1.2.1 网络的网络	4
1.2.2 互联网基础结构发展的三个阶段	5
1.2.3 互联网的标准化工作	8
1.3 互联网的组成	10
1.3.1 互联网的边缘部分	10
1.3.2 互联网的核心部分	12
1.4 计算机网络在我国的发展	17
1.5 计算机网络的类别	19
1.5.1 计算机网络的定义	19
1.5.2 几种不同类别的计算机网络	20
1.6 计算机网络的性能	21
1.6.1 计算机网络的性能指标	21
1.6.2 计算机网络的非性能特征	26
1.7 计算机网络体系结构	27
1.7.1 计算机网络体系结构的形成	27
1.7.2 协议与划分层次	28
1.7.3 具有五层协议的体系结构	30
1.7.4 实体、协议、服务和访问点	34
1.7.5 TCP/IP 的体系结构	35
本章的重要概念	37
习题	38
第 2 章 物理层	41
2.1 物理层的基本概念	41
2.2 数据通信的基础知识	42
2.2.1 数据通信系统的模型	42
2.2.2 有关信道的几个基本概念	43
2.2.3 信道的极限容量	45
2.3 物理层下面的传输媒体	46
2.3.1 导引型传输媒体	47
2.3.2 非导引型传输媒体	51
2.4 信道复用技术	53

2.4.1	频分复用、时分复用和统计时分复用.....	53
2.4.2	波分复用	56
2.4.3	码分复用	57
2.5	数字传输系统	59
2.6	宽带接入技术	60
2.6.1	ADSL 技术.....	60
2.6.2	光纤同轴混合网 (HFC 网)	63
2.6.3	FTTx 技术	65
	本章的重要概念.....	66
	习题	67
第 3 章	数据链路层.....	69
3.1	使用点对点信道的数据链路层.....	70
3.1.1	数据链路和帧	70
3.1.2	三个基本问题	71
3.2	点对点协议 PPP.....	76
3.2.1	PPP 协议的特点.....	76
3.2.2	PPP 协议的帧格式.....	78
3.2.3	PPP 协议的工作状态.....	80
3.3	使用广播信道的数据链路层.....	82
3.3.1	局域网的数据链路层	82
3.3.2	CSMA/CD 协议	85
3.3.3	使用集线器的星形拓扑	90
3.3.4	以太网的信道利用率	92
3.3.5	以太网的 MAC 层	93
3.4	扩展的以太网	97
3.4.1	在物理层扩展以太网	97
3.4.2	在数据链路层扩展以太网	99
3.4.3	虚拟局域网	101
3.5	高速以太网	103
3.5.1	100BASE-T 以太网	103
3.5.2	吉比特以太网	104
3.5.3	10 吉比特以太网(10GE)和更快的以太网.....	106
3.5.4	使用以太网进行宽带接入	107
	本章的重要概念.....	108
	习题	109
第 4 章	网络层.....	113
4.1	网络层提供的两种服务.....	113
4.2	网际协议 IP.....	115
4.2.1	虚拟互连网络	116

4.2.2	分类的 IP 地址.....	118
4.2.3	IP 地址与硬件地址.....	122
4.2.4	地址解析协议 ARP.....	124
4.2.5	IP 数据报的格式.....	128
4.2.6	IP 层转发分组的流程.....	132
4.3	划分子网和构造超网	135
4.3.1	划分子网	135
4.3.2	使用子网时分组的转发	140
4.3.3	无分类编址 CIDR (构造超网)	141
4.4	网际控制报文协议 ICMP.....	147
4.4.1	ICMP 报文的种类.....	148
4.4.2	ICMP 的应用举例.....	149
4.5	互联网的路由选择协议.....	151
4.5.1	有关路由选择协议的几个基本概念.....	151
4.5.2	内部网关协议 RIP	153
4.5.3	内部网关协议 OSPF.....	159
4.5.4	外部网关协议 BGP.....	163
4.5.5	路由器的构成	167
4.6	IPv6	171
4.6.1	IPv6 的基本首部.....	171
4.6.2	IPv6 的地址.....	173
4.6.3	从 IPv4 向 IPv6 过渡	176
4.6.4	ICMPv6	177
4.7	IP 多播.....	178
4.7.1	IP 多播的基本概念.....	178
4.7.2	在局域网上进行硬件多播	179
4.7.3	网际组管理协议 IGMP 和多播路由选择协议.....	180
4.8	虚拟专用网 VPN 和网络地址转换 NAT.....	185
4.8.1	虚拟专用网 VPN	185
4.8.2	网络地址转换 NAT.....	187
4.9	多协议标记交换 MPLS	189
4.9.1	MPLS 的工作原理.....	190
4.9.2	MPLS 首部的位置与格式	192
	本章的重要概念.....	193
	习题	195
第 5 章	运输层.....	203
5.1	运输层协议概述	203
5.1.1	进程之间的通信	203
5.1.2	运输层的两个主要协议	205
5.1.3	运输层的端口	206

5.2	用户数据报协议 UDP	208
5.2.1	UDP 概述	208
5.2.2	UDP 的首部格式	209
5.3	传输控制协议 TCP 概述	210
5.3.1	TCP 最主要的特点	210
5.3.2	TCP 的连接	212
5.4	可靠传输的工作原理	212
5.4.1	停止等待协议	213
5.4.2	连续 ARQ 协议	216
5.5	TCP 报文段的首部格式	217
5.6	TCP 可靠传输的实现	221
5.6.1	以字节为单位的滑动窗口	221
5.6.2	超时重传时间的选择	225
5.6.3	选择确认 SACK	226
5.7	TCP 的流量控制	227
5.7.1	利用滑动窗口实现流量控制	227
5.7.2	TCP 的传输效率	228
5.8	TCP 的拥塞控制	229
5.8.1	拥塞控制的一般原理	229
5.8.2	TCP 的拥塞控制方法	232
5.8.3	主动队列管理 AQM	236
5.9	TCP 的运输连接管理	238
5.9.1	TCP 的连接建立	238
5.9.2	TCP 的连接释放	240
5.9.3	TCP 的有限状态机	241
	本章的重要概念	242
	习题	244
第 6 章	应用层	251
6.1	域名系统 DNS	252
6.1.1	域名系统概述	252
6.1.2	互联网的域名结构	253
6.1.3	域名服务器	255
6.2	文件传送协议	260
6.2.1	FTP 概述	260
6.2.2	FTP 的基本工作原理	261
6.2.3	简单文件传送协议 TFTP	262
6.3	远程终端协议 TELNET	263
6.4	万维网 WWW	264
6.4.1	万维网概述	264
6.4.2	统一资源定位符 URL	266

6.4.3	超文本传送协议 HTTP	267
6.4.4	万维网的文档	274
6.4.5	万维网的信息检索系统	279
6.4.6	博客和微博	281
6.4.7	社交网站	283
6.5	电子邮件	284
6.5.1	电子邮件概述	284
6.5.2	简单邮件传送协议 SMTP	287
6.5.3	电子邮件的信息格式	289
6.5.4	邮件读取协议 POP3 和 IMAP	289
6.5.5	基于万维网的电子邮件	291
6.5.6	通用互联网邮件扩充 MIME	291
6.6	动态主机配置协议 DHCP	295
6.7	简单网络管理协议 SNMP	297
6.7.1	网络管理的基本概念	297
6.7.2	管理信息结构 SMI	299
6.7.3	管理信息库 MIB	302
6.7.4	SNMP 的协议数据单元和报文	304
6.8	应用进程跨越网络的通信	307
6.8.1	系统调用和应用编程接口	307
6.8.2	几种常用的系统调用	309
6.9	P2P 应用	311
6.9.1	具有集中目录服务器的 P2P 工作方式	311
6.9.2	具有全分布式结构的 P2P 文件共享程序	313
6.9.3	P2P 文件分发的分析	315
6.9.4	在 P2P 对等方中搜索对象	316
	本章的重要概念	319
	习题	321
第 7 章	网络安全	324
7.1	网络安全问题概述	324
7.1.1	计算机网络面临的安全性威胁	324
7.1.2	安全的计算机网络	326
7.1.3	数据加密模型	327
7.2	两类密码体制	328
7.2.1	对称密钥密码体制	328
7.2.2	公钥密码体制	329
7.3	数字签名	330
7.4	鉴别	332
7.4.1	报文鉴别	332
7.4.2	实体鉴别	335

7.5 密钥分配	337
7.5.1 对称密钥的分配	337
7.5.2 公钥的分配	339
7.6 互联网使用的安全协议	340
7.6.1 网络层安全协议	340
7.6.2 运输层安全协议	344
7.6.3 应用层安全协议	347
7.7 系统安全：防火墙与入侵检测	349
7.7.1 防火墙	349
7.7.2 入侵检测系统	351
7.8 一些未来的发展方向	351
本章的重要概念	352
习题	353
第 8 章 互联网上的音频/视频服务	355
8.1 概述	355
8.2 流式存储音频/视频	359
8.2.1 具有元文件的万维网服务器	360
8.2.2 媒体服务器	360
8.2.3 实时流式协议 RTSP	362
8.3 交互式音频/视频	364
8.3.1 IP 电话概述	364
8.3.2 IP 电话所需要的几种应用协议	368
8.3.3 实时运输协议 RTP	369
8.3.4 实时运输控制协议 RTCP	371
8.3.5 H.323	372
8.3.6 会话发起协议 SIP	373
8.4 改进“尽最大努力交付”的服务	375
8.4.1 使互联网提供服务质量	375
8.4.2 调度和管制机制	377
8.4.3 综合服务 IntServ 与资源预留协议 RSVP	380
8.4.4 区分服务 DiffServ	383
本章的重要概念	385
习题	386
第 9 章 无线网络和移动网络	390
9.1 无线局域网 WLAN	390
9.1.1 无线局域网的组成	391
9.1.2 802.11 局域网的物理层	395
9.1.3 802.11 局域网的 MAC 层协议	397
9.1.4 802.11 局域网的 MAC 帧	403

9.2 无线个人区域网 WPAN	406
9.3 无线城域网 WMAN	409
9.4 蜂窝移动通信网	411
9.4.1 蜂窝无线通信技术简介	411
9.4.2 移动 IP.....	414
9.4.3 蜂窝移动通信网中对移动用户的路由选择.....	418
9.4.4 GSM 中的切换.....	419
9.4.5 无线网络对高层协议的影响	420
9.5 两种不同的无线上网	421
本章的重要概念.....	422
习题	423
附录 A 部分习题的解答	426
附录 B 英文缩写词	442
附录 C 参考文献与网址	452

表)。该警察既不指明到下一个岔口以后再应当如何走，也不指明还要经过几个岔口才到达机场。他仅仅指出下一个岔口的方向。其回答可能是：“向左转。”到了下一个岔口，再讯问到机场该走哪个方向？回答可能是：“直行。”这样，每到一个岔口，就询问下一步该如何走。这样，即使我们没有地图，但最终一定可以到达目的地——机场。

上面所讨论的是 IP 层怎样根据路由表的内容进行分组转发，而没有涉及到路由表一开始是如何建立的以及路由表中的内容应如何进行更新。但是在进一步讨论路由选择之前，我们还要先介绍划分子网和构造超网这两个非常重要的概念。

4.3 划分子网和构造超网

4.3.1 划分子网

1. 从两级 IP 地址到三级 IP 地址

在今天看来，在 ARPANET 的早期，IP 地址的设计确实不够合理。

第一，IP 地址空间的利用率有时很低。

每一个 A 类地址网络可连接的主机数超过 1000 万，而每一个 B 类地址网络可连接的主机数也超过 6 万。有的单位申请到了一个 B 类地址网络，但所连接的主机数并不多，可是又不愿意申请一个足够使用的 C 类地址，理由是考虑到今后可能的发展。IP 地址的浪费，还会使 IP 地址空间的资源过早地被用完。

第二，给每一个物理网络分配一个网络号会使路由表变得太大因而使网络性能变坏。

每一个路由器都应当能够从路由表查出应怎样到达其他网络的下一跳路由器。因此，互联网中的网络数越多，路由器的路由表的项目数也就越多。这样，即使我们拥有足够多的 IP 地址资源可以给每一个物理网络分配一个网络号，也会导致路由器的路由表中的项目数过多。这不仅增加了路由器的成本（需要更多的存储空间），而且使查找路由时耗费更多的时间，同时也使路由器之间定期交换的路由信息急剧增加，因而使路由器和整个互联网的性能都下降了。

第三，两级 IP 地址不够灵活。

有时情况紧急，一个单位需要新的地点马上开通一个新的网络。但是在申请到一个新的 IP 地址之前，新增加的网络是不可能连接到互联网上工作的。我们希望有一种方法，使一个单位能随时灵活地增加本单位的网络，而不必事先到互联网管理机构去申请新的网络号。原来的两级 IP 地址无法做到这一点。

为解决上述问题，从 1985 年起在 IP 地址中又增加了一个“子网号字段”，使两级 IP 地址变成为三级 IP 地址，它能够较好地解决上述问题，并且使用起来也很灵活。这种做法叫做划分子网(subnetting) [RFC 950]，或子网寻址或子网路由选择。划分子网已成为互联网的正式标准协议。

划分子网的基本思路如下：

(1) 一个拥有许多物理网络的单位，可将所属的物理网络划分为若干个子网(subnet)。划分子网纯属一个单位内部的事情。本单位以外的网络看不见这个网络是由多少个子网组成，因为这个单位对外仍然表现为一个网络。

(2) 划分子网的方法是从网络的主机号借用若干位作为子网号(subnet-id), 当然主机号也就相应减少了同样的位数。于是两级 IP 地址在本单位内部就变为三级 IP 地址: 网络号、子网号和主机号。也可以用以下记法来表示:

$$\text{IP 地址} ::= \{ \langle \text{网络号} \rangle, \langle \text{子网号} \rangle, \langle \text{主机号} \rangle \} \quad (4-2)$$

(3) 凡是从其他网络发送给本单位某台主机的 IP 数据报, 仍然是根据 IP 数据报的目的网络号找到连接在本单位网络上的路由器。但此路由器在收到 IP 数据报后, 再按目的网络号和子网号找到目的子网, 把 IP 数据报交付目的主机。

下面用例子说明划分子网的概念。图 4-18 表示某单位拥有一个 B 类 IP 地址, 网络地址是 145.13.0.0 (网络号是 145.13)。凡目的地址为 145.13.x.x 的数据报都被送到这个网络上的路由器 R₁。

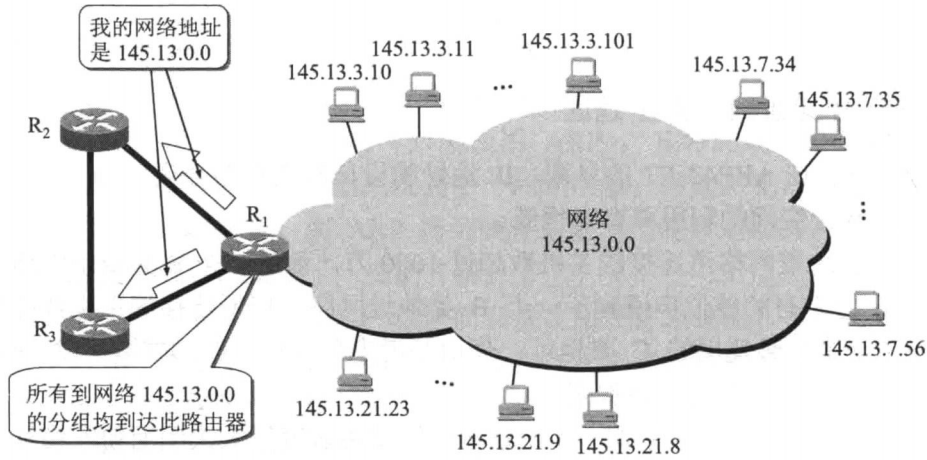


图 4-18 一个 B 类网络 145.13.0.0

现把图 4-18 的网络划分为三个子网 (图 4-19)。这里假定子网号占用 8 位, 因此在增加了子网号后, 主机号就只有 8 位。所划分的三个子网分别是: 145.13.3.0, 145.13.7.0 和 145.13.21.0。在划分子网后, 整个网络对外部仍表现为一个网络, 其网络地址仍为 145.13.0.0。但网络 145.13.0.0 上的路由器 R₁ 在收到外来的数据报后, 再根据数据报的目的地址把它转发到相应的子网。

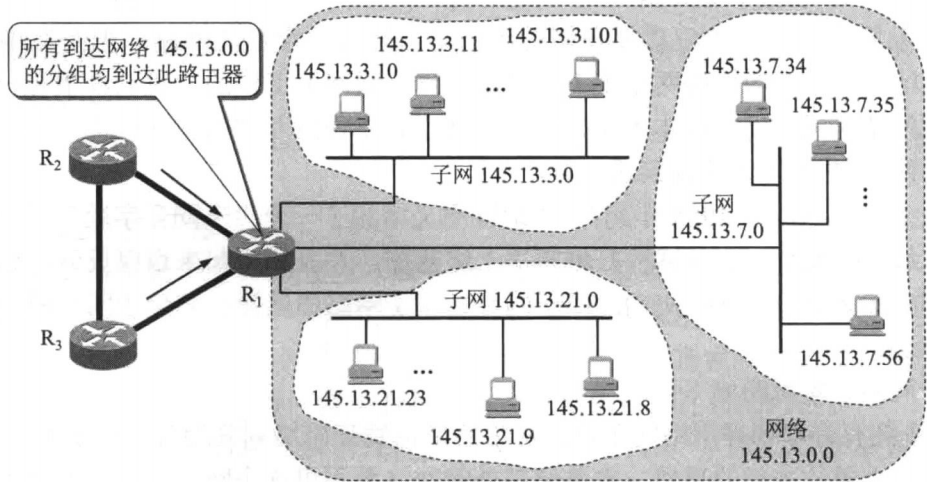


图 4-19 把图 4-18 的网络 145.13.0.0 划分为三个子网, 但对外仍是一个网络

总之，当没有划分子网时，IP 地址是两级结构。划分子网后 IP 地址变成了三级结构。划分子网只是把 IP 地址的主机号这部分进行再划分，而不改变 IP 地址原来的网络号。

2. 子网掩码

现在剩下的问题就是：假定有一个数据报（其目的地址是 145.13.3.10）已经到达了路由器 R_1 。那么这个路由器如何把它转发到子网 145.13.3.0 呢？

我们知道，从 IP 数据报的首部无法看出源主机或目的主机所连接的网络是否进行了子网的划分。这是因为 32 位的 IP 地址本身以及数据报的首部都没有包含任何有关子网划分的信息。因此必须另外想办法，这就是使用子网掩码(subnet mask)（见图 4-20）。

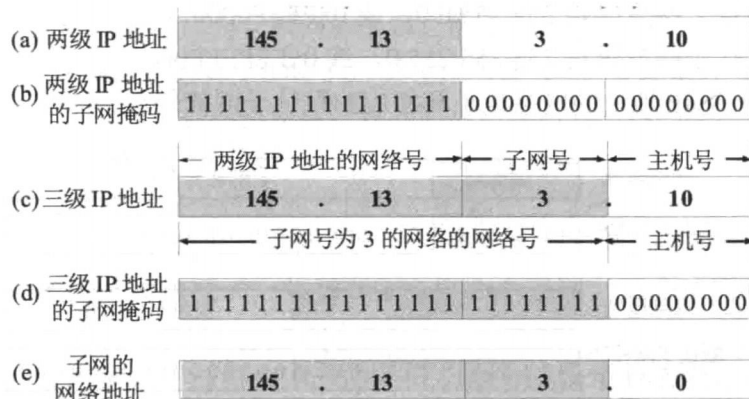


图 4-20 IP 地址的各字段和子网掩码（以 145.13.3.30 为例）

图 4-20(a)是 IP 地址为 145.13.3.10 的主机本来的两级 IP 地址结构。图 4-20(b)是这个两级 IP 地址的子网掩码。图 4-20(c)是同一地址的三级 IP 地址结构，也就是说，现在从原来 16 位的主机号中拿出 8 位作为子网号，而主机号由 16 位减少到 8 位。请注意，现在子网号为 3 的网络的网络地址是 145.13.3.0（既不是原来两级 IP 地址的网络地址 145.13.0.0，也不是简单的子网号 3）。为了使路由器 R_1 能够很方便地从数据报中的目的 IP 地址中提取出所要找的子网的网络地址，路由器 R_1 就要使用三级 IP 地址的子网掩码。图 4-20(d)是三级 IP 地址的子网掩码，它也是 32 位，由一串 24 个 1 和跟随的一串 8 个 0 组成。子网掩码中的 1 对应于 IP 地址中原来二级地址中的 16 位网络号加上新增加的 8 位子网号，而子网掩码中的 0 对应于现在的 8 位主机号。虽然 RFC 文档中没有规定子网掩码中的一串 1 必须是连续的，但却极力推荐在子网掩码中选用连续的 1，以免出现可能发生的差错。

图 4-20(e)表示 R_1 把三级 IP 地址的子网掩码和收到的数据报的目的 IP 地址 145.13.3.10 逐位相“与”(AND)（计算机进行这种逻辑 AND 运算是很容易的），得出了所要找的子网的网络地址 145.13.3.0。

使用子网掩码的好处就是：不管网络有没有划分子网，只要把子网掩码和 IP 地址进行逐位的“与”运算(AND)，就立即得出网络地址来。这样在路由器处理到来的分组时就可采用同样的算法。

归纳一下上述的要点。从网络 145.13.0.0 外面看，这这就是一个普通的 B 类网络，其子网掩码为 16 个连 1 后面跟上 16 个连 0。但进入到这个网络后（即到了路由器 R_1 ），就看到了还有许多网络（即划分了子网后的许多网络），其网络地址是 145.13.x.0（这里 x 可以表示不同的数值），而这些网络的子网掩码都是 24 个连 1 后面跟上 8 个连 0。总之，在这个 B 类网

络的外面和里面，看到的网络是不一样的。

这里还要弄清一个问题，这就是：在不划分子网时，既然没有子网，为什么还要使用子网掩码？这就是为了更便于查找路由表。现在互联网的标准规定：所有的网络都必须使用子网掩码，同时在路由器的路由表中也必须要有子网掩码这一栏。如果一个网络不划分子网，那么该网络的子网掩码就使用**默认子网掩码**。默认子网掩码中 1 的位置和 IP 地址中的网络号字段 net-id 正好相对应。因此，若用默认子网掩码和某个不划分子网的 IP 地址逐位相“与”(AND)，就应当能够得出该 IP 地址的网络地址来。这样做可以不用查找该地址的类别位就能知道这是哪一类的 IP 地址。显然，

A 类地址的默认子网掩码是 255.0.0.0，或 0xFF000000。

B 类地址的默认子网掩码是 255.255.0.0，或 0xFFFF0000。

C 类地址的默认子网掩码是 255.255.255.0，或 0xFFFFFF00。

图 4-21 是这三类 IP 地址的网络地址和相应的默认子网掩码。

A 类 地 址	网络地址	网络号	主机号为全 0
	默认子网掩码 255.0.0.0	11111111	000000000000000000000000
B 类 地 址	网络地址	网络号	主机号为全 0
	默认子网掩码 255.255.0.0	1111111111111111	0000000000000000
C 类 地 址	网络地址	网络号	主机号为全 0
	默认子网掩码 255.255.255.0	111111111111111111111111	00000000

图 4-21 A 类、B 类和 C 类 IP 地址的默认子网掩码

子网掩码是一个网络或一个子网的重要属性。在 RFC 950 成为互联网的正式标准后，路由器在和相邻路由器交换路由信息时，必须把自己所在网络（或子网）的子网掩码告诉相邻路由器。在路由器的路由表中的每一个项目，除了要给出目的网络地址外，还必须同时给出该网络的子网掩码。若一个路由器连接在两个子网上就拥有两个网络地址和两个子网掩码。

我们以一个 B 类地址为例，说明可以有多少种子网划分的方法。在采用固定长度子网时，所划分的所有子网的子网掩码都是相同的（见表 4-6）。

表 4-6 B 类地址的子网划分选择（使用固定长度子网）

子网号的位数	子网掩码	子网数	每个子网的主机数
2	255.255.192.0	2	16382
3	255.255.224.0	6	8190
4	255.255.240.0	14	4094
5	255.255.248.0	30	2046
6	255.255.252.0	62	1022
7	255.255.254.0	126	510
8	255.255.255.0	254	254
9	255.255.255.128	510	126
10	255.255.255.192	1022	62

续表

子网号的位数	子网掩码	子网数	每个子网的主机数
11	255.255.255.224	2046	30
12	255.255.255.240	4094	14
13	255.255.255.248	8190	6
14	255.255.255.252	16382	2

在表 4-6 中，子网数是根据子网号(subnet-id)计算出来的。若 subnet-id 有 n 位，则共有 2^n 种可能的排列。除去全 0 和全 1 这两种情况，就得出表中的子网数。

表中的“子网号的位数”中没有 0, 1, 15 和 16 这四种情况，因为这没有意义。

请读者注意，虽然根据已成为互联网标准协议的 RFC 950 文档，子网号不能为全 1 或全 0，但随着无分类域间路由选择 CIDR 的广泛使用（在 4.3.3 节讨论），现在全 1 和全 0 的子网号也可以使用，但一定要谨慎使用，要弄清你的路由器所用的路由选择软件是否支持全 0 或全 1 的子网号这种较新的用法。

我们可以看出，若使用较少位数的子网号，则每一个子网上可连接的主机数就较多。反之，若使用较多位数的子网号，则子网的数目较多但每个子网上可连接的主机数就较少。因此我们可根据网络的具体情况（一共需要划分多少个子网，每个子网中最多有多少台主机）来选择合适的子网掩码。

通过简单的计算，读者不难得到这样的结论：划分子网增加了灵活性，但却减少了能够连接在网络上的主机总数。例如，本来一个 B 类地址最多可连接 65534 台主机，但表 4-6 中任意一行的最后两项的乘积一定小于 65534。

对 A 类和 C 类地址的子网划分也可得出类似的表格，读者可自行算出。

【例 4-2】 已知 IP 地址是 141.14.72.24，子网掩码是 255.255.192.0。试求网络地址。

【解】 子网掩码是 11111111 11111111 11000000 00000000。请注意，掩码的前两个字节都是全 1，因此网络地址的前两个字节可写为 141.14。子网掩码的第四字节是全 0，因此网络地址的第四字节是 0。可见本题仅需对地址中的第三字节进行计算。我们只要把 IP 地址和子网掩码的第三字节用二进制表示，就可以很容易地得出网络地址（图 4-22）。

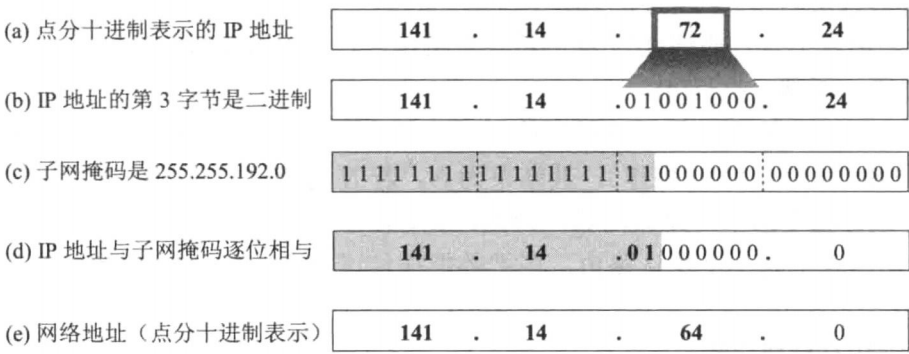


图 4-22 网络地址的计算

请注意，在一个 IP 地址中不允许把十进制和二进制混合使用。图 4-22 中(b)和(d)的写法，仅仅为了说明解题的步骤，而并非表示平时可以这样书写 IP 地址。

【例 4-3】 在上例中，若子网掩码改为 255.255.224.0。试求网络地址，并讨论所得结果。

【解】 用同样方法可得出网络地址是 141.14.64.0，和上例的结果相同（图 4-23）。

(a) 点分十进制表示的 IP 地址	141 . 14 . 72 . 24
(b) IP 地址的第 3 字节是二进制	141 . 14 . 01001000 . 24
(c) 子网掩码是 255.255.224.0	11111111 11111111 11100000 00000000
(d) IP 地址与子网掩码逐位相与	141 . 14 . 01000000 . 0
(e) 网络地址（点分十进制表示）	141 . 14 . 64 . 0

图 4-23 不同的子网掩码得出相同的网络地址

这个例子说明，同样的 IP 地址和不同的子网掩码可以得出相同的网络地址。但是，不同的掩码的效果是不同的。在例 4-2 中，子网号是 2 位，主机号是 14 位。在例 4-3 中，子网号是 3 位，主机号是 13 位。因此这两个例子中可划分的子网数和每一个子网中的最大主机数都是不一样的。

下面进一步讨论使用了子网掩码后应怎样查找路由表。

4.3.2 使用子网时分组的转发

在划分子网的情况下，分组转发的算法必须做相应的改动。

我们应当注意到，使用子网划分后，路由表必须包含以下三项内容：目的网络地址、子网掩码和下一跳地址。

在划分子网的情况下，路由器转发分组的算法如下：

(1) 从收到的数据报的首部提取目的 IP 地址 D 。

(2) 先判断是否为直接交付。对路由器直接相连的网络逐个进行检查：用各网络的子网掩码和 D 逐位相“与”（AND 操作），看结果是否和相应的网络地址匹配。若匹配，则把分组进行直接交付（当然还需要把 D 转换成物理地址，把数据报封装成帧发送出去），转发任务结束。否则就是间接交付，执行(3)。

(3) 若路由表中有目的地址为 D 的特定主机路由，则把数据报传送给路由表中所指明的下一跳路由器；否则，执行(4)。

(4) 对路由表中的每一行（目的网络地址，子网掩码，下一跳地址），用其中的子网掩码和 D 逐位相“与”（AND 操作），其结果为 N 。若 N 与该行的目的网络地址匹配，则把数据报传送给该行指明的下一跳路由器；否则，执行(5)。

(5) 若路由表中有一个默认路由，则把数据报传送给路由表中所指明的默认路由器；否则，执行(6)。

(6) 报告转发分组出错。

【例 4-4】图 4-24 有三个子网，两个路由器，以及路由器 R_1 中的部分路由表。现在源主机 H_1 向目的主机 H_2 发送分组。试讨论 R_1 收到 H_1 向 H_2 发送的分组后查找路由表的过程。

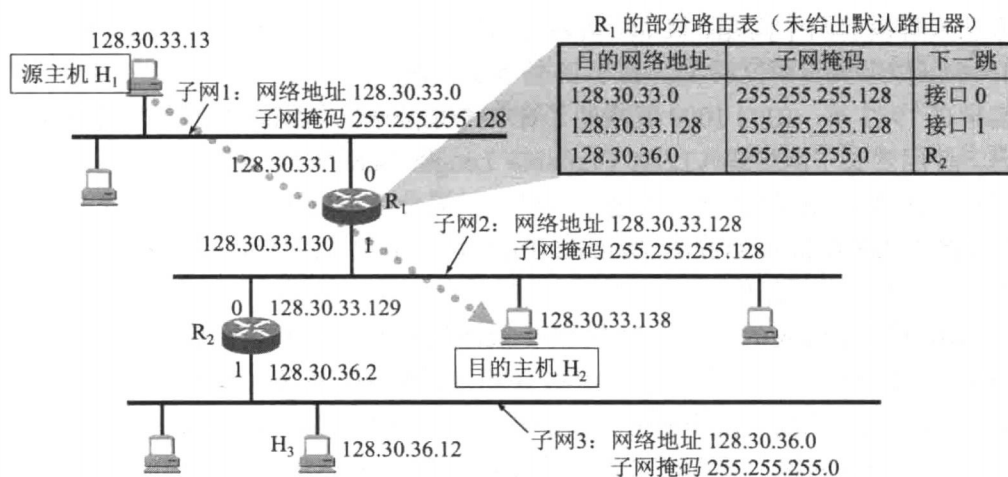


图 4-24 主机 H₁ 向 H₂ 发送分组

【解】源主机 H₁ 向目的主机 H₂ 发送的分组的地址是 H₂ 的 IP 地址 128.30.33.138。

源主机 H₁ 首先要进行的操作是要判断：发送的这个分组，是在本子网上进行直接交付还是要通过本子网上的路由器进行间接交付？

源主机 H₁ 把本子网的“子网掩码 255.255.255.128”与目的主机 H₂ 的“IP 地址 128.30.33.138”逐位相“与”（即逐位进行 AND 操作），得出 128.30.33.128，它不等于 H₁ 的网络地址（128.30.33.0）。这说明 H₂ 与 H₁ 不在同一个子网上。因此 H₁ 不能把分组直接交付 H₂，而必须交给子网上的默认路由器 R₁，由 R₁ 来转发。

路由器 R₁ 在收到一个分组后，就在其路由表中逐行寻找有无匹配的网络地址。

先看 R₁ 路由表中的第一行。用这一行的“子网掩码 255.255.255.128”和收到的分组的“目的地址 128.30.33.138”逐位相“与”（即逐位进行 AND 操作），得出 128.30.33.128。然后和这一行给出的目的网络地址 128.30.33.0 进行比较。但比较的结果不一致（即不匹配）。

用同样方法继续往下找第二行。用第二行的“子网掩码 255.255.255.128”和该分组的“目的地址 128.30.33.138”逐位相“与”（即逐位进行 AND 操作），结果也是 128.30.33.128。这个结果和第二行的目的网络地址 128.30.33.128 相匹配，说明这个网络（子网 2）就是收到的分组所要寻找的目的网络。于是不要再继续查找下去。R₁ 把分组从接口 1 直接交付主机 H₂（它们都在一个子网上）。

4.3.3 无分类编址 CIDR（构造超网）

1. 网络前缀

划分子网在一定程度上缓解了互联网在发展中遇到的困难。然而在 1992 年互联网仍然面临三个必须尽早解决的问题，这就是：

- (1) B 类地址在 1992 年已分配了近一半，眼看很快就将全部分配完毕！
- (2) 互联网主干网上的路由表中的项目数急剧增长（从几千个增长到几万个）。
- (3) 整个 IPv4 的地址空间最终将全部耗尽。在 2011 年 2 月 3 日，IANA 宣布 IPv4 地址已经耗尽了。

当时预计前两个问题将在 1994 年变得非常严重。因此 IETF 很快就研究出采用无分类

编址的方法来解决前两个问题。IETF 认为上面的第三个问题属于更加长远的问题，因此专门成立 IPv6 工作组负责研究解决新版本 IP 协议的问题。

其实早在 1987 年，RFC 1009 就指明了在一个划分子网的网络中可同时使用几个不同的子网掩码。使用变长子网掩码 VLSM (Variable Length Subnet Mask)可进一步提高 IP 地址资源的利用率。在 VLSM 的基础上又进一步研究出无分类编址方法，它的正式名字是无分类域间路由选择 CIDR (Classless Inter-Domain Routing, CIDR 的读音是“sider”)。

CIDR 最主要的特点有两个：

(1) CIDR 消除了传统的 A 类、B 类和 C 类地址以及划分子网的概念，因而能更加有效地分配 IPv4 的地址空间，并且在新的 IPv6 使用之前容许互联网的规模继续增长。CIDR 把 32 位的 IP 地址划分为前后两个部分。前面部分是“网络前缀”(network-prefix) (或简称为“前缀”)，用来指明网络，后面部分则用来指明主机。因此 CIDR 使 IP 地址从三级编址(使用子网掩码)又回到了两级编址，但这已是无分类的两级编址。其记法是：

$$\text{IP 地址} ::= \{ \langle \text{网络前缀} \rangle, \langle \text{主机号} \rangle \} \tag{4-3}$$

CIDR 还使用“斜线记法”(slash notation)，或称为 CIDR 记法，即在 IP 地址后面加上斜线“/”，然后写上网络前缀所占的位数。

(2) CIDR 把网络前缀都相同的连续的 IP 地址组成一个“CIDR 地址块”。我们只要知道 CIDR 地址块中的任何一个地址，就可以知道这个地址块的起始地址(即最小地址)和最大地址，以及地址块中的地址数。例如，已知 IP 地址 128.14.35.7/20 是某 CIDR 地址块中的一个地址，现在把它写成二进制表示，其中的前 20 位是网络前缀(用粗体和下划线表示出)，而前缀后面的 12 位是主机号：

$$128.14.35.7/20 = \underline{10000000 \ 00001110 \ 00100011 \ 00000111}$$

这个地址所在的地址块中的最小地址和最大地址可以很方便地得出：

最小地址	128.14.32.0	<u>10000000 00001110 00100000 00000000</u>
最大地址	128.14.47.255	<u>10000000 00001110 0010</u> 1111 11111111

当然，以上这两个特殊地址的主机号是全 0 和全 1 的地址。一般并不使用。通常只使用在这两个特殊地址之间的地址。不难看出，这个地址块共有 2^{12} 个地址。我们可以用地址块中的最小地址和网络前缀的位数指明这个地址块。例如，上面的地址块可记为 128.14.32.0/20。在不需要指出地址块的起始地址时，也可把这样的地址块简称为“/20 地址块”。

为了更方便地进行路由选择，CIDR 使用 32 位的地址掩码(address mask)。地址掩码由一串 1 和一串 0 组成，而 1 的个数就是网络前缀的长度。虽然 CIDR 不使用子网了，但由于目前仍有一些网络还使用子网划分和子网掩码，因此 CIDR 使用的地址掩码也可继续称为子网掩码。例如，/20 地址块的地址掩码是：11111111 11111111 11110000 00000000 (20 个连续的 1)。斜线记法中，斜线后面的数字就是地址掩码中 1 的个数。

请读者注意，“CIDR 不使用子网”是指 CIDR 并没有在 32 位地址中指明若干位作为子网字段。但分配到一个 CIDR 地址块的单位，仍然可以在本单位内根据需要划分出一些子网。这些子网也都只有一个网络前缀和一台主机号字段，但子网的网络前缀比整个单位的网

络前缀要长些。例如，某单位分配到地址块/20，就可以再继续划分为 8 个子网（即需要从主机号中借用 3 位来划分子网）。这时每一个子网的网络前缀就变成 23 位（原来的 20 位加上从主机号借来的 3 位），比该单位的网络前缀多了 3 位。

斜线记法还有一个好处就是它除了表示一个 IP 地址外，还提供了其他一些重要信息。我们举例说明如下。

例如，地址 192.199.170.82/27 不仅表示 IP 地址是 192.199.170.82，而且还表示这个地址块的网的前缀有 27 位（剩下的 5 位是主机号），因此这个地址块包含 32 个 IP 地址（ $2^5 = 32$ ）。通过简单的计算还可得出，这个地址块的最小地址是 192.199.170.64，最大地址是 192.199.170.95。具体的计算方法是这样的。找出地址掩码中 1 和 0 的交界处发生在地址中的哪一个字节。现在是在第四字节。因此只要把这**一个字节**的十进制 82 用二进制表示即可。十进制 82 的二进制是 01010010，取其前 3 位（这 3 位加上前 3 个字节的 24 位等于前缀的 27 位），再把后面 5 位都写成 0，即 01000000，等于十进制的 64。这就找出了地址块的最小地址 192.199.170.64。再把地址的第四字节的最后 5 位都置 1，即 01011111，等于十进制的 95，这就找出了地址块中的最大地址 192.199.170.95。

由于一个 CIDR 地址块中有很多地址，所以在路由表中就利用 CIDR 地址块来查找目的网络。这种地址的聚合常称为**路由聚合(route aggregation)**，它使得路由表中的一个项目可以表示原来传统分类地址的很多个（例如上千个）路由。路由聚合也称为**构成超网(supernetting)**。如果没有采用 CIDR，则在 1994 年和 1995 年，互联网的一个路由表就会超过 7 万个项目，而使用了 CIDR 后，在 1996 年一个路由表的项目数才只有 3 万多个。路由聚合有利于减少路由器之间的路由选择信息的交换，从而提高了整个互联网的性能。

CIDR 记法有多种形式，例如，地址块 10.0.0.0/10 可简写为 10/10，也就是把点分十进制中低位连续的 0 省略。另一种简化表示方法是在网络前缀的后面加一个星号*，如：

00001010 00*

意思是：在星号 * 之前是网络前缀，而星号 * 表示 IP 地址中的主机号，可以是任意值。

前缀位数不是 8 的整数倍时，需要进行简单的计算才能得到一些地址信息。

表 4-7 给出了最常用的 CIDR 地址块。表中的 K 表示 2^{10} 即 1024。网络前缀小于 13 或大于 27 都较少使用。在“包含的地址数”中没有把全 1 和全 0 的主机号除外。

表 4-7 常用的 CIDR 地址块

CIDR 前缀长度	点分十进制	包含的地址数	相当于包含分类的网络数
/13	255.248.0.0	512 K	8 个 B 类或 2048 个 C 类
/14	255.252.0.0	256 K	4 个 B 类或 1024 个 C 类
/15	255.254.0.0	128 K	2 个 B 类或 512 个 C 类
/16	255.255.0.0	64 K	1 个 B 类或 256 个 C 类
/17	255.255.128.0	32 K	128 个 C 类
/18	255.255.192.0	16 K	64 个 C 类
/19	255.255.224.0	8 K	32 个 C 类
/20	255.255.240.0	4 K	16 个 C 类

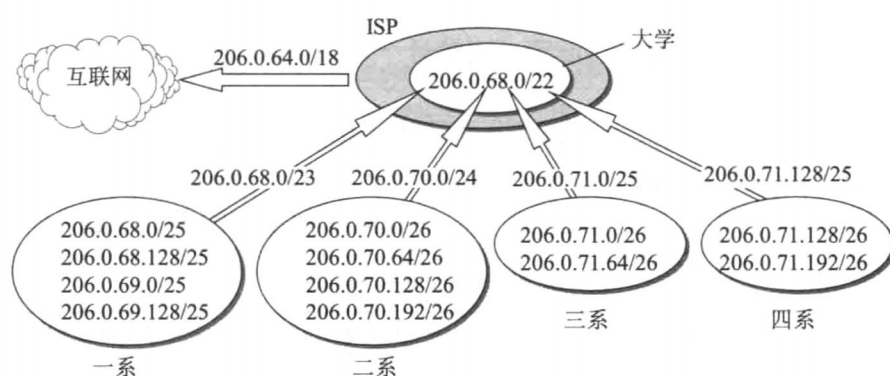
续表

CIDR 前缀长度	点分十进制	包含的地址数	相当于包含分类的网络数
/21	255.255.248.0	2 K	8 个 C 类
/22	255.255.252.0	1 K	4 个 C 类
/23	255.255.254.0	512	2 个 C 类
/24	255.255.255.0	256	1 个 C 类
/25	255.255.255.128	128	1/2 个 C 类
/26	255.255.255.192	64	1/4 个 C 类
/27	255.255.255.224	32	1/8 个 C 类

从表 4-7 可看出，每一个 CIDR 地址块中的地址数一定是 2 的整数次幂。除最后几行外，CIDR 地址块都包含了多个 C 类地址（是一个 C 类地址的 2^n 倍， n 是整数），这就是“构成超网”这一名词的来源。

使用 CIDR 的一个好处就是可以更加有效地分配 IPv4 的地址空间，可根据客户的需要分配适当大小的 CIDR 地址块。然而在分类地址的环境中，向一个部门分配 IP 地址，就只能以 /8、/16 或 /24 为单位来分配。这就很不灵活。

图 4-25 给出的是 CIDR 地址块分配的例子。假定某 ISP 已拥有地址块 206.0.64.0/18（相当于有 64 个 C 类网络）。现在某大学需要 800 个 IP 地址。ISP 可以给该大学分配一个地址块 206.0.68.0/22，它包括 1024（即 2^{10} ）个 IP 地址，相当于 4 个连续的 C 类 /24 地址块，占该 ISP 拥有的地址空间的 1/16。这个大学然后可自由地对本校的各系分配地址块，而各系还可再划分本系的地址块。CIDR 的地址块分配有时不易看清，这是因为网络前缀和主机号的界限不是恰好出现在整数字节处。只要写出地址的二进制表示（从图中的地址块的二进制表示中可看出，实际上只需要将其中的一个关键字节转换为二进制的表示即可），弄清网络前缀的位数，就不会把地址块的范围弄错。



单位	地址块	二进制表示	地址数
ISP	206.0.64.0/18	11001110.00000000.01*	16384
大学	206.0.68.0/22	11001110.00000000.010001*	1024
一系	206.0.68.0/23	11001110.00000000.0100010*	512
二系	206.0.70.0/24	11001110.00000000.01000110.*	256
三系	206.0.71.0/25	11001110.00000000.01000111.0*	128
四系	206.0.71.128/25	11001110.00000000.01000111.1*	128

图 4-25 CIDR 地址块划分举例

从图 4-25 可以清楚地看出**地址聚合**的概念。这个 ISP 共拥有 64 个 C 类网络。如果不采用 CIDR 技术, 则在与该 ISP 的路由器交换路由信息的每一个路由器的路由表中, 就需要有 64 个项目。但采用地址聚合后, 就只需用路由聚合后的一个项目 206.0.64.0/18 就能找到该 ISP。同理, 这个大学共有 4 个系。在 ISP 内的路由器的路由表中, 也需使用 206.0.68.0/22 这个项目。这个项目好比是大学的收发室。凡寄给这个大学任何一个系的邮件, 邮递员都不考虑大学各个系的地址, 而是把这些邮件集中投递到大学的收发室, 然后由大学的收发室再进行下一步的投递。这样就减轻了邮递员的工作量 (相当于简化了路由表的查找)。

从图 4-25 下面表格中的二进制地址可看出, 把四个系的路由聚合为大学的一个路由 (即构成超网), 是将网络前缀**缩短**。**网络前缀越短, 其地址块所包含的地址数就越多**。而在三级结构的 IP 地址中, 划分子网是使网络前缀变长。

2. 最长前缀匹配

在使用 CIDR 时, 由于采用了网络前缀这种记法, IP 地址由网络前缀和主机号这两个部分组成, 因此在路由表中的项目也要有相应的改变。这时, 每个项目由“**网络前缀**”和“**下一跳地址**”组成。但是在查找路由表时可能会得到不止一个**匹配结果**。这样就带来一个问题: 我们应当从这些匹配结果中选择哪一条路由呢?

正确的答案是: 应当从匹配结果中选择具有**最长网络前缀**的路由。这叫做**最长前缀匹配**(longest-prefix matching), 这是因为网络前缀越长, 其地址块就越小, 因而路由就越具体 (more specific)。最长前缀匹配又称为**最长匹配**或**最佳匹配**。为了说明最长前缀匹配的概念, 我们仍以前面的例子来讨论。

假定大学下属的四系希望 ISP 把转发给四系的数据报直接发到四系而不要经过大学的路由器, 但又不愿意改变自己使用的 IP 地址块。因此, 在 ISP 的路由器的路由表中, 至少要有以下两个项目, 即 206.0.68.0/22 (大学) 和 206.0.71.128/25 (四系)。现在假定 ISP 收到一个数据报, 其目的 IP 地址为 $D = 206.0.71.130$ 。把 D 分别和路由表中这两个项目的掩码逐位相“与”(AND 操作)。将所得的逐位 AND 操作的结果按顺序写在下面。

D 和 11111111 11111111 11111100 00000000 逐位相“与” = 206.0.68.0/22 匹配

D 和 11111111 11111111 11111111 10000000 逐位相“与” = 206.0.71.128/25 匹配

不难看出, 现在同一个 IP 地址 D 可以在路由表中找到两个目的网络 (大学和四系) 和该地址相匹配。根据最长前缀匹配的原理, 应当选择后者, 把收到的数据报转发到后一个目的网络 (四系), 即选择两个匹配的地址中更具体的一个。

从以上的讨论可以看出, 如果 IP 地址的分配一开始就采用 CIDR, 那么我们可以按网络所在的地理位置来分配地址块, 这样就可大大减少路由表中的路由项目。例如, 可以将世界划分为四大地区, 每一地区分配一个 CIDR 地址块:

地址块 194/7 (194.0.0.0 至 195.255.255.255) 分配给欧洲;

地址块 198/7 (198.0.0.0 至 199.255.255.255) 分配给北美洲;

地址块 200/7 (200.0.0.0 至 201.255.255.255) 分配给中美洲和南美洲;

地址块 202/7 (202.0.0.0 至 203.255.255.255) 分配给亚洲和太平洋地区。

上面的每一个地址块包含约有 3200 万个地址。这种分配地址的方法就使得 IP 地址与地理位置相关联。它的好处是可以大大压缩路由表中的项目数。例如, 凡是从中国发往北美的

IP 数据报（不管它是地址块 198/7 中的哪一个地址）都先送交位于美国的一个路由器，因此在路由表中使用一个项目就行了。

但是，在使用 CIDR 之前，互联网的地址管理机构没有按地理位置来分配 IP 地址。现在要把已分配出的 IP 地址收回再重新分配是件十分困难的事，因为这牵涉到很多正在工作的主机必须改变其 IP 地址。尽管这样，CIDR 的使用已经推迟了 IP 地址耗尽的日期。

3. 使用二叉线索查找路由表

使用 CIDR 后，由于要寻找最长前缀匹配，使路由表的查找过程变得更加复杂了。当路由表的项目数很大时，怎样设法减小路由表的查找时间就成为一个非常重要的问题。例如，连接路由器的线路的速率为 10 Gbit/s，而分组的平均长度为 2000 bit，那么路由器就应当平均每秒钟能够处理 500 万个分组（常记为 5 Mpps）。或者说，路由器处理一个分组的平均时间只有 200 ns（1 ns = 10^{-9} 秒）。因此，查找每一个路由所需的时间是非常短的。可见在路由表中必须使用很好的数据结构和使用先进的快速查找算法，这一直是人们积极研究的热门课题。

对无分类编址的路由表的最简单的查找算法就是对所有可能的前缀进行循环查找。例如，给定一个目的地址 D 。对每一个可能的网络前缀长度 M ，路由器从 D 中提取前 M 个位成一个网络前缀，然后查找路由表中的网络前缀。所找到的最长匹配就对应于要查找的路由。

这种最简单的算法的明显缺点就是查找的次数太多。最坏的情况是路由表中没有这个路由。在这种情况下，算法仍要进行 32 次（具有 32 位的网络前缀是一个特定主机路由）。就是要找到一个传统的 B 类地址（即/16），也要查找 16 次。对于经常使用的默认路由，这种算法都要经历 31 次不必要的查找。

为了进行更加有效的查找，通常是把无分类编址的路由表存放在一种层次的数据结构中，然后自上而下地按层次进行查找。这里最常用的就是**二叉线索(binary trie)**^①，它是一种特殊结构的树。IP 地址中从左到右的比特值决定了从根节点逐层向下层延伸的路径，而二叉线索中的各个路径就代表路由表中存放的各个地址。

图 4-26 用一个例子来说明二叉线索的结构。图中给出了 5 个 IP 地址。为了简化二叉线索的结构，可以先找出对应于每一个 IP 地址的**唯一前缀(unique prefix)**。所谓唯一前缀就是在表中所有的 IP 地址中，该前缀是唯一的。这样就可以用这些唯一前缀来构造二叉线索。在进行查找时，只要能够和唯一前缀相匹配就行了。

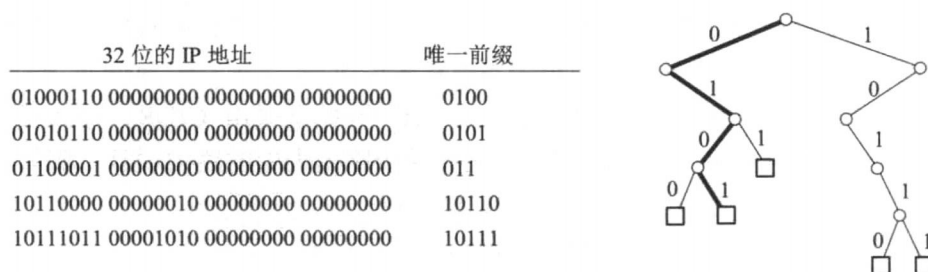


图 4-26 用 5 个前缀构成的二叉线索

① 注：线索(trie)来自 retrieval（检索），读音与“try”相同。

从二叉线索的根节点自顶向下的深度最多有 32 层，每一层对应于 IP 地址中的一位。一个 IP 地址存入二叉线索的规则很简单。先检查 IP 地址左边的第一位，如为 0，则第一层的节点就在根节点的左下方；如为 1，则在右下方。然后再检查地址的第二位，构造出第二层的节点。依此类推，直到唯一前缀的最后一位。由于唯一前缀一般都小于 32 位，因此用唯一前缀构造的二叉线索的深度往往不到 32 层。图中较粗的折线就是前缀 0101 在这个二叉线索中的路径。二叉线索中的小圆圈是中间节点，而在路径终点的小方框是叶节点（也叫做外部节点）。每个叶节点代表一个唯一前缀。节点之间的连线旁边的数字表示这条边在唯一前缀中对应的比特是 0 或 1。

假定有一个 IP 地址是 10011011 01111010 00000000 00000000，需要查找该地址是否在此二叉线索中。我们从最左边查起。很容易发现，查到第三个字符（即前缀 10 后面的 0）时，在二叉线索中就找不到匹配的，说明这个地址不在这个二叉线索中。

以上只是给出了二叉线索这种数据结构的用法，而并没有说明“与唯一前缀匹配”和“与网络前缀匹配”的关系。显然，要将二叉线索用于路由表中，还必须使二叉线索中的每一个叶节点包含所对应的网络前缀和子网掩码。当搜索到一个叶节点时，就必须将寻找匹配的目的地址和该叶节点的子网掩码进行逐位“与”运算，看结果是否与对应的网络前缀相匹配。若匹配，就按下一跳的接口转发该分组。否则，就丢弃该分组。

总之，二叉线索只是提供了一种可以快速在路由表中找到匹配的叶节点的机制。但这是否和网络前缀匹配，还要和子网掩码进行一次逻辑与的运算。

为了提高二叉线索的查找速度，广泛使用了各种压缩技术。例如，在图 4-26 中的最后两个地址，其最前面的 4 位都是 1011。因此，只要一个地址的前 4 位是 1011，就可以跳过前面 4 位（即压缩了 4 个层次）而直接从第 5 位开始比较。这样就可以减少查找的时间。当然，制作经过压缩的二叉线索需要更多的计算，但由于每一次查找路由表时都可以提高查找速度，因此这样做还是值得的。

4.4 网际控制报文协议 ICMP

为了更有效地转发 IP 数据报和提高交付成功的机会，在网际层使用了网际控制报文协议 ICMP (Internet Control Message Protocol) [RFC 792]。ICMP 允许主机或路由器报告差错情况和提供有关异常情况的报告。ICMP 是互联网的标准协议。但 ICMP 不是高层协议（看起来好像是高层协议，因为 ICMP 报文是装在 IP 数据报中，作为其中的数据部分），而是 IP 层的协议。ICMP 报文作为 IP 层数据报的数据，加上数据报的首部，组成 IP 数据报发送出去。ICMP 报文格式如图 4-27 所示。

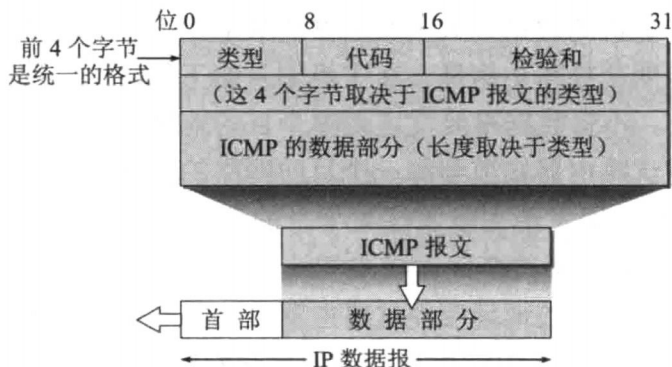


图 4-27 ICMP 报文的格式

计算机网络

(第7版)

与本教材配套的辅导书:

《计算机网络释疑与习题解答》

作者: 谢希仁



策划编辑: 郝志恒
责任编辑: 牛晓丽
责任美编: 秦 靖

ISBN 978-7-121-30295-4



9 787121 302954 >

定价: 45.00元