

Administering Cisco QoS in IP Networks

IP 网络的

Cisco QoS

管理

(美) Benoit Durand 等著
宁 科 等译



机械工业出版社
China Machine Press

SYNGRESS

Cisco 专业技术丛书

IP网络的Cisco QoS管理

(美) Benoit Durand 等著

宁 科 等译

前导工作室 审校



机械工业出版社
China Machine Press

本书全面深入地讲述了IP网络中的Cisco QoS管理和配置,对IP网络中的QoS一般性原理和技术以及最新发展动态也做了广泛的介绍。主要包括: Cisco IOS基本性能、QoS基本原理、流量分类、排队和阻塞回避技术、高级QoS管理、MPLS等。并在对每项技术的讲述之后,介绍了在Cisco路由器中的配置方法。

本书内容详尽且具权威性,将技术原理与实际应用有机地结合在一起。本书不仅适用于Cisco网络管理员,也适合广大网络技术工程和管理人员。

Benoit Durand, et al: Administering Cisco QoS in IP Networks.

Original English language edition published by Syngress Media, Inc.

Copyright © 2001 by Syngress Publishing, Inc. All rights reserved.

本书中文简体字版由美国Syngress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内容。

版权所有,侵权必究。

本书版权登记号: 图字: 01-2001-4411

图书在版编目(CIP)数据

IP网络的Cisco QoS管理/(美)杜兰(Durand, B.)等著;宁科等译. —北京:机械工业出版社, 2002.1

(Cisco专业技术丛书)

书名原文: Administering Cisco QoS in IP Networks

ISBN 7-111-09432-8

I. I… II. ①杜… ②宁… III. 计算机网络—管理 IV. TP393.07

中国版本图书馆CIP数据核字(2001)第071603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: 谢君英 张鸿斌

北京第二外国语学院印刷厂印刷·新华书店北京发行所发行

2002年1月第1版第1次印刷

787mm×1092mm 1/16·21.75印张

印数: 0 001-4 000册

定价: 36.00元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换

译者序

传统Internet是一种尽力而为传输的网络体系，随着Internet的迅猛发展，特别是以多媒体传输为代表的电信和广播电视网与IP网络的汇合，这种体系的弊端逐渐显露，已经无法满足人们对网络的性能和多样化的需要。因此，人们对服务质量(QoS)的要求与日俱增，QoS技术成为了当今网络技术一个新的热点。在网络厂商中，Cisco系统公司无疑是该行业的巨人，它领导和代表着网络技术的最前沿。因此，不管你是Cisco网络管理员，还是普通的网络技术工程或管理人员，都有必要了解Cisco的技术和产品。

国内尽管已出版了大量的关于QoS和Cisco网络技术的书，但全面介绍QoS和针对Cisco网络中QoS配置的书几乎没有。本书的好处在于，它不是一本简单的Cisco路由器管理手册，事实上，如果只是想要一本单纯的配置说明书，读者可以从Cisco路由器配套的用户手册和Cisco网站上获取。本书侧重于一般性原理的介绍，同时介绍了相关的各方面知识。在结构上，每介绍一个方面的内容，接着就进行配置实践，利于读者更牢固地掌握它。即使在介绍配置时，本书也不是简单地说教，而是结合比较实际的例子来说明，这有利于读者适应实际的情况，真正做到所学即所用。如果有条件，读者可以在Cisco设备上实践这些例子，这样会对它们有更好的理解。

本书由几位资深的Cisco技术专家共同编写。他们有的是网络技术的高级顾问，有的是长期从事网络工程设计的有着丰富经验的工程师。可以说，本书集众家之所长，既有很强的知识性，又有专家的经验之谈。

全书由宁科、王俊驹、赵娟、张靓、何志权、刘利、张节、王小东、王林、李小林、陈翔、刘晓明、黄学细、刘瑛、宋涛、于伯、张勇、李毅、孟伟等翻译，最后由前导工作室进行审校。由于专业水平及驾驭中英文的能力有限，译文中难免出现不当或欠妥之处，恳请广大读者批评和指正。

如果您在阅读中碰到了什么问题，请同我们工作室联系：qiandao@263.net。我们会尽力解决您的问题。

2001年9月15日

前 言

本书将讨论IP服务质量（QoS）以及如何将它应用到企业和服务供应商的环境中。本书将帮助你温习当前的Cisco网络设备（路由器、交换机等）中使用的路由协议和服务质量机制。本书将为你提供一些示例和练习，以帮助你在实际网络中获得这些能力，并为你提供详细的背景知识和必要的预备知识。

在今天的大多数企业中采用QoS可以确保商业应用成功地在某个底线的基础上将信息传递到正确的地方。QoS的目的就是满足要求苛刻的应用的需要，并与要求并不苛刻的应用共享网络资源。今天，采用Cisco的产品，QoS通过高效的算法最终可以有效地确保它所承诺的传输质量。

在过去的几年里，在网络设备中设置服务质量（QoS）的方法和协议的数量飞速地增长着。高级排队算法、流量整形、访问表过滤等，使得对QoS策略的选择非常艰难。无论网络是用于小型公司、企业还是Internet服务供应商，所有的网络都可以充分利用QoS的方方面面以获取最适宜的性能。

本书的作者是受过专业训练的技术、操作和项目管理专家，可以帮助现代企业和服务供应商来设计和部署网络，以更充分地创造商业价值。我们通过策略业务规划、网络设计以及实现服务来帮助客户在电子商务的市场环境中获取更高的性能和收益。

作者介绍

Benoit Durand (CCIE #5754、CCNA、CCDA、CCNP、CCDP)是Tivoli Systems (www.tivoli.com) 公司的中西部网络工程师，这家公司位于印第安纳州的印第安纳波利斯。Ben在维持自己的支持Cisco的网络的同时，为Tivoli的世界范围运作设计高档次的网络解决方案。他拥有在各种环境工作过的10年工作经验。为Tivoli工作之前，Ben为加拿大空军从事过许多部队里的高水准的工程项目，还在南斯拉夫的Kuwait和其他国际地区为维和部队设计广域网解决方案。最近从事的项目有Voice-over-ATM，虚拟专用网和广域网交换。Ben和他的妻子Dr. Chrity Snider住在加蓬州的Kingston。

Ron Fuller(EEIE #5851，CCNP-ATM，CCNP-Voice，CCNP-Security，CCDP，MCNE)是3X公司的高级系统工程师。他目前正为美国西部的3X公司客户进行网络设计和提供各种服务。他的专长包括了Cisco LAN/WAN设计、安全顾问和Novell网络设计。在过去的9年里，他曾做过另外两个网络咨询公司的工程师。Ron还参加编写了Syngress的《Building Cisco Remote Access Networks》(1-928994-13-X)。

Jerry Sommerville (CCIE #1293)是Callisma的高级顾问。他的工作背景包括网络管理、系统管理、系统集成、网络支持和策划、用户培训、程序自动化和程序分析。Jerry获得了东密歇根大学的计算机辅助设计和计算机辅助制造专业的工科硕士学位和德克萨斯A&M大学的工业技术与工程专业的工科学士学位。

James Placer (CCDP，CCNP Security，Voice Access，NNCDS，NNCSS，MCSE)是企业网络工作组 (www.ibsentg.com) (Enterprise Networking Group) 下的交互式办公系统公司的高级网络设计工程师。他设计、解决网络故障和构建大规模的LAN/WAN网络，这种网络主要基于Cisco网络系统和Nortel网络平台。James以前编写过Syngress CCNP的《Support Study Guide for Exam》(640-506)，有超过14年在网络与计算机系统领域工作的经验。

Kevin Davis (CCNA，MCSE，MCP+I)是Callisma的顾问，为服务提供商和公司客户在各种网络问题上做咨询。以前，Kevin是在新喀里多尼亚州Raleigh的国际网络服务公司的顾问，曾在Research Triangle Park (RTP)做过服务提供商。他在College Station，TX的德克萨斯A&M大学Dwight Look College of Engineering获得计算机工程专业学位。Kevin还参加编写了Syngress的《Building Cisco Remote Access Networks (1-928994-13-X)》，写了几本关于在网络环境和浏览器安全中的最小化计算机病毒的白皮书。他住在阿拉斯加州麦金利。

Paul Salas (CCNA，MCT，MCSE，Networks+)是Fleet Mortgage公司的高级网络工程师。他设计和管理着Fleet的Internet网络的下部构造，该构造由来自于各部件提供商的各种网络设备组成。他现在参与了设计高级Web网络解决方案。他还兼职为Microstaff公司做技术讲师，主要讲授Windows 2000系列的Microsoft官方教程。

Jaff Corcoran(CCNA，MCSE，CNE)是西门子公司网络部的顾问，在那里，他是福特汽车公

司高级网络技术小组的网络策划者。他负责为福特Intranet整体上的网络策划和网络新技术与应用测试。他特别关注VoIP、QoS、结构的高有效性和多播。他获得了Toledo大学的物理和应用数学专业的理工科硕士学位。他住在密歇根州的迪尔伯恩市。

Lisa Giebehaus (CCNA)是Callisma的高级顾问。她在电信领域工作了8年。她主要的工作是为大规模的公司设计和维护网络工程。在加入Callisma之前, Lisa是密歇根州底特律市的朗讯网络维护专业服务(以前是INS)的高级网络顾问。她在密歇根州立大学获得工程设计专业的理工学士学位。她住在密歇根州的Roal Oak。

Richard Hamilton是Callisma的高级顾问。他现在负责领导一个工程小组, 为服务提供商设计和实现复杂网络。Richard被业界公认为在MPLS、ATM和帧中继交换领域的专家。他在14年中为金融和网络服务提供商提供技术服务, 其中包括NatWest Bank、Fleet Bank、International Network Service、朗讯、Cisco、Sprint、WorldCom、South Western Bell、GTE、CapRock、CTC 通讯、ILD电信和Triton PCS。Richard还参加过Syngress出版的《Building Cisco Remote Access Networks》(1-928994-13-X)。他住在德克萨斯州的Flower Mound。

Robert Melancon是Callisma的顾问。他最近的项目是维护400多个LAN/WAN站点的TCP/IP实现设备, 包括帧中继(设备)、3COM集线器、Cisco Catalyst 1900系列交换机和Cisco 2500系列路由器。他还参加对xDSL与WAN技术和网络设备(包括Promatory与Pairgain的DSLAM和北电与朗讯的WAN交换机)的检验与论证工作。Robert还编写了许多学习软件与文档。他获得了Southern Methodist大学的工程学位, 现在住在得克萨斯州的达拉斯市。

目 录

译者序

前言

作者介绍

第1章 Cisco IOS特征回顾1

1.1 概述1

1.2 IP地址分类和有类IP路由1

1.2.1 A、B、C类地址3

1.2.2 D类地址5

1.2.3 RIPv1和IGRP6

1.3 可变长度子网掩码回顾11

1.4 标准访问控制表14

1.5 扩展访问控制表17

1.6 网络地址转换23

1.7 路由映射表27

1.8 小结29

1.9 常见问题解答29

第2章 EIGRP详细介绍31

2.1 概述31

2.2 IGRP基本概念回顾31

2.3 定义EIGRP四个基本构成部分37

2.3.1 建立协议相关模块37

2.3.2 建立邻居发现/恢复38

2.3.3 管理可靠传输协议38

2.3.4 建立DUAL有限状态机39

2.4 实现数据包类型39

2.5 配置基本的EIGRP49

2.6 配置高级 EIGRP59

2.7 重分布EIGRP和OSPF66

2.8 需要注意的问题74

2.8.1 在激活中阻塞75

2.8.2 自动汇总75

2.9 EIGRP的故障诊断76

2.9.1 在激活中阻塞路由的故障诊断76

2.9.2 自动汇总的故障诊断79

2.9.3 不在共同子网的故障诊断81

2.10 小结82

2.11 常见问题解答83

第3章 服务质量介绍84

3.1 概述84

3.2 定义服务质量84

3.2.1 服务质量是什么85

3.2.2 服务质量的应用85

3.3 理解拥塞管理87

3.4 定义普通排队概念88

3.4.1 漏桶88

3.4.2 尾部丢弃89

3.4.3 标记桶90

3.4.4 先进先出排队90

3.4.5 公平排队91

3.4.6 优先权排队92

3.4.7 自定义排队93

3.5 理解拥塞回避95

3.6 介绍监视和流量整形96

3.7 小结97

3.8 常见问题解答98

第4章 流量分类概述99

4.1 概述99

4.2 介绍服务类型99

4.2.1 ToS服务描述100

4.2.2 定义IP优先权的7个级别101

4.3 解释综合服务102

4.4 定义QoS的参数103

4.5 介绍资源预留协议104

4.5.1 RSVP流量类型104

4.5.2	RSVP的运行过程	105	5.11.2	给SNA流量以优先权	141
4.5.3	RSVP消息	105	5.12	小结	142
4.6	介绍区分服务	107	5.13	常见问题解答	143
4.7	扩展QoS: Cisco内容组网	111	第6章	排队和拥塞回避概述	144
4.7.1	应用识别分类: Cisco NBAR	111	6.1	概述	144
4.7.2	PDLM	114	6.2	使用FIFO排队	144
4.8	小结	115	6.2.1	高速与低速链路	145
4.9	常见问题解答	116	6.2.2	什么时候该使用FIFO	146
第5章	配置流量分类	118	6.3	使用优先权排队	146
5.1	概述	118	6.3.1	优先权排队是怎样工作的	146
5.2	配置基于策略的路由	118	6.3.2	为什么在网络上需要优先权排队	147
5.3	定义交付访问速率	121	6.4	使用自定义排队	148
5.4	标记和传输Web流量	122	6.4.1	自定义排队是怎样工作的	148
5.4.1	标记优先权位和传输Web流量	123	6.4.2	协议同自定义排队的交互	150
5.4.2	标记和传输多级别的CAR	124	6.4.3	为什么在网络中需要自定义排队	150
5.5	标记和限制ISP速率	125	6.5	使用加权公平排队	151
5.5.1	通过访问表来限制速率	126	6.5.1	加权公平排队是怎样工作的	151
5.5.2	通过MAC地址使用CAR来匹配和限制	127	6.5.2	权因子在什么地方起作用	152
5.5.3	监视CAR	128	6.5.3	资源预留协议	153
5.6	配置Cisco特快转发	129	6.5.4	为什么在网络中需要加权公平排队	153
5.6.1	启用CEF	130	6.6	使用随机早期检测	153
5.6.2	监视CEF	130	6.6.1	随机早期检测是怎样工作的	154
5.6.3	Cisco特快转发警告和错误的故障诊断	132	6.6.2	TCP/IP滑窗	154
5.7	配置基本的基于网络的应用识别	132	6.6.3	为什么网络中需要随机早期检测	156
5.7.1	建立NBAR类映射表	134	6.7	小结	156
5.7.2	建立策略映射表	134	6.8	常见问题解答	156
5.7.3	将策略映射表应用于接口	134	第7章	配置排队和拥塞回避	159
5.8	配置复杂的NBAR	134	7.1	概述	159
5.9	综合NBAR和基于类的加权公平排队	137	7.2	配置FIFO排队	159
5.9.1	创建类映射以标识NBAR	137	7.3	配置优先权排队	162
5.9.2	在策略映射表里配置类策略	137	7.3.1	启用优先权排队	163
5.9.3	将策略应用于接口	138	7.3.2	将优先权列表应用于接口	165
5.10	利用随机早期检测来配置NBAR	139	7.4	配置自定义排队	169
5.11	配置系统网络结构服务类型	140	7.4.1	启用自定义排队	169
5.11.1	将SNA CoS映射到IP ToS	141	7.4.2	将配置应用于接口	171
			7.5	配置加权公平排队	175
			7.6	配置随机早期检测	178

7.7 小结	182	8.9.2 什么时候我需要RTP包头压缩	210
7.8 常见问题解答	182	8.10 小结	211
第8章 高级QoS概述	184	8.11 常见问题解答	213
8.1 概述	184	第9章 配置高级QoS	215
8.2 使用资源预留协议	184	9.1 概述	215
8.2.1 什么是RSVP	185	9.2 启用、验证和故障诊断资源预留协议	215
8.2.2 RSVP不是什么	186	9.2.1 启用RSVP	216
8.2.3 RSVP是怎样工作的	186	9.2.2 验证RSVP配置	217
8.2.4 为什么在网络上需要RSVP	191	9.2.3 RSVP的故障诊断	219
8.2.5 使用RSVP的优点	191	9.3 启用、验证和故障诊断基于类别的加权	
8.2.6 使用RSVP的缺点	191	公平排队	219
8.3 使用基于类别的加权公平排队	191	9.3.1 启用CBWFQ	219
8.3.1 CBWFQ是怎样工作的	192	9.3.2 验证CBWFQ配置	223
8.3.2 为什么在网络上需要CBWFQ	193	9.3.3 CBWFQ的故障诊断	225
8.3.3 RSVP与CBWFQ的结合	195	9.4 配置、验证和故障诊断低延迟排队	225
8.4 使用低延迟排队	195	9.4.1 配置LLQ	225
8.4.1 LLQ是怎样工作的	196	9.4.2 验证LLQ配置	226
8.4.2 为什么在网络上需要LLQ	197	9.4.3 LLQ的故障诊断	227
8.5 使用加权随机早期检测	197	9.5 配置、验证和故障诊断加权随机早期	
8.5.1 WRED是怎样工作的	198	检测	227
8.5.2 为什么在网络上需要WRED	200	9.5.1 配置WRED	227
8.6 使用一般流量整形和帧中继流量整形	200	9.5.2 验证WRED配置	229
8.6.1 标记桶	200	9.5.3 WRED的故障诊断	234
8.6.2 GTS是怎样工作的	202	9.6 配置和验证一般流量整形和帧中继流	
8.6.3 为什么网络上需要GTS	202	量整形	235
8.6.4 FRTS是怎样工作的	203	9.6.1 配置GTS	235
8.6.5 为什么在网络里需要FRTS	204	9.6.2 验证GTS配置	236
8.7 在分布式模式中运行	205	9.6.3 配置FRTS	238
8.7.1 分布式模式里支持的特性	205	9.6.4 验证FRTS配置	240
8.7.2 IOS版本	206	9.7 理解分布式技术	242
8.7.3 操作的不同	206	9.7.1 DCEF	242
8.7.4 限制	206	9.7.2 DWRED	243
8.8 使用链路分段和交叉	207	9.8 配置、验证和故障诊断链路分段和交叉	244
8.8.1 LFI是怎样工作的	208	9.8.1 配置LFI	244
8.8.2 它在网络上有什么用?	209	9.8.2 验证LFI配置	245
8.9 理解RTP包头压缩	210	9.8.3 MLP的故障诊断	247
8.9.1 RTP包头压缩是怎样工作的	210	9.9 配置、验证和故障诊断RTP包头压缩	248

9.9.1 配置RTP包头压缩	248	11.4 配置EBGP邻居	286
9.9.2 验证RTP包头压缩	248	11.4.1 定义远程版本	287
9.9.3 RTP包头压缩的故障诊断	249	11.4.2 删除专用AS编号	288
9.10 小结	250	11.5 配置IBGP邻居	290
9.11 常见问题解答	251	11.6 配置反射器	291
第10章 边界网关协议概述	253	11.7 配置联合体	293
10.1 概述	253	11.8 权、MED、LOCAL PREF和其他高 级选项	296
10.2 BGP历史	253	11.8.1 路由映射表、匹配和设置的命令	297
10.2.1 外部网关协议	253	11.8.2 权属性	298
10.2.2 最初的实现	254	11.8.3 多出口鉴别属性	300
10.3 最大化BGP的功能	255	11.8.4 本地优先权属性	301
10.4 外部BGP和Internet	264	11.8.5 AS路径属性	303
10.5 BGP路径选择过程	270	11.8.6 起源属性	303
10.6 将BGP重分布到IGP中	274	11.8.7 下一跳属性	303
10.7 定义内部BGP、路由、反射器和 联合体	276	11.8.8 其他高级选项：BGP多协议扩展	304
10.8 高级BGP网络设计	278	11.9 小结	306
10.9 小结	280	11.10 常见问题解答	307
10.10 常见问题解答	282	第12章 多协议标签交换	308
第11章 配置边界网关协议	283	12.1 概述	308
11.1 概述	283	12.2 理解MPLS	308
11.2 相关的RFC	283	12.3 集成MPLS到QoS	318
11.3 启用BGP路由	284	12.4 标准化MPLS以获得最高效率	319
11.3.1 定义一个自治系统的BGP	284	12.5 使用流量工程控制MPLS流量	320
11.3.2 定义远程AS	285	12.6 集成MPLS和虚拟专用网	322
11.3.3 公众与专用自治系统比较	285	12.7 小结	335
11.3.4 启用BGP路由	286	12.8 常见问题解答	335

第1章 Cisco IOS特征回顾

本章主要内容：

- IP地址分类和有类IP路由
- 变长度子网掩码（VLSM）回顾
- 标准访问控制表（ACL）
- 扩展访问控制表（ACL）
- 网络地址转换（NAT）
- 路由映射表

1.1 概述

为了理解和配置Cisco IOS服务质量（QoS）机制，你有必要完全理解IP寻址、可变长度子网掩码和所有类型的访问表。本书中介绍的大部分QoS机制将与访问表有关，所以，即使是有经验的网络管理者，我也要强烈建议你密切关注本章的内容。

IP编址看起来是一件很简单的事，但是，如果你考虑在网络上的QoS，就需要注意寻址方案了。为过滤或分类基于源和目标地址的数据包而生成地址列表时，编址显得尤为重要。如果正确地配置了IP地址，你会发现在细节上区分数据包是件很容易的事。

本章还介绍了网络地址转换（NAT）。虽然NAT是一个有用的工具，但是当你为分类或排队数据包而匹配访问表时，它会造成很多困难。使用NAT之前有很多东西需要考虑，但是掌握了正确的信息以后，你就能做出最好的设计决策。

1.2 IP地址分类和有类IP路由

TCP/IP地址如同一个城市的街道地址一样，它定义了网络里面节点的位置。TCP/IP网络中的每个节点必须拥有一个地址，以便能加入到网络中。与街道地址一样，TCP/IP地址必须是惟一的。想像一下，如果两个房子有同一个街道地址，将会发生什么情况。这种情况会使邮差的工作变得困难，你将不可能收到邮件。这种基本的概念也适用于从最简单到最复杂的网络，如Internet。

要理解TCP/IP寻址，首先必须理解二进制概念。二进制数据位只有两种取值：1或0。1和0（分别对应开和关）是二元世界里仅有的两个可接收的数值。每8个比特构成一个字节或八元组。一个八元组就像这样：10111011。注意，八元组由八个比特或占位符构成。

在一个八元组里，每个比特或占位符都有一个特定的值。为了确定一个八元组对应的十进制数值，你首先必须建立二进制表。从最左边的位（值是128）开始，向右边移动，每个数刚好是它左边数的一半（见表1-1）。一旦你完成了将每个数分半的任务，你会注意到你已经有了8个数值。每个数由八元组里各位所代表。例如，如果一个八元组以1开始，随后全是0，那么这个

八元组的值就是128。如果一个八元组除了最右边位以外都是0，那么这个八元组值为1。

表1-1 二进制转换表

128	64	32	16	8	4	2	1
-----	----	----	----	---	---	---	---

要计算一个八元组的十进制值，你必须加上所有是1的二进制位的值。例如，如果最左边的两位是1，其他位都为0，也就是11000000，那么这个八元组的值是192（128+64）。如果八元组是10101010，那么它的值是170（128+32+8+2）。表1-2给出了几个二到十进制转换的例子。

在学习二到十进制转换的概念时，最好多进行转换练习，一直到你完全地适应它。大部分科学计算器有二到十进制转换的功能，这样就能确保计算练习的正确性。在继续学习更复杂的TCP/IP概念之前，这个基本技能必须掌握。

表1-2 二到十进制转换例子

二进制值	十进制值
11101010	234
00010010	18
00000101	5
01101001	105
11111111	255
00000000	0

TCP/IP的编址方案使用了4个八元组（以十进制格式表示）来标识一个主机或网络的地址，地址的每个八元组是一个0到255的二进制数值（请参见表1-3）。例如，考虑地址120.89.234.23，二进制表示为：01111000.01011001.11101010.00010111。

将TCP/IP地址在十进制与二进制之间来回转换的能力是一项重要的技能。例如，许多子网问题以十进制格式很难看出来，而研究二进制地址就会使问题明朗化。

表1-3 TCP/IP地址二进制转换例子

二进制值	TCP/IP 地址
00000101.10100100.11110000.01010010	5.164.240.82
01010010.01000100.01000100.01101001	82.68.68.105
11101010.01000100.01010010.11111111	234.68.82.255
00000000.00000000.11111111.11111111	0.0.255.255

TCP/IP地址由两个不同的部分构成。第一部分确定了主机所在的网络，第二部分确定主机的地址。子网掩码确定哪一部分地址是用来鉴别网络的，哪一部分是用来鉴别主机的。

确定地址是在本地子网还是在远程网络中的过程称作AND（与）操作。将TCP/IP地址与子网掩码进行二进制逻辑与的操作，从而确定网络标号。如果TCP/IP地址和子网掩码对应位都是1，那么子网值是1。如果对应位的地址、子网掩码或者两者都是0，那么值就是0（请参见表1-4）。

表1-4 与操作转换例子

变 量	二进制值	十进制值
IP地址	00000101.10100100.11110100.01010010	5.164.244.82
子网掩码	11111111.11111111.11111100.00000000	255.255.252.0
子网	00000101.10100100.11110100.00000000	5.164.244.0
IP地址	11001000.10011110.00010000.00100011	200.158.16.35
子网掩码	11111111.11111111.11111111.00000000	255.255.255.0
子网	11001000.10011110.00010000.00000000	200.158.16.0

TCP/IP寻址的实现要么使用有类编址，要么使用无类编址。两种方法都有各自的用途，理解它们之间的差别是很重要的。有类编址由三种不同的默认的网络地址类A到C构成。每个类的地址使用默认的子网掩码（请参见表1-5）。此外，还有两种地址类别：D类和E类，它们在普通的数据传输中没有使用。D类地址保留用做组播，本章的后面将要讨论它。E类地址保留用作开发和实验用。

有类编址是传统的IP寻址方式。在有类路由中，路由器在网络更新中不须通告掩码。这意味着每个路由器假定掩码与分配给它们的接口的一样（请参见表1-5）。另外，当将一个网络的地址通告到另外一个网络地址中时，路由器会自动汇总网络地址类的边界，而不是子网域。

有类路由要求在同一个网络下的所有子网有同样的子网掩码，这严重地限制了网络的灵活性。

无类编址通过包括路由通告的子网掩码弥补了这个缺陷。在无类编址中，没有指定默认的子网掩码。无类编址的另一个好处是它支持可变长度子网掩码（VLSM）。

在VLSM中，子网掩码是可以操作的，这样就能为网络提供更多的主机或网络。无类编址可以给你提供更多的灵活性，这样就能使网络地址表适应网络的逻辑或物理层网络的布局。在本章的VLSM部分我们将作更进一步的讨论。

1.2.1 A、B、C类地址

上面提到过，有类编址采用三种不同的地址类，A到C类。这些不同的地址类能够提供不同的主机和网络数目，而这经常与网络需求有关。

在一个有类的网络地址表的边界里，每个地址类使用地址的不同部分来确定网络以及主机地址使用的位。每个地址类都指定有一个默认的子网掩码（表1-5）。

表1-5 地址类范围和默认的子网掩码表

地址类	第一个八元组范围	子网掩码
A类	1-126*	255.0.0.0
B类	128-191	255.255.0.0
C类	192-223	255.255.255.0
D类	224-239	255.255.255.255
E类	240-254	255.255.255.255

* 地址127.0.0.1保留用作回送地址，不能用在网络中。

A类地址使用起始的8位来确定网络。B类地址使用起始的16位，C类地址使用起始的24位来确定网络（见图1-1）。

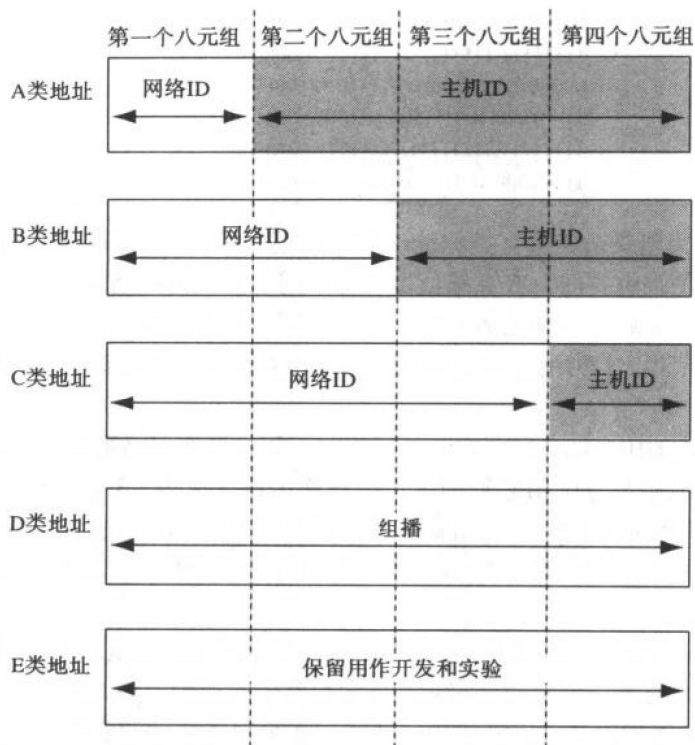


图1-1 IP地址八元组表示

127.0.0.1地址是普通TCP/IP地址的又一个特例。虽然整个地址范围被保留，但是只有127.0.0.1可以正常地使用。在一台计算机既作客户端又作服务器端时，此地址用作特殊用途。这个地址不能用作主机地址。

地址的分类和默认子网掩码的思想不能应用到可变长度子网掩码（VLSM）中去，VLSM将在本章后面讨论。但是，在继续学习更复杂的VLSM之前，它非常有助于我们理解有类编址。

虽然大部分的Internet的A、B、C类地址已经分配，但是从地址供应商那儿得到这些类的地址块也是件普通的事情。地址供应商已经将这些地址按照最节约的方式细分为子网。

Internet使用一种称作无类域间路由（CIDR，发音为“cider”）的技术。CIDR技术描述多个网络的聚合而不是子网，该聚合在Internet中通告为一个路由表入口（图1-2）。可以想像，Internet路由表是非常庞大的，而CIDR减小了它的规模。这一点很容易理解，CIDR减小了网络掩码大小，而子网化使子网掩码的大小增大了。CIDR也称为“超网联网”。

使用VLSM或CIDR的网络经常称为“斜杆X”网络。斜杆后的数字代表了子网掩码的位数，本章后面将讨论它。

不同类的网络要求不同的网络或主机部分的位数，每个网络类提供不同数目的主机以及网

络。A类地址范围提供了126个网络，每个网络可有16 777 216个主机。A类地址（或/8）最适合于拥有数量巨大的主机和少量网络的网络。B类地址可容纳65 384个网络，每个网络有同样数目的主机。C类地址提供了16 777 216个网络，每个网络可有254个主机。

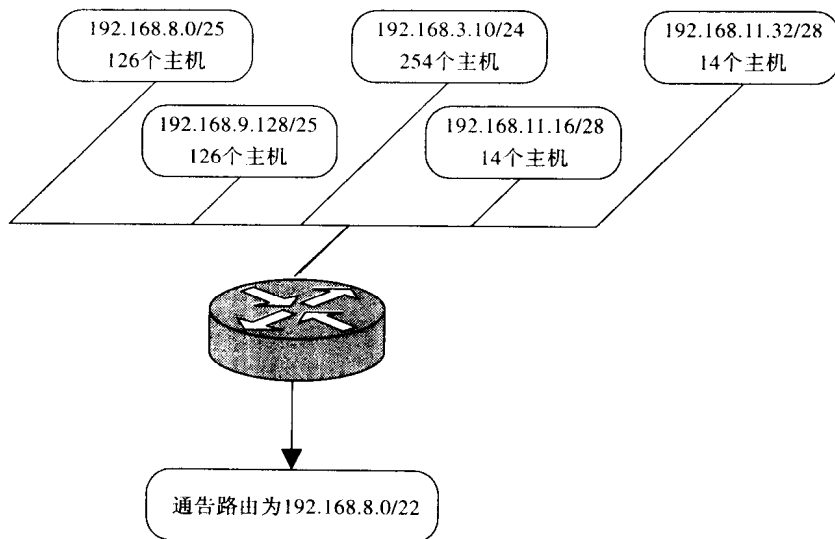


图1-2 无类域间路由（CIDR）示例

提示 记住，主机地址的数目直接关系到可使用主机的数目。例如，主机地址越多，网络地址就越少。必须考虑为网络容量的增长和可扩展性留出一定的空间。

1.2.2 D类地址

D类地址是保留用做组播的。组播指从一个服务器发送数据包到网络里的许多主机（点—多点）。图1-3说明了这个概念。与之相比较，广播是从一个服务器发送数据包到网络里所有的用户，而不是一部分（点—所有点）。

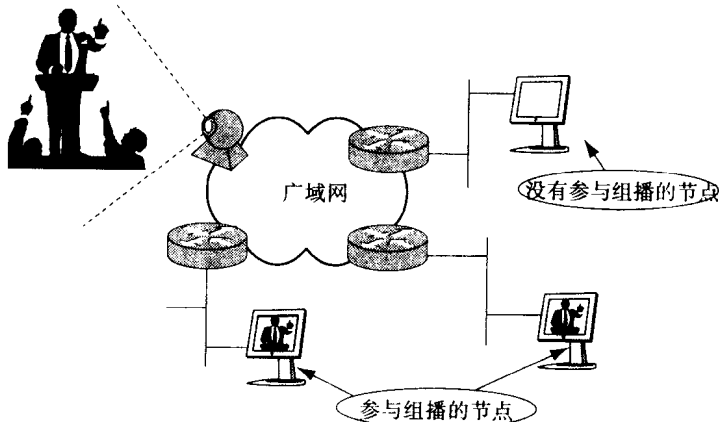


图1-3 组播示例

主机一旦加入了组播组,就开始接收数据流,而离开时就停止接收。除此之外,客户端能够在对网络组播造成的影响很少或没有的情况下动态地加入或离开组播组。

组播能够实时地传送音频或视频数据,这一点很受公司或网络服务供应商的欢迎。因为该技术不像广播那样涉及到整个网络,所以组播也被广泛用做计算机映像工具。

D类地址的范围是224.0.0.0到239.255.255.255。在二进制表示中,以1110开始的地址是组播地址,它允许后面的28位用作特定组播的组地址。

在以太网中,只有参与同一个组播组的设备才会侦听和处理目的地址是本组的数据包。在同一广播域而没有参与组播的计算机不会查看到数据包,但不会发出中断让CPU来处理。广播要求广播域的每台计算机不仅要查看而且要处理数据包,因此,组播比广播更有效。

组播与广播在某些方面类似。与广播类似,大部分组播提供无连接的传输,这意味着你接收来自组播服务器的数据包而不需要确认接收。广播和组播都不需要目的主机的确认。

虽然组播与广播有类似的一面,但某些特征却是组播所独有的。上面提到过,只有同一组的终端才会侦听和处理组播数据包。每一个组播使用不同的地址,这使终端能同时参加多个不同的组播。

Internet地址编号管理局(Internet Assigned Numbers Authority, IANA)给需要在Internet上进行组播应用的网络服务商分配组播地址。更特别的是,IANA还分配已注册的组播地址。

有两种基本的组播路由协议可选择:密集和稀疏模式。在网络里的大部分或全部路由器参与了组播的情况下使用密集模式。稀疏模式协议不认为所有的路由器参与了组播,而是使用加入消息来建立参与组播路由器的树结构。密集模式组播协议有两种最基本的类型:距离向量组播路由协议(Distance Vector Multicast Routing Protocol, DVMRP)和组播开放式最短路径优先(Multicast Open Shortest Path First, MOSPF)。DVMRP是第一个组播协议,它实际上来自于RIP路由协议。与RIP类似,DVMRP使用跳计数来作出判断,这使得它的规模不能扩大。MOSPF基于开放式最短路径优先(OSPF)协议,因此,在使用OSPF作为路由协议的环境中它能工作得很好。

基于核心树协议(Core-Based Tree Protocol, CBT)是目前最受欢迎的稀疏模式组播协议。CBT是一个开放的标准,它在RFC 2201中得到定义。作为一个稀疏模式协议,CBT建立一个分布树,它的网络开销很小。在CBT里,汇合点(rendezvous point)定义为所有其他树枝都能向它传输数据包的点。稀疏模式协议的扩展性比密集模式协议要好些。

还有另一种组播协议:协议无关组播协议(Protocol Independent Multicast, PIM)。它是一个新的协议,还没有被IETF标准化。PIM是一个特别的协议,它支持稀疏和密集两种模式协议。PIM与协议无关,这使得它有很大的灵活性。

1.2.3 RIPv1和IGRP

距离向量协议是最原始的动态路由协议。最普通的距离向量协议是RIPv1(Routing Information Protocol, 路由信息协议)和IGRP(Interior Gateway Routing Protocol, 内部网关路由协议)。虽然这两个协议非常相似,但是它们路由的方式完全不同。

我们注意到,RIPv1和IGRP都是有类协议,这意味着它们不随着TCP/IP地址发送子网掩码。