



IP 多播网络的设计与部署 (第1卷)

Developing IP Multicast Networks

Volume I

The definitive guide to designing and deploying
Cisco IP multicast networks

[美] Beau Williamson, CCIE #1346 著

徐鹏飞 郝记生 译

IP多播网络的设计与部署 (第1卷)

**Developing IP
Multicast Networks**
Volume I

Williamson, CCIE #1346 著
徐鹏飞 郝记生 译

人 民 邮 电 出 版 社
北 京

图书在版编目 (CIP) 数据

IP多播网络的设计与部署. 第1卷 / (美) 威廉姆森
(Williamson, B.) 著; 徐鹏飞, 郝记生译. — 北京:
人民邮电出版社, 2011.3
ISBN 978-7-115-24642-4

I. ①I… II. ①威… ②徐… ③郝… III. ①计算机
通信网 IV. ①TP393

中国版本图书馆CIP数据核字(2010)第250418号

版 权 声 明

Beau Williamson: Developing IP Multicast Networks (ISBN: 1578700779)

Copyright © 2000 Cisco Systems, Inc.

Authorized translation from the English language edition published by Cisco Press.

All rights reserved.

本书中文简体字版由美国 Cisco Press 授权人民邮电出版社出版。未经出版者书面许可, 对本书任何部分不得以任何方式复制或抄袭。

版权所有, 侵权必究。

IP 多播网络的设计与部署 (第 1 卷)

- ◆ 著 [美] Beau Williamson CCIE#1346
译 徐鹏飞 郝记生
责任编辑 傅道坤
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷
- ◆ 开本: 800×1000 1/16
印张: 32.5
字数: 768 千字 2011 年 3 月第 1 版
印数: 1—3 000 册 2011 年 3 月河北第 1 次印刷

著作权合同登记号 图字: 01-2010-3638 号

ISBN 978-7-115-24642-4

定价: 79.00 元

读者服务热线: (010)67132705 印装质量热线: (010)67129223

反盗版热线: (010)67171154

广告经营许可证: 京崇工商广字第 0021 号

内容提要

本书是 Cisco Press 出版的设计和部署 Cisco IP 多播网络的权威指南，也是目前为止市面上难得一见的多播网络的学习圣经。

本书总共分为 17 章和 1 个附录，主要内容有：IP 多播的基础理论、Internet 组管理协议、多媒体多播应用、距离向量多播路由协议、PIM 密集模式/稀疏模式的基本概念和具体使用、有核树、多播开放最短路径优先、PIM 集中点、连接 DVMRP 网络，以及园区网/NBMA 网络上的多播。此外，本书还讲解了多播流量工程和域间多播路由这两个高级主题。通过附录，读者可以对 PIM 数据包的格式有所了解。

本书适合即将部署多播网络或者已经部署了多播网络，但对其工作机制不甚明了的工程维护人员阅读，也适合备考 Cisco 认证考试的人员阅读。此外，Cisco 网络技术爱好者，以及各高校、科研院所的相关科研人员也可以从本书中受益。

关于作者

Beau Williamson 是 Cisco 系统公司 CTO 办公室的咨询工程师。他的专业领域是普通 IP 网络，当前主要研究 IP 多播。他在 1984 年从得克萨斯大学（达拉斯）获得了数学学士学位（计算机专业），然后进入计算机和网络技术领域工作，并长达 20 年之久。世界范围内的 Cisco 客户以及 Cisco 内部的工程师经常会拜访他，向他咨询设计、实施和调试 IP 多播网络服务的建议。Beau 也是 Cisco 内部 IP 多播培训课程的作者和开发人员，而且经常在（美国）国内和国外的 Cisco Networkers 与 Cisco Certified Internetwork Expert (CCIE) 会议上讲解与 IP 多播相关的主题。他与他的妻子、儿子现在住在得克萨斯州达拉斯市。在 IP 多播的工作之外，他还有广泛的爱好，这包括业余无线电、打高尔夫、做木工活，以及驾驶自己的飞机。

关于技术审稿人

Dino Farinacci 部署和实施网络协议已经有 18 年的时间了。他在距离向量和链路状态协议以及多播路由协议的实施方面具有广泛的经验。而在过去的 5 年里，他一直重点研究多播路由协议。Dino 目前在 Cisco 系统公司的多媒体组工作。他作为 IETF 的积极会员已经有 10 年之久，在 IETF，他参与了开放最短路径优先（OSPF）、协议无关多播（PIM）和不同的 IPng 候选协议的设计工作。他曾经做过短时间的 IPng 理事会成员，并帮助 IETF 解决了一个 IPng 方案的汇聚问题。当前，他专注于多播标记交换、基于策略的域间多播路由和可靠的多播协议。他是将 Cisco 多播路由器部署到 MBone 和多个本地 ISP 网络基础设施之上的首席工程师之一。

Kevin Almeroth 是圣巴巴拉加利福尼亚大学的助理教授。他的研究兴趣包括计算机网络和协议、多播通信、大规模多媒体系统以及性能评估。除了这些研究活动之外，Almeroth 博士还是几个 IETF 工作组的积极参与人员，并作为网络运维中心团队的一员来管理 NetworkWorld+Interop 的多播。他是 IP 多播倡导（Multicast Initiative）的高级技术专家，也是 Internet 2 多播工作组的主席。

Erick Mar 是 Cisco 系统公司的高级系统工程师，具有路由与交换方向的 CCIE 认证（CCIE #3882）。作为一名系统工程师，在过去的 7 年里，他在多家不同的网络生产厂商工作过，并为世界 500 强公司提供过设计和实施支持。Erick 从圣克拉拉大学获得了 MBA 学位，从州立圣弗朗西斯科大学获得了商业管理的学士学位。

Bob Quinn 是 Stardust.com 的高级技术专家，在那里，他为 IP 多播倡导和 QoS 论坛编写了白皮书，并跟踪 IETF 在这些方面的开发进展。他是广受专注的《Windows Sockets Network Programming》（Addison-Wesley）一书的主要作者，还是 WinSock 2 编辑委员会的主席，该委员会负责监督 WinSock 应用编程接口（API）的新开发和相关问题。可以通过 rcq@stardust.com 与他取得联系。

致 谢

如果没有众人的支持，本书不可能出版。尽管我不能一一列出所有人，但是我仍然深深地感激他们。我要特别感谢 Dino Farinacci、Liming Wei 以及他们的来自 Cisco IP 多播开发组的经理 Achutha Rao。Dino 和 Liming 在 IP 多播主题的异议和探讨等方面所表现出来的支持和耐心远远超出了他们应有的职责。我还要特别感谢我的开发编辑 Kathy Trace，除了在我需要征求他人意见时而向她倾诉之外，她还要忍受我延期交稿，以及我对稿件的奇怪格式的要求。此外，还要感谢我的技术审稿人员 Dino Farinacci、Manoj Leelanivas、Kevin Almeroth、Erick Mar 和 Bob Quinn，他们给本书的技术内容做出了卓越的贡献。

最后，我还要感谢我的妻子和儿子，感谢他们的支持和耐心，以及每当我的 PC 上的字处理软件产生意外结果而导致我对着 PC 大声吼叫时，他们所表现出的包容。

目 录

第 1 部分 IP 多播的基本原理

第 1 章 IP 多播介绍	3
1.1 IP 多播简史	4
1.2 IP 多播的优点	5
1.2.1 带宽	6
1.2.2 服务器负载	7
1.2.3 网络负载	8
1.3 IP 多播的缺点	9
1.3.1 不可靠的信息包交付	9
1.3.2 数据包复制	10
1.3.3 网络阻塞	11
1.4 多播应用	12
1.4.1 多媒体会议	13
1.4.2 数据分发	13
1.4.3 实时数据多播	14
1.4.4 游戏和仿真	14
1.5 Internet 的多播主干 (MBone)	15
1.5.1 MBone 会话	16
1.5.2 MBone 历史	16
1.5.3 今天的 MBone 体系结构	18
1.5.4 明天的 MBone 体系结构	20
1.6 小结	20
第 2 章 多播基础	23
2.1 多播地址	23
2.1.1 IP D 类地址	24

2.1.2	已分配的多播地址	24
2.1.3	管理范围的多播地址	27
2.2	多播 MAC 地址	27
2.2.1	以太网多播 MAC 地址映射	28
2.2.2	FDDI 多播 MAC 地址映射	29
2.2.3	令牌环网多播 MAC 地址映射	30
2.3	多播分发树	32
2.3.1	源树	32
2.3.2	共享树	33
2.4	多播转发	37
2.4.1	逆向路径转发	37
2.4.2	多播转发缓存	39
2.4.3	TTL 阈值	39
2.4.4	管理范围边界	41
2.5	多播路由协议分类	42
2.5.1	密集协议模式	43
2.5.2	稀疏模式协议	45
2.5.3	链路状态协议	48
2.6	小结	48
第 3 章	Internet 组管理协议	51
3.1	IGMP 版本 1	52
3.1.1	IGMPv1 消息格式	52
3.1.2	IGMPv1 查询—响应过程	53
3.1.3	报告抑制机制	55
3.1.4	IGMPv1 查询器	55
3.1.5	IGMPv1 加入过程	55
3.1.6	IGMPv1 离开过程	56
3.2	IGMP 版本 2	58
3.2.1	IGMPv2 消息格式	58
3.2.2	查询—响应调整	60
3.2.3	IGMPv2 离开组消息	62
3.2.4	IGMPv2 指定组查询信息	62
3.2.5	IGMPv2 离开过程	62
3.2.6	查询器选举过程	65

3.2.7 早期的 IGMPv2 实现	65
3.3 IGMPv1—IGMPv2 互操作性	66
3.3.1 IGMPv2 主机/IGMPv1 路由器互操作性	66
3.3.2 IGMPv1 主机/IGMPv2 路由器	67
3.3.3 混合的 IGMPv1 和 IGMPv2 路由器的互操作性	68
3.4 IGMPv3 的可能性	68
3.5 小结	70
第 4 章 多媒体多播应用	73
4.1 实时协议	73
4.1.1 在音频会议中使用 RTP 和 RTCP: 一个案例	74
4.1.2 RTP 控制协议	75
4.2 会话通告协议	76
4.2.1 SAP 通告	76
4.2.2 SAP 带宽限制	77
4.3 会话描述协议	77
4.3.1 SDP 信息格式	77
4.3.2 SDP 信息描述类型	78
4.3.3 SDP 会话描述示例	79
4.4 MBone 多媒体会议应用	80
4.4.1 SDR——会话目录工具	81
4.4.2 VAT——BMone 多媒体音频工具	84
4.4.3 VIC——MBone 多媒体视频工具	86
4.4.4 WB——共享的白板工具	91
4.5 小结	92

第 2 部分 多播路由协议概述

第 5 章 距离向量多播路由协议	97
5.1 DVMRP 邻居发现	97
5.2 DVMRP 路由表	98
5.3 交换 DVMRP 路由报告	100
5.4 DVMRP 截断广播树	102
5.5 DVMRP 多播转发	106

5.6	DVMRP 剪枝	107
5.7	DVMRP 嫁接	111
5.8	DVMRP 可扩展性	113
5.9	小结	114
第 6 章	PIM 密集模式	117
6.1	PIM 邻居发现	118
6.1.1	PIM Hello 消息	118
6.1.2	PIM-DM 源分发树	120
6.2	PIM-DM 多播转发	121
6.3	PIM-DM 剪枝	123
6.3.1	剪枝否决	125
6.3.2	剪枝延迟积累	127
6.4	PIM-DM 断言	128
6.5	PIM-DM 嫁接	130
6.6	将来的 PIM 增强——状态刷新	132
6.7	PIM-DM 可扩展性	133
6.8	小结	133
第 7 章	PIM 稀疏模式	135
7.1	显式加入模型	136
7.2	PIM-SM 共享树	136
7.2.1	共享树加入	137
7.2.2	共享树剪枝	140
7.3	PIM-SM 最短路径树	142
7.3.1	最短路径树加入	143
7.3.2	最短路径树剪枝	145
7.4	PIM 加入/剪枝消息	147
7.5	PIM-SM 状态刷新	148
7.6	源注册	149
7.6.1	PIM 注册消息	150
7.6.2	PIM 注册停止消息	151
7.6.3	源注册示例	151
7.7	最短路径树切换	154
7.7.1	SPT 切换示例	154

7.7.2 从共享树上剪枝源	156
7.8 PIM-SM 指定路由器	158
7.8.1 指定路由器的作用	158
7.8.2 指定路由器失效	159
7.9 RP 发现	159
7.10 PIM-SM 适用性/可扩展性	159
7.11 小结	160
第 8 章 有核树	163
8.1 CBT 概述	163
8.2 加入共享树	166
8.2.1 暂态	167
8.2.2 转发缓存	167
8.2.3 多播转发	168
8.2.4 非成员的发送	168
8.3 CBT 状态维护	170
8.3.1 Echo-Request 消息	170
8.3.2 Echo-Response 消息	170
8.3.3 Flush-Tree 消息	171
8.4 剪枝共享树	171
8.5 CBT 指定路由器	172
8.5.1 CBT Hello 协议	172
8.5.2 DR 加入代理	173
8.6 核心路由器发现	174
8.7 CBT 版本 3	174
8.8 CBT 适用性/可扩展性	175
8.9 小结	175
第 9 章 多播开放最短路径优先	177
9.1 MOSPF 区域内多播路由	178
9.1.1 组成员关系链路状态通告	178
9.1.2 区域内最短路径树	180
9.1.3 MOSPF 转发缓存	181
9.2 MOSPF 区间多播路由	183
9.2.1 多播边界路由器	183

9.2.2 区间组成员关系汇总	183
9.2.3 通配符多播接收者	185
9.3 MOSPF AS 间多播路由	188
9.4 MOSPF 适用性/可扩展性	190
9.5 小结	192

第 3 部分 Cisco 多播网络的实施

第 10 章 使用 PIM 密集模式	197
10.1 配置 PIM-DM	197
10.2 PIM-DM 状态规则	199
10.2.1 PIM-DM (*, G) 状态规则	199
10.2.2 PIM-DM (S, G) 状态规则	200
10.2.3 PIM-DM 状态维护规则	202
10.3 PIM-DM 状态条目	203
10.3.1 PIM-DM 状态标记	203
10.3.2 PIM-DM 状态示例	204
10.4 PIM 转发	207
10.5 PIM-DM 泛洪	208
10.6 PIM-DM 剪枝	209
10.7 密集模式嫁接	212
10.8 新的 PIM 邻居邻接性	215
10.9 小结	218
第 11 章 使用 PIM 稀疏模式	221
11.1 配置 PIM-SM	221
11.2 PIM-SM 状态规则	222
11.2.1 PIM-SM (*, G) 状态规则	223
11.2.2 PIM-SM (S, G) 状态规则	224
11.2.3 PIM-SM 出口规则	224
11.2.4 PIM-SM 外出接口计时器	226
11.2.5 PIM-SM 状态维护规则	226
11.2.6 特殊的 PIM-SM (S, G) RP 位状态规则	227
11.3 PIM-SM 状态条目	231

11.4 加入共享树	234
11.5 PIM 注册过程	238
11.5.1 接收者先加入	238
11.5.2 源先注册	246
11.5.3 沿着 SPT 的接收者	254
11.6 SPT 切换	257
11.6.1 超过 SPT 阈值	257
11.6.2 SPT 切换过程	258
11.6.3 SPT-Switchback 过程	266
11.7 剪枝	266
11.7.1 剪枝共享树	266
11.7.2 剪枝源树	271
11.8 PIM-SM 的特殊情况	277
11.8.1 意外数据的到达	278
11.8.2 单臂 RP	280
11.8.3 调头路由器	284
11.8.4 代理加入计时器	285
11.9 小结	289
第 12 章 PIM 集中点	291
12.1 Auto-RP	291
12.1.1 Auto-RP 概述	292
12.1.2 配置 Auto-RP 的候选 RP	294
12.1.3 配置 Auto-RP 的映射代理	294
12.1.4 使用多个映射代理以提供冗余	295
12.1.5 使用多个 RP 以提供冗余	296
12.1.6 稀疏—密集模式的产生	296
12.1.7 简单的 Auto-RP 配置	298
12.1.8 Auto-RP 网络中的 RP 切换	300
12.1.9 抑制 Auto-RP 消息	300
12.1.10 防止候选 RP 的欺骗	303
12.2 PIMv2 Bootstrap 路由器机制	304
12.2.1 PIMv2 Bootstrap 路由器概述	304
12.2.2 配置 PIMv2 候选 RP	306
12.2.3 配置 PIMv2 候选 BSR	307

12.2.4	用多个 RP 实现冗余和 RP 负载均衡	308
12.2.5	RP 选择——RP 哈希算法	308
12.2.6	使用多个候选 BSR 以实现冗余	309
12.2.7	PIMv2 BSR 网络中的 RP 切换	311
12.2.8	抑制 BSR 消息	312
12.3	RP 的放置和调节	313
12.3.1	选择 RP 的安放位置	313
12.3.2	RP 上的资源需求	314
12.3.3	强迫组停留在密集模式	317
12.3.4	强迫组停留在稀疏模式	324
12.4	小结	325
第 13 章	连接 DVMRP 网络	329
13.1	Cisco DVMRP 的互操作性	329
13.1.1	启用 DVMRP 互操作性	330
13.1.2	PIM-DVMRP 交互	331
13.2	DVMRP 路由交换	342
13.2.1	通告已连接的路由（默认行为）	344
13.2.2	DVMRP 路由的有类汇总	345
13.2.3	控制 DVMRP 路由通告	347
13.2.4	控制 DVMRP 路由接受	353
13.2.5	调整默认的 DVMRP 距离	354
13.2.6	调整 DVMRP 度量	354
13.2.7	特殊的 MBone 特性	355
13.3	PIM-DVMRP 边界问题	357
13.3.1	单播-多播一致性	357
13.3.2	PIM-SM 问题	362
13.4	DVMRP 网络连接示例	365
13.4.1	物理上一致的网络	365
13.4.2	单独的 MBone 路由器	368
13.5	调试技巧	371
13.5.1	检验 DVMRP 隧道状态	372
13.5.2	检查 DVMRP 路由交换	374
13.6	小结	377

第 4 部分 第二层的多播

第 14 章 园区网上的多播	381
14.1 地平协会	381
14.2 LAN 交换机的特征	382
14.2.1 广播/多播泛洪	383
14.2.2 抑制多播泛洪	384
14.3 IGMP 窃听	385
14.3.1 使用 IGMP 窃听加入一个组	385
14.3.2 IGMP 窃听的性能影响	387
14.3.3 使用 IGMP 窃听离开组	390
14.3.4 用 IGMP 窃听维护组	393
14.3.5 IGMP 窃听和只能发送的源	395
14.3.6 用 IGMP 窃听来检测路由器	396
14.3.7 IGMP 窃听小结	397
14.4 Cisco 组管理协议	398
14.4.1 CGMP 消息	398
14.4.2 用 CGMP 加入组	400
14.4.3 用 CGMP 维护组	401
14.4.4 用 CGMP 离开组	402
14.4.5 CGMP 的本地离开处理	403
14.4.6 CGMP 的性能影响	405
14.4.7 CGMP 和只发送的源	405
14.4.8 用 CGMP 检测路由器	405
14.4.9 CGMP 小结	406
14.5 LAN 交换的其他问题	406
14.5.1 IGMP 离开延迟问题	407
14.5.2 交换机之间的链路问题	408
14.5.3 路由器核心交换机问题	409
14.6 小结	411
第 15 章 NBMA 网络上的多播	413
15.1 传统的 NBMA 网络	413

15.2 传统 NBMA 网上的多播	415
15.2.1 伪广播	417
15.2.2 PIM 和部分互连的 NBMA 网络	419
15.2.3 PIM NBMA 模式	421
15.2.4 NBMA 网络上的 Auto-RP	424
15.3 ATM NBMA 网络云上的多播	427
15.3.1 ATM 点到多点广播虚电路	428
15.3.2 每组的 ATM 点到多点虚电路	429
15.3.3 PIM 多点信令	431
15.3.4 限制 PIM 多点虚电路的数量	431
15.3.5 调试 ATM 点到多点虚电路	432
15.4 ATM 网络上的传统 IP	435
15.5 小结	437

第 5 部分 高级多播主题

第 16 章 多播流量工程	441
16.1 控制多播带宽的使用	441
16.1.1 使用速率限制的带宽控制	442
16.1.2 用作用区域进行带宽控制	445
16.1.3 配置作用区域	448
16.1.4 作用区域和 BSR	453
16.2 控制多播流量路径	454
16.2.1 其他 RPF 信息源	454
16.2.2 流量工程示例	461
16.3 使用 GRE 隧道进行多播负载分割	463
16.3.1 配置多播负载分割	464
16.3.2 进程交换与快速交换的对比	465
16.4 广播到多播的流量转换	466
16.5 小结	470
第 17 章 域间多播路由	473
17.1 域间多播路由中的问题	473
17.2 多协议 BGP	475