

Cisco专业技术丛书

HZ BOOKS

网络技术



CISCO

组播路由

与交换技术

Cisco Multicast Routing And Switching

(美) William R. Parkhurst (CCIE) 著

京京工作室 译



机械工业出版社
China Machine Press



McGraw-Hill

Cisco 专业技术丛书

Cisco 组播路由 与交换技术

(美) William R. Parkhurst (CCIE) 著

京京工作室 译



机械工业出版社
China Machine Press

2001 9

本书全面介绍了单播、广播和组播IP定址技术，重点讲解了IP组播路由协议，包括互联网组管理协议(IGMP)、Cisco组管理协议(CGMP)、与协议无关的组播协议(PIM)及资源预留协议(RSVP)。通过本书的学习，您将在Cisco路由器和交换机环境中，成功配置IP组播。

本书适合于网络工程人员、大学生及各种互联网资格认证(考试)人员，且可作为Cisco环境中配置IP组播的指南。

William R. Parkhurst: Cisco Multicast Routing and Switching.

Authorized translation from the English language edition published by The McGraw-Hill Companies, Inc.

Copyright©1999 by The McGraw-Hill Companies, Inc.

All rights reserved. For sale in Mainland China Only.

本书中文简体字版由机械工业出版社出版，未经出版者书面许可，本书的任何部分不得以任何方式复制或抄袭。

版权所有，翻印必究。

本书版权登记号：图字：01-1999-2370

图书在版编目(CIP)数据

Cisco组播路由与交换技术/(美)帕克赫斯特(Parkhurst, W. R.)著；京京工作室译。-北京：机械工业出版社，1999.11

(Cisco专业技术丛书)

书名原文：Cisco Multicast Routing and Switching

ISBN 7-111-07519-6

I. C… II. ①帕… ②京… III. ①多信道-路由选择-技术②多信道-信息交换-技术 IV. TN913.1

中国版本图书馆CIP数据核字(1999)第61274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：郭东青

北京忠信诚胶印厂印刷·新华书店北京发行所发行

1999年11月第1版·1999年12月第2次印刷

787mm×1092mm 1/16·15.5印张

印数 5 001 - 8 000册

定价：26.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

目 录

第1章 IP组播入门	1	5.5 DVMRP的工作原理	63
1.1 单播IP通信模型	3	5.5.1 DVMRP邻居查找	64
1.2 广播通信模型	4	5.5.2 DVMRP路由交换	66
1.3 组播通信模型	5	5.5.3 基于资源的组播树	70
1.4 本书提要	6	5.5.4 DVMRP修剪和嫁接	72
1.5 IP路由协议推荐读物	7	5.5.5 跟踪与故障诊断	74
第2章 IP协议地址	8	5.5.6 DVMRP通道和互联网组 播骨干网	75
2.1 IP地址格式	8	5.5.7 DVMRP路由器命令	77
2.2 IP子网	11	第6章 与协议无关的组播——密集 模式	81
2.2.1 子网示例	14	6.1 PIM-DM版本1: 协议工作原理	82
2.2.2 IP地址设计示例1	17	6.2 查找邻居	83
2.2.3 变长子网掩码	18	6.3 PIM-DM包转发	85
第3章 互联网组管理协议	24	6.4 PIM-DM接口修剪	87
3.1 RFC 1112、IP组播的主机扩展	25	6.5 PIM-DM接口嫁接	88
3.1.1 以太网组播定址	25	6.6 PIM-DM声明消息	91
3.1.2 令牌环组播定址	27	6.7 PIM-DM版本2	93
3.2 互联网组管理协议, IGMP版本1	28	6.8 PIM-DM路由器配置	97
3.3 互联网组管理协议, IGMP版本2	31	第7章 与协议无关的组播——分散 模式	101
3.4 IGMP版本2: 计时器和计数器	32	7.1 PIM-SM: 协议工作原理与邻 居的查找	102
3.5 IGMP路由器状态	35	7.1.1 PIM-SM包转发	104
3.5.1 配置IGMP	38	7.1.2 PIM-SM接口的修剪	107
3.5.2 IGMP的显示与调试命令	40	7.1.3 PIM-SM声明消息	109
3.5.3 与IGMP连接的组成员	41	7.1.4 PIM-SM版本2	111
第4章 Cisco组管理协议	43	7.1.5 集合点——在什么地方?	115
4.1 CGMP的监视	49	7.1.6 SPT替换	119
4.2 CGMP命令总结	52	7.2 PIM-SM路由器配置命令	119
第5章 距离向量组播路由协议	54	7.2.1 集合点配置与静态RP配置	120
5.1 单播与组播路由	54	7.2.2 自动RP配置	120
5.2 反向转发	54	7.3 PIM-SM版本2的RP选择	121
5.3 DVMRP和RIP	56	第8章 PIM-DVMRP网络	131
5.4 路由信息协议	57		
5.4.1 计数无限问题	60		
5.4.2 RIP和VLSM	61		
5.4.3 RIP版本2	62		

8.1 路由交换	133	10.1.1 预约样式	155
8.2 路由选择	136	10.1.2 通配过滤器样式	156
8.3 DVMRP配置命令	139	10.1.3 固定过滤器样式	157
第9章 组播支持命令	142	10.1.4 共享显式样式	158
9.1 组播边界	142	10.2 预约样式总结	159
9.2 广播/组播转换	144	10.2.1 RSVP协议消息	159
9.3 会话目录	145	10.2.2 RSVP消息格式	161
9.4 IP组播速率限制	146	10.3 RSVP的配置与监视	168
9.5 根组播路由	147	10.3.1 RSVP配置命令	168
9.6 负载均衡	148	10.3.2 示例	168
9.7 组播静态路由	150	10.4 RSVP方案	171
9.8 组播和非广播多访问网络	151	附录A Cisco组播命令索引	188
9.9 通过ATM实现组播	152	附录B 已分配的组播地址	206
第10章 资源预约协议	154	附录C 参考资源	212
10.1 RSVP预约模型	155		

第1章 IP组播入门

在我们开始探索IP组播及相关协议之前，首先要了解对一个内联网(intranet)或者因特网(Internet)来说，在两个或多个主机之间，存在着哪些通信模型。现在市面上任何一本与联网有关的书籍，都无一例外地提供了对OSI分层通信模型的一个简介，如图1-1所示。在OSI分层模型中，活跃于各个级别的通信协议协作得相当好，这一切当属层次化协议模型的功劳。OSI建立的原始模型对网络的不同功能进行了一个逻辑性划分。只要遵守公共的标准，这个模型就能使来自不同厂商的软件模块很好地沟通，它们能够和平共处，并完全正常地发挥出自己的功用。

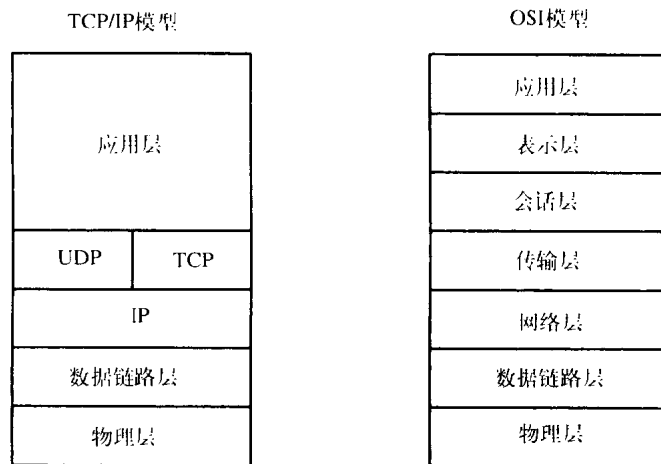


图1-1 TCP/IP和OSI分层网络模型

在OSI模型中，最低的一层是物理层。物理层负责控制的是某种通信媒体及其相关接口的电子与机械规格。例如，10和100兆位以太网、同步和异步串行链路以及ATM均属“物理层”的典型例子。物理层关心的只是怎样以电子或光学形式，将二进制位(比特)从A处传到B处。对于自己收发的数据，它并不关心它们的结构或格式，只在乎如何将无数的1和0从起点弄到终点。

在物理层之上，OSI模型的下一级是数据链路层。这一层负责创建数据帧(或数据报)，其中除包含起点与终点地址外，还增加了与侦错甚至纠错有关的字段。当然，帧内还集成了用户需要传送的正式数据。数据链路层的协议是不可进行路由(选择)的，以太网(Ethernet)和令牌环(Token Ring)便是其中两个典型的例子。

网络设计人员大量的时间都花在网络层上。这一层控制着穿越因特网的路由；讨论组播技术时，这无疑是最重要的一个层。对一种准备路由的协议来说，在其定址方案中，必须包括一个网络地址和一个主机地址。这一点对“普通”的IP通信来说是完全成立的，但对组播通信则不然。正如我们将会学到的那样，组播地址并不采用“网络/主机”形式，而是代表着一个组地址。尽管“网络/主机”地址未在组播地址中出现，但组播通信仍然是可以路由的。

目前有多种可路由的协议，比如IP，IPX，AppleTalk和DECnet等等。

在各个上层应用进程之间，传送层负责着对数据流的多路复用及多路分解，如图1-2所示。我们注意到OSI模型的三个上层，包括应用层、表示层与会话层，被合并到了TCP/IP分层模型的“应用层”。通常，我们很难判断一个特定的上层应用是否应当逻辑性地置入。但是，在不必知晓用户欲采用什么应用的前提下，我们完全能够顺利地设计好网络。因此，具体什么应用并不重要，重要的只是应用打算采用的协议。事实上，无论OSI还是TCP/IP模型，我们只需关心最下面的四层便足够了。

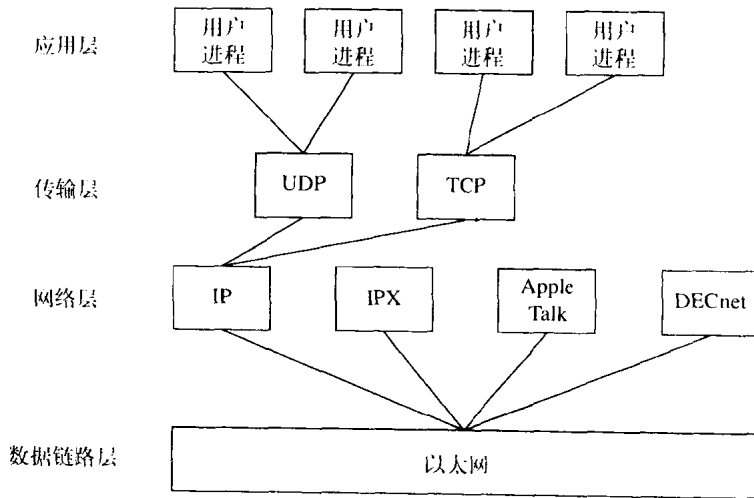


图1-2 TCP/IP模型中的多路复用和多路分解

一个网络应用(如Telnet)希望发出数据时，数据会首先发给传输层的TCP模块。随后，TCP为本地和远程Telnet会话分配一个编号，使TCP能够判断出数据要投递到哪个会话。IP要么接收数据，要么将数据投递给UDP或者TCP模块——具体由应用的类型决定。

最后，一个Ethernet帧包含了一个标识符，它标识着从什么网络层协议收到协议，或者应将数据传给哪个网络层协议。

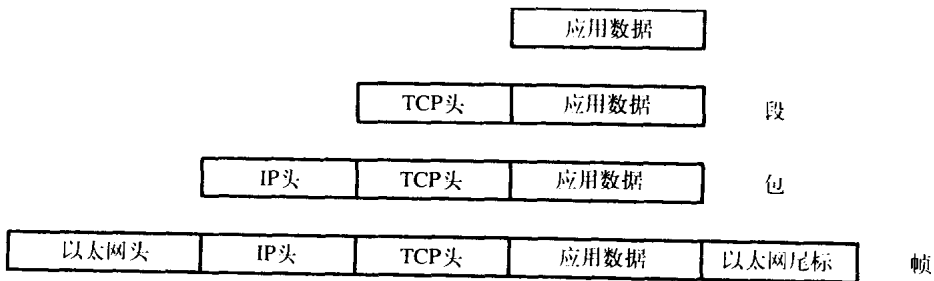


图1-3 数据封装

为具体阐明OSI模型中各个层之间如何实现沟通，我们将跟随数据从一个主机到达另一个主机(参见图1-3)。假定我们在两个主机之间运行一个Telnet会话。用户数据产生于应用层，然后经过协议堆栈，向下投递给传送层的TCP模块。TCP层检查与此次会话对应的一个标识符(它应包含于TCP头内)，并将TCP段传递给网络层的IP模块。随后，IP将这个包标记成TCP或

者UDP包。数据链路层收到这个包时，会构建一个Ethernet帧，在其中加入Ethernet帧头与帧尾。在帧头内，至少应包含一个字段，指出该帧负载的是IP数据，而非其他。最后，这个帧被传至物理层，准备送入网络通信媒体。

远程主机收到这个Ethernet帧后，数据链路层模块通过检查，知道它负载的是IP数据，所以剔除其中的Ethernet帧头和帧尾。随后，将数据传给网络层的IP模块。IP判断它是一个TCP包，还是一个UDP包，并分别传至传输层对应的模块。最后，TCP提取出用户数据，并将其发给合适的用户进程。

1.1 单播IP通信模型

无论内联网还是因特网，网络中不同主机的通信都可选用三种模型。第一种便是“单播”(Unicast)模型，它属于“一对一”通信。在图1-4a到1-4c中，一个主机希望将数据发给同一IP子网内的另一个主机(IP定址和子网的概念将在第2章详述)。对于Ethernet LAN(局域网)来说，主机必须同两种不同的地址方案打交道。第一种方案是“Ethernet地址”，已烧录并固化到网卡(NIC)。Ethernet地址是一种6字节(48位)链路层地址，肯定是独一无二的，不可更改。由于Ethernet地址已固化到网卡，所以只要更换网卡，它的Ethernet地址也会相应发生改变。通过前面的学习，我们知道在一个Ethernet LAN中，网上传输的所有数据都要封装到“帧”(Frame)里。即使主机将数据发给一个IP地址，IP包也必须先封装到Ethernet帧内。为完成这个封装，作为发送方的主机必须将接收方主机的IP地址解析成Ethernet地址，即完成地址的“映射”。这种地址映射是通过ARP(地址解析协议)完成的。

在图1-4a到1-4c中，主机A希望将一个包发给主机B。主机A已经知道主机B的IP地址，但不知道它的Ethernet地址。此时便需ARP的介入，它将完成以下操作：

- 1) 主机A发出一个ARP广播包(参见图1-4a)，网络上的所有主机都能收到，包括路由器。
- 2) 主机B收到这个ARP包，发现ARP内包含的IP地址正是自己的地址。因此，主机B发出一个ARP回应包，在其中加入自己的Ethernet地址(参见图1-4b)。
- 3) 现在，主机A可将IP包封装到一个Ethernet帧内，并将帧传给主机B(参见图1-4c)。
 - a. 主机A向IP地址172.16.1.2发送一个ARP请求。
 - b. 主机B用自己的Ethernet地址响应这个请求。
 - c. 主机A与主机B建立正常通信。

如果主机A想把一个包传给另一个IP子网的主机，就必须将包发给路由器。主机A会配置一个默认网关，指向用于连接主机A所在LAN的那个路由器接口。由于目标IP地址位于一个不同的子网上，所以主机A能自动分辨应将帧发给路由器，并向路由器的Ethernet地址发出一个ARP。路由器收到这个帧后，会提取出IP包，并根据其中包含的目标IP地址，判断目标主机是否在一个同自己直接连接的网络内。若答案是肯定的，路由器就向那个网络送入一个ARP，解析出目标主机的Ethernet地址。从目标主机收到ARP回应后，路由器就可构建起一个Ethernet帧，将IP包包括其中，然后将整个帧发给目标主机。若答案是否定的，目标主机不在同自己直连的一个网络内，路由器就必须查询自己的路由表，决定下一步应将帧发给哪个路由器。IP单播路由协议并非本书重点，但书末仍然列出了相关的参考资源，便于您进一步学习。

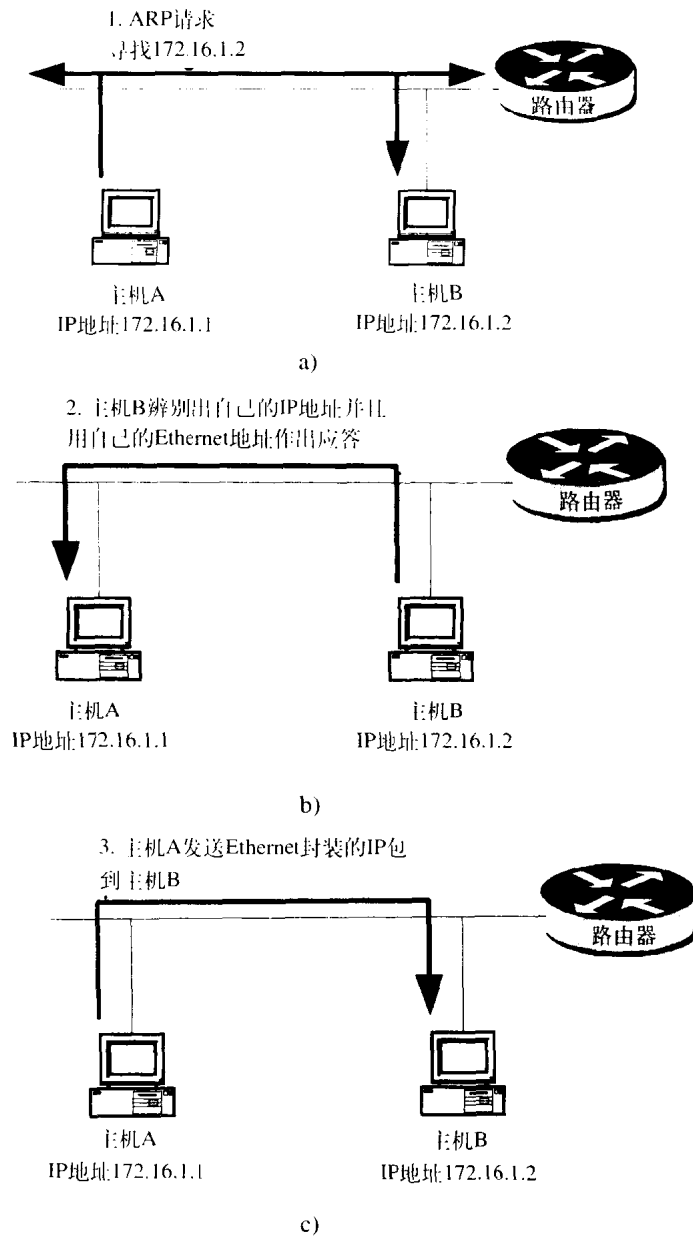


图1-4 将IP解析成Ethernet地址

- a) 将IP解析成Ethernet地址, 第一步
- b) 将IP解析成Ethernet地址, 第二步
- c) 将IP解析成Ethernet地址, 第三步

1.2 广播通信模型

在广播模型中, 一个主机会把数据发给子网内的所有主机。ARP在此不再需要, 因为Ethernet广播地址成为一个众所周知的地址, 即0xFF FF FF FF FF FF(也存在广播IP地址, 第2章将详述)。在单播模型中, 一个主机可将IP包发给任何网络内的任何主机(假定我们有通向目标主机的路由)。而在广播模型中, 广播的范围只能是本地的这个子网。路由器会封锁广播通

信，所以广播范围仅限于本地子网，如图1-5所示。

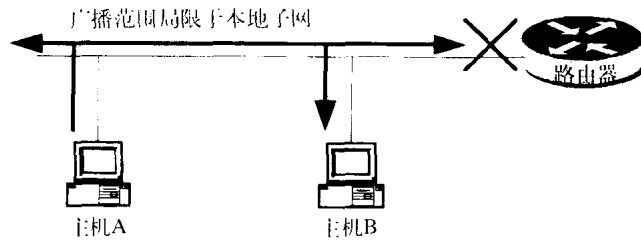


图1-5 广播通信模型

1.3 组播通信模型

现在进入本书的正题，也是有趣的一部分。在这儿，我们要解决的是“一对多”的通信问题。如果主机想把相同的包发给多个接收者，但又不是网络内的全部主机，该如何实现呢？我们可试着使用单播通信模型，而且可能取得成功，但同时也会遇到一定的问题。假定在单播模型中，主机A希望将一个包发给五个主机。这暗示着主机A知道每个接收者的IP地址。此时，主机A需将相同的包发给五个不同的IP地址，如图1-6所示。

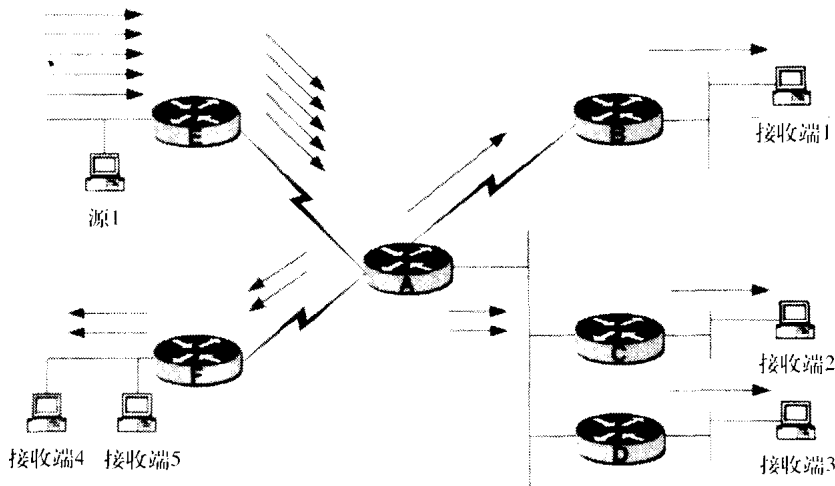


图1-6 用单播通信模型实现组播功能

随着接收者数量的增多，需要发出的包数也会线性增加。换言之，对于 n 个接收者，便要发出同一个包的 n 份拷贝。试想主机发送的是一段实时声音或影像剪辑，那么在接收者数量较少的情况下，尽管仍然可以接受，但随着接收者数量的增多，发送主机必然不堪重负，各个包之间的延迟也会令人不堪忍受。另外，源路由器(图1-6的路由器E)上的链路也会严重消耗宝贵的带宽。

这一方案另一个主要的问题是主机并不知道接收者位于何处。如果需要通信的接收者不发生改变，数据就可正常进入。但这也造成了很大的限制，因为新的接收者不可动态加入或者离开当前的组。

那么广播模型的情况又如何呢？凡是在本地子网上的主机，都会接收到数据，而且每个包只需发送一次。那么问题在哪里呢？最主要的问题有两个。第一个是只有同一个子网上的

接收者才能收到数据，其他子网上的接收者不能收到，因为路由器会阻止广播通信。如果我们不希望每次广播都传遍整个世界，这样设置显然是有利的。当然，有些人的想法会与此相反，但在正常情况下，都没必要广播到整个世界。

广播模型的第二个问题是每个主机都需要处理Ethernet广播，判断数据是不是发给自己的。IP包必须从Ethernet帧里提取出来，而且由于目标IP地址也是一个广播地址，所以包的UDP或者TCP部分也需要提取出来，并在协议堆栈中向上传递。如果有一个进程在等待着数据，就必须传至应用层；如果没有进程在等待数据，数据就必须抛弃。显然，对那些“无辜”的、没有数据要求的主机，这个过程是对宝贵的主机处理时间的一个极大浪费，也会成为许多用户抱怨的对象。看起来，我们还需要另一个模型，来解决前两种模型的限制！

在组播通信模型中，我们需要两种新型地址：一个IP组播地址和一个Ethernet组播地址。其中，IP组播地址标识着一组接收者，它们要接收发给整个组的数据。由于所有IP包都封装在Ethernet帧内，所以还需要一个组播Ethernet地址。为使组播模型正常工作，主机应能同时接收单播和组播数据，这就意味着主机需要多个IP和Ethernet地址。一个单播IP和Ethernet地址用于单播通信，并用于零个或多个IP；而Ethernet组播地址用于组播通信。如果主机不准备接收组播通信，就需要零组播地址。总之，对于接收者希望加入的每个组播组，都需要一对组播地址，分别为IP和Ethernet地址。单播和组播地址之间的主要差异在于每个主机都有一个唯一的单播地址，组播地址则不然。例如，假设有五个主机希望接收发给A组的组播数据，那么所有主机都要监听发给同一个组播地址的数据，无论IP还是Ethernet地址。如图1-7所示，来自单播环境的通信量被大大地削减了。

采用组播模型后，我们获得的另一项能力是“动态组成员”。只有一个活动的应用正在运行，而且本身提出了数据要求，一个主机才应接收传给一个特定组播组的数据。主机应有能力根据自己的愿望，自由加入和离开组播组，从而排除“静态”分组的必要。为了最有效地利用现有带宽，要求路由器自行判断是否应将组播数据路由给组成员。所以，路由器必须随时注意动态组成员的信息，而且必须有相应的路由协议，对组播通信进行控制。

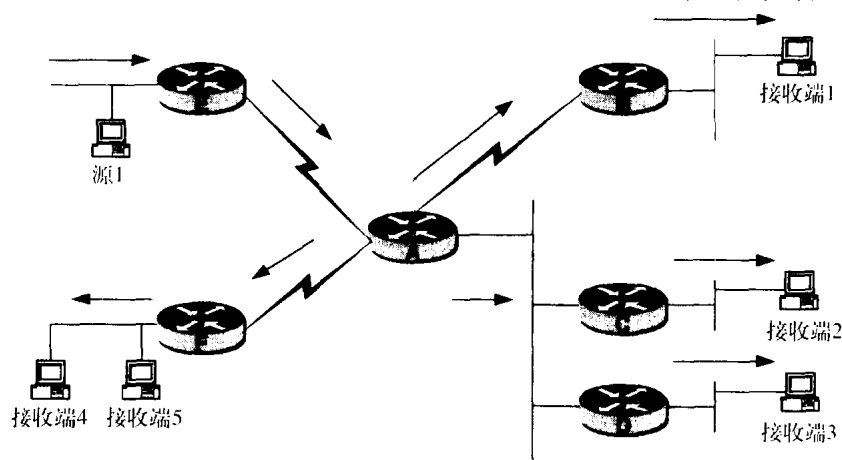


图1-7 组播通信模型

1.4 本书提要

上面提出的问题和解决方案正是本书剩下部分要全力讲述的内容。第2章详细解释了单播

和组播IP地址方案。在第3章“互联网组管理协议(IGMP)”里，将讨论IGMP协议，它在主机与路由器之间使用，负责报告动态组播组成员的信息。

第4章“Cisco组管理协议(CGMP)”讨论了一种专利性的Cisco协议，用于判断交换机处的组成员。使用CGMP协议，可将一个虚拟LAN(VLAN)内的组播通信限制到那些希望接收它的主机。

在第5章“距离向量组播路由协议”里，我们将开始组播路由协议的学习。首先介绍的是“距离向量组播路由协议”，即DVMRP，它在互联网组播干线(MBONE)上使用。Cisco并没有完整地实现DVMRP，但可与DVMRP实现沟通，将来自MBONE的路由交换到本地环境。

在第6章“与协议无关的组播——密集模式”和第7章“与协议无关的组播——分散模式”中，讨论了两种形式的“与协议无关组播”(Protocol Independent Multicast, PIM)协议。第一种叫做“PIM密集模式”(PIM-DM)。PIM-DM通常在局域网(LAN)环境中使用；而第二种形式“PIM分散模式”(PIM-SM)适用于广域网(WAN)。无论PIM-DM还是PIM-SM都已在Cisco路由器上实现。

与DVMRP和PIM网络的连接将在第8章“PIM-DVMRP网络”中讨论。由于MBONE运行的是DVMRP，而Cisco实现的是PIM，所以必须利用一种机制，实现DVMRP和PIM之间的沟通。

在第9章“组播支持命令”中，我们将讨论可随Cisco支持的任何一种组播路由协议使用的组播配置命令。而到第10章“资源预约协议”，我们将暂停组播路由协议的讨论，转为讨论一种在组播发送方到组播接收方的路径之间，用来对资源进行预约(或保留)的协议。“资源预约协议”(RSVP)是一种互联网控制协议，由组播接收者使用，针对来自单播或组播源的数据流，对一种特定的“服务质量”(QOS)进行请求。

在对Cisco支持协议进行讨论的每个章节，我们都会具体的网络环境下，向大家介绍用于配置、监视和调试协议的所有Cisco路由器命令，并阐述它们的用法。相信这正是本书最重要的价值所在！尽管与特定路由协议相关的信息已包括在相应的“注释请求”(RFC)文件内，Cisco公司本身也提供了与组播路由器配置有关的大量文档、资料，但我还是希望本书的解释与示范，能够帮助大家更容易地理解，并有效地填补其他方面的欠缺。

1.5 IP路由协议推荐读物

Cisco Router OSPF Design and Implementation Guide, William R. Parkhurst, CCIE #2969, McGraw-Hill

Advanced IP Routing in Cisco Networks, Terry Slattery, CCIE #1026, and Bill Burton, CCIE #1119, McGraw-Hill

Cisco TCP/IP Routing Professional Reference, Chris Lewis, McGraw-Hill

Internet Routing Architectures, Bassam Halabi, Cisco Press. This book is an excellent presentation of the Border Gateway Protocol (BGP)

第2章 IP协议地址

如欲设计与实现出真正“健壮”的IP网络，对单播及组播IP定址的完整理解是必不可少的一环。读者应掌握诸如“子网”和“变长子网掩码”(VLSM)这样的概念，以便IP定址计划能够有效利用分配给你的地址空间。IP单播路由协议、RIP、IGRP、EIGRP和OSPF的概念、运作与配置也需要完全地掌握，因为大多数组播路由协议都要依赖基层的单播路由配置。

2.1 IP地址格式

一个IP地址由32个二进制位构成，可表达成多种格式。路由器和计算机只有使用二进制数据，才能有效工作，所以对IP地址的保存与操作来讲，二进制是再自然不过的一种形式。一个典型的32位IP地址如下：

10011100000110100001111000111100

对路由器来说，这是再好不过的一种表示方法。但对我们来说，却不是最具吸引力的方法。所以，先让我们来观察一下二进制方法，再看看是否有更好的办法来表达相同的东西。最容易想到的一个办法就是将IP地址转换成十进制。上例的二进制数字换算成十进制后，为：

2 618 957 372

尽管更易阅读，但数位仍嫌过多，仍然很难掌握。另一个办法就是将二进制数字分隔为不同的片断，将每个片断都表达为十进制数字。二进制最自然的分段方法就是每8位为一段，成为我们熟知的“字节”，或者不那么熟悉的“八位元”(Octet，原来是远程通信的一个术语，两者可通用)。因此，让我们来划分一下前面的二进制数字，将其划分为8位一组(4个八位元)，再将每个组换算为十进制：

10011100	00011010	00011110	00111100
156	26	30	60

表2-1 IP地址范围

	低	高
二进制	00000000000000000000000000000000	11111111111111111111111111111111
十进制	0	4 294 967 295
点-十进制	0.0.0	255.255.255.255

我们不需要所有数字间的空格，所以让我们换成一个小数点，用它作为分隔符。现在，新的IP地址形式如下：

156.26.30.60

对于这种地址，我们认为它采用的是“点-十进制”记号法。总共可以有多少个IP地址呢？在表2-1中，总结了在我们的所有表示方案里，可以使用的IP地址的范围。

理论上，总共可以使用4 294 967 295个IP地址。但通过本章的学习，读者会知道实际能用的IP地址数要少得多。

可分类IP定址

为使一种协议能被路由，它的地址结构必须为十六进制。这意味着地址内至少应包含两个部分。对于IP地址，这就是网络部分与主机部分。所谓“主机”，是指网络通信的一个终点位置，比如计算机工作站、路由器接口或打印机等等。在一个网络中，可能包含了一个或多个主机。图2-1显示的是一个简单网络，由两个更小的网络构成，之间通过一个双端口的路由器连接。对于网络上的每个主机，包括路由器的接口，它的地址是由它的网络及主机编号指定的。

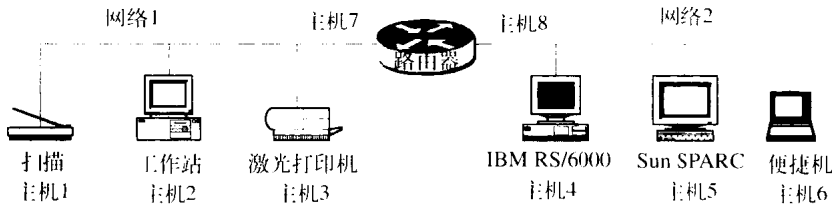


图2-1 分级定址

设计好IP地址方案后，我们决定创建五类IP地址，简单地命名为A类、B类、C类、D类和E类。对于头三类地址，IP定址方案的应用范围如下：

- A类地址：应用于数量最少的大型网络，其中包含了大量主机。
- B类地址：应用于中等数量的网络，包含了中等数量的主机。
- C类地址：应用于数量最多的网络，包含了少量主机。

而D类地址用于本书重点讲述的“组播”，E类地址则保留为将来使用，现在没什么用处。

为使三类IP地址能够应付不同规模的网络，要求每类地址的网络与主机部分都采用不同的大小。在图2-2中，分别展示了对于头四类IP地址来说，允许采用的网络与主机部分分配方式。

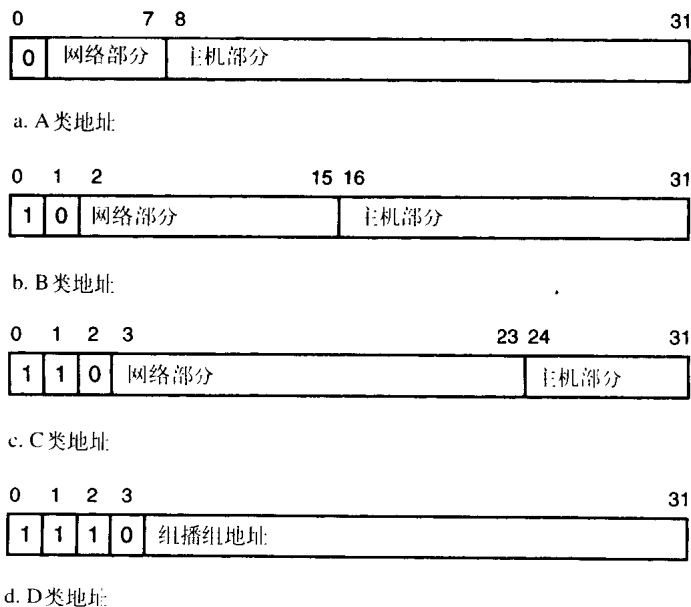


图2-2 可分类IP地址结构

其中，A类地址用8位来标识一个网络，而用24位来标识该网络内一个特定的主机。区分这类地址最明显的方法是观察头1位是否为0。

B类地址用16位来标识网络，而用另16位来标识主机。区分方法是观察头2位是否为10。

C类地址用24位来标识网络，用8位来标识主机，并将头3位固定为110。D类或组播地址则与单播地址完全不同。

A、B和C类地址都用于标识一个网络和网络内的一个主机，而D类组播地址用于标识通信双方的一组接收方和发送方。此外，组播发送和接收方可存在于任何网络。

观察每类的第一个八位元，就可发现这四类地址允许的取值范围为：

00000001(1)~01111110(126)	A类
10000000(128)~10111111(191)	B类
11000000(192)~11011111(223)	C类
11100000(224)~11101111(239)	D类

调查IP地址的第一个八位元，便可轻易辨别出网络是属于哪一类的。例如，前面使用的地址156.26.30.60是一个B类地址，因为第一个八位元在128和191之间。辨别网络类的另一个办法(不推荐)是将第一个八位元换算成二进制，观察头两个位的设置。例如，156等于二进制的10011100。头两个位是10，所以根据图2-1，该是一个B类地址。

那末，总共存在多少个A类、B类和C类网络呢？A类网络用7位来表示网络ID，所以总共只有128个A类网络。B类地址则要用到第一个八位元的6位，以及第二个八位元的8位，所以总共有16384个网络(64×256)，64由第一个八位元提供，256由第二个八位元提供。C类网络则要用到第一个八位元的5位，第二个八位元的8位以及第三个八位元的8位，所以总共可以有2 097 152个C类网络($32 \times 256 \times 256$)。D类地址则与网络无关，只与“组播组”有关。

A类、B类和C类地址均属单播地址。这三类中的每个IP地址都只能标识一个具体和唯一的Internet主机。D类地址则不同，它用于标识一组主机，这些主机均从属于一个特定的IP组播组。组播地址的范围在224.0.0.0~239.255.255.255之间(附录B总结了目前已经分配好的组播地址)。其中，对于224.0.0.0~224.0.0.255之间的地址，它们保留为路由协议以及其他低级拓扑查找及维护协议使用，比如用于进行网关查找和组成员报告等等。对于目标地址在这一范围之内任何一个组播数据报，除TTL以外，组播路由器均不得进行转发。

每个网络里可以存在多少个主机呢？A类网络用24位来标识一个主机，所以每个网络可以有1 677 216个主机；B类网络用16位来标识主机，所以能有65 536个主机；而C类网络仅用8位来标识不同的主机，所以最多可以有256个主机。表2-2对A、B和C类网络的容量进行了总结。

大家可能已经注意到了，表2-2中列出的主机数量总要比前面计算出来的少2个。原因在于有两个特殊的地址是不可分配给一个主机的，一个是全部为1的地址，另一个是全部为0的地址。全部为1的主机地址是一个特定网络的广播地址，而全为零的主机地址由一个主机用来临时性地标识自己(即“这个主机”)，除

非已经为它分配了一个IP地址。之所以只有126个A类网络，是由于网络0是不能使用的；而网络127保留作为打转(Loopback)地址使用，用于测试不同进程间的通信。如果主机将一个包发

表2-2 IP可分类地址容量

类	网络	主机
A	126	16 777 214
B	16 384	65 534
C	2 097 152	254

给127.0.0.1, 那么数据根本不会进入网络, 而是立即打转, 返回发送方主机。

下列IP地址范围保留作为内部使用, 不可在Internet上公开使用:

10.0.0.0~10.255.255.255
172.16.0.0~172.31.255.255
192.168.0.0!~192.168.255.255

这些都属于私有IP地址, 绝对不能在Internet上使用, 因为任何一个私有或内部互连网络都有可能用到它。使用这些地址时, 为了能与整个公共Internet通信, 必须先经过一道“网络地址翻译”的手续。

正如以下解释的理由, 某些情况下, 可分类IP地址分配方案的效率会变得相当低。假定我们为一所大学设计网络, 总共大致有1500个节点或者终端站。另外假定在可以预见的将来, 节点数量会在5年内超过5 000个。乍一看来, 似乎一个B类网络足以应付当前的网络需求, 另外也留下了足够的升级空间, 可满足未来规模扩大的需要。但对1 500多个节点来说(将来会有5 000多), 会成为一个非常大的Ethernet冲突域。如果希望将一个Ethernet网段上的节点数限制到100以下, 那么需要50个网络来完成我们的设计。无论决定使用哪个类别的IP地址(假定可以挑选自己希望的任何地址), 都会造成IP地址的巨大浪费, 如表2-3所示。

表2-3 低效率的IP地址设计

网络类	需要的地址	可用的地址	浪费的地址
A	100	16 777 214	16 777 114
B	100	65 534	65 434
C	100	254	154

现在, 将表2-3的每一项都乘以需要的50个网络。从结果可轻易看出, 无论选取哪一类地址, 都会浪费掉数量巨大的IP地址。此外, 假如准备接入Internet, 网络还必须向Internet的路由器广播50个网络。将数字乘以全世界的校园数量, 会得到多么可怕的一个数字! 这对Internet的路由表来说是毫不现实的, 根本无法管理。那么, 我们怎样来解决这个问题呢? 一句话, 用子网!

2.2 IP子网

对于上一节遇到的设计问题, 解决的方案就是将我们分配的无论哪一类IP地址都分割为多个更小的网络, 同时令每个网络只容纳较少的主机。为达此目的, 需要依赖来自IP地址“主机”部分的“借用”位, 将这些位转移到“网络”部分使用。那么, 具体该怎样做? 更重要的是, 路由器如何知道该为网络使用多少位, 以及为主机使用多少位呢? 答案就在于“子网掩码”的运用!

所谓“子网掩码”(Subnet Mask), 是指总共32位(比特)的一个数字, 它指出地址中的哪些位用于主机, 哪些位用于网络。掩码中的一个二进制1标志着IP地址中对应的位是一个“网络”位; 而一个0标志着对应的位是一个“主机”位。路由器做出判断的方法就是对IP地址和子网掩码执行一次“按位AND”或者“按位与”运算。

$$0 \text{ AND } 0 = 0 \quad 0 \text{ AND } 1 = 0$$

$$1 \text{ AND } 0 = 0 \quad 1 \text{ AND } 1 = 1$$

以下面这个“IP地址/子网掩码”对为例:

$$156.26.30.60/255.255.240.0$$

它们换算成二进制后变成:

地址: 10111100 00011010 00011110 00111100

掩码: 11111111 11111111 11110000 00000000

对上面两组号字执行AND运算, 得到:

$$10111100 \quad 00011010 \quad 00010000 \quad 00000000$$

将结果换算成点-十进制格式, 得到IP地址的网络部分为:

$$156.26.16.0$$

对子网掩码的一项限制在于: 掩码中的所有1必须是连贯的! 考虑到这一点, 我们可换用一种较简单的方法来表示掩码, 那就是单单指出掩码中有多少个1! 例如, 前例中的“IP地址/子网掩码”对可改写为156.26.30.60/20, 亦即掩码中有20个1。在图2-3中, 我们总结了用于非子网化网络的子网掩码。

Class A
11111111.00000000.00000000.00000000 255.0.0.0
Class B
11111111.11111111.00000000.00000000 255.255.0.0
Class C
11111111.11111111.11111111.00000000 255.255.255.0

图2-3 标准的IP子网掩码

表2-4 A类子网掩码

子网位数量	子网掩码	子网数	每个子网的主机数	主机总数
1	—	—	—	—
2	255.192.0.0	2	4194302	8388604
3	255.224.0.0	6	2097150	12582900
4	255.240.0.0	14	1048574	14680036
5	255.248.0.0	30	524286	15728580
6	255.252.0.0	62	262142	16252804